# ALGORITHMIC INFORMATION, FRACTAL GEOMETRY, AND DISTRIBUTED DYNAMICS

**By**

**NEIL J. LUTZ**

**A dissertation submitted to the**

**Graduate School–New Brunswick**

**Rutgers, The State University of New Jersey**

**In partial fulfillment of the requirements**

**For the degree of**

**Doctor of Philosophy**

**Graduate Program in Computer Science**

**Written under the direction of**

**Rebecca N. Wright**

**And approved by**

_____

_____

_____

_____

**New Brunswick, New Jersey**

**October, 2017**

## ABSTRACT OF THE DISSERTATION

# Algorithmic Information, Fractal Geometry, and Distributed Dynamics

**By NEIL J. LUTZ**

**Dissertation Director:**

**Rebecca N. Wright**

This dissertation applies two distinct algorithmic perspectives to questions in the field of fractal geometry and dynamics.

In Part I, we establish connections between algorithmic information theory and classical fractal geometry. Working in Euclidean spaces, we characterize Hausdorff and packing dimensions in terms of relativized Kolmogorov complexity, and we develop conditional dimensions. These tools give rise to new dimensional bounding techniques, which we apply to problems in fractal geometry. Most significantly, we prove that a classical dimension bound for intersections of Borel sets holds for arbitrary sets, and we give a new lower bound on the Hausdorff dimension of generalized Furstenberg sets.

In Part II, we use ideas from distributed computing and game theory to study dynamic and decentralized environments in which computational nodes interact strategically and with limited information. We exhibit a general non-convergence result for a broad class of dynamics in asynchronous settings. For uncoupled game dynamics, in which preferences are private inputs, we give new bounds on the recall necessary for self stabilization to an equilibrium.

# Acknowledgments

I am very thankful for the guidance and support of my advisor, Rebecca Wright. Many thanks are also due to Michael Saks, who supervised my early research, and to my thesis committee members Eric Allender, Shubhangi Saraf, and Mark Braverman. Financially, the work in this dissertation was supported in part by NSF grants 1101690 and 1445755.

Much of this work was completed away from Rutgers. I am grateful to Michael Schapira and his group, especially Doron Zarchy, for welcoming me during two extended visits to the Hebrew University of Jerusalem. I am also grateful to Aaron Jaggard for hosting me at the U.S. Naval Research Laboratory, and to the theory group at Iowa State University, including Don Stull, for making the time I spent in Ames even more productive and enjoyable.

Most of all, I want to thank my parents, Robyn and Jack, and my sister, Laura, who have helped me in more ways than I can count. None of this would have been possible without their extraordinary understanding, love, and encouragement.

## Acknowledgement of Previous Publications

Most of Chapter 2 and all of Chapters 3 and 4 are based on a paper with Jack H. Lutz; a preliminary version of that work appears in *Proceedings of the 34th International Symposium on Theoretical Aspects of Computer Science (STACS 2017)* [66], and an extended version is under review for journal publication. Chapter 5 is based on a paper with D. M. Stull; a preliminary version of that work appears in *Proceedings of the 14th Annual Conference on Theory and Applications of Models of Computation (TAMC 2017)* [70], and an extended version is under review for journal publication. Chapter 6 is based on a single-author paper that is currently under review for conference proceedings publication [69]. Chapter 7 is

based on parts of a joint paper with Aaron D. Jaggard, Michael Schapira, and Rebecca N. Wright [52], which is to appear in *ACM Transactions on Economics and Computation.* I am not an author on the preliminary version of that work [53], in which the main theorem, a non-convergence result, is shown via a complex axiomatic framework. My contributions to the extended version include a direct proof that yields a slightly stronger non-convergence result (Theorem 43). Chapter 8 is based on another paper with Jaggard, Schapira, and Wright, which appears in *Proceedings of the 7th International Symposium on Algorithmic Game Theory (SAGT 2014)* [51]. The introduction (Chapter 1) and conclusion (Chapter 9) also contain text from some of these papers.

# Table of Contents

# Chapter 1

# Introduction

For studying the intricate, irregular patterns exhibited by real-world systems—both natural and computational—fractal geometry and dynamical system theory constitute a powerful framework that has yielded advances throughout the sciences [89, 32]. Fractal geometry provides tools for analyzing sets that have non-trivial structure at every scale, for which "standard" geometry is inadequate. Dynamical system theory considers the *trajectories* over time of a system's state, under iterative (or continuous) application of some *transition function*. Even simple transition functions often induce complex dynamics that map out highly irregular, fragmented sets—i.e., *fractal* sets. In this dissertation, we use ideas and methods from theoretical computer science to examine both fractal geometry and dynamic behavior.

First, we view the points in fractal sets as infinite data objects with varying degrees of *incompressibility* or *algorithmic information density*, which allows us to analyze those sets using tools from algorithmic information theory. Second, we regard certain classes of dynamical systems as distributed computational systems, in which the system state reflects the local states of many computational (and potentially *strategic*) agents, whose individual behaviors induce a global transition function. Drawing on techniques from distributed computing and game theory, we show how environmental uncertainty can prevent the systems from *stabilizing* to a single state.

*Fractal dimensions* are critical tools for studying fractal sets and dynamical systems [89]. Classical fractal dimensions, among which *Hausdorff dimension* [48] is the first and most important, emerged from geometric measure theory nearly a century ago as refinements of

notions of measure that enable quantitative classifications of sets with zero volume [33]. In 2000, J. H. Lutz [64] showed that Hausdorff dimension can be simply characterized using betting strategies called *gales* and that this characterization can be *effectivized* in order to quantitatively classify non-random infinite data objects according to their level of algorithmic "unpredictability." The same dimension was later characterized by Mayordomo [79] using *Kolmogorov complexity* [57], a measure of algorithmic information. This *effective Hausdorff dimension* and other, related *algorithmic dimensions* have been applied to multiple areas of computer science and have proven especially useful in algorithmic information theory [39, 83, 28].

In Part I, we exploit the connection between algorithmic and classical dimensions in the other direction, i.e., to apply algorithmic information theoretic methods and intuition to classical fractal geometry in Euclidean spaces. Our primary tool in this effort is a *point-to-set principle*, Theorem 2, in which we use relativization to characterize the classical Hausdorff dimension of any set in $\mathbb{R}^n$ in terms of the effective Hausdorff dimensions of its individual points. In Chapter 2, we define these concepts, present our point-to-set principle (along with its counterpart for *packing dimension*), and show that this principle gives rise to a new, pointwise technique for dimensional lower bounds. As a second major tool for algorithmic information theoretic investigations of classical fractal geometry, we develop *conditional dimensions* in Chapter 3, thereby filling a gap in effective dimensional theory. These dimensions are analogous to conditional Kolmogorov complexity and conditional entropy, and they allow us to analyze the part of a point's dimension that is "independent" of another point. We show that our conditional dimensions are robust, obey useful chain rules, and interact correctly with mutual dimensions.

Our first application of the tools developed in Chapters 2 and 3 is to planar *Kakeya sets* in Chapter 4. A Kakeya set in $\mathbb{R}^n$ is one that contains a unit-length line segment in every direction. Besicovitch [8, 9] showed that such a set can have zero volume, and the famous *Kakeya conjecture* asserts that every Kakeya set in $\mathbb{R}^n$ must have Hausdorff dimension $n$.

We apply our point-to-set principle to give a new, algorithmic information theoretic proof of Davies's 1971 theorem [22] stating that this conjecture holds for $n = 2$. Our proof is very different from Davies's proof and entails showing that if $a \in \mathbb{R}$ is algorithmically random, then for all $b \in \mathbb{R}$ and almost all $x \in \mathbb{R}$, the point $(x, ax + b)$ has effective Hausdorff dimension 2.

In Chapter 5, we study a more general class of sets and use the point-to-set principle and conditional dimension to derive the first *new* theorem in classical fractal geometry to be proven using algorithmic dimensions, an improved lower bound on the Hausdorff dimension of *generalized Furstenberg sets*. Sets of Furstenberg type generalize Kakeya sets in $\mathbb{R}^2$; instead of containing segments in every direction, they contain $\alpha$-(Hausdorff)-dimensional subsets of lines in every direction, for some parameter $\alpha \in (0, 1]$. While Davies's theorem gives the minimum Hausdorff dimension of Kakeya sets in $\mathbb{R}^2$, the minimum Hausdorff dimension of Furstenberg sets is an important open question [107, 55]. Molter and Rela [81] generalized the class further by requiring $\alpha$-dimensional subsets of lines in only a $\beta$-dimensional set of directions, for some second parameter $\beta \in (0, 1]$. They showed that any such set has Hausdorff dimension at least $\alpha + \max\{\beta/2, \alpha + \beta - 1\}$. In Theorem 37, we give a lower bound of $\alpha + \min\{\beta, \alpha\}$, which constitutes an improvement whenever $\alpha, \beta < 1$ and $\beta/2 < \alpha$.

This theorem corresponds to a more general lower bound on the effective Hausdorff dimension of the point $(x, ax + b)$. Proving this pointwise bound is the main technical challenge of Chapter 5, requiring us to introduce a new technique for bounding effective dimensions and is, on its own, a significant result in computable analysis. As a corollary, it resolves a well-known open question by showing that, although the set of all points in $\mathbb{R}^2$ with effective Hausdorff dimension 1 is connected [104], that set does not contain any line.

We apply the point-to-set principle and conditional dimensions once again in Chapter 6 to derive dimensional bounds on intersections and Cartesian products of arbitrary sets in $\mathbb{R}^n$. In doing so, we exhibit another new theorem, Theorem 38: a general lower bound on

Hausdorff dimension for the intersection of two sets. This theorem was previously known to hold for Borel sets [30], but our algorithmic information theoretic approach allows us to eliminate that restriction. It also greatly simplifies the proof of a known bound on the Hausdorff dimension of product sets. The proofs in Chapter 6 are short and intuitive, further demonstrating the strength and versatility of our approach for reasoning about fractal geometry.

In part II, we turn our attention to decentralized dynamics. Dynamic environments where decision makers repeatedly interact arise in a variety of settings, such as Internet protocols, large-scale markets, social networks, and multi-processor computer architectures [53]. Study of these environments lies at the boundary of game theory and distributed computing, as the decision makers are both strategic entities with individual economic preferences and computational entities with limited resources, working in a decentralized and uncertain environment. To understand the long-term global behaviors that result from these interactions—the dynamics of these systems—we draw on ideas from both disciplines. We focus primarily on systems in which the decision makers, or *computational nodes*, are deterministic and have *bounded recall*, meaning that their behavior is based only on the "recent history" of system interaction. We investigate these dynamics under two types of distributed environmental uncertainty: *asynchrony* and *uncoupledness*.

To study the effects of asynchrony, we describe in Chapter 7 a formal model of asynchronous interaction and present a general impossibility result for asynchronous environments, showing that a large and natural class of bounded-recall dynamics can fail to converge whenever the dynamics possess at least two possible points of convergence the dynamics. To prove this result, we employ a valency argument (a now-standard technique in distributed computing theory [72, 35]); we also show that this theorem is essentially tight. We then discuss the connections of this work to the impossibility of resilient consensus (the famous result of Fisher, Lynch, and Paterson [36]) and its implications for game dynamics.

We consider bounded-recall dynamics in environments that are uncoupled [45], meaning

that nodes' behaviors are determined by decentralized private inputs, in Chapter 8. For game dynamics, these inputs represent each player's private preferences, of which the other players are initially ignorant. The global objective is for these players to self stabilize, reaching equilibrium from arbitrary initial conditions despite having no prior knowledge of each other's preferences. We show that deterministic historyless dynamics fail to self stabilize over any non-trivial state space in this setting, then present deterministic bounded-recall dynamics that guarantee self stabilization. We also extend results of [46] to give a complete characterization of the state spaces over which randomized historyless dynamics can self-stabilize.

We summarize the results of this dissertation in Chapter 9 and describe several specific directions for future investigations, including some that are already in progress [71, 25]. In particular, there is reason to believe that the research agendas addressed in Parts I and II will soon intersect more directly by geometrically analyzing decentralized dynamics. One natural link is provided by recent work of Babichenko, et al. [5] on constructing Kakeya sets—which provided the original motivation for and first application of our algorithmic dimensional methods—using strategies in discrete *pursuit games*. Among the future directions we describe is a geometric approach to asynchronous schedules, yielding an alternative path to applying the techniques of Part I to the systems of Part II.

# Part I

# Algorithmic Information and Fractal Geometry

# Chapter 2

# Classical and Algorithmic Fractal Dimensions

This chapter describes connections between algorithmic and classical dimensions in Euclidean spaces. These connections will form the basis of our applications of algorithmic information theory to fractal geometry in Chapters 4, 5, and 6. The central message of this chapter is a pair of useful *point-to-set principles* by which the existence of a single high-dimensional point in a set $E \subseteq \mathbb{R}^n$ implies that the set $E$ has high dimension.

Roughly fifteen years after the mid-twentieth century development of the *Shannon information theory* [96] of probability spaces, Kolmogorov [57] recognized that Turing's mathematical theory of computation could be used to refine the Shannon theory to enable the amount of information in individual data objects to be quantified. The resulting theory of *Kolmogorov complexity*, or *algorithmic information theory*, is now a large enterprise with many applications in computer science, mathematics, and other sciences [62]. Kolmogorov [57] proved the first version of the fundamental relationship between the Shannon and algorithmic theories of information, and this relationship was made precise by Levin's coding theorem [60, 61]. (Solomonoff [99] and Chaitin [17, 18]) independently developed Kolmogorov complexity at around the same time as Kolmogorov with somewhat different motivations.)

At the turn of the present century, J. H. Lutz [63, 64] showed that Hausdorff's [48] theory of fractal dimension is an older theory of information that can also be refined using Turing's mathematical theory of computation, thereby enabling the *density* of information in individual infinite data objects, such as infinite binary sequences or points in Euclidean spaces, to be quantified. The resulting theory of *effective fractal dimensions* is now an

active enterprise with a growing array of applications [28].

How can the dimensions of individual points—dimensions that are defined using the theory of computing—have any bearing on classical problems of geometric measure theory? The problems that we have in mind here are problems in which one seeks to establish lower bounds on the classical Hausdorff $\dim_H(E)$ (or packing dimension $\dim_P(E)$) of a set $E \subseteq \mathbb{R}^n$. Such problems involve global properties of sets and make no mention of algorithms.

The key to bridging this gap is relativization. Specifically, our point-to-set principle for Hausdorff dimension states that, in order to prove a lower bound $\dim_H(E) \geq \alpha$, it suffices to show that, for every $A \subseteq \mathbb{N}$ and every $\varepsilon > 0$, there is a point $x \in E$ such that $\dim^A(x) \geq \alpha - \varepsilon$, where $\dim^A(x)$ is the effective Hausdorff dimension of $x$ relative to the oracle $A$. We also present an analogous principle for classical packing dimension and relativized effective packing dimension. Before proving these principles, we begin by defining classical fractal dimensions and their algorithmic counterparts.

## 2.1 Classical Fractal Dimensions

We now define, for sets $E \subseteq \mathbb{R}^n$, the Hausdorff dimension $\dim_H(E)$, and its dual, the packing dimension $\dim_P(E)$. Although there are more primitive alternatives with simpler definitions (e.g., box-counting dimension), three desirable properties have made Hausdorff dimension the most standard notion of fractal dimension since it was introduction by Hausdorff in 1919. First, it is defined on every set in $\mathbb{R}^n$. Second, it is *monotone*: if $E \subseteq F$, then $\dim_H(E) \leq \dim_H(F)$. Third, it is *countably stable*: if $E = \bigcup_{i\in\mathbb{N}} E_i$, then $\dim_H(E) = \sup_{i\in\mathbb{N}} \dim_H(E_i)$. These three properties also hold for packing dimension, which was defined much later, independently by Tricot [103] and by Sullivan [101].

**Notation.** For $E \subseteq \mathbb{R}^n$, $\mathrm{diam}(X) = \sup_{x,y\in X} |x - y|$ is the diameter of $E$. For $x \in \mathbb{R}^n$ and $\rho > 0$, $B_\rho(x) = \{y \in \mathbb{R}^n : |x - y| < \rho\}$ is the open ball of radius $\rho$ around $x$.

**Definition** (Hausdorff [48])**.** Let $E \subseteq \mathbb{R}^n$. For $\delta > 0$, define $\mathcal{U}_\delta(E)$ to be the collection of

all countable covers of $E$ by sets of positive diameter at most $\delta$. That is, for every cover $\{U_i\}_{i\in\mathbb{N}} \in \mathcal{U}_\delta(E)$, we have $E \subseteq \bigcup_{i\in\mathbb{N}} U_i$ and $\operatorname{diam}(U_i) \in (0,\delta]$ for all $i \in \mathbb{N}$. For all $s \geq 0$, define

$$H^s_\delta(E) = \inf\left\{ \sum_{i\in\mathbb{N}} \operatorname{diam}(U_i)^s \,:\, \{U_i\}_{i\in\mathbb{N}} \in \mathcal{U}_\delta(E) \right\}.$$

1. The *s-dimensional Hausdorff outer measure* of $E$ is

$$H^s(E) = \lim_{\delta\to 0^+} H^s_\delta(E)\,.$$

2. The *Hausdorff dimension* of $E$ is

$$\dim_H(E) = \inf\{s > 0 : H^s(E) = 0\}\,.$$

**Definition** (Tricot [103] and Sullivan [101])**.** Let $E \subseteq \mathbb{R}^n$. For $\delta > 0$, define $\mathcal{V}_\delta(E)$ to be the collection of all countable *packings* of $E$ by disjoint open balls of diameter at most $\delta$. That is, for every packing $\{V_i\}_{i\in\mathbb{N}} \in \mathcal{V}_\delta(E)$ and every $i \in \mathbb{N}$, we have $V_i = B_{\varepsilon_i}(x_i)$ for some $x_i \in E$ and $\varepsilon_i \in (0,\delta/2]$, and for every $j \in \mathbb{N}\setminus\{i\}$, $V_i \cap V_j = \emptyset$.

For all $s \geq 0$, define

$$P^s_\delta(E) = \sup\left\{ \sum_{i\in\mathbb{N}} \operatorname{diam}(V_i)^s \,:\, \{V_i\}_{i\in\mathbb{N}} \in \mathcal{V}_\delta(E) \right\},$$

and let

$$P^s_0(E) = \lim_{\delta\to 0^+} P^s_\delta(E)\,.$$

1. The *s-dimensional packing measure of* $E$ is

$$P^s(E) = \inf\left\{ \sum_{i\in\mathbb{N}} P^s_0(E_i) \,:\, E \subseteq \bigcup_{i\in\mathbb{N}} E_i \right\}.$$

2. The *packing dimension of* $E$ is

$$\dim_P(E) = \inf\{s : P^s(E) = 0\}\,.$$

Notice that defining packing dimension in this way requires an extra step of optimization compared to Hausdorff dimension. For every set $E \subseteq \mathbb{R}^n$, we have $\dim_H(E) \leq \dim_P(E)$, with equality when $E$ is very "regular." More properties and details about classical fractal dimensions may be found in standard references such as [78, 32, 100].

## 2.2 Kolmogorov Complexity and Dimensions of Sequences

Hausdorff and packing dimensions are defined similarly in the Cantor space of infinite binary sequences, and it was shown by J. H. Lutz [63] that Hausdorff dimension can be characterized in this space using betting strategies called *gales*, which generalize martingales. By requiring these gales to be constructive (i.e., upper semicomputable), the same author defined a constructive version of Hausdorff dimension. Unlike classical Hausdorff dimension, the constructive version is nonzero on some singleton sets and thus assigns a meaningful dimension to *individual* infinite binary sequences [64]. In this work, instead of using the original gale definition, we use a later characterization based on Kolmogorov complexity, which we now define.

**Definition.** The *conditional Kolmogorov complexity* of a string $\sigma \in \{0,1\}^*$ given a string $\tau \in \{0,1\}^*$ is

$$K(\sigma|\tau) = \min_{\pi \in \{0,1\}^*} \{|\pi| : U(\pi, \tau) = \sigma\},$$

where $U$ is a fixed universal prefix machine and $|\pi|$ is the length of $\pi$. The *Kolmogorov complexity* of $\sigma$ is $K(\sigma) = K(\sigma|\lambda)$, where $\lambda$ is the empty string. We write $U(\pi)$ for $U(\pi, \lambda)$. When $U(\pi) = \sigma$, the string $\pi$ is called a *program* for $\sigma$.

The quantity $K(\sigma)$ is also called the *algorithmic information content* of $\sigma$. Routine coding extends this definition from $\{0,1\}^*$ to other discrete domains, so that the Kolmogorov complexities of natural numbers, rational numbers, tuples of these, etc., are well defined up to additive constants. Detailed discussions of prefix Turing machines and Kolmogorov complexity appear in the books [62, 83, 28] and many papers.

The first identity below, which we take as a definition here, was proven as a theorem by Mayordomo [79]. The second was proven by Athreya, Hitchcock, J. H. Lutz, and Mayordomo [2], after first constructivizing packing dimension using gales.

**Definition** (Mayordomo [79] and Athreya, et al. [2])**.** Let $w$ be an infinite binary sequence, and for $r \in \mathbb{N}$, let $w{\upharpoonright}r$ denote the length-$r$ prefix of $w$.

1. The *effective Hausdorff dimension* of $w$ is

$$\dim(w) = \liminf_{r\to\infty} \frac{K(w\upharpoonright r)}{r}\,.$$

2. The *effective packing dimension* of $w$ is

$$\mathrm{Dim}(x) = \limsup_{r\to\infty} \frac{K(w\upharpoonright r)}{r}\,.$$

Recall that $K(\sigma)$ is the *algorithmic information content* of $\sigma \in \{0,1\}^*$. Similarly, the ratio $K(w\upharpoonright r)/r$ may be described as the *algorithmic information density* of $w \in \{0,1\}^\infty$ at precision $r \in \mathbb{N}$. This ratio might not converge as $r \to \infty$, but it is easy to see that it has asymptotes in the interval $[0,1]$. Hence, $\dim(w)$ and $\mathrm{Dim}(w)$ are the lower and upper asymptotic densities of the algorithmic information in $w$.

Note that these dimensions are more precisely called *constructive dimensions*, as they originally were, because other effectivizations (e.g., computable dimensions, polynomial time dimensions, and finite-state dimensions) have also been investigated. Since only the $\Sigma^0_1$ effectivizations are discussed in this work, we use the above terminology, which has become common in the literature.

## 2.3 Dimensions of Points in Euclidean Spaces

Effective Hausdorff and packing dimensions were extended to Euclidean spaces by J. H. Lutz and Mayordomo [67]. In this setting, the precision parameter $r$ defines a neighborhood in which we find a rational point of minimal complexity. Applications of these dimensions in Euclidean spaces appear in [67, 40, 68, 27, 41].

**Definition.** For $x \in \mathbb{R}^n$ and $r \in \mathbb{N}$, the *Kolmogorov complexity* of $x$ at *precision* $r$ is

$$K_r(x) = \min\{K(q) \,:\, q \in \mathbb{Q}^n \cap B_{2^{-r}}(x)\}\,, \tag{2.3.1}$$

where $B_{2^{-r}}(x)$ is the open ball with radius $2^{-r}$ and center $x$.

We note that $K_r(x) = K(x\restriction r) + o(r)$, where $x\restriction r$ is the binary expansion of $x$, truncated $r$ bits to the right of the binary point. However, it has been known since Turing's famous correction [105] that binary notation is not a suitable representation for the arguments and values of computable functions on the reals. (See also [106].) Hence, in order to make our definitions useful for further work in computable analysis, we formulate complexities and dimensions in terms of rational approximations. This relationship between these two formulations is addressed more formally in Section 5.1.

We will frequently use the fact that $K_r(x)$ is only linearly sensitive to the precision $r$.

**Lemma 1.** (Case and J. Lutz [15]) *There is a constant $c \in \mathbb{N}$ such that for all $n, r, s \in \mathbb{N}$ and $x \in \mathbb{R}^n$,*

$$K_r(x) \leq K_{r+s}(x) \leq K_r(x) + K(r) + ns + a_s + c\,, \tag{2.3.2}$$

*where $a_s = K(s) + 2\log(\lceil \frac{1}{2}\log n\rceil + s + 3) + (\lceil \frac{1}{2}\log n\rceil + 3)n + K(n) + 2\log n$.*

Usually, we will treat $n$ as a constant and only require a less precise version of the inequalities (2.3.2), namely,

$$K_{r+s}(x) = K_r(x) + O(s + \log r)\,.$$

**Definition** (J. H. Lutz and Mayordomo [67])**.** Let $x \in \mathbb{R}^n$.

1. The *effective Hausdorff dimension*, or simply *dimension* of $x$ is

$$\dim(x) = \liminf_{r\to\infty} \frac{K_r(x)}{r}\,.$$

2. The *effective packing dimension*, or *strong dimension* of $x$ is

$$\mathrm{Dim}(x) = \limsup_{r\to\infty} \frac{K_r(x)}{r}\,.$$

For $x \in \mathbb{R}^n$, it is easy to see that

$$0 \leq \dim(x) \leq \mathrm{Dim}(x) \leq n\,,$$

and it is known that, for any two reals $0 \leq \alpha \leq \beta \leq n$, there exist uncountably many points $x \in \mathbb{R}^n$ satisfying $\dim(x) = \alpha$ and $\mathrm{Dim}(x) = \beta$ [2]. We will also consider points in $\mathbb{R}^n$ that are (Martin-Löf) random [75].

**Definition.** A point $x \in \mathbb{R}^n$ is *random* if there exists $c \in \mathbb{N}$ such that, for all $r \in \mathbb{N}$,

$$K_r(x) \geq nr - c\,.$$

It is well known that (Lebesgue) almost every point $x \in \mathbb{R}^n$ is random [62], and it is immediate from definitions that if $x \in \mathbb{R}^n$ is random, then $\dim(x) = \mathrm{Dim}(x) = n$.

## 2.4 From Points to Sets

We now establish our point-to-set principles relating algorithmic dimensions of points to classical dimensions of sets. To formulate these principles, we use relativization. All the algorithmic information concepts in Sections 2.2–2.3 can be relativized to an arbitrary oracle $A \subseteq \mathbb{N}$ by giving the Turing machine in their definitions oracle access to $A$. Relativized Kolmogorov complexity $K_r^A(x)$ and relativized dimensions $\dim^A(x)$ and $\mathrm{Dim}^A(x)$ are thus well defined. Moreover, the results of Section 2.3 hold relative to any oracle $A$.

**Theorem 2** (Point-to-set principle for Hausdorff dimension)**.** *For every set $E \subseteq \mathbb{R}^n$,*

$$\dim_H(E) = \min_{A \subseteq \mathbb{N}} \sup_{x \in E} \dim^A(x)\,.$$

Three things should be noted about this principle. First, while the left-hand side is the *classical* Hausdorff dimension, which is a global property of $E$ that does not involve the theory of computing, the right-hand side is a pointwise property of the set that makes essential use of relativized algorithmic information theory. Second, as the proof shows, the right-hand side is a minimum, not merely an infimum. Third, and most crucially, this principle implies that, in order to prove a lower bound $\dim_H(E) \geq \alpha$, it suffices to show that, for every $A \subseteq \mathbb{N}$ and every $\varepsilon > 0$, there is a point $x \in E$ such that $\dim^A(x) \geq \alpha - \varepsilon$.[1]

[1]The $\varepsilon$ here is useful in general but is not needed in some cases, including our proof of Theorem 22 below.

For the ($\geq$) direction of this principle, we construct the minimizing oracle $A$. The oracle encodes, for a carefully chosen sequence of increasingly refined covers for $E$, the approximate locations and diameters of all cover elements. Using this oracle, a point $x \in \mathbb{R}^n$ can be approximated by specifying an appropriately small cover element that it belongs to, which requires an amount of information that depends on the number of similarly-sized cover elements. We use the definition of Hausdorff dimension to bound that number. The ($\leq$) direction can be shown using results from [67], but in the interest of self-containment we prove it directly.

*Proof of Theorem 2.* Let $E \subseteq \mathbb{R}^n$, and let $d = \dim_H(E)$. For every $s > d$ we have $H^s(E) = 0$, so there is a sequence $\{\{U_i^{t,s}\}_{i\in\mathbb{N}}\}_{t\in\mathbb{N}}$ of countable covers of $E$ such that $\operatorname{diam}(U_i^{t,s}) \leq 2^{-t})$ for every $i, t \in \mathbb{N}$, and for every sufficiently large $t$ we have

$$\sum_{i\in\mathbb{N}} \operatorname{diam}(U_i^{t,s})^s < 1\,. \tag{2.4.1}$$

Let $D = \mathbb{N}^3 \times (\mathbb{Q} \cap (d, \infty))$. Our oracle $A$ encodes functions $f_A : D \to \mathbb{Q}^n$ and $g_A : D \to \mathbb{Q}$ such that for every $(i, t, r, s) \in D$, we have

$$f_A(i, t, r, s) \in B_{2^{-r-1}}(u)$$

for some $u \in U_i^{t,s}$ and

$$\left|g_A(i, t, r, s) - \operatorname{diam}(U_i^{t,s})\right| < 2^{-r-4}\,. \tag{2.4.2}$$

We will show, for every $x \in E$ and rational $s > d$, that $\dim^A(x) \leq s$.

Fix $x \in E$ and $s \in \mathbb{Q} \cap (d, \infty)$. If for any $i_0, t_0 \in \mathbb{N}$ we have $x \in U_{i_0}^{t_0,s}$ and $\operatorname{diam}(U_{i_0}^{t_0,s}) = 0$, then $U_{i_0}^{t_0,s} = \{x\}$, so $f_A(i_0, t_0, r, s) \in B_{2^{-r}}(x)$ for every $r \in \mathbb{N}$. In this case, let $M$ be a prefix-free Turing machine with oracle access to $A$ such that, whenever $U(\iota) = i \in \mathbb{N}$, $U(\tau) = t \in \mathbb{N}$, $U(\rho) = r \in \mathbb{N}$, and $U(\sigma) = q \in \mathbb{Q} \cap (d, \infty)$,

$$M(\iota\tau\rho\sigma) = f_A(i, t, r, q)\,.$$

Now for any $r \in \mathbb{N}$, let $\iota$, $\tau$, $\rho$, and $\sigma$ be witnesses to $K(i_0)$, $K(t_0)$, $K(r)$, and $K(s)$, respectively. Since $i_0$, $t_0$, and $s$ are all constant in $r$ and $|\rho| = o(r)$, we have $|\iota\tau\rho\sigma| = o(r)$.

Thus $K_r^A(x) = o(r)$, and $\dim^A(x) = 0$. Hence assume that every cover element containing $x$ has positive diameter.

Fix sufficiently large $t$, and let $U_{i_x}^{t,s}$ be some cover element containing $x$. Let $M'$ be a prefix-free Turing machine with oracle access to $A$ such that whenever $U(\kappa) = k \in \mathbb{N}$, $U(\tau) = \ell \in \mathbb{N}$, $U(\rho) = r \in \mathbb{N}$, and $U(\sigma) = q \in \mathbb{Q} \cap (d, \infty)$,

$$M'(\kappa\tau\rho\sigma) = f_A(p, \ell, r, q)\,,$$

where $p$ is the $k^{\text{th}}$ index $i$ such that $g_A(i, t, r, q) \geq 2^{-r-3}$.

Now fix $r \geq t - 1$ such that

$$\operatorname{diam}(U_{i_x}^{t,s}) \in \left[2^{-r-2}, 2^{-r-1}\right)\,.$$

Notice that $g_A(i_x, t, r, s) \geq 2^{-r-3}$. Hence there is some $k$ such that, letting $\kappa$, $\tau$, $\rho$, and $\sigma$ be witnesses to $K(k)$, $K(t)$, $K(r)$, and $K(s)$, respectively,

$$M'(\kappa\tau\rho\sigma) \in B_{2^{-r-1}}(u)\,,$$

for some $u \in U_{i_x}^{t,s}$. Because $\operatorname{diam}(U_{i_x}^{t,s}) < 2^{-r-1}$ and $x \in U_{i_x}^{t,s}$, we have

$$M'(\kappa\tau\rho\sigma) \in B_{2^{-r}}(x)\,.$$

Thus

$$K_r^A(x) \leq K(k) + K(t) + K(s) + K(r) + c\,,$$

where $c$ is a machine constant for $M'$. Since $s$ is constant in $r$ and $t < r$, Observation 10 tells us that this expression is $K(k) + o(r) \leq \log(k) + o(r)$. By (2.4.1), there are fewer than $2^{(r+4)s}$ indices $i \in \mathbb{N}$ such that

$$\operatorname{diam}(U_i^{t,s}) \geq 2^{-r-4}\,,$$

hence by (2.4.2) there are fewer than $2^{(r+4)s}$ indices $i \in \mathbb{N}$ such that $g_A(i, t, r, s) \geq 2^{-r-3}$, so $\log(k) < (r + 4)s$. Therefore $K_r^A(x) \leq rs + o(r)$.

There are infinitely many such $r$, which can be seen by replacing $t$ above with $r+2$. We have shown

$$\dim^A(x) = \liminf_{r\to\infty} \frac{K_r^A(x)}{r} \leq s\,,$$

for every rational $s > d$, hence $\dim^A(x) \leq d$. It follows that

$$\min_{A\subseteq\mathbb{N}} \sup_{x\in E} \dim^A(x) \leq d\,.$$

For the other direction, assume for contradiction that there is some oracle $A$ and $d' < d$ such that

$$\sup_{x\in E} \dim^A(x) = d'\,.$$

Then for every $x \in E$, $\dim^A(x) \leq d'$. Let $s \in (d', d)$. For every $r \in \mathbb{N}$, define the sets

$$\mathcal{B}_r = \left\{B_{2^{-r}}(q) \,:\, q \in \mathbb{Q} \text{ and } K^A(q) \leq rs\right\}$$

and

$$\mathcal{W}_r = \bigcup_{k=r}^{\infty} \mathcal{B}_k\,.$$

There are at most $2^{ks+1}$ balls in each $\mathcal{B}_k$, so for every $r \in \mathbb{N}$ and $s' \in (s, d)$,

$$\begin{aligned}\sum_{W\in\mathcal{W}_r} \operatorname{diam}(W)^{s'} &= \sum_{k=r}^{\infty} \sum_{W\in\mathcal{B}_k} \operatorname{diam}(W)^{s'} \\ &\leq \sum_{k=r}^{\infty} 2^{ks+1}(2^{1-k})^{s'} \\ &= 2^{1+s'} \cdot \sum_{k=r}^{\infty} 2^{(s-s')k}\,,\end{aligned}$$

which approaches 0 as $r \to \infty$. As every $\mathcal{W}_r$ is a cover for $E$, we have $H^{s'}(E) = 0$, so $\dim_H(E) \leq s' < d$, a contradiction. □

Using similar techniques, we now establish the point-to-set principle for packing dimension.

**Theorem 3** (Point-to-set principle for packing dimension)**.** *For every set $E \subseteq \mathbb{R}^n$,*

$$\dim_P(E) = \min_{A\subseteq\mathbb{N}} \sup_{x\in E} \operatorname{Dim}^A(x)\,.$$

*Proof.* Let $E \subseteq \mathbb{R}^n$, and let $d = \dim_P(E)$. For every $s > d$ we have $P^s(E) = 0$, so there is a cover $\left\{E_j^s\right\}_{j\in\mathbb{N}}$ for $E$ such that

$$\sum_{j\in\mathbb{N}} \lim_{\delta\to 0^+} P_\delta^s(E_j^s) < 1\,. \tag{2.4.3}$$

For every $r, j \in \mathbb{N}$, let

$$\left\{V_i^{r,s,j}\right\}_{i\in\mathbb{N}} \in \mathcal{V}_{2^{-r-2}}(E_j^s)$$

be a maximal packing of $E_j^s$ by open balls of radius exactly $2^{-r-2}$ (and higher-indexed balls of radius 0).

Let $D = \mathbb{N}^3 \times (\mathbb{Q} \cap (d, \infty))$. Our oracle $A$ encodes a function $f_A : D \to \mathbb{Q}^n$ such that for every $(i, j, r, s) \in D$ we have

$$f_A(i, j, r, s) \in V_i^{r,s,j}\,.$$

We will show, for every $x \in E$ and rational $s > d$, that $\mathrm{Dim}^A(x) \leq s$.

Let $M$ be a prefix-free Turing machine with oracle access to $A$ such that, whenever $U(\iota) = i \in \mathbb{N}$, $U(\kappa) = j \in \mathbb{N}$, $U(\rho) = r \in \mathbb{N}$, and $U(\sigma) = q \in \mathbb{Q} \cap (d, \infty)$,

$$M(\iota\kappa\rho\sigma) = f_A(i, j, r, s)\,.$$

Fix $x \in E$ and $s \in \mathbb{Q} \cap (0, \infty)$, and let $k \in \mathbb{N}$ be such that $x \in E_k^s$. Notice that by our choice of packing, for every $r \in \mathbb{N}$ there must be some $i_r \in \mathbb{N}$ such that

$$V_{i_r}^{r,s,k} \subseteq B_{2^{-r}}(x)\,.$$

Thus, for every $r \in \mathbb{N}$, letting $\iota$, $\kappa$, $\rho$, $\sigma$ testify to $K(i_r)$, $K(k)$, $K(r)$, and $K(s)$, respectively,

$$\begin{aligned} M(\iota\kappa\rho\sigma) &= f_A(i_r, k, r, s) \\ &\in V_{i_r}^{r,s,k} \\ &\subseteq B_{2^{-r}}(x)\,, \end{aligned}$$

hence $K_r^A(x) \leq K(i_r) + K(k) + K(r) + K(s) + c$, where $c$ is a machine constant for $M$. Because $k$ and $s$ are constant in $r$, $K(r) = o(r)$, and $K(i_r) \leq \log i_r + o(r)$, we have

$$K_r^A(x) \leq \log i_r + o(r)\,.$$

By (2.4.3),

$$\lim_{\delta \to 0^+} P^s_\delta(E^s_k) < 1\,,$$

so there is some $R \in \mathbb{N}$ such that, for every $r > R$, $P^s_{2^{-r}}(E^s_k) < 1$. Then for every $r > R$,

$$\sum_{i \in \mathbb{N}} \operatorname{diam}(V_i^{r,s,k})^s < 1\,,$$

hence there are fewer than $2^{(r+2)s}$ balls of radius $2^{-r-2}$ in the packing, and $\log i_r < (r+2)s$. We conclude that $K^A_r(x) \leq rs + o(r)$ for every $r > R$, so

$$\dim^A(x) = \limsup_{r \to \infty} \frac{K^A_r(x)}{r} \leq s\,.$$

Since this holds for every rational $s > d$, we have shown $\operatorname{Dim}^A(x) \leq d$ and thus

$$\min_{A \subseteq \mathbb{N}} \sup_{x \in E} \operatorname{Dim}^A(x) \leq d\,.$$

For the other direction, assume for contradiction that there is some oracle $A$ and $d' < d$ such that

$$\sup_{x \in E} \operatorname{Dim}^A(x) = d'\,.$$

Then for every $x \in E$, $\operatorname{Dim}^A(x) \leq d'$. Let $s \in (d', d)$. For every $k \in \mathbb{N}$, define the set

$$C_k = \bigcup \left\{ B_{2^{-k}}(q) \,:\, q \in \mathbb{Q} \text{ and } K^A(q) \leq ks \right\}\,,$$

and for every $i \in \mathbb{N}$, define

$$E_i = \bigcap_{k=i}^{\infty} C_k\,.$$

For $r \geq i$, consider any packing in $\mathcal{V}_{2^{-r}}(E_i)$. Let $B_\varepsilon(x)$ be an element of the packing, and let $k = \lceil -\log \varepsilon \rceil$. Then $k \geq r + 1 > i$, so $B_\varepsilon(x) \subseteq E_i \subseteq C_k$. In particular $x \in C_k$, meaning that there is some $q \in \mathbb{Q}$ such that $K^A(q) \leq ks$ and $x \in B_{2^{-k}}(q)$. As $2^{-k} \leq \varepsilon$, we also have $q \in B_\varepsilon(x)$. Thus, every packing element of radius at least $2^{-k}$ contains a (distinct) member of the set $\{q \in \mathbb{Q} : K^A(q) \leq ks\}$. It follows that for every $k \geq r + 1$, the packing includes at most $2^{ks+1}$ elements with diameters in the range $[2^{1-k}, 2^{2-k})$.

Now let $s' \in (s, d)$. For every $i \in \mathbb{N}$ and $r \geq i$, we have

$$\begin{aligned} P^{s'}_{2^{-r}}(E_i) &= \sup\left\{ \sum_{j\in\mathbb{N}} \operatorname{diam}(V_j)^{s'} \,:\, \{V_j\}_{j\in\mathbb{N}} \in \mathcal{V}_{2^{-r}}(E_i) \right\} \\ &\leq \sum_{k=r+1}^{\infty} 2^{ks+1}(2^{2-k})^{s'} \\ &= 2^{1+2s'} \cdot \sum_{k=r+1}^{\infty} 2^{(s-s')k} \,. \end{aligned}$$

This approaches 0 as $r \to \infty$, so $P^{s'}_0(E_i) = 0$. Observe now that

$$E \subseteq \bigcup_{i\in\mathbb{N}} E_i \,.$$

Thus,

$$P^{s'}(E) \leq \sum_{i\in\mathbb{N}} P^{s'}_0(E_i) = 0 \,,$$

meaning that $\dim_P(E) \leq s' < d$, a contradiction. We conclude that for every oracle $A$,

$$\sup_{x\in E} \operatorname{Dim}^A(x) \geq d \,.$$

□

## 2.5 Classical Pointwise Dimensions of Measures

As we have seen, effective Hausdorff and packing dimension are fundamentally pointwise. There is also a classical notion of pointwise dimensions of measures, which is central to the study of fractals and dynamics. In this section, we describe a connection between these two formulations of pointwise dimension. We then use this connection to compare Theorems 2 and 3 to classical pointwise characterizations of Hausdorff and packing dimensions.

In the classical setting, pointwise dimensions are defined for a given measure according to its (lower and upper asymptotic) rate of decay around $x$.

**Definition.** For any locally finite measure $\mu$ on $\mathbb{R}^n$, the *lower* and *upper pointwise dimension* of $\mu$ at $x \in \mathbb{R}^n$ are

$$\dim_\mu(x) = \liminf_{\rho \to 0} \frac{\log \mu(B_\rho(x))}{\log \rho}$$
$$\operatorname{Dim}_\mu(x) = \limsup_{\rho \to 0} \frac{\log \mu(B_\rho(x))}{\log \rho},$$

respectively [31].

As Young [111] notes, these limits are unchanged if $\rho$ is replaced by any sequence $\{\rho_r\}_{r \in \mathbb{N}}$ satisfying $\rho_r \downarrow 0$ and $\log \rho_{r+1} / \log \rho_r \to 1$. In particular, the sequence $\{2^{-r}\}_{r \in \mathbb{N}}$ may be used. Also, this definition in no way relies on additivity, so it applies equally well to outer measures and semimeasures.

We relate these two notions of pointwise dimension by defining an outer measure $\kappa^A$ on $\mathbb{R}^n$ for any given oracle set $A \subseteq \mathbb{N}$. For every $E \subseteq \mathbb{R}^n$,

$$\kappa^A(E) = 2^{-K^A(E)},$$

where, following Shen and Vereschagin [97],

$$K^A(E) = \min_{q \in E \cap \mathbb{Q}^n} K^A(q).$$

This minimum is taken to be infinite when $E \cap \mathbb{Q}^n = \emptyset$. It is easy to see that $\kappa^A$ is also subadditive and monotonic, and that $\kappa^A(\emptyset) = 0$. Since $K^A$ is non-negative, $\kappa^A$ is finite.

**Observation 4.** *For every oracle set $A \subseteq \mathbb{N}$ and all $x \in \mathbb{R}^n$,*

$$\dim_{\kappa^A}(x) = \dim^A(x).$$
$$\operatorname{Dim}_{\kappa^A}(x) = \operatorname{Dim}^A(x).$$

This fact is closely related to (and was observed independently of) an unpublished remark by Reimann [91] stating that $\dim(x)$ is equal to the pointwise dimension at $x$ of Levin's [61] universal lower semicomputable continuous semimeasure.

Pointwise dimensions of measures give rise to global dimensions of measures, which we now briefly comment on. In classical fractal geometry, the global dimensions of Borel measures play a substantial role in studying the interplay between local and global properties of fractal sets and measures.

**Definition.** For any locally finite Borel measure $\mu$ on $\mathbb{R}^n$ and $x \in \mathbb{R}^n$, the *lower* and *upper Hausdorff and packing dimension* of $\mu$ are

$$\dim_H(\mu) = \sup\{\alpha : \mu(\{x : \dim_\mu(x) < \alpha\}) = 0\}$$

$$\mathrm{Dim}_H(\mu) = \inf\{\alpha : \mu(\{x : \dim_\mu(x) > \alpha\}) = 0\}$$

$$\dim_P(\mu) = \sup\{\alpha : \mu(\{x : \mathrm{Dim}_\mu(x) < \alpha\}) = 0\}$$

$$\mathrm{Dim}_P(\mu) = \inf\{\alpha : \mu(\{x : \mathrm{Dim}_\mu(x) > \alpha\}) = 0\}\,,$$

respectively [31].

Extending these definitions to outer measures, we may consider global dimensions of the outer measures $\kappa^A$. For every $A \subseteq \mathbb{N}$, $\kappa^A$ is supported on $\mathbb{Q}^n$ and $\dim^A(p) = 0$ for all $p \in \mathbb{Q}^n$, which implies the following.

**Observation 5.** *For every $A \subseteq \mathbb{N}$,*

$$\dim_H(\kappa^A) = \mathrm{Dim}_H(\kappa^A) = \dim_P(\kappa^A) = \mathrm{Dim}_P(\kappa^A) = 0\,.$$

In light of Observation 4, the point-to-set principles of Theorems 2 and 3 may be considered members of the family of results, such as Billingsley's lemma [10] and Frostman's lemma [37], that relate the local decay of measures to global properties of measure and dimension. Useful references on such results include [12, 50, 78].

Among classical results, this principle is most directly comparable to the *weak duality principle* of Cutler [20] (see also [31]), which expresses Hausdorff and packing dimensions in terms of lower and upper pointwise dimensions of measures. For nonempty $E \subseteq \mathbb{R}^n$, let $\mathcal{P}(E)$ be the collection of Borel probability measures on $\mathbb{R}^n$ such that the $E$ is measurable and has measure 1, and let $\overline{E}$ be the closure of $E$.

**Theorem 6** (Cutler [20])**.** *For every nonempty $E \subseteq \mathbb{R}^n$,*

1. $\dim_H(E) = \inf_{\mu \in \mathcal{P}(\overline{E})} \sup_{x \in E} \dim_\mu(x)$.

2. $\dim_P(E) = \inf_{\mu \in \mathcal{P}(\overline{E})} \sup_{x \in E} \mathrm{Dim}_\mu(x)$.

By defining the family of outer measures $\mathcal{A} = \{\kappa^A : A \subseteq \mathbb{N}\}$, Theorems 2 and 3 can be restated in a form that is even more similar to Theorem 6, as

1. $\dim_H(E) = \inf_{\mu \in \mathcal{A}} \sup_{x \in E} \dim_\mu(x)$.

2. $\dim_P(E) = \inf_{\mu \in \mathcal{A}} \sup_{x \in E} \mathrm{Dim}_\mu(x)$.

Notice, however, that the collections over which the infima are taken in these two results, $\mathcal{A}$ and $\mathcal{P}(\overline{E})$, are disjoint and qualitatively very different. In particular, $\mathcal{A}$ does not depend on $E$. Whereas the global dimensions of the measures in $\mathcal{P}(\overline{E})$ are closely tied to the dimensions of $E$ [31], Observation 5 shows that the outer measures in $\mathcal{A}$ all have trivial global dimensions.

## 2.6 Mutual Dimensions

Mutual dimensions were developed very recently, and Kolmogorov complexity was the starting point. The relationship between effective dimensions and Kolmogorov complexity led to the development of a dimensional analogue to a second algorithmic information theoretic quantity, mutual information.

**Definition.** The *mutual (algorithmic) information* between two strings $\sigma, \tau \in \{0,1\}^*$ is

$$I(\sigma : \tau) = K(\sigma) - K(\sigma|\tau)\,.$$

Again, routine coding extends $K(\sigma|\tau)$ and $I(\sigma : \tau)$ to other discrete domains. Discussions of $K(\sigma|\tau)$, $I(\sigma : \tau)$, and the correspondence of $K(\sigma)$, $K(\sigma|\tau)$, and $I(\sigma : \tau)$ with Shannon entropy, Shannon conditional entropy, and Shannon mutual information appear in [62].

In parallel with the definitions of Kolmogorov complexity and effective dimensions in Euclidean spaces, Case and J. H. Lutz [15] lifted the definition of $I(p:q)$ for rational points $p$ and $q$ in Euclidean spaces in two steps to define the mutual dimensions between two arbitrary points in (possibly distinct) Euclidean spaces.

**Definition** (Case and J. H. Lutz [15])**.** Let $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$, and $r \in \mathbb{N}$.

1. The *mutual information* between $x$ and $y$ at *precision* $r$ is

$$I_r(x:y) = \min\left\{I(p:q) \,:\, p \in B_{2^{-r}}(x) \cap \mathbb{Q}^m \text{ and } q \in B_{2^{-r}}(y) \cap \mathbb{Q}^n\right\}.$$

2. The *lower mutual dimension* between $x$ and $y$ is

$$\mathrm{mdim}(x:y) = \liminf_{r\to\infty} \frac{I_r(x:y)}{r}.$$

3. The *upper mutual dimension* between $x$ and $y$ is

$$\mathrm{Mdim}(x:y) = \limsup_{r\to\infty} \frac{I_r(x:y)}{r}.$$

**Observation 7.** *For all* $x \in \mathbb{R}^n$*,*

*1.* $\dim(x) = \mathrm{mdim}(x:x)$.

*2.* $\mathrm{Dim}(x) = \mathrm{Mdim}(x:x)$.

Case and J. H. Lutz also showed that mutual dimensions are preserved under bi-Lipschitz computable bijections. Combined with Observation 7, this implies that effective Hausdorff and packing dimensions are preserved by such bijections. This is a Euclidean-space version of a fact that was shown in Cantor space by Reimann.

**Lemma 8** (Reimann [92], Case and J. H. Lutz [15])**.** *If* $f : \mathbb{R}^m \to \mathbb{R}^n$ *is computable and bi-Lipschitz, then* $\dim(x) = \dim(f(x))$ *and* $\mathrm{Dim}(x) = \mathrm{Dim}(f(x))$ *for all* $x \in \mathbb{R}^m$.

Other useful properties of these mutual dimensions, especially including data processing inequalities, appear in [15].

# Chapter 3

# Conditional Dimensions

In this chapter we develop *conditional dimensions*, which fill a gap in effective dimension theory. The fundamental quantities in Shannon information theory are the *entropy* (information content) $H(X)$ of a probability space $X$, the *conditional entropy* $H(X|Y)$ of a probability space $X$ given a probability space $Y$, and the *mutual information* (shared information) $I(X;Y)$ between two probability spaces $X$ and $Y$ [19]. The analogous quantities in Kolmogorov complexity theory are the *Kolmogorov complexity* $K(u)$ of a finite data object $u$, the *conditional Kolmogorov complexity* $K(u|v)$ of a finite data object $u$ given a finite data object $v$, and the *algorithmic mutual information* $I(u : v)$ between two finite data objects $u$ and $v$ [62]. The above-described dimensions $\dim(x)$ and $\mathrm{Dim}(x)$ of a point $x$ in Euclidean space (or an infinite sequence $x$ over a finite alphabet) are analogous by limit theorems [79, 2] to $K(u)$ and hence to $H(X)$. Case and J. H. Lutz have recently developed and investigated the *lower* and *upper mutual dimensions* $\mathrm{mdim}(x : y)$ and $\mathrm{Mdim}(x : y)$, which are densities of the algorithmic information shared by points $x$ and $y$ in Euclidean spaces [15] or sequences $x$ and $y$ over a finite alphabet [16]. These mutual dimensions are analogous to $I(u : v)$ and $I(X;Y)$.

What is conspicuously missing from the above account is a notion of conditional dimension. To remedy this, we first define the *conditional Kolmogorov complexity* $K_r(x|y)$ of $x \in \mathbb{R}^m$ *given* $y \in \mathbb{R}^n$ at *precision* $r \in \mathbb{N}$. This is a "conditional version" of the Kolmogorov complexity $K_r(x)$ of $x$ at precision $r$ defined in Chapter 2. We use this conditional Kolmogorov complexity to develop the *lower conditional dimension* $\dim(x|y)$ of $x$ given $y$ and its dual, the *upper conditional dimension* $\mathrm{Dim}(x|y)$ of $x$ given $y$, where $x$ and $y$ are points

in Euclidean spaces. We prove that these conditional dimensions are well behaved and that they have the correct information theoretic relationships with the previously defined dimensions and mutual dimensions. These conditional dimensions are used in Chapters 4 and 5 and (like the conditional entropy and conditional Kolmogorov complexity that motivate them) are very likely to be useful in future investigations.

## 3.1 Definitions

**Definition.** Let $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$.

1. For $q \in \mathbb{Q}^n$ and $r \in \mathbb{N}$, the *conditional Kolmogorov complexity* of $x$ at precision $r$ given $q$ is
$$\hat{K}_r(x|q) = \min\{K(p|q) \,:\, p \in \mathbb{Q}^m \cap B_{2^{-r}}(x)\} \,. \tag{3.1.1}$$

2. For $r, s \in \mathbb{N}$, the *conditional Kolmogorov complexity* of $x$ at precision $r$ given $y$ at precision $s$ is
$$K_{r,s}(x|y) = \max\left\{\hat{K}_r(x|q) \,:\, q \in \mathbb{Q}^n \cap B_{2^{-s}}(y)\right\}. \tag{3.1.2}$$

Intuitively, the maximizing argument $q$ is the point near $y$ that is least helpful in the task of approximating $x$. Note that $K_{r,s}(x|y)$ is finite, because $\hat{K}_r(x|q) \leq K_r(x) + O(1)$. When the precision parameters are equal, we abbreviate $K_{r,r}(x|y)$ by $K_r(x|y)$.

We define conditional dimensions in parallel with the definitions of effective dimensions and mutual dimensions in Chapter 2.

**Definition.** Let $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$.

1. The *lower conditional dimension* of $x$ given $y$ is
$$dim(x|y) = \liminf_{r\to\infty} \frac{K_r(x|y)}{r} \,.$$

2. The *upper conditional dimension* of $x$ given $y$ is
$$Dim(x|y) = \limsup_{r\to\infty} \frac{K_r(x|y)}{r} \,.$$

The use of the same precision bound $r$ for both $x$ and $y$ makes our conditional dimensions appear arbitrary and "brittle." We will show in Theorem 15 that this is not the case.

In the remainder of this chapter, we derive properties of conditional Kolmogorov complexity and conditional dimensions in Euclidean spaces. In particular, we show:

- $\hat{K}_r(x|q)$ and $K_{r,s}(x|y)$ are only linearly sensitive to the precision parameters $r$ and $s$ (Lemmas 13 and 14), and our definitions of conditional dimensions are therefore robust (Theorem 15).
- The correct relationships to mutual information and mutual dimension hold (Theorem 16 and Corollary 17).
- These quantities obey *chain rules* (Theorem 18 and Corollary 19).
- Conditional Kolmogorov complexity and dimension are bounded below by *relative* Kolmogorov complexity and dimension (Lemma 20 and Corollary 21).

In deriving these properties, we will make use of the following four observations about Kolmogorov complexity in Euclidean spaces.

**Observation 9.** *For every open ball $B \subseteq \mathbb{R}^m$ of radius $2^{-r}$,*

$$B \cap 2^{-\left(r+\left\lfloor \frac{1}{2}\log m\right\rfloor+1\right)}\mathbb{Z}^m \neq \emptyset\,.$$

For $a \in \mathbb{Z}^m$, let $|a|$ denote the distance from the origin to $a$.

**Observation 10.** *There is a constant $c_0 \in \mathbb{N}$ such that, for all $j \in \mathbb{N}$,*

$$K(j) \leq \log(1+j) + 2\log\log(2+j) + c_0\,.$$

**Observation 11.** *There is a constant $c \in \mathbb{N}$ such that, for all $a \in \mathbb{Z}^m$,*

$$K(a) \leq m\log(1+|a|) + \varepsilon(|a|)\,,$$

*where $\varepsilon(t) = c + 2\log\log(2+t)$.*

Observation 10 holds by a routine technique [62]. The proof of Observation 11 is also routine:

*Proof.* Fix a computable, nonrepeating enumeration $a_0, a_1, a_2, \ldots$ of $\mathbb{Z}^m$ in which tuples $a_j$ appear in nondecreasing order of $|a_j|$. Let $M$ be a Turing machine such that, for all $\pi \in \{0,1\}^*$, if $U(\pi) \in \mathbb{N}$, then $M(\pi) = a_{U(\pi)}$. Let $c = c_0 + c_M + m + \lceil 2 \log m \rceil + 2$, where $c_0$ is as in Observation 10 and $c_M$ is an optimality constant for $M$.

To see that $c$ affirms Observation 11, let $a \in \mathbb{Z}^m$. Let $j \in \mathbb{N}$ be the index for which $a_j = a$, and let $\pi \in \{0,1\}^*$ testify to the value of $K(j)$. Then $M(\pi) = a_{U(\pi)} = a_j = a$, so

$$K(a) \leq K_M(a) + c_M \leq |\pi| + c_M = K(j) + c_M \,.$$

It follows by Observation 10 that

$$K(a) \leq \log(1+j) + 2\log\log(2+j) + c + c_M \,. \tag{3.1.3}$$

We thus estimate $j$.

Let $B$ be the closed ball of radius $|a|$ centered at the origin in $\mathbb{Z}^m$, and let $Q$ be the solid, axis-parallel $m$-cube circumscribed about $B$. Let $B' = B \cap \mathbb{Z}^M$ and $Q' = Q \cap \mathbb{Z}^m$. Then

$$j \leq |B'| - 1 \leq |Q'| - 1 \leq (2|a|+1)^m - 1 \,,$$

so (3.1.3) tells us that

$$\begin{aligned} K(a) &\leq m\log(2|a|+1) + 2\log\log(1 + (2|a|+1)^m) + c + c_M \\ &\leq m\log(2|a|+2) + 2\log(m\log(2|a|+4)) + c + c_M \,. \end{aligned}$$

Since

$$m\log(2|a|+2) = m + m\log(1+|a|)$$

and

$$\begin{aligned} \log(m\log(2|a|+4)) &= \log m + \log(1 + \log(2+|a|)) \\ &\leq \log m + 1 + \log\log(2+|a|) \,, \end{aligned}$$

it follows that $K(a) \leq m \log(1 + |a|) + \varepsilon(|a|)$. □

**Observation 12.** *For every $r, n \in \mathbb{N}$, $x \in \mathbb{R}^n$, and $q \in \mathbb{Q}^n$,*

$$K_r(x + q) = K_r(x) + O(1)\,.$$

*Proof.* Let $M$ be a self-delimiting Turing machine such that $M(\pi\kappa) = U(\pi)+U(\kappa)$ whenever $U(\pi), U(\kappa) \in \mathbb{Q}^n$. If $\pi$ is a witness to $K_r(x)$ and $\kappa$ is a witness to $q$, then $M(\pi\kappa) = p + q$ for some $p \in B_{2^{-r}}(x)$, so $M(\pi\kappa) \in B_{2^{-r}}(x + q)$. Thus

$$K_r(x + q) \leq K_r(x) + K(q) + c\,,$$

where $c$ is a machine constant for $M$. Since $K(q)$ is constant in $r$, we have $K_r(x + q) \leq K_r(x) + O(1)$. Applying the same argument with $-q$ replacing $q$ completes the proof. □

## 3.2 Linear Sensitivity to Precision Parameters

**Lemma 13** (Linear Sensitivity of $\hat{K}_r(x|q)$ to $r$)**.** *There is a constant $c_1 \in \mathbb{N}$ such that, for all $x \in \mathbb{R}^m$, $q \in \mathbb{Q}^n$, and $r, \Delta r \in \mathbb{N}$,*

$$\hat{K}_r(x|q) \leq \hat{K}_{r+\Delta r}(x|q) \leq \hat{K}_r(x|q) + m\Delta r + \varepsilon_1(r, \Delta r)\,,$$

*where $\varepsilon_1(r, \Delta r) = 2 \log(1 + \Delta r) + K(r, \Delta r) + c_1$.*

*Proof.* Let $M$ be a Turing machine such that, for all $\pi_1, \pi_2, \pi_3 \in \{0, 1\}^*$ and $q \in \mathbb{Q}^n$, if $U(\pi_1, q) = p \in \mathbb{Q}^m$, $U(\pi_2) = (r, \Delta r) \in \mathbb{N}^2$, and $U(\pi_3) = a \in \mathbb{Z}^m$, then $M(\pi_1\pi_2\pi_3, q) = p + 2^{-r^*}a$, where $r^* = r + \Delta r + \left\lfloor \frac{1}{2} \log m \right\rfloor + 1$. Let $c_1 = c + c_M + 3m + m \left\lfloor \frac{1}{2} \log m \right\rfloor + \left\lceil 2 \log(3 + \left\lfloor \frac{1}{2} \log m \right\rfloor) \right\rceil$, where $c$ is the constant from Observation 11 and $c_M$ is an optimality constant for $M$.

To see that $c_1$ affirms the lemma, let $x$, $q$, $r$, and $\Delta r$ be as given. The first inequality holds trivially. To see that the second inequality holds, let $\pi_1, \pi_2 \in \{0, 1\}^*$ testify to the values of $\hat{K}_r(x|q)$ and $K(r, \Delta r)$, respectively. Let $B = B_{2^{-r}}(x)$, $B' = B_{2^{-(r+\Delta r)}}(x)$ and

$p = U(\pi_1, q)$, noting that $p \in \mathbb{Q}^m \cap B$. Applying Observation 9 to the ball $B' - p$ tells us that

$$(B' - p) \cap 2^{-r^*}\mathbb{Z}^m \neq \emptyset\,,$$

i.e., that

$$B' \cap (p + 2^{-r^*}\mathbb{Z}^m) \neq \emptyset\,.$$

So fix a point $p' \in B' \cap (p + 2^{-r^*}\mathbb{Z}^m)$, say, $p' = p + 2^{-r^*}a$, where $a \in \mathbb{Z}^m$, and let $\pi_3 \in \{0,1\}^*$ testify to the value of $K(a)$. Then

$$M(\pi_1\pi_2\pi_3, q) = p' \in \mathbb{Q} \cap B'\,,$$

so

$$\begin{aligned}\hat{K}_{r+\Delta r}(x|q) &\leq K(p'|q)\\ &\leq \hat{K}_M(p'|q) + c_M\\ &\leq |\pi_1\pi_2\pi_3| + c_M\,.\end{aligned}$$

By our choice of $\pi_1$, $\pi_2$, and $\pi_3$, this implies that

$$\hat{K}_{r+\Delta r}(x|q) \leq \hat{K}_r(x|q) + K(r, \Delta r) + K(a) + c_M\,. \tag{3.2.1}$$

We thus estimate $K(a)$.

Since

$$\begin{aligned}|a| &= 2^{r^*}|p' - p|\\ &\leq 2^{r^*}(|p' - x| + |p - x|)\\ &< 2^{r^*}\left(2^{-(r+\Delta r)} + 2^{-r}\right)\\ &= 2^{1+\lfloor \frac{1}{2}\log m\rfloor}\left(1 + 2^{\Delta r}\right)\,,\end{aligned}$$

Observation 11 tells us that

$$\begin{aligned}K(a) &\leq m\log\left(1 + 2^{\lfloor \frac{1}{2}\log m\rfloor}\left(1 + 2^{\Delta r}\right)\right) + \varepsilon(|a|)\\ &\leq m\log\left(2^{\Delta r + 3 + \lfloor \frac{1}{2}\log m\rfloor}\right) + \varepsilon(|a|)\,,\end{aligned}$$

i.e., that

$$K(a) \leq m\Delta r + 3m + m\left\lfloor \tfrac{1}{2}\log m \right\rfloor + \varepsilon(|a|)\,, \tag{3.2.2}$$

where

$$\begin{aligned}
\varepsilon(|a|) &\leq c + 2\log\log\left(2 + 2^{1+\left\lfloor \frac{1}{2}\log m\right\rfloor}\left(1+2^{\Delta r}\right)\right)\\
&\leq c + 2\log\log\left(2^{\Delta r+3+\left\lfloor \frac{1}{2}\log m\right\rfloor}\right)\\
&= c + 2\log\left(\Delta r + \left\lfloor \tfrac{1}{2}\log m\right\rfloor + 3\right)\\
&\leq c + 2\log\left((1+\Delta r)\left(3 + \left\lfloor \tfrac{1}{2}\log m\right\rfloor\right)\right)\\
&= c + 2\log(1+\Delta r) + 2\log\left(3 + \left\lfloor \tfrac{1}{2}\log m\right\rfloor\right)\,.
\end{aligned}$$

It follows by (3.2.1) and (3.2.2) that

$$\hat{K}_{r+\Delta r}(x|q) \leq \hat{K}_r(x|q) + m\Delta r + \varepsilon_1(r,\Delta r)\,.$$

□

**Lemma 14** (Linear Sensitivity of $K_{r,s}(x|y)$ to $s$)**.** *There is a constant $c_2 \in \mathbb{N}$ such that, for all $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$, and $r, s, \Delta s \in \mathbb{N}$,*

$$K_{r,s}(x|y) \geq K_{r,s+\Delta s}(x|y) \geq K_{r,s}(x|y) - n\Delta s - \varepsilon_2(s,\Delta s)\,,$$

*where $\varepsilon_2(s,\Delta s) = 2\log(1+\Delta s) + K(s,\Delta s) + c_2$.*

*Proof.* Let $M$ be a Turing machine such that, for all $\pi_1, \pi_2, \pi_3 \in \{0,1\}^*$ and $q \in \mathbb{Q}^n$, if $U(\pi_1) = (s,\Delta s) \in \mathbb{N}^2$ and $U(\pi_2) = a \in \mathbb{Z}^m$, then $M(\pi_1\pi_2\pi_3, q) = U(\pi_3, q + 2^{-s^*}a)$, where $s^* = s + \Delta s + \left\lceil \frac{1}{2}\log n\right\rceil$. Let $c_2 = c + c_M + 3n + n\left\lfloor \frac{1}{2}\log n\right\rfloor + 2\left\lceil 2\log(3 + \left\lfloor \frac{1}{2}\log n\right\rfloor)\right\rceil$, where $c$ is the constant from Observation 11 and $c_M$ is an optimality constant for $M$.

To see that $c_2$ affirms the lemma, let $x$, $y$, $r$, $s$, and $\Delta s$ be as given The first inequality holds trivially. To see that the second inequality holds, let $B = B_{2^{-}s}(y)$, $B' = B_{2^{-(s+\Delta s)}}(y)$, and $q \in \mathbb{Q}^n \cap B$. It suffices to prove that

$$\hat{K}_r(x|q) \leq K_{r,s+\Delta s}(x|y) + n\Delta s + \varepsilon_2(s,\Delta s)\,. \tag{3.2.3}$$

Let $\pi_1 \in \{0,1\}^*$ testify to the value of $K(s, \Delta s)$. Applying Observation 9 to the ball $B' - q$ tells us that

$$(B' - q) \cap 2^{-s^*}\mathbb{Z}^n \neq \emptyset\,,$$

i.e., that

$$B' \cap (q + 2^{-s^*}\mathbb{Z}^n) \neq \emptyset\,.$$

So fix a point $q' \in B' \cap (q + 2^{-s^*}\mathbb{Z}^n)$, say, $q' = q + 2^{-s^*}a$, where $a \in \mathbb{Z}^n$. Note that

$$\hat{K}_r(x|q') \leq K_{r,s+\Delta s}(x|y). \tag{3.2.4}$$

Let $\pi_2, \pi_3 \in \{0,1\}^*$ testify to the values of $K(a)$ and $\hat{K}_r(x|q')$, respectively, noting that $U(\pi_3, q') = p$ for some $p \in \mathbb{Q}^m \cap B_{2^{-r}}(x)$. Then

$$M(\pi_1\pi_2\pi_3, q) = U(\pi_3, q') = p \in \mathbb{Q}^m \cap B_{2^{-r}}(x)\,,$$

so

$$\begin{aligned}\hat{K}_r(x|q) &\leq K(p|q)\\ &\leq K_M(p|q) + c_M\\ &\leq |\pi_1\pi_2\pi_3| + c_M\,.\end{aligned}$$

By our choice of $\pi_1$, $\pi_2$, and $\pi_3$, and by (3.2.4), this implies that

$$\hat{K}_r(x|q) \leq K_{r,s+\Delta s}(x|y) + K(a) + K(s, \Delta s) + c_M\,. \tag{3.2.5}$$

We thus estimate K(a).

Since

$$\begin{aligned}|a| &= 2^{s^*}|q' - q|\\ &\leq 2^{s^*}(|q' - y| + |q - y|)\\ &< 2^{s^*}(s^{-(s+\Delta s)} + 2^{-s})\\ &= 2^{1+\lfloor \frac{1}{2}\log n\rfloor}(1 + 2^{\Delta s})\,,\end{aligned}$$

Observation 11 tells us that

$$\begin{aligned} K(a) &\leq n \log(1 + 2^{1+\lfloor \frac{1}{2} \log n \rfloor}(1 + 2^{\Delta s})) + \varepsilon(|a|) \\ &\leq n \log(2^{\Delta s + 3 + \lfloor \frac{1}{2} \log n \rfloor}) + \varepsilon(|a|)\,, \end{aligned}$$

i.e., that

$$K(a) \leq n\Delta s + 3n + n \left\lfloor \tfrac{1}{2} \log n \right\rfloor + \varepsilon(|a|), \tag{3.2.6}$$

where

$$\begin{aligned} \varepsilon(|a|) &\leq c + 2 \log\log(2 + 2^{1+\lfloor \frac{1}{2} \log n \rfloor}(1 + 2^{\Delta s})) \\ &\leq c + 2 \log\log(2^{\Delta s + 3 + \lfloor \frac{1}{2} \log n \rfloor}) \\ &= c + 2 \log(\Delta s + 3 + \left\lfloor \tfrac{1}{2} \log n \right\rfloor) \\ &\leq c + 2 \log\left((1 + \Delta s)(3 + \left\lfloor \tfrac{1}{2} \log n \right\rfloor)\right) \\ &= c + 2 \log(1 + \Delta s) + 2 \log(3 + \left\lfloor \tfrac{1}{2} \log n \right\rfloor)\,. \end{aligned}$$

It follows by (3.2.5) and (3.2.6) that (3.2.3) holds. □

**Theorem 15.** *Let $s : \mathbb{N} \to \mathbb{N}$. If $|s(r) - r| = o(r)$, then, for all $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$,*

$$\dim(x|y) = \liminf_{r \to \infty} \frac{K_{r,s(r)}(x|y)}{r}\,,$$

*and*

$$\mathrm{Dim}(x|y) = \limsup_{r \to \infty} \frac{K_{r,s(r)}(x|y)}{r}\,.$$

*Proof.* Assume the hypothesis. Define $s^-, s^+ : \mathbb{N} \to \mathbb{N}$ by

$$s^-(r) = \min\{r, s(r)\},\ s^+(r) = \max\{r, s(r)\}\,.$$

Lemma 14 tells us that, for all $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$,

$$\begin{aligned} K_{r,s^-(r)}(x|y) &\geq K_{r,r}(x|y) \\ &\geq K_{r,s^+(r)}(x|y) \\ &\geq K_{r,s^-(r)}(x|y) - O(s^+(r) - s^-(r)) - o(r) \\ &= K_{r,s^-(r)}(x|y) - O(|s(r) - r|) - o(r) \\ &= K_{r,s^-(r)}(x|y) - o(r)\,. \end{aligned}$$

Since

$$K_{r,s^-(r)}(x|y) \geq K_{r,s(r)}(x|y) \geq K_{r,s^+(r)}(x|y)\,,$$

it follows that

$$\left|K_{r,s(r)}(x|y) - K_{r,r}(x|y)\right| = o(r)\,.$$

The theorem follows immediately. □

## 3.3 Relationships to Mutual Information and Dimensions

**Theorem 16.** *For all $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$,*

$$I_r(x : y) = K_r(x) - K_r(x|y) + o(r)\,.$$

*Proof.* Let $B_x = B_{2^{-r}}(x) \cap \mathbb{Q}^m$ and $B_y = B_{2^{-r}}(y) \cap \mathbb{Q}^n$. Let $p_0$ and $q_0$ be $K$-minimizers for $B_x$ and $B_y$, respectively, such that

$$I_r(x : y) = I(p_0 : q_0) + o(r)\,. \tag{3.3.1}$$

These exist by Theorem 4.6 of [15]. Then

$$\begin{aligned} K_r(x) - K_r(x|y) &= \min_{p \in B_x} K(p) - \max_{q \in B_y} \min_{p \in B_x} K(p|q) \\ &\geq \min_{p \in B_x} K(p) - \min_{p \in B_x} \max_{q \in B_y} K(p|q) \\ &= \min_{p \in B_x} K(p) - \min_{p \in B_x} K(p|q_0) + o(r)\,, \end{aligned}$$

by Lemma 4.2 and Observation 3.7 of [15].

$$\begin{aligned} &= K(p_0) - \min_{p \in B_x} K(p|q_0) + o(r) \\ &\geq K(p_0) - K(p_0|q_0) + o(r) \\ &= I(p_0 : q_0) + o(r) \\ &= I_r(x : y) + o(r)\,. \end{aligned}$$

For the other direction, let $p_1 \in B_x$ be such that

$$K(p_1|q_0) = \min_{p \in B_x} K(p|q_0)\,.$$

By Lemma 4.5 of [15],

$$\begin{aligned} I(p_0 : q_0) &\geq K(p_1) - K(p_1|p_0, K(p_0)) - K(p_1|q_0, K(q_0)) + o(r) \\ &\geq K(p_1) - K(p_1|p_0, K(p_0)) - K(p_1|q_0) + o(r)\,. \end{aligned} \tag{3.3.2}$$

Now

$$\begin{aligned} K(p_0) + K(p_1|p_0, K(p_0)) + o(r) &= K(p_0, p_1) \\ &= K(p_1) + K(p_0|p_1, K(p_1)) + o(r) \\ &\leq K(p_1) + K(p_0|p_1) + o(r) \\ &= K(p_1) + o(r)\,, \end{aligned}$$

by Corollary 4.4 of [15]. So

$$K(p_1) - K(p_1|p_0, K(p_0)) \geq K(p_o) + o(r)\,,$$

thus by (3.3.2),

$$\begin{aligned} I(p_0 : q_0) &\geq K(p_0) - K(p_1|q_0) + o(r) \\ &= K(p_0) - \min_{p \in B_x} K(p|q_0) + o(r) \\ &= K_r(x) - \min_{p \in B_x} K(p|q_0) + o(r) \\ &\geq K_r(x) - \max_{q \in B_y} \min_{p \in B_x} K(p|q) + o(r) \\ &= K_r(x) - K_r(x|y) + o(r)\,. \end{aligned}$$

Then by (3.3.1), $I_r(x : y) \geq K_r(x) - K_r(x|y) + o(r)$, so equality holds. □

**Corollary 17.** *For all $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$, the following hold.*

1. $\mathrm{mdim}(x : y) \geq \dim(x) - \mathrm{Dim}(x|y)$.
2. $\mathrm{Mdim}(x : y) \leq \mathrm{Dim}(x) - \dim(x|y)$.

## 3.4 Chain Rules

**Theorem 18** (Chain rule for $K_r$)**.** *For all $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$,*

$$K_r(x, y) = K_r(x|y) + K_r(y) + o(r)\,.$$

*Proof.* Theorem 4.10 of [15] tells us that

$$I_r(x : y) = K_r(x) + K_r(y) - K_r(x, y) + o(r)\,.$$

Combining this with Theorem 16, we have

$$K_r(x) + K_r(y) - K_r(x, y) + o(r) = K_r(x) - K_r(x|y) + o(r)\,.$$

The theorem follows immediately. □

**Corollary 19** (Chain rule for dimension)**.** *For all $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$,*

$$\begin{aligned}\dim(x) + \dim(y|x) &\leq \dim(x, y) \\ &\leq \dim(x) + \mathrm{Dim}(y|x) \\ &\leq \mathrm{Dim}(x, y) \\ &\leq \mathrm{Dim}(x) + \mathrm{Dim}(y|x)\,.\end{aligned}$$

## 3.5 Relationships to Relative Kolmogorov Complexity and Dimensions

Recall from Chapter 2 that the Kolmogorov complexity of $x \in \mathbb{R}^m$ at precision $r$ *relative* to another point $y \in \mathbb{R}^n$ is defined as $K_r^y(x) := K_r^{A_y}(x)$, where $A_y \subseteq \mathbb{N}$ is an oracle that encodes the binary expansions of $y$'s coordinates, and $\dim^y(x)$ and $\mathrm{Dim}^y(x)$ are defined accordingly. The following lemma reflects the intuition that oracle access to $y$ is at least as useful as any bounded-precision estimate for $y$.

**Lemma 20.** *For each $m, n \in \mathbb{N}$ there is a constant $c \in \mathbb{N}$ such that, for all $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$, and $r, s \in \mathbb{N}$,*

$$K_r^y(x) \leq K_{r,s}(x|y) + K(s) + c\,.$$

*In particular, $K_r^y(x) \leq K_r(x|y) + K(r) + c$.*

*Proof.* Let $m, n \in \mathbb{N}$, and let $U$ be the optimal Turing machine fixed for the definition of conditional Kolmogorov complexity. Let $M$ be an oracle Truing machine that, on input $\pi \in \{0,1\}^*$ with oracle $g : \mathbb{N} \to \mathbb{Q}^n$, does the following. If $\pi$ is of the form $\pi = \pi_1\pi_2$, where $U(\pi_1, \lambda) = t \in \mathbb{N}$, then $M$ simulates $U(\pi_2, g(t))$. Let $c$ be an optimality constant for the oracle Turing machine $M$.

To see that $c$ affirms the lemma, let $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$, and $r, s \in \mathbb{N}$. Let $q = y \upharpoonright (s + \log\sqrt{n})$, the truncation of the binary expansions of each of $y$'s coordinates to $s + \log\sqrt{n}$ bits to the right of the binary point. Let $\pi_s \in \{0,1\}^*$ testify to the value of $K(s)$, and let

$\pi_x$ testify to the value of $\hat{K}_r(x|q)$. Then

$$q \in \mathbb{Q}^n \cap B_{2^{-s}}(y)$$

and

$$M^y(\pi_s \pi_x) = U(\pi_x, q) \in \mathbb{Q}^m \cap B_{2^{-r}}(x)\,,$$

so

$$\begin{aligned} K_r^y(x) &\leq K_{M,r}^y(x) + c \\ &\leq |\pi_s \pi_x| + c \\ &= \hat{K}_r(x|q) + K(s) + c \\ &\leq K_{r,s}(x|y) + K(s) + c\,. \end{aligned}$$

□

**Corollary 21.** *For all $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$,*

$$\dim^y(x) \leq \dim(x|y) \quad \textit{and} \quad \mathrm{Dim}^y(x) \leq \mathrm{Dim}(x|y)\,.$$

# Chapter 4

# Points on Random Lines and Plane Kakeya Sets

As a first illustration of the power of the point-to-set principle for Hausdorff dimension (Theorem 2), we use it to give a new proof of a known theorem in geometric measure theory.

**Definition.** A *Kakeya set* in a Euclidean space $\mathbb{R}^n$ is a set $K \subseteq \mathbb{R}^n$ that contains a unit line segment in every direction. That is, for every point $z$ on the unit sphere $S^{n-1}$, there exists $t \in \mathbb{R}^n$ such that $\{\alpha z + t : \alpha \in (0,1)\} \subseteq K$.

Besicovitch [8, 9] proved that Kakeya sets can have Lebesgue measure 0 and asked whether Kakeya sets in the Euclidean plane can have dimension less than 2 [22]. The famous Kakeya conjecture asserts a negative answer to this and to the analogous question in higher dimensions, i.e., states that every Kakeya set in a Euclidean space $\mathbb{R}^n$ has Hausdorff dimension $n$.[1] This conjecture holds trivially for $n = 1$ and was proven by Davies [22] for $n = 2$. A version of the conjecture in finite fields has been proven by Dvir [29]. For Euclidean spaces of dimension $n \geq 3$, it is an important open problem with deep connections to other problems in analysis [107, 102].

In this chapter we use Theorem 2 to give a new proof of Davies's theorem. This proof does not resemble the classical proof, which is not difficult but relies on Marstrand's projection theorem [74] and point-line duality. Instead of analyzing the set $K$ globally, our proof focuses on the information content of a single, judiciously chosen point in $K$. Given a

[1] Statements of the Kakeya conjecture vary in the literature. For example, the set is sometimes required to be compact or Borel, and the dimension used may be Minkowski instead of Hausdorff. Since the Hausdorff dimension of a set is never greater than its Minkowski dimension, our formulation is at least as strong as those variations.

Kakeya set $K \subseteq \mathbb{R}^2$ and an oracle $A \subseteq \mathbb{N}$, we first choose a particular line segment $L \subseteq K$ and a particular point $(x, mx + b) \in L$, where $y = mx + b$ is the equation of the line containing $L$. One might naïvely expect that for independently random $m$ and $x$, the point $(x, mx + b)$ must be random. In fact, this is not the case: in every direction, there is a line that contains no random point [65]. Nevertheless, we show that $\dim^A(x, mx + b) \geq 2$. By our point-to-set principle, this implies that $\dim_H(K) \geq 2$.

**Theorem 22** (Davies [22])**.** *Every Kakeya set in $\mathbb{R}^2$ has Hausdorff dimension 2.*

Our new proof of Theorem 22 uses a relativized version of the following lemma.

**Lemma 23.** *Let $m \in [0, 1]$ and $b \in \mathbb{R}$. Then for almost every $x \in [0, 1]$,*

$$\liminf_{r\to\infty} \frac{K_r(m, b, x) - K_r(b|m)}{r} \leq \dim(x, mx + b)\,. \tag{4.0.1}$$

*Proof.* We build a program that takes as input a precision level $r$, an approximation $p$ of $x$, an approximation $q$ of $mx+b$, a program $\pi$ that will approximate $b$ given an approximation for $m$, and a natural number $h$. In parallel, the program considers each multiple of $2^{-r}$ in [0,1] as a possible approximate value $u$ for the slope $m$, and it checks whether each such $u$ is consistent with the program's inputs. If $u$ is close to $m$, then $\pi(u)$ will be close to $b$, so $up + \pi(u)$ will be close to $mx + b$. Any $u$ that satisfies this condition is considered a "candidate" for approximating $m$.

Some of these candidates may be "false positives," in that there can be values of $u$ that are far from $m$ but for which $up + \pi(u)$ is still close to $mx + b$. Thus the program is also given an input $h$ so that it can choose the correct candidate; it selects the $h^{\text{th}}$ candidate that arises in its execution. We will show that this $h$ is often not large enough to significantly affect the total input length.

Formally, let $M$ be a Turing machine that runs the following algorithm on input $\rho\pi\sigma\eta$ whenever $U(\rho) = r \in \mathbb{N}$, $U(\eta) = h \in \mathbb{N}$, and $U(\sigma) = (p, q) \in \mathbb{Q}^2$:

$candidate \leftarrow 0$

for $i = 0, 1, \ldots, 2^r$, in parallel:

$u_i \leftarrow 2^{-r} i$

$v_i \leftarrow U(\pi, u_i)$

do atomically:

if $v_i \in \mathbb{R}$ and $|u_i p + v_i - q| < 2^{2-r}$, then $\mathit{candidate} \leftarrow \mathit{candidate} + 1$

if $\mathit{candidate} = h$, then return $(u_i, v_i, p)$ and halt

Fix $m \in [0, 1]$ and $b \in \mathbb{R}$. For each $r \in \mathbb{N}$, let $m_r = 2^{-r} \lfloor m \cdot 2^r \rfloor$, and fix $\pi_r$ testifying to the value of $\hat{K}_r(b|m_r)$ and $\sigma_r$ testifying to the value of $K_r(x, mx + b)$.

We complete the proof with the following four claims. Intuitively, Claim 1 says that no point in $B_{2^{-r}}(m)$ gives much less information about $b$ than $m_r$ does. Claim 2 states that there is always some value of $h$ that causes this machine to return the desired output. Claim 3 says that for almost every $x$, this value does not grow too quickly with $r$, and Claim 4 says that (4.0.1) holds for every such $x$.

*Claim* 1. For every $r \in \mathbb{N}$, $K_r(b|m) = \hat{K}_r(b|m_r) + o(r)$.

*Proof.* $K_r(b|m) \geq \hat{K}_r(b|m_r)$ by definition, since $m_r \in B_{2^{-r}}(m)$.

Let $\hat{b} \in B_{2^{-r}}(b)$ be such that $K(\hat{b}|m_r) = \hat{K}_r(b|m_r)$. Then

$$|(\hat{b}, m_r) - (b, m)| \leq \sqrt{2} \cdot 2^{-r} < 2^{1-r},$$

so

$$K(\hat{b}, m_r) \geq K_{r-1}(b, m) = K_r(b, m) + o(r),$$

by Corollary 3.9 of [15].

Let $\mu$ testify to the value of $K_r(m)$, and let $\hat{m} = U(\mu)$. Then $|\hat{m} - m| < 2^{-r}$, so $|\hat{m} - m_r| < 2^{1-r}$. Thus once $\hat{m}$ and $r$ have been specified, there are at most four possible values for $m_r$. Therefore there is a self-delimiting Turing machine that takes as input $\mu$, an encoding of $r$ of length $o(r)$, and $O(1)$ additional bits and outputs $m_r$. We conclude that

$K(m_r) \leq K_r(m) + o(r)$. Therefore we have

$$\begin{aligned}
\hat{K}_r(b|m_r) &= K(\hat{b}|m_r) \\
&= K(\hat{b}, m_r) - K(m_r) + o(r) \\
&\geq K_r(b, m) + o(r) - (K_r(m) + o(r)) + o(r) \\
&= K_r(b|m) + o(r)\,,
\end{aligned}$$

by Theorem 18. □

*Claim* 2. For each $x \in [0,1]$ and $r \in \mathbb{N}$, there exists an $h \in \mathbb{N}$ such that

$$M(\rho\pi_r\sigma_r\eta) \in B_{2^{1-r}}(m, b, x)\,,$$

where $U(\rho) = r$ and $U(\eta) = h$.

*Proof.* Fix $x \in [0,1]$ and $r \in \mathbb{N}$. It is clear that for some $j \in \{0, 1, \ldots, 2^r\}$, $|u_j - m| < 2^{-r}$. By the definition of $K_r(b|m)$, $u_j \in \mathbb{Q} \cap B_{2^{-r}}(m)$ implies that $U(\pi_r, u_j)$ halts and outputs $v_j \in \mathbb{Q} \cap B_{2^{-r}}(b)$. $U(\sigma_r) \in B_{2^{-r}}(x, mx + b)$ by the definition of $\sigma_r$, so $|p - x| < 2^{-r}$. It follows that

$$|(u_j, v_j, p) - (m, x, b)| < \sqrt{3(2^{-r})^2} < 2^{1-r}\,.$$

It remains to show that $|u_i p + v_j - q| < 2^{2-r}$. To do so, we repeatedly apply the triangle inequality and use the fact that $x, m \in [0,1]$:

$$\begin{aligned}
|u_i p + v_j - q| &\leq |u_i p + v_j - (mx + b)| + |mx + b - q| \\
&< |u_j p - mx + v_j - b| + 2^{-r} \\
&\leq |u_j p - mx| + |v_j - b| + 2^{-r} \\
&< |u_j p - u_j x| + |u_j x - mx| + 2^{1-r} \\
&\leq |p - x| + |u_j - m| + 2^{1-r} \\
&< 2^{2-r}\,.
\end{aligned}$$

□

For every $x \in [0,1]$ and $r \in \mathbb{N}$, define $h(x,r)$ to be the minimal $h$ satisfying the conditions of Claim 2.

*Claim* 3. For almost every $x \in [0,1]$, $\log(h(x,r)) = o(r)$.

*Proof.* By the countable additivity of Lebesgue measure, it suffices to show for every $k \in \mathbb{N}$ that the set

$$D_k = \{x \in [0,1] : \exists \text{ infinitely many } r \in \mathbb{N} \text{ such that } \log(h(x,r)) > r/k\}$$

has Lebesgue measure 0. For each $r \in \mathbb{N}$, let $D_{k,r} = \{x : h(x,r) > 2^{r/k}\}$. We now estimate $\lambda(D_{k,r})$, the Lebesgue measure of $D_{k,r}$.

For fixed $x$ and $r$, the algorithm run by the Turing machine $M$ entails

$$h(x,r) \leq \left|\left\{i : |u_i p + v_i - q| < 2^{2-r}\right\}\right| .$$

For fixed $i$,

$$\begin{aligned} |u_i p + v_i - q| &> |u_i x - u_i p| + |u_i p + v_i - q| - 2^{-r} \\ &\geq |u_i x + v_i - q| - 2^{-r} \\ &> |u_i x + v_i - q| + |q - (mx+b)| - 2^{1-r} \\ &\geq |u_i x + v_i - (mx+b)| - 2^{1-r} . \end{aligned}$$

That is,

$$\left\{i : |u_i p + v_i - q| < 2^{2-r}\right\} \subseteq \left\{i : |u_i x + v_i - (mx+b)| - 2^{1-r} < 2^{2-r}\right\} ,$$

so

$$h(x,r) \leq \left|\left\{i : |u_i x + v_i - (mx+b)| < 2^{3-r}\right\}\right| .$$

For fixed $r$ and $i = 0, 1, \ldots, 2^r$, define

$$C_i^r = \{x \in [0,1] : |u_i x + v_i - (mx+b)| < 2^{3-r}\} ,$$

For each $i$, if $m = u_i$, then $C_i^r$ is either $[0,1]$ or empty; otherwise, $C_i^r$ is an interval of length

$$\lambda(C_i^r) \leq \min\left\{\frac{2^{3-r}}{|u_i - m|}, 1\right\} .$$

Notice that for each $k = 0, \ldots, 2^r$, there are at most 2 values of $i$ for which

$$2^{-r}k \leq |u_i - m| < 2^{-r}(k+1)\,,$$

so we have

$$\begin{aligned}\int_0^1 h(x,r)dx &\leq \sum_{i=0}^{2^r} \lambda(C_i^r) \\ &\leq 2 + \sum_{k=1}^{2^r} 2\frac{2^{3-r}}{2^{-r}k} \\ &= 2 + 2^4 \sum_{k=1}^{2^r} \frac{1}{k} \\ &< r2^6\,.\end{aligned}$$

Thus, as $h(x,r) > 2^{r/k}$ for all $x \in D_{k,r}$,

$$\lambda(D_{k,r}) < \frac{r2^6}{2^{r/k}} = r2^{6-r/k}\,.$$

This implies that

$$\sum_{r=1}^{\infty} \lambda(D_{k,r}) < \infty\,,$$

so the Borel-Cantelli Lemma tells us that $\lambda(D_k) = 0$. □

*Claim* 4. For every $x \in [0,1]$, if $\log(h(x,r)) = o(r)$, then

$$\liminf_{r\to\infty} \frac{K_r(m,b,x) - K_r(b|m)}{r} \leq \dim(x, mx+b)\,.$$

*Proof.* For fixed $r$, Claim 2 gives

$$K_{r-1}(m,b,x) \leq K(u_i, v_i, p) \leq K_M(u_i, v_i, p) + c_M\,,$$

where $c_M$ is an optimality constant for $M$. Let $\rho$ and $\eta$ testify to the values of $K(r)$ and $K(h(x,r))$, respectively. Then $K_M(u_i, v_i, p) \leq |\rho\pi_r\sigma_r\eta|$. By our choices of $\rho, \pi_r, \sigma_r$, and $\eta$,

$$\begin{aligned}|\rho\pi_r\sigma_r\eta| &= K(r) + \hat{K}_r(b|m_r) + K_r(x, mx+b) + K(h(x,r)) \\ &= K(r) + K_r(b|m) + K_r(x, mx+b) + K(h(x,r)) + o(r)\,,\end{aligned}$$

by Claim 1. By Corollary 3.9 of [15],

$$\begin{aligned}
&\liminf_{r\to\infty} \frac{K_r(m,b,x) - K_r(b|m)}{r} \\
&= \liminf_{r\to\infty} \frac{K_{r-1}(m,b,x) - K_r(b|m) + o(r)}{r} \\
&\leq \liminf_{r\to\infty} \frac{K(r) + K_r(x, mx+b) + K(h(x,r)) + o(r)}{r} \\
&\leq \liminf_{r\to\infty} \frac{K_r(x,mx+b)}{r} + \limsup_{r\to\infty} \frac{K(r) + K(h(x,r)) + o(r)}{r} \\
&= \dim(x, mx+b) + \limsup_{r\to\infty} \frac{K(h(x,r))}{r}\,.
\end{aligned}$$

Applying Observation 10, for some constant $c$,

$$\begin{aligned}
\limsup_{r\to\infty} \frac{K(h(x,r))}{r} &\leq \limsup_{r\to\infty} \frac{\log(1+h(x,r)) + 2\log\log(2+h(x,r)) + c}{r} \\
&= \limsup_{r\to\infty} \frac{\log(h(x,r)) + 2\log\log(h(x,r))}{r}\,.
\end{aligned}$$

If $\log(h(x,r)) = o(r)$, then this is

$$\limsup_{r\to\infty} \frac{o(r) + 2\log(o(r))}{r} = 0\,.$$

□

The lemma follows immediately from Claims 3 and 4. □

*Proof of Theorem 22.* Let $K$ be a Kakeya set in $\mathbb{R}^2$. By Theorem 2, there exists an oracle $A$ such that

$$\dim_H(K) = \sup_{p\in K} \dim^A(p)\,.$$

Let $m \in [0,1]$ such that $\dim^A(m) = 1$; such an $m$ exists by Theorem 4.5 of [64]. $K$ contains a unit line segment $L$ of slope $m$. Let $(x_0, y_0)$ be the left endpoint of such a segment. Let $q \in \mathbb{Q} \cap [x_0, x_0 + 1/8]$, and let $L'$ be the unit segment of slope $m$ whose left endpoint is $(x_0 - q, y_0)$. Let $b = y_1 + qm$, the $y$-intercept of $L'$.

By a relativized version of Lemma 23, there is some $x \in [0, 1/2]$ such that $\dim^{A,m,b}(x) = 1$ and

$$\liminf_{r\to\infty} \frac{K_r^A(m,b,x) - K_r^A(b|m)}{r} \leq \dim^A(x, mx+b)\,.$$

This holds because almost every $x \in [0, 1/2]$ is algorithmically random relative to $(A, m, b)$ and hence satisfies $\dim^{A,m,b}(x) = 1$.

Fix such an $x$, and notice that $(x, mx + b) \in L'$. Now, applying a relativized version of Theorem 18,

$$\begin{aligned}
\dim^A(x, mx + b) &\geq \liminf_{r\to\infty} \frac{K_r^A(m, b, x) - K_r^A(b|m)}{r} \\
&= \liminf_{r\to\infty} \frac{K_r^A(m, b, x) - K_r^A(b, m) + K_r^A(m)}{r} \\
&= \liminf_{r\to\infty} \frac{K_r^A(x|b, m) + K_r^A(m)}{r} \\
&\geq \liminf_{r\to\infty} \frac{K_r^A(x|b, m)}{r} + \liminf_{r\to\infty} \frac{K_r^A(m)}{r} .
\end{aligned}$$

By Lemma 20, $K_r^A(x|b, m) \geq K_r^{A,b,m}(x) + o(r)$, so we have

$$\begin{aligned}
\dim^A(x, mx + b) &\geq \liminf_{r\to\infty} \frac{K_r^{A,b,m}(x)}{r} + \liminf_{r\to\infty} \frac{K_r^A(m)}{r} \\
&= \dim^{A,b,m}(x) + \dim^A(m) ,
\end{aligned}$$

which is 2 by our choices of $m$ and $x$.

By Observation 12,

$$\dim^A(x, mx + b) = \dim^A(x + q, mx + b) .$$

Hence, there exists a point $(x + q, mx + b) \in K$ such that $\dim^A(x + q, mx + b) \geq 2$. By Theorem 2, the point-to-set principle for Hausdorff dimension, this completes the proof. □

It is natural to ask what prevents us from extending this proof to $\mathbb{R}^n$ for all $n \geq 2$. The point of failure in a direct extension would be Claim 3 in the proof of Lemma 23. Speaking informally, the problem is that the total number of candidates may grow as $2^{(n-1)r}$, meaning that $\log(h(x, r))$ could be $\Omega((n - 2)r)$ for every $x$.

# Chapter 5

# Dimension Spectra and Generalized Furstenberg Sets

This chapter focuses once again on the effective Hausdorff dimension $\dim(z)$ of individual points $z \in \mathbb{R}^n$. Given the pointwise nature of this quantity, it is natural to investigate the *dimension spectrum* of a set $E \subseteq \mathbb{R}^n$, i.e., the set $\{\dim(z) : z \in E\}$. Even for apparently simple sets, the structure of the dimension spectrum may not be obvious, as exemplified by a longstanding open question originally posed by J. H. Lutz [80]: *Is there a straight line $L \subseteq \mathbb{R}^2$ such that every point on $L$ has effective Hausdorff dimension 1?*

J. H. Lutz and Weihrauch [68] have shown that the set of points in $\mathbb{R}^n$ with dimension less than 1 is totally disconnected, as is the set of points with dimension greater than $n-1$. Turetsky has shown that the set of points in $\mathbb{R}^n$ of dimension exactly 1 is connected [104], which implies that every line in $\mathbb{R}^2$ contains a point of dimension 1. As we saw in Chapter 4, J. H. Lutz and N. Lutz have shown that almost every point on any line with random slope has dimension 2 [66], despite the surprising fact, shown by the same authors, that there are lines in every direction that contain no random points [65]. These results give insight into the dimension spectra of lines, but they also leave open the question of whether or not a line in $\mathbb{R}^2$ can have a singleton dimension spectrum.

We resolve this question in the negative with the following theorem, a general lower bound on the dimension of points on lines in $\mathbb{R}^2$. Our bound depends only on the dimension of the description $(a, b)$ of the line (i.e., the ordered pair giving the line's slope and vertical intercept) and the dimension of the coordinate $x$ relative to $(a, b)$.

**Theorem 24.** *For all $a, b, x \in \mathbb{R}$,*

$$\dim(x, ax+b) \geq \dim^{a,b}(x) + \min\left\{ \dim(a,b),\, \dim^{a,b}(x) \right\}.$$

*In particular, for almost every $x \in \mathbb{R}$, $\dim(x, ax+b) = 1 + \min\{\dim(a,b), 1\}$.*

Since $\dim(0,b) \leq \min\{\dim(a,b), 1\}$, the second statement implies that every line contains two points whose dimensions differ by at least 1, and therefore that the dimension spectrum cannot be a singleton.

Furthermore, we apply the point-to-set principle for Hausdorff dimension (Theorem 2) to derive a new result in classical fractal geometry from the above theorem. Recall from Chapter 1 that Molter and Rela [81] defined $(\alpha, \beta)$-generalized sets of Furstenberg type, sets in $\mathbb{R}^2$ which contain $\alpha$-dimensional subsets of lines in all of a $\beta$-dimensional set of directions, for some parameters $\alpha, \beta \in (0, 1]$. They also showed that any such set has Hausdorff dimension at least $\alpha + \max\{\beta/2, \alpha + \beta - 1\}$. In Theorem 37, we give a lower bound of $\alpha + \min\{\beta, \alpha\}$, which constitutes an improvement whenever $\alpha, \beta < 1$ and $\beta/2 < \alpha$.

The arguments in this chapter require us to work with binary expansions of reals, so we begin in Section 5.1 by showing how those representations interact with the definitions of Chapters 2 and 3. In Section 5.2, we use the results of Section 5.1 to give chain rules that are slightly more precise than Theorem 18. We discuss and prove our bound on $\dim(x, ax+b)$ in Section 5.3, and we apply it to generalized Furstenberg sets in Section 5.4.

## 5.1 Initial Segments versus $K$-optimizing Rationals

In this section we formalize the relationship between $K_r(x)$ and the *initial segment complexity* $K(x\restriction r)$. The three lemmas in this section are proved by standard techniques.

For $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and $r \in \mathbb{N}$, let $x\restriction r = (x_1\restriction r, \ldots, x_n\restriction r)$, where each $x_i\restriction r = 2^{-r}\lfloor 2^r x_i \rfloor$, the truncation of $x_i$ to $r$ bits to the right of the binary point. For $r \in (0, \infty)$, let $x\restriction r = x\restriction \lceil r \rceil$.

**Lemma 25.** *For every $m, n \in \mathbb{N}$, there is a constant $c$ such that for all $x \in \mathbb{R}^m$, $p \in \mathbb{Q}^n$, and $r \in \mathbb{N}$,*

$$|\hat{K}_r(x|p) - K(x\restriction r \,|\, p)| \leq K(r) + c\,.$$

*Proof.* Let $m, n, r \in \mathbb{N}$, $x \in \mathbb{R}^m$, and $p \in \mathbb{Q}^n$. Observe that $x\restriction r \in B_{2^{-r}\sqrt{m}}(x)$, and therefore $K(x\restriction r \,|\, p) \geq \hat{K}_{r-\log(m)/2}(x|p)$. Thus, by Lemma 13, there exists $c_1 \in \mathbb{N}$ depending only on $m$ such that

$$\hat{K}_r(x|p) \leq K(x\restriction r \,|\, p) + K(r) + c_1\,.$$

For the other direction, observe that for every $q \in \mathbb{Q}^n \cap B_{2^{-r}}(x)$, we have $x\restriction r \in B_{2^{-r}(1+\sqrt{m})}(q)$, and that $B_{2^{-r}(1+\sqrt{m})}(q)$ contains at most $(2(1+\sqrt{m}))^m$ $r$-dyadic points, i.e., points in the set

$$\mathcal{Q}_t^m = \{2^{-r}z : z \in \mathbb{Z}^m\}\,.$$

Let $M$ be a Turing machine that, on input $(\pi, p') \in \{0,1\}^* \times \mathbb{Q}^n$, does the following. If $\pi = \pi_1\pi_2\pi_3$, with $U(\pi_1, p') = q \in \mathbb{Q}^m$, $U(\pi_2) = t \in \mathbb{N}$, and $U(\pi_3) = k \in \mathbb{N}$, then $M$ outputs the (lexicographically) $k^{\text{th}}$ point in $\mathcal{Q}_r^m \cap B_{2^{-t}(1+\sqrt{m})}(q)$.

Now let $\pi_q$ testify to $\hat{K}_r(x|p)$, let $\pi_r$ testify to $K(r)$, and let $q = U(\pi_q, p)$. There is some $k \leq (2(1+\sqrt{m}))^m$ such that $x\restriction r$ is the $k^{\text{th}}$ point in $\mathcal{Q}_r^m \cap B_{2^{-r}(1+\sqrt{m})}(q)$; let $\pi_k$ testify to $K(k)$. Then $M(\pi_q\pi_r\pi_k, p) = x\restriction r$, so there is some machine constant $c_M$ for $M$ such that

$$\begin{aligned} K(x\restriction r \,|\, p) &\leq \ell(\pi_q) + \ell(\pi_r) + \ell(\pi_k) + c_M \\ &= \hat{K}_r(x|p) + K(x) + K(k) + c_M \end{aligned}$$

It is well known (see, e.g., [28]) that there is some constant $c_2$ such that

$$\begin{aligned} K(k) &\leq \log k + 2\log\log k + c_2 \\ &\leq m\log(2(1+\sqrt{m})) + 2\log(m\log(2(1+\sqrt{m}))) + c_2\,. \end{aligned}$$

The above value depends only on $m$, as does $c_M$; let $c_3$ be their sum. Then

$$K(x\restriction r \,|\, p) \leq \hat{K}_r(x|p) + K(r) + c_3\,,$$

so $c = \max\{c_1, c_3\}$ affirms the lemma. □

Observing that there exists a constant $c_0$ such that, for all $m \in \mathbb{N}$ and $q^m \in \mathbb{Q}$, $|K(q) - K(q|0)| \leq c_0$, we also have the following.

**Corollary 26.** *For every $m \in \mathbb{N}$, there is a constant $c$ such that for every $x \in \mathbb{R}^m$ and $r \in \mathbb{N}$,*

$$|K_r(x) - K(x\restriction r)| \leq K(r) + c\,.$$

**Corollary 27.** *For every $m, n \in \mathbb{N}$, there is a constant $c$ such that for all $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$, and $r, s \in \mathbb{N}$,*

$$|K_{r,s}(x|y) - K(x\restriction r \,|\, y\restriction s)| \leq K(r) + K(s) + c\,.$$

*Proof.* Let $m, n, r, s \in \mathbb{N}$, $x \in \mathbb{R}^m$, and $y \in \mathbb{R}^n$. Let $p \in \mathbb{Q}^2 \cap B_{2^{-s}}(y)$ be such that $K_{r,s}(x|y) = \hat{K}_r(x|p)$. Since $y\restriction s \in B_{2^{-s}\sqrt{n}}(y)$, we have $\hat{K}_r(x \,|\, y\restriction s) \geq K_{r,s-\log(n)/2}(x|y)$. Thus, by Lemma 14 there is a constant $c_1$ (depending on $n$) such that $\hat{K}_r(x \,|\, y\restriction s) \geq K_{r,s}(x|y) - K(s) - c_1$. Lemma 25 tells us that there is a constant $c_2$ (depending on $m$) such that $K(x\restriction r \,|\, y\restriction s) \geq \hat{K}_r(x \,|\, y\restriction s) - K(r) - c_2$, so we have

$$K_{r,s}(x|y) \leq K(x\restriction r \,|\, y\restriction s) + K(r) + K(s) + c_1 + c_2\,.$$

For the other direction, we use essentially the same technique as was used in the proof of Lemma 25, and we describe a Turing machine $M'$ that is very similar to the machine $M$ used above. On every input $(\pi, p') \in \{0,1\}^* \times \mathbb{Q}^n$ such that $\pi = \pi_1\pi_2\pi_3$, $U(\pi_1, p') = q \in \mathbb{Q}$, $U(\pi_2) = t \in \mathbb{N}$, and $U(\pi_3) = k \in \mathbb{N}$, $M'$ outputs $U(\pi_1, q')$, where $q'$ is the $k^{\text{th}}$ point in $\mathcal{Q}^n_t \cap B_{2^{-t}(1+\sqrt{n})}(p')$.

Much as before, let $\pi_x$ testify to $K(x\restriction r \,|\, y\restriction s)$, let $\pi_s$ testify to $K(s)$, and let $\pi_k$ testify to $K(k)$, where $y\restriction s$ is the $k^{\text{th}}$ point in $\mathcal{Q}^n_s \cap B_{2^{-t}(1+\sqrt{n})}(p)$. Then

$$M'(\pi_x, \pi_s, \pi_k) = U(\pi_x, y\restriction s) = x\restriction r\,,$$

As $k \leq |\mathcal{Q}_s^n \cap B_{2^{-t}(1+\sqrt{n})}(p)| \leq (2(1+\sqrt{n}))^n$, there exist constants $c_{M'}$ and $c_k$ (depending on $n$) such that

$$\begin{aligned} K(x\restriction r \,|\, p) &\leq \ell(\pi_x) + \ell(\pi_s) + \ell(\pi_k) + c_{M'} \\ &= K(x\restriction r \,|\, y\restriction s) + K(s) + K(k) + c_{M'} \\ &= K(x\restriction r \,|\, y\restriction s) + K(s) + c_k + c_{M'}\,, \end{aligned}$$

Applying Lemma 25 again, there is a constant $c_3$ (depending on $m$) such that $K(x\restriction r\,|p) \leq \hat{K}_r(x|p) + K(r) + c_3$. We conclude that

$$K(x\restriction r \,|\, y\restriction s) \leq K(r) + K(s) + c_k + c_{M'} + c_3\,,$$

therefore $c = \max\{c_1 + c_2, c_k + c_{M'} + c_3\}$ affirms the lemma. □

## 5.2 Approximate Symmetry of Information

Using the results of Section 5.1, it is straightforward to show that approximate symmetry of information holds for Kolmogorov complexity in Euclidean spaces.

**Lemma 28.** *For every $m, n \in \mathbb{N}$, $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$, and $r, s \in \mathbb{N}$ with $r \geq s$,*

(i) $\big|K_r(x|y) + K_r(y) - K_r(x,y)\big| \leq O_{m,n}(\log r) + O_n(\log\log \|y\|) + O_{m,n}(1)\,.$

(ii) $|K_{r,s}(x|x) + K_s(x) - K_r(x)| \leq O_m(\log r) + O_m(\log\log \|x\|) + O_m(1)\,.$

*Proof.* For (i), let $m, n, r \in \mathbb{N}$, $x \in \mathbb{R}^m$, and $y \in \mathbb{R}^n$. By Corollary 26,

$$|K_r(y) - K(y\restriction r)| \leq K(r) + O_n(1)$$

and

$$|K_r(x,y) - K((x,y)\restriction r)| \leq K(r) + O_{m,n}(1)\,.$$

Notice that $K((x,y)\restriction r) = K(x\restriction r, y\restriction r)$. By Corollary 27,

$$|K_r(x|y) - K(x\restriction r \,|\, y\restriction r)| \leq 2K(r) + O_{m,n}(1)\,.$$

By the symmetry of information,

$$K(x\restriction r \,|\, y\restriction r, K(y\restriction r)) + K(y\restriction r) - K(x\restriction r, y\restriction r) = O_{m,n}(1)\,.$$

It is also true that

$$\begin{aligned} |K(x\restriction r \,|\, y\restriction r) - K(x\restriction r \,|\, y\restriction r, K(y\restriction r))| &\leq K(K(y\restriction r)) + O_{m,n}(1) \\ &\leq \log K(y\restriction r) + 2\log\log K(y\restriction r) + O_{m,n}(1) \\ &= O_n(\log r) + O_n(\log\log \|y\|) + O_{m,n}(1)\,. \end{aligned}$$

The second term is necessary because the integer part of $y$ is not included in the truncation length $r$. In sum,

$$\begin{aligned} |K_r(x|y) + K_r(y) - K_r(x,y)| &\leq 4K(r) + K(K(y\restriction r)) + O_{m,n}(1) \\ &\leq O_{m,n}(\log r) + O_n(\log\log \|y\|) + O_{m,n}(1)\,. \end{aligned}$$

The argument for (ii) is nearly identical; the only additional error is due to

$$\begin{aligned} K(x\restriction r, x\restriction s) - K(x\restriction r) &\leq K(s) + O_m(1) \\ &\leq \log r + 2\log\log r + O_m(1)\,, \end{aligned}$$

as $s \leq r$. □

## 5.3 Bounding the Dimension of $(x, ax+b)$

In this section we prove Theorem 24, our general bound on the dimension of points on arbitrary lines in $\mathbb{R}^2$. We first restate the theorem in the form we will prove, which is slightly stronger than its statement above. The dimension of $x$ in the first term is conditioned on—instead of relative to—$(a,b)$, and even when working relative to an arbitrary oracle $A$, the last term $\dim^{a,b}(x)$ remains unchanged.

**Theorem 24.** (Restated) *For every $a, b, x \in \mathbb{R}$ and $A \subseteq \mathbb{N}$,*

$$\dim^A(x, ax+b) \geq \dim^A(x|a,b) + \min\left\{\dim^A(a,b),\, \dim^{a,b}(x)\right\}.$$

*In particular, for almost every $x \in \mathbb{R}$, $\dim(x, ax+b) = 1 + \min\{\dim(a,b), 1\}$.*

To prove this theorem, we proceed in three major steps, which we first sketch at a very high level here. In Section 5.3.1, we give sufficient conditions, at a given precision $r$, for a point $(x, ax+b)$ to have information content $K_r(x, ax+b)$ approaching $K_r(a, b, x)$. Notice that this is essentially the maximum possible value for $K_r(x, ax+b)$, since an estimate for $(a, b, x)$ has enough information to estimate $(x, ax+b)$ to similar precision. Informally, the conditions are

(i) $K_r(a, b)$ is small.

(ii) If $ux + v = ax + b$, then either $K_r(u, v)$ is large or $(u, v)$ is close to $(a, b)$.

We show in Lemma 29 that when these conditions hold, we can algorithmically estimate $(a, b, x)$ given an estimate for $(x, ax+b)$. In Section 5.3.2, we give a lower bound, Lemma 31, on $K_r(u, v)$ in terms of $\|(u, v) - (a, b)\|$, essentially showing that condition (ii) holds. In Section 5.3.3 we construct oracles with some desirable properties. Finally, we prove Theorem 24 in Section 5.3.4 by showing that these oracles allow $(a, b)$ to satisfy condition (i) without disrupting condition (ii) or too severely lowering $K_r(x, a, b)$.

### 5.3.1 Sufficient Conditions for a High-complexity Point

Suppose that $x$, $a$, and $b$ satisfy conditions (i) and (ii) above. Then, given an estimate $q$ for the point $(x, ax+b)$, a machine can estimate $(a, b)$ by simply running all short programs until some output approximates a pair $(u, v)$ such that the line $L_{u,v} = \{(x, ux+v) : x \in \mathbb{R}\}$ passes near $q$. Since $(u, v)$ was approximated by a short program, it has low information density and is therefore close to $(a, b)$ by condition (ii). We formalize this intuition in the following lemma.

**Lemma 29.** *Suppose that $a, b, x \in \mathbb{R}$, $r \in \mathbb{N}$, $\delta \in \mathbb{R}_+$, and $\varepsilon, \eta \in \mathbb{Q}_+$ satisfy $r \geq \log(2|a| + |x| + 7) + 1$ and the following conditions.*

(i) $K_r(a, b) \leq (\eta + \varepsilon)\, r$.

(ii) *For every* $(u, v) \in B_1(a, b)$ *such that* $ux + v = ax + b$,

$$K_r(u, v) \geq (\eta - \varepsilon)\, r + \delta \cdot (r - t)\,,$$

*whenever* $t = -\log \|(a, b) - (u, v)\| \in (0, r]$.

*Then for every oracle set* $A \subseteq \mathbb{N}$,

$$K_r^A(x, ax + b) \geq K_r^A(a, b, x) - \frac{4\varepsilon}{\delta} r - K(\varepsilon) - K(\eta) - O_{a,b,x}(\log r)\,.$$

To prove this lemma, we need the following geometric observation.

**Observation 30.** *Let* $a, x, b \in \mathbb{R}$, $r \in \mathbb{N}$, *and* $(q_1, q_2) \in B_{2^{-r}}(x, ax + b)$.

(i) *If* $(p_1, p_2) \in B_{2^{-r}}(a, b)$, *then* $|p_1 q_1 + p_2 - q_2| < 2^{-r}(|p_1| + |q_1| + 3)$.

(ii) *If* $|p_1 q_1 + p_2 - q_2| \leq 2^{-r}(|p_1| + |q_1| + 3)$, *then there is some* $(u, v) \in B_{2^{-r}(2|a|+|x|+7)}(p_1, p_2)$ *such that* $ax + b = ux + v$.

*Proof.* Only the triangle inequality is needed. If $(p_1, p_2) \in B_{2^{-r}}(a, b)$, then

$$\begin{aligned}
|p_1 q_1 + p_2 - q_2| &\leq |p_1 q_1 + p_2 - (ax + b)| + |ax + b - q_2| \\
&< |p_1 q_1 - ax| + |b - p_2| + 2^{-r} \\
&< |p_1 q_1 - p_1 x| + |p_1 x - ax| + 2^{1-r} \\
&= |p_1| \cdot |q_1 - x| + |x| \cdot |p_1 - a| + 2^{1-r} \\
&\leq 2^{-r}|p_1| + 2^{-r}|x| + 2^{1-r} \\
&\leq 2^{-r}(|p_1| + |x - q_1| + |q_1| + 2) \\
&< 2^{-r}(|p_1| + |q_1| + 3)\,.
\end{aligned}$$

If $|p_1q_1 + p_2 - q_2| < 2^{-r}(|p_1| + |q_1| + 3)$, then

$$\begin{aligned}
|p_1x + p_2 - (ax + b)| &\leq |p_1| \cdot |x - q_1| + |p_1q_1 + p_2 - (ax + b)| \\
&\leq 2^{-r}|p_1| + |p_1q_1 + p_2 - q_2| + |q_2 - (ax + b)| \\
&< 2^{-r}|p_1| + 2^{-r}(|p_1| + |q_1| + 3) + 2^{-r} \\
&= 2^{-r}(2|p_1| + |q_1| + 4) \\
&\leq 2^{-r}(2|a| + |x| + 7)\,.
\end{aligned}$$

so $(u, v) \in (p_1, ax + b - p_1x)$ affirms (ii). □

*Proof of Lemma 29.* Let $a$, $b$, $x$, $r$, $\delta$, $\varepsilon$, $\eta$, and $A$ be as described in the lemma statement.

Define an oracle Turing machine $M$ that does the following given oracle $A$ and input $\pi = \pi_1\pi_2\pi_3\pi_4\pi_5$ such that $U^A(\pi_1) = (q_1, q_2) \in \mathbb{Q}^2$, $U(\pi_2) = h \in \mathbb{Q}^2$, $U(\pi_3) = s \in \mathbb{N}$, $U(\pi_4) = \zeta \in \mathbb{Q}$, and $U(\pi_5) = \iota \in \mathbb{Q}$.

For every program $\sigma \in \{0,1\}^*$ with $\ell(\sigma) \leq (\iota+\zeta)s$, in parallel, $M$ simulates $U(\sigma)$. If one of the simulations halts with some output $(p_1, p_2) \in \mathbb{Q}^2 \cap B_{2^{-1}}(h)$ such that $|p_1q_1 + p_2 - q_2| < 2^{-s}(|p_1| + |q_1| + 3)$, then $M$ halts with output $(p_1, p_2, q_1)$. Let $c_M$ be a constant for the description of $M$.

Now let $\pi_1$, $\pi_2$, $\pi_3$, $\pi_4$, and $\pi_5$ testify to $K_r^A(x, ax + b)$, $K_1(a, b)$, $K(r)$, $K(\varepsilon)$, and $K(\eta)$, respectively, and let $\pi = \pi_1\pi_2\pi_3\pi_4\pi_5$.

By condition (i), there is some $(\hat{p}_1, \hat{p}_2) \in B_{2^{-r}}(a, b)$ such that $K(\hat{p}_1, \hat{p}_2) \leq (\eta + \varepsilon)r$, meaning that there is some $\hat{\sigma} \in \{0,1\}^*$ with $\ell(\hat{\sigma}) \leq (\eta + \varepsilon)r$ and $U(\hat{\sigma}) = (\hat{p}_1, \hat{p}_2)$. A routine calculation (Observation 30(i)) shows that

$$|\hat{p}_1q_1 + \hat{p}_2 - q_2| < 2^{-r}(|\hat{p}_1| + |q_1| + 3)\,,$$

for every $(q_1, q_2) \in B_{2^{-r}}(x, ax + b)$, so $M$ is guaranteed to halt on input $\pi$. Hence, let $(p_1, p_2, q_1) = M(\pi)$. Another routine calculation (Observation 30(ii)) shows that there is some

$$(u, v) \in B_{2^{\gamma-r}}(p_1, p_2) \subseteq B_{2^{-1}}(p_1, p_2) \subseteq B_{2^0}(a, b)$$

such that $ux + v = ax + b$, where $\gamma = \log(2|a| + |x| + 7)$.

We have $\|(p_1, p_2) - (u, v)\| < 2^{\gamma - r}$ and $|q_1 - x| < 2^{-r}$, so

$$(p_1, p_2, q_1) \in B_{2^{\gamma+1-r}}(u, v, x)\,.$$

It follows that

$$\begin{aligned} K^A_{r-\gamma-1}(u, v, x) &\leq \ell(\pi_1\pi_2\pi_3\pi_4\pi_5) + c_M \\ &\leq K^A_r(x, ax + b) + K_1(a, b) + K(r) + K(\varepsilon) + K(\eta) + c_M \\ &= K^A_r(x, ax + b) + K(\varepsilon) + K(\eta) + O_{a,b}(\log r)\,. \end{aligned}$$

Rearranging and applying Lemma 1,

$$K^A_r(x, ax + b) \geq K^A_r(u, v, x) - K(\varepsilon) - K(\eta) - O_{a,b,x}(\log r)\,. \tag{5.3.1}$$

By the definition of $t$, if $t > r$ then $B_{2^{-r}}(u, v, x) \subseteq B_{2^{1-r}}(a, b, x)$, which implies $K^A_r(u, v, x) \geq K^A_{r-1}(a, b, x)$. Applying Lemma 1 gives

$$K^A_r(u, v, x) \geq K^A_r(a, b, x) - O_{a,x}(\log r)\,.$$

Otherwise, when $t \leq r$, we have $B_{2^{-r}}(u, v, x) \subseteq B_{2^{1-t}}(a, b, x)$, which implies $K^A_r(u, v, x) \geq K_{t-1}(a, b, x)$, so by Lemma 1,

$$K^A_r(u, v, x) \geq K^A_r(a, b, x) - 2(r - t) - O_{a,x}(\log r)\,. \tag{5.3.2}$$

We now bound $r - t$. By our construction of $M$ and Lemma 1,

$$\begin{aligned} (\eta + \varepsilon)r &\geq K(p_1, p_2) \\ &\geq K_{r-\gamma}(u, v) \\ &\geq K_r(u, v) - O_{a,x}(\log r)\,. \end{aligned}$$

Combining this with condition (ii) in the lemma statement and simplifying yields

$$r - t \leq \frac{2\varepsilon}{\delta} r + O_{a,x}(\log r)\,,$$

which, together with (5.3.1) and (5.3.2), gives the desired result. □

### 5.3.2 Bounding the Complexity of Lines through a Point

In this section we bound the information content of any pair $(u, v)$ such that the line $L_{u,v}$ intersects $L_{a,b}$ at $x$. Intuitively, an estimate for $(u, v)$ gives significant information about $(a, b)$ whenever $L_{u,v}$ and $L_{a,b}$ are nearly coincident. On the other hand, estimates for $(a, b)$ and $(u, v)$ passing through $x$ together give an estimate of $x$ whose precision is greatest when $L_{a,b}$ and $L_{u,v}$ are nearly orthogonal. We make this dependence on $\|(a, b) - (u, v)\|$ precise in the following lemma.

**Lemma 31.** *Let $a, b, x \in \mathbb{R}$. For all $u, v \in B_1(a, b)$ such that $ux + v = ax + b$, and for all $r \geq t := -\log \|(a, b) - (u, v)\|$,*

$$K_r(u, v) \geq K_t(a, b) + K_{r-t,r}(x|a, b) - O_{a,b,x}(\log r)\,.$$

In proving this lemma, we will use the following geometric observation.

**Observation 32.** *If $x \in \mathbb{R}$ and $a, b, p, q \in \mathbb{R}^2$ satisfy $(p_1, p_2) \in B_{2^{-r}}(a_1, a_2)$, $(q_1, q_2) \in B_{2^{-r}}(b_1, b_2)$, and $a_1 x + a_2 = b_1 x + b_2$, then*

$$\left|\frac{p_2 - q_2}{p_1 - q_1} - \frac{a_2 - b_2}{a_1 - b_1}\right| < 2^{4+2|x|+t-r}\,,$$

*whenever $t = -\log \|a - b\|$ and $r \geq t + |x| + 2$.*

*Proof.* From $a_1 x + a_2 = b_1 x + b_1$ we have $(b_2 - a_2) = (a_1 - b_1)x$, so

$$\begin{aligned} 2^{-t} &\leq \|(a_1, a_2) - (b_1, b_2)\| \\ &= \sqrt{(a_1 - b_1)^2(1 + x^2)} \\ &= |a_1 - b_1|\sqrt{1 + x^2} \\ &\leq |a_1 - b_1|2^{|x|}\,. \end{aligned}$$

Applying this fact and the triangle inequality several times,

$$\begin{aligned}
&\left|\frac{p_2-q_2}{p_1-q_1}-\frac{a_2-b_2}{a_1-b_1}\right| \\
&=\left|\frac{(a_1-b_1)(p_2-q_2)-(a_2-b_2)(p_1-q_1)}{(a_1-b_1)(p_1-q_1)}\right| \\
&\leq \frac{|a_1-b_1|(|p_2-a_2|+|q_2-b_2|)+|a_2-b_2|(|p_1-a_1|+|q_1-b_1|)}{|a_1-b_1|(|a_1-b_1|-|a_1-p_1|-|b_1-q_1|)} \\
&< \frac{2^{1-t}(2^{-r}+2^{-r})+2^{1-t}(2^{-r}+2^{-r})}{2^{-t-|x|}\cdot(2^{-t-|x|}-2^{-t-|x|-2}-2^{-t-|x|-2})} \\
&= \frac{2^{3-t-r}}{2^{-2t-2|x|-1}} \\
&= 2^{4+2|x|+t-r}\,.
\end{aligned}$$

□

*Proof.* Fix $a, b, x \in \mathbb{R}$. By Lemma 28(i), for all $(u, v) \in B_1(a, b)$ and every $r \in \mathbb{N}$,

$$K_r(u,v) \geq K_r(u,v|a,b) + K_r(a,b) - K_r(a,b|u,v) - O_{a,b}(\log r)\,. \tag{5.3.3}$$

We bound $K_r(a,b) - K_r(a,b|u,v)$ first. Since $(u,v) \in B_{2^{-t}}(a,b)$, for every $r \geq t$ we have $B_r(u,v) \subseteq B_{2^{1-t}}(a,b)$, so

$$K_r(a,b|u,v) \leq K_{r,t-1}(a,b|a,b)\,.$$

By Lemma 28(ii), then,

$$\begin{aligned}
K_r(a,b) - K_r(a,b|u,v) &\geq K_r(a,b) - K_{r,t-1}(a,b|a,b) \\
&\geq K_{t-1}(a,b) - O_{a,b}(\log r)\,.
\end{aligned}$$

Lemma 1 tells us that

$$K_{t-1}(a,b) \geq K_t(a,b) - O(\log t)\,.$$

Therefore we have, for every $u, v \in B_1(a,b)$ and every $r \geq t$,

$$K_r(a,b) - K_r(a,b|u,v) \geq K_t(a,b) - O_{a,b}(\log r)\,. \tag{5.3.4}$$

We now bound the term $K_r(u,v|a,b)$. Let $(u,v) \in \mathbb{R}^2$ be such that $ux + v = ax + b$. If $t \leq r < t + |x| + 2$, then $r - t = O_x(1)$, so by Lemma 14, $K_{r-t,r}(x|a,b) = O_x(1)$. In this case, $K_r(u,v|a,b) \geq K_{r-t,r}(x|a,b) - O_{a,b,x}(\log r)$ holds trivially. Hence, assume $r \geq t + |x| + 2$.

Let $M$ be a Turing machine such that, whenever $q = (q_1, q_2) \in \mathbb{Q}^2$ and $U(\pi, q) = p = (p_1, p_2) \in \mathbb{Q}^2$, with $p_1 \neq q_1$,

$$M(\pi, q) = \frac{p_2 - q_2}{p_1 - q_1}\,.$$

For each $q \in B_{2^{-r}}(a, b) \cap \mathbb{Q}^2$, let $\pi_q$ testify to $\hat{K}_r(u, v|q)$. Then

$$U(\pi_q, q) \in B_{2^{-r}}(u, v) \cap \mathbb{Q}^2\,.$$

It follows by a routine calculation (Observation 32) that

$$|M(\pi_q, q) - x| = \left|\frac{p_2 - q_2}{p_1 - q_1} - \frac{b - v}{a - u}\right| < 2^{4+2|x|+t-r}\,.$$

Thus, $M(\pi_q, q) \in B_{2^{4+2|x|+t-r}}(x) \cap \mathbb{Q}^2$. For some constant $c_M$, then,

$$\begin{aligned}\hat{K}_{r-4-2|x|-t}(x|q) &\leq \ell(\pi_q) + c_M \\ &= \hat{K}_r(u, v|q) + c_M\,.\end{aligned}$$

Taking the maximum of each side over $q \in B_{2^{-r}}(a, b) \cap \mathbb{Q}^2$ and rearranging,

$$K_r(u, v|a, b) \geq K_{r-4-2|x|-t,r}(x|a, b) - c_M\,.$$

Then since Lemma 14 implies that

$$K_{r-4-2|x|-t,r}(x|a, b) \geq K_{r-t,r}(x|a, b) - O_x(\log r)\,,$$

we have shown, for every $(u, v)$ satisfying $ux + v = ax + b$ and every $r \geq t$,

$$K_r(u, v|a, b) \geq K_{r-t,r}(x|a, b) - O_{a,b,x}(\log r)\,. \tag{5.3.5}$$

The lemma follows immediately from (5.3.3), (5.3.4), and (5.3.5). □

### 5.3.3 Oracle Construction

To prove Theorem 24, we will show at every precision $r$ that there is an oracle relative to which the hypotheses of Lemma 29 hold and $K_r(a, b, x)$ is still relatively large. These oracles will be based on the following lemma.

**Lemma 33.** *Let $n, r \in \mathbb{N}$, $z \in \mathbb{R}^n$, and $\eta \in \mathbb{Q} \cap [0, \dim(z)]$. Then there is an oracle $D = D(n, r, z, \eta)$ satisfying*

(i) *For every $t \leq r$, $K_t^D(z) = \min\{\eta r, K_t(z)\} + O(\log r)$.*

(ii) *For every $m, t \in \mathbb{N}$ and $y \in \mathbb{R}^m$, $K_{t,r}^D(y|z) = K_{t,r}(y|z) + O(\log r)$ and $K_t^{z,D}(y) = K_t^z(y) + O(\log r)$.*

Informally, for some $s \leq r$ such that $K_s(z)$ is near $\eta r$, the oracle $D$ encodes $r$ bits of $z$ conditioned on $s$ bits of $z$. Unsurprisingly, access to this oracle lowers $K_t(z)$ to $K_s(z)$ whenever $t \geq s$ and has only a negligible effect when $t \leq s$, or when $r$ bits of $z$ are already known.

Our proof of this lemma uses the fact that conditional Kolmogorov complexity is essentially equivalent to Kolmogorov complexity relative to a finite oracle set.[1]

**Observation 34.** *For every $k \in \mathbb{N}$ and $\tau = (\tau_1, \ldots, \tau_k) \in \{0,1\}^k$, define the oracle set*

$$C(\tau) = \left\{ j \leq 2k : \tau_{\lfloor j/2 \rfloor} = 1 \right\} \cup \{2k+1\} \subseteq \mathbb{N}\,.$$

*Then there is a constant $c$ such that for every $\sigma, \tau \in \{0,1\}^*$,*

$$\left| K(\sigma|\tau) - K^{C(\tau)}(\sigma) \right| \leq c\,.$$

*Proof.* Let $\pi \in \{0,1\}^*$ be such that $U(\pi, \tau) = \sigma$. Then given the oracle $C(\tau)$ and input $\pi$, a machine can discern $\tau$ from $2\ell(\tau) + 2$ queries to $C(\tau)$ and use it to simulate $U(\pi, \tau)$. Let $\pi \in \{0,1\}^*$ such that $U^{C(\tau)}(\pi) = \sigma$. Likewise, given input $(\pi', \tau)$, a machine can compute any bit $C(\tau)$ queried in a simulation of $U^{C(\tau)}(\pi)$. □

*Proof of Lemma 33.* Let $s = \max\{t \leq r : K_{t-1}(z) < \eta r\}$. Observe that

$$\eta r \leq K_s(z) \leq \eta r + K(s) + c\,.$$

[1] In fact, [28] defines conditional Kolmogorov complexity in terms of a finite oracle, using a construction similar to the one described here.

Let $\sigma$ be the lexicographically first time-minimizing witness to $K(z\restriction r \,|\, z\restriction s)$, and let $A = C(\sigma)$, as defined in Observation 34.

Suppose $s \leq t \leq r$. Then applying a relativized version of Corollary 26 and Observation 34,

$$\begin{aligned} K_t^A(z) &\leq K_r^A(z) \\ &\leq K^A(z\restriction r) + K(r) + O(1) \\ &\leq K(z\restriction r \,|\, \sigma) + K(r) + O(1)\,. \end{aligned}$$

There exists a Turing machine $M_1$ that, on input $(\pi, \sigma)$, for $\pi \in \{0,1\}^*$, simulates $U(\sigma, U(\pi, \sigma))$. If $\pi$ is a witness to $K(z\restriction s \,|\, \sigma)$, then

$$M(\pi, \sigma) = U(\sigma, U(\pi, \sigma)) = U(\sigma, z\restriction s) = z\restriction r\,.$$

Thus, $K(z\restriction r \,|\, \sigma) \leq K(z\restriction s \,|\, \sigma) + c_{M_1}$, where $c_{M_1}$ is a constant for the description length of $M_1$. We now have

$$\begin{aligned} K_t^A(z) &\leq K(z\restriction s \,|\, \sigma) + K(r) + O(1) \\ &\leq K(z\restriction s) + K(r) \\ &\leq K_s(z) + 2K(r) + O(1) \\ &\leq \eta r + 2K(r) + K(s) + O(1)\,. \end{aligned}$$

For the other direction, since $K_t^A(z) \geq K_s^A(z)$ whenever $t \geq s$, it is sufficient to show that

$K_s^A(z) \geq \eta r$. We use Corollary 26, Observation 34, and the symmetry of information:

$$\begin{aligned}
K_s^A(z) &\geq K^A(z{\restriction}s) - K(s) - O(1) \\
&\geq K(z{\restriction}s \,|\, \sigma) - K(s) - O(1) \\
&\geq K(z{\restriction}r \,|\, \sigma) - K(s) - O(1) \\
&\geq K(z{\restriction}r) - K(\sigma) - K(s) - O(1) \\
&= K(z{\restriction}r) - K(z{\restriction}r \,|\, z{\restriction}s) - K(s) - O(1) \\
&\geq K(z{\restriction}r, z{\restriction}s) - K(z{\restriction}r \,|\, z{\restriction}s, K(z{\restriction}s)) - K(K(z{\restriction}s)) - 2K(s) - O(1) \\
&= K(z{\restriction}s) - K(K(z{\restriction}s)) - 2K(s) - O(1) \\
&\geq K_s(z) - K(K(z{\restriction}s)) - 3K(s) - O(1) \\
&= K_s(z) - O(\log r)\,.
\end{aligned}$$

Since $K_s(z) \geq \eta r$, property (i) holds in this case.

Now suppose instead that $t \leq s \leq r$. We again use Corollary 26, Observation 34, and

the symmetry of information.

$$\begin{aligned}
K_t^A(z) =& K^A(z\restriction t) - K(t) - O(1)\\
=& K(z\restriction t \,|\, \sigma) - K(t) - O(1)\\
\geq& K(z\restriction t \,|\, \sigma, K(\sigma)) - K(t) - O(1)\\
=& K(\sigma \,|\, z\restriction t, K(z\restriction t)) + K(z\restriction t) - K(\sigma) - K(t) - O(1)\\
\geq& K(\sigma \,|\, z\restriction t) - K(K(z\restriction t)) + K(z\restriction t) - K(\sigma) - K(t) - O(1)\\
\geq& K(\sigma \,|\, z\restriction s, t) - K(K(z\restriction t)) + K(z\restriction t) - K(\sigma) - K(t) - O(1)\\
\geq& K(z\restriction t) + K(\sigma \,|\, z\restriction s, K(z\restriction s)) - K(\sigma) - K(K(z\restriction t)) - 2K(t) - O(1)\\
=& K(z\restriction t) + K(z\restriction s \,|\, \sigma, K(\sigma)) - K(z\restriction s) - K(K(z\restriction t)) - 2K(t) - O(1)\\
\geq& K(z\restriction t) + K(z\restriction s \,|\, \sigma) - K(z\restriction s) - K(K(\sigma)) - K(K(z\restriction t))\\
& - 2K(t) - O(1)\\
\geq& K_t(z) + K_s^A(z) - K_s(z) - K(K(\sigma)) - K(K(z\restriction t))\\
& - 3K(t) - 2K(s) - O(1)\\
=& K_t(z) + K_s^A(z) - K_s(z) - O(\log r)\,.
\end{aligned}$$

As we have already shown that $K_s^A(z) - K_s(z) = O(\log r)$, we conclude that property (i) holds in this case as well.

For property (ii), we again apply Corollary 26, relativized to $(z, A)$, and Observation 34, relativized to $z$, to see that

$$\begin{aligned}
K_t^{z,A}(y) &\geq K^{z,A}(y\restriction t) - K(t) - O(1)\\
&= K^z(y\restriction t \,|\, \sigma) - K(t) - O(1)\\
&\geq K^z(y\restriction t) - K^z(\sigma) - K(t) - O(1)\\
&\geq K_t^z(y) - K^z(\sigma) - 2K(t) - O(1)\\
&\geq K_t^z(y) - K(\sigma \,|\, z\restriction r) - 2K(t) - O(1)\,,
\end{aligned}$$

where the last inequality is due to Lemma 20. We argue that $K(\sigma \mid z\restriction r)$ is at most logarithmic in $r$.

$$\begin{aligned} K(\sigma \mid z\restriction r) &\leq K(\sigma, s, \ell(\sigma) \mid z\restriction r) + O(1) \\ &\leq K(\sigma \mid s, \ell(\sigma), z\restriction r) + K(s) + K(\ell(\sigma)) + O(1) \\ &\leq K(\sigma \mid s, \ell(\sigma), z\restriction r) + O(\log r)\,. \end{aligned}$$

To see that the first term is constant, define a Turing machine $M_2$ that does the following. Given input $(j, k, x)$, $M_2$ simulates, for every $\pi \in \{0,1\}^k$ in parallel, $U(\pi, x\restriction j)$. It outputs the first such $\pi$ whose simulation halts with output $x$. We defined $\sigma$ in such a way that $M_2^z(s, \ell(\sigma), z\restriction r) = \sigma$, so

$$K(\sigma \mid s, \ell(\sigma), z\restriction r) \leq c_{M_2}\,,$$

where $c_{M_2}$ is a constant for the length of $M_2$'s description. We conclude that $K(\sigma \mid z\restriction r) = O(\log r)$, so $K_t^{z,A}(y) \geq K_t^z(y) - O(\log r)$.

The argument for conditional complexity is essentially identical. By a relativized version of Corollary 27 and Observation 34,

$$\begin{aligned} K^A_{t,r}(y|z) &\geq K^{z,A}(y\restriction t \mid z\restriction r) - K(t) - O(1) \\ &= K(y\restriction t \mid z\restriction r, \sigma) - K(t) - O(1) \\ &\geq K(y\restriction t \mid z\restriction r) - K(\sigma \mid z\restriction r) - K(t) - O(1) \\ &\geq K_{t,r}(y|z) - K(\sigma \mid z\restriction r) - 2K(t) - O(1) \\ &\geq K_{t,r}(y|z) - K(\sigma \mid z\restriction r) - O(\log r)\,, \end{aligned}$$

and we have already shown that $K(\sigma \mid z\restriction r) = O(\log r)$. □

### 5.3.4 Proof of Theorem 24

**Theorem 24.** *For every $a, b, x \in \mathbb{R}$ and $A \subseteq \mathbb{N}$,*

$$\dim^A(x, ax+b) \geq \dim^A(x|a,b) + \min\left\{ \dim^A(a,b),\, \dim^{a,b}(x) \right\}.$$

*In particular, for almost every $x \in \mathbb{R}$, $\dim(x, ax+b) = 1 + \min\{\dim(a,b), 1\}$.*

*Proof.* Let $a, b, x \in \mathbb{R}$, and treat them as constant for the purposes of asymptotic notation here. Let $A \subseteq \mathbb{N}$,

$$H = \mathbb{Q} \cap \left[0, \dim^A(a,b)\right] \cap \left[0, \dim^{a,b}(x)\right),$$

and $\eta \in H$. Let $\delta = \dim^{a,b}(x) - \eta > 0$ and $\varepsilon \in \mathbb{Q}_+$. For each $r \in \mathbb{N}$, let $D_r = D(2, r, (a,b), \eta)$, as defined in Lemma 33. We claim that for every sufficiently large $r$, the conditions of Lemma 29, relativized to oracle $D_r$, are satisfied by these choices of $a$, $b$, $x$, $r$, $\delta$, $\varepsilon$, $\eta$.

Property (i) of Lemma 33 guarantees that $K_r^{D_r}(a,b) \leq \eta r + O(\log r)$, so condition (i) of Lemma 29 is satisfied for every sufficiently large $r$.

To see that condition (ii) of Lemma 29 is also satisfied, let $(u,v) \in B_1(a,b)$ such that $ax + b = ux + v$ and $t = -\log \|(a,b) - (u,v)\| \leq r$. Then by Lemma 31, relativized to $D_r$, we have

$$K_r^{D_r}(u,v) \geq K_t^{D_r}(a,b) + K_{r-t,r}^{D_r}(x|a,b) - O(\log r).$$

Therefore, by Lemma 33 and Lemma 20,

$$\begin{aligned}
K_r^{D_r}(u,v) &\geq \min\{\eta r, K_t(a,b)\} + K_{r-t,r}(x|a,b) - O(\log r) \\
&\geq \min\{\eta r, K_t(a,b)\} + K_{r-t}^{a,b}(x) - O(\log r) \\
&\geq \min\{\eta r, \dim(a,b)t - o(t)\} + \dim^{a,b}(x)(r-t) - o(r) \\
&\geq \min\{\eta r, \eta t - o(t)\} + (\eta + \delta)(r-t) - o(r) \\
&= \eta t - o(t) + (\eta + \delta)(r-t) - o(r) \\
&= \eta r + \delta \cdot (r-t) - o(r) \\
&\geq (\eta - \varepsilon) r + \delta \cdot (r-t),
\end{aligned}$$

whenever $r$ is large enough.

For every sufficiently large $r$, then, the conclusion of Lemma 29 applies here. Thus, for

constant $a$, $b$, $\varepsilon$, and $\eta$,

$$\begin{aligned} K_r^A(x, ax+b) &\geq K_r^{A,D_r}(x, ax+b) - O(1) \\ &\geq K_r^{A,D_r}(a, b, x) - 4\varepsilon r/\delta - O(\log r) \\ &= K_r^{A,D_r}(x|a,b) + K_r^{A,D_r}(a,b) - 4\varepsilon r/\delta - O(\log r) \\ &= K_r^A(x|a,b) + \eta r - 4\varepsilon r/\delta - O(\log r)\,, \end{aligned}$$

where the last equality is due to the properties of $D_r$ guaranteed by Lemma 33.

Dividing by $r$ and taking limits inferior,

$$\begin{aligned} \dim^A(x, ax+b) &\geq \liminf_{r\to\infty} \frac{K_r^A(x|a,b) + \eta r - 4\varepsilon r/\delta - O(\log r)}{r} \\ &= \dim^A(x|a,b) + \eta - \frac{4\varepsilon}{\delta}\,. \end{aligned}$$

Since this holds for every $\eta \in H$ and $\varepsilon \in \mathbb{Q}_+$, we have

$$\dim^A(x, ax+b) \geq \dim^A(x|a,b) + \min\left\{ \dim^A(a,b),\, \dim^{a,b}(x) \right\}.$$

The second part of the theorem statement follows easily, as relative to any given oracle for $(a,b)$, almost every $x \in \mathbb{R}$ is Martin-Löf random and therefore has dimension 1. Applying Corollary 21, then, almost every $x \in \mathbb{R}$ has $\dim(x|a,b) \geq \dim^{a,b}(x) = 1$. □

We can now easily answer the motivating question of whether or not there is a line in $\mathbb{R}^2$ on which every point has effective Hausdorff dimension 1.

**Corollary 35.** *For every $a, b \in \mathbb{R}$, there exist $x, y \in \mathbb{R}$ such that*

$$\dim(x, ax+b) - \dim(y, ay+b) \geq 1\,.$$

*In particular, there is no line in $\mathbb{R}^2$ on which every point has dimension 1.*

*Proof.* Theorem 24 tells us that $\dim(x, ax+b) \geq 1 + \min\{\dim(a,b), 1\}$ for almost every $x \in \mathbb{R}$. For $y = 0$, we have $\dim(y, ay+b) = \dim(b) \leq \min\{\dim(a,b), 1\}$. □

There are lines for which the inequality in Corollary 35 is strict. Consider, for example, a line through the origin whose slope $a$ is random. For every $x$ that is random relative to $a$, the point $(a, ax)$ has dimension $\dim(x) + \dim(a) = 2$, but the origin itself has dimension 0.

## 5.4 Generalized Sets of Furstenberg Type

**Definition.** A *set of Furstenberg type* with parameter $\alpha$ is a set $E \subseteq \mathbb{R}^2$ such that, for every $e \in S^1$ (the unit circle in $\mathbb{R}^2$), there is a line $\ell_e$ in the direction $e$ satisfying $\dim_H(E \cap \ell_e) \geq \alpha$.

Finding the minimum possible dimension of such a set is an important open problem with connections to Falconer's distance set conjecture and to Kakeya sets [55, 107]. The the best known lower bound is $\alpha + \max\{1/2, \alpha\}$, and the lower bound cannot be greater than $(3 + \alpha)/2$. According to Wolff [107], these results are due, "in all probability," to Furstenberg and Katznelson. The study of Furstenberg sets is also related to questions in dynamical systems about sets that are invariant under maps of the form $x \mapsto px \bmod 1$. A major conjecture of Furstenberg on this topic was resolved very recently, independently by Wu [108] and by Shmerkin [98]. The papers [93, 86] contain surveys of progress on conjectures about Furstenberg sets.

Molter and Rela introduced a natural generalization of Furstenberg sets, in which the set of directions may itself have fractal dimension.

**Definition** (Molter and Rela [81])**.** A set $E \subseteq \mathbb{R}^2$ is in the class $F_{\alpha\beta}$ if there is some set $J \subseteq S^1$ such that $\dim_H(J) \geq \beta$ and for every $e \in J$, there is a line $\ell_e$ in the direction $e$ satisfying $\dim_H(E \cap \ell_e) \geq \alpha$.

They proved the following lower bound on the dimension of such sets.

**Theorem 36.** (Molter and Rela [81]) *For all $\alpha, \beta \in (0, 1]$ and every set $E \in F_{\alpha\beta}$,*

$$\dim_H(E) \geq \alpha + \max\left\{\frac{\beta}{2}, \alpha + \beta - 1\right\}.$$

We now show that Theorem 24 yields an improvement on this bound whenever $\alpha, \beta < 1$ and $\beta/2 < \alpha$.

**Theorem 37.** *For all $\alpha, \beta \in (0,1]$ and every set $E \in F_{\alpha\beta}$,*

$$\dim_H(E) \geq \alpha + \min\{\beta, \alpha\}\,.$$

*Proof.* Let $\alpha, \beta \in (0,1]$, $\varepsilon \in (0, \beta)$, and $E \in F_{\alpha\beta}$. Using Theorem 2, let $A$ satisfy

$$\sup_{z \in E} \dim^A(z) = \dim_H(E)\,.$$

and $e \in S^1$ satisfy $\dim^A(e) = \beta - \varepsilon > 0$. Let $\ell_e$ be a line in direction $e$ such that $\dim_H(\ell_e \cap E) \geq \alpha$. Since $\dim(e) > 0$, we know $e \notin \{(0,1),(0,-1)\}$, so we may let $a, b \in \mathbb{R}$ be such that $L_{a,b} = \ell_e$. Notice that $\dim^A(a) = \dim^A(e)$ because the mapping $e \mapsto a$ is computable and bi-Lipschitz in some neighborhood of $e$. Let $S = \{x : (x, ax+b) \in E\}$, which is similar to $\ell_e \cap E$, so $\dim_H(S) \geq \alpha$ also. We now have

$$\begin{aligned}
\dim_H(E) &= \sup_{z \in E} \dim^A(z) \\
&\geq \sup_{z \in \ell_e \cap E} \dim^A(z) \\
&= \sup_{x \in S} \dim^A(x, ax+b)\,.
\end{aligned}$$

By Theorem 24 and Corollary 21, both relativized to $A$,

$$\begin{aligned}
\sup_{x \in S} \dim^A(x, ax+b) &\geq \sup_{x \in S} \Big\{ \dim^{A,a,b}(x) + \min\{\dim^A(a,b), \dim^A(x|a,b)\} \Big\} \\
&\geq \sup_{x \in S} \Big\{ \dim^{A,a,b}(x) + \min\{\dim^A(a,b), \dim^{A,a,b}(x)\} \Big\} \\
&\geq \sup_{x \in S} \dim^{A,a,b}(x) + \min\left\{ \dim^A(a), \sup_{x \in S} \dim^{A,a,b}(x) \right\}.
\end{aligned}$$

Theorem 2 gives

$$\sup_{x \in S} \dim^{A,a,b}(x) \geq \dim_H(S) \geq \alpha\,,$$

so we have shown, for every $\varepsilon \in (0, \beta)$, that $\dim_H(E) \geq \alpha + \min\{\beta - \varepsilon, \alpha\}$. □

# Chapter 6

# Intersections and Products of Fractals

In this chapter, we apply algorithmic dimensional techniques to bound the Hausdorff and packing dimensions of intersections and products of fractals. Most significantly, we extend the following intersection formula, previously shown to hold when $E$ and $F$ are Borel sets [30], to arbitrary sets $E$ and $F$.[1]

**Theorem 38.** *For all $E, F \subseteq \mathbb{R}^n$, and for almost every $z \in \mathbb{R}^n$,*

$$\dim_H(E \cap (F + z)) \leq \max\{0, \dim_H(E \times F) - n\},$$

*where $F + z = \{x + z : x \in F\}$.*

This approach also yields a simplified proof of the following known product formula for general sets.

**Theorem 39** (Marstrand [74])**.** *For all $E \subseteq \mathbb{R}^m$ and $F \subseteq \mathbb{R}^n$,*

$$\dim_H(E) + \dim_H(F) \leq \dim_H(E \times F).$$

Both of these formulas are prominent, fundamental results in fractal geometry that are taught in typical undergraduate courses on the subject. We also use symmetric arguments to derive the known corresponding statements about packing dimension [103, 34]. These results are included here to showcase the versatility of this technique and its ability to capture the exact duality between Hausdorff and packing dimensions.

[1] This result is closely related to the Marstrand Slicing Theorem, as stated in the excellent recent book by Bishop and Peres [12]. The proof given there assumes that a set is Borel, but this assumption was inadvertently omitted from the theorem statement [11].

## 6.1 Intersections of Fractals

In this section we prove Theorem 38. We then use a symmetric argument to prove the corresponding statement for packing dimension, which is known [34]. For the case where $E, F \subseteq \mathbb{R}^n$ are Borel sets, Theorem 38 was shown in its present form by Falconer [30]. Closely related results, which also place restrictions on $E$ and $F$, were proven earlier by Mattila [76, 77] and Kahane [54].

**Theorem 38.** *For all $E, F \subseteq \mathbb{R}^n$, and for almost every $z \in \mathbb{R}^n$,*

$$\dim_H(E \cap (F + z)) \leq \max\{0, \dim_H(E \times F) - n\}\,, \tag{6.1.1}$$

*where $F + z = \{x + z : x \in F\}$.*

*Proof.* Let $E, F \subseteq \mathbb{R}^n$ and $z \in \mathbb{R}^n$. If $E \cap (F + z) = \emptyset$, then (6.1.1) holds trivially, so assume that the intersection is nonempty. Theorem 2 guarantees that there is some oracle set $A \subseteq \mathbb{N}$ satisfying

$$\dim_H(E \times F) = \sup_{(x,y) \in E \times F} \dim^A(x, y)\,. \tag{6.1.2}$$

It also guarantees, given any $\varepsilon > 0$, that there is an $x \in E \cap (F + z)$ such that

$$\dim^{A,z}(x) \geq \dim_H(E \cap (F + z)) - \varepsilon\,. \tag{6.1.3}$$

Since $(x, x - z) \in E \times F$, we have

$$\begin{aligned}
\dim_H(E \times F) &\geq \dim^A(x, x - z) \\
&= \dim^A(x, z) \\
&\geq \dim^A(z) + \dim^A(x|z) \\
&\geq \dim^A(z) + \dim^{A,z}(x) \\
&\geq \dim^A(z) + \dim_H(E \cap (F + z)) - \varepsilon\,.
\end{aligned}$$

The above lines follow from (6.1.2), Lemma 8, Corollary 19, Corollary 21, and (6.1.3), respectively. Letting $\varepsilon \to 0$, we have

$$\dim_H(E \cap (F + z)) \leq \dim_H(E \times F) - \dim^A(z)\,.$$

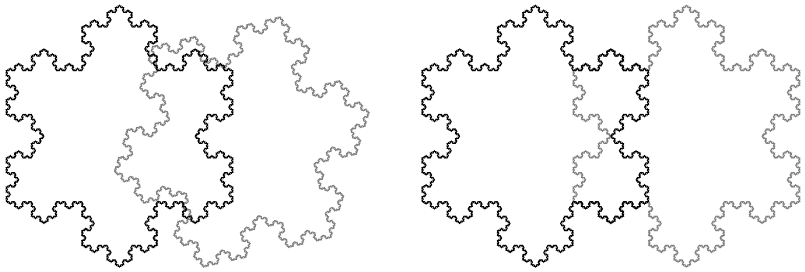

Figure 6.1: Let $E$ and $F$ each be Koch snowflakes, which have Hausdorff dimension $\log_3 4 \approx 1.26$. Left: For almost all rigid motions $\sigma$, the intersection $E \cap \sigma(F)$ has Hausdorff dimension at most $2\log_3 4 - 2 \approx 0.52$. Right: For a measure zero set of rigid motions, the Hausdorff dimension of the intersection may be as large as $\log_3 4$. Note that Koch curves are Borel sets, so the new generality given by Theorem 38 and Corollary 40 is not required for this example.

Thus, (6.1.1) holds whenever $\dim^A(z) = n$. In particular, it holds when $z$ is Martin-Löf random relative to $A$, i.e., for Lebesgue almost every $z \in \mathbb{R}^n$ [62, 75]. □

For the case that $E$ and $F$ are Borel sets, Falconer [32] notes that the intersection formula is readily extended to rigid motions and similarities. The same argument applies in the general case, so Theorem 38 has the following corollary.

**Corollary 40.** *Let $E, F \subseteq \mathbb{R}^n$. Let $G$ be the group of rigid motions or the group of similarities on $\mathbb{R}^n$. Then, for almost all $\sigma \in G$,*

$$\dim_H(E \cap \sigma(F)) \leq \max\{0, \dim_H(E \times F) - n\}. \tag{6.1.4}$$

*Proof (Following Falconer [32]).* For all rotations (and all scalings) of $F$, Theorem 38 tells us that (6.1.4) holds for almost all translations. Thus, (6.1.4) holds for almost all rigid motions and almost all similarities. □

A corresponding intersection formula for packing dimension has been shown for arbitrary $E, F \subseteq \mathbb{R}^n$ by Falconer [34]. That proof is not difficult or long, but an algorithmic dimensional proof is presented here as an instance where this technique applies symmetrically to both Hausdorff and packing dimension.

**Theorem 41** (Falconer [34])**.** *For all $E, F \subseteq \mathbb{R}^n$, and for almost every $z \in \mathbb{R}^n$,*

$$\dim_P(E \cap (F + z)) \leq \max\{0, \dim_P(E \times F) - n\}.$$

*Proof.* As in Theorem 38, we may assume that the intersection is nonempty. Apply Theorem 3 to choose an oracle set $B \subseteq \mathbb{N}$ such that

$$\dim_P(E \times F) = \sup_{(x,y)\in E\times F} \mathrm{Dim}^B(x,y) \tag{6.1.5}$$

and, given $\varepsilon > 0$, a point $y \in E \cap (F + z)$ satisfying

$$\mathrm{Dim}^{B,z}(y) \geq \dim_P(E \cap (F+z)) - \varepsilon\,. \tag{6.1.6}$$

Then $(y, y - z) \in E \times F$, and we may proceed much as before:

$$\begin{aligned}
\dim_P(E \times F) &\geq \mathrm{Dim}^B(y, y-z) \\
&= \mathrm{Dim}^B(y,z) \\
&\geq \dim^B(z) + \mathrm{Dim}^B(y|z) \\
&\geq \dim^B(z) + \mathrm{Dim}^{B,z}(y) \\
&\geq \dim^B(z) + \dim_P(E \cap (F+z)) - \varepsilon\,.
\end{aligned}$$

These lines follow from (6.1.5), Lemma 8, Corollary 19, Corollary 21, and (6.1.6). Again, $\dim^B(z) = n$ for almost every $z \in \mathbb{R}^n$, so this completes the proof. □

## 6.2 Products of Fractals

In this section we prove four known product inequalities for fractal dimensions. Inequality (6.2.1), which was stated in the introduction as Theorem 39, is due to Marstrand [74]. When $E$ and $F$ are Borel sets, it is simple to prove (6.2.1) by using Frostman's Lemma, but the argument for general sets using net measures is considerably more difficult [78, 33]. The other three inequalities are due to Tricot [103]. Reference [78] gives a more detailed account of this history.

**Theorem 42** (Marstrand [74], Tricot [103])**.** *For all $E \subseteq \mathbb{R}^m$ and $F \subseteq \mathbb{R}^n$,*

$$\dim_H(E) + \dim_H(F) \leq \dim_H(E \times F) \tag{6.2.1}$$

$$\leq \dim_H(E) + \dim_P(F) \tag{6.2.2}$$

$$\leq \dim_P(E \times F) \tag{6.2.3}$$

$$\leq \dim_P(E) + \dim_P(F)\,. \tag{6.2.4}$$

Notice the superficial resemblance of this theorem to Corollary 19, the chain rule for dimension. This similarity is not a coincidence; each inequality in Theorem 42 follows from the corresponding line in Corollary 19. The arguments given here for (6.2.1–6.2.4) are each similar in length to the proof of (6.2.1) for Borel sets. That is, they are quite short.

*Proof.* Theorem 2 guarantees, for every $\varepsilon > 0$, that there exist an oracle set $A \subseteq \mathbb{N}$ and points $x \in E$ and $y \in F$ such that

$$\dim_H(E \times F) = \sup_{z \in E \times F} \dim^A(z)\,, \tag{6.2.5}$$

$$\dim^A(x) \geq \dim_H(E) - \varepsilon\,,$$

$$\dim^{A,x}(y) \geq \dim_H(F) - \varepsilon\,.$$

Then by (6.2.5), Corollary 19 relative to $A$, and Corollary 21 relative to $A$, we have

$$\begin{aligned}
\dim_H(E \times F) &\geq \dim^A(x, y) \\
&\geq \dim^A(x) + \dim^A(y|x) \\
&\geq \dim^A(x) + \dim^{A,x}(y) \\
&\geq \dim_H(E) + \dim_H(F) - 2\varepsilon\,,
\end{aligned}$$

by our choice of $x$ and $y$. Since $\varepsilon > 0$ was arbitrary, we conclude that (6.2.1) holds.

For (6.2.2), let $\varepsilon > 0$ and use the point-to-set principles for Hausdorff and packing

dimension to find $B, C \subseteq \mathbb{N}$, $u \in E$, and $v \in F$ such that

$$\begin{aligned}
\dim_H(E) &= \sup_{x \in E} \dim^B(x)\,, \\
\dim_P(F) &= \sup_{y \in E} \mathrm{Dim}^C(y)\,, \\
\dim^{B,C}(u,v) &\geq \dim_H(E \times F) - \varepsilon\,.
\end{aligned}$$

Since $B$ and $C$ minimize their respective expressions, we also have

$$\begin{aligned}
\dim_H(E) &= \sup_{x \in E} \dim^{B,C}(x)\,, \\
\dim_P(F) &= \sup_{y \in E} \mathrm{Dim}^{B,C}(y)\,.
\end{aligned}$$

Thus, we can apply Corollary 19 relative to $B, C$, after first noticing that conditioning on another point never increases dimension.

$$\begin{aligned}
\dim_H(E) + \dim_P(F) &\geq \dim^{B,C}(u) + \mathrm{Dim}^{B,C}(v) \\
&\geq \dim^{B,C}(u|v) + \mathrm{Dim}^{B,C}(v) \\
&\geq \dim^{B,C}(u,v) \\
&\geq \dim_H(E \times F) - \varepsilon\,.
\end{aligned}$$

Again, $\varepsilon$ was arbitrary, so (6.2.2) holds.

For (6.2.3) and (6.2.4), we use essentially the same arguments as above. By Theorem 3, there are $A', B' \subseteq \mathbb{N}$, $x', u' \in E$, $y', v' \in F$, and $\varepsilon > 0$ that satisfy

$$\begin{aligned}
\dim_P(E \times F) &= \sup_{z \in E \times F} \mathrm{Dim}^{A'}(z)\,, \\
\dim_H(E) &= \sup_{z \in E} \mathrm{Dim}^{B'}(z)\,, \\
\dim^{A'}(x') &\geq \dim_H(E) - \varepsilon\,, \\
\mathrm{Dim}^{A',x'}(y') &\geq \dim_P(F) - \varepsilon\,, \\
\mathrm{Dim}^{B',C}(u',v') &\geq \dim_P(E \times F) - \varepsilon\,,
\end{aligned}$$

where $x$ and $C$ are as above. We once again apply relativized versions of Corollary 19 and Corollary 21:

$$\begin{aligned}
\dim_P(E) + \dim_P(F) &\geq \mathrm{Dim}^{B',C}(u') + \mathrm{Dim}^{B',C}(v') \\
&\geq \mathrm{Dim}^{B',C}(u'|v') + \mathrm{Dim}^{B',C}(v') \\
&\geq \mathrm{Dim}^{B',C}(u', v') \\
&\geq \dim_P(E \times F) - \varepsilon \\
&\geq \mathrm{Dim}^{A'}(x', y') - \varepsilon \\
&\geq \dim^{A'}(x') + \mathrm{Dim}^{A'}(y'|x') - \varepsilon \\
&\geq \dim^{A'}(x') + \mathrm{Dim}^{A',x'}(y') - \varepsilon \\
&\geq \dim_H(E) + \dim_P(F) - 3\varepsilon\,.
\end{aligned}$$

Letting $\varepsilon \to 0$ completes the proof. □

# Part II

# Decentralized Dynamics

# Chapter 7

# Asynchronous Dynamics

We now turn our attention to the convergence of dynamics in distributed environments, with an emphasis on *game dynamics*, in which individual agents have underlying incentives. The notion of *self stabilization* to a "legitimate" state in a distributed system parallels that of convergence to an equilibrium in a game. The foci, however, are very different. In game theory, there is extensive research on dynamics that result from what is perceived as natural strategic decision making (e.g., best- or better-response dynamics, fictitious play, or regret minimization). Even simple heuristics that require little information or computational resources can yield sophisticated behavior, such as the convergence of best-response dynamics to equilibrium points (see [43] and references therein). These positive results for simple game dynamics are, with few exceptions (see below), based on the sometimes implicit and often unrealistic premise of a controlled environment in which actions are synchronous and coordinated. Distributed computing research emphasizes the environmental uncertainty that results from decentralization, but has no notion of "natural" rules of behavior. It has long been known that environmental uncertainty—in the form of both asynchrony [36, 72] and arbitrary initialization [26]—introduces substantial difficulties for protocol termination in distributed systems. Our work bridges the gap between these two approaches by initiating the study of game dynamics in distributed computing settings. Before describing our model, we discuss the relationship of this work to several other areas.

**Algorithmic game theory.** Since our work draws on both game theory and computer science, it may be considered part of the broader research program of *algorithmic game theory* (AGT), which merges concepts and ideas from those two fields [85]. Three main

areas of study in AGT have been algorithmic mechanism design, which applies concepts from computer science to economic mechanism design [84]; the "price of anarchy," which describes the efficiency of equilibria and draws on approximability research [58]; and algorithmic and complexity research on the computation of equilibria [87]. Analyzing the computational power of learning dynamics in games has been of particular interest (see, e.g., [21, 56, 6, 88]). Our work creates another link between game theory and computer science by drawing on two previously disjoint areas, self-stabilization in distributed computing theory and game dynamics, to explore broader classes of dynamics operating in adversarial distributed environments.

**Adaptive heuristics.** Much work in game theory and economics deals with *adaptive heuristics* (see [43] and references therein). Generally speaking, this long line of research explores the "convergence" of simple and myopic rules of behavior (e.g., best-response, fictitious-play, or no-regret dynamics) to an "equilibrium". However, with few exceptions (see below), such analysis has so far primarily concentrated on synchronous environments in which steps take place simultaneously or in some other predetermined order. In this work, we explore dynamics of this type in asynchronous environments, which are more realistic for many applications.

**Game-theoretic work on asynchronous environments.** Some game-theoretic work on repeated games considers "asynchronous moves." Often, as in [73], this asynchrony merely indicates that players are not all activated at each time step, and thus is used to describe environments where only one player is activated at a time ("alternating moves"), or where there is a probability distribution that determines which player(s) are activated at each timestep. Other work does not explore the behavior of dynamics, but has other research goals (e.g., characterizing equilibria, establishing folk theorems); see [59], [109], among others, and references therein. To the best of our knowledge, we are the first to study the effects of asynchrony (in the broad distributed computing sense) on the convergence of *game dynamics* to equilibria.

**Fault-tolerant computation.** We use ideas and techniques from work in distributed computing on protocol termination in asynchronous computational environments where nodes and communication channels are possibly faulty. Protocol termination in such environments, initially motivated by multi-processor computer architectures, has been extensively studied in the past three decades [36, 7, 24, 13, 49, 95], as nicely surveyed in [72, 35]. Fischer, Lynch and Paterson [36] showed, in a landmark paper, that a broad class of failure-resilient consensus protocols cannot provably terminate. Intuitively, the risk of protocol non-termination in that work stems from the possibility of failures; a computational node cannot tell whether another node is silent due to a failure or is simply taking a long time to react. Our non-convergence result, by contrast, applies to failure-free environments. In game-theoretic work that incorporated fault tolerance concepts, [1] studied equilibria that are robust to defection and collusion.

**Self stabilization.** The concept of self stabilization is fundamental to distributed computing and dates back to [23] (see [26] and references therein). Convergence of dynamics to an "equilibrium" in our model can be viewed as the self stabilization of such dynamics (where the "equilibrium points" are the legitimate configurations). Our formulation draws ideas from work in distributed computing (e.g., Burns' distributed daemon model [14]) and in networking research [38] on self stabilization.

## 7.1 Asynchronous Dynamic Interaction

In this section we present our model of asynchronous dynamic interaction. Intuitively, an *interaction system* consists of a collection of computational nodes, each capable of selecting *actions* that are visible to the other nodes. The *state* of the system at any time consists of each node's current action. Each node has a deterministic *reaction function* that maps system histories to actions. At every discrete timestep, each node activated by a schedule simultaneously applies its deterministic reaction function to select a new action, which is immediately visible to all other nodes.

**Definition.** An *interaction system* is characterized by a tuple $(n, A, \mathbf{f})$:

- The system has $n \in \mathbb{Z}_+$ *computational nodes*, labeled $1, \ldots, n$.
- $A = A_1 \times \ldots \times A_n$, where each $A_i$ is a finite set called the *action space* of node $i$. $A$ is called the *state space* of the system, and a *state* is an $n$-tuple $\mathbf{a} = (a_1, \ldots, a_n) \in A$. A *history* of the system is a nonempty finite sequence of states, $H \in A^\ell$, for some $\ell \in \mathbb{Z}_+$. The set of all histories is $A^+ = \bigcup_{\ell \in \mathbb{Z}_+} A^\ell$.
- $\mathbf{f} : A^+ \to A$ is a function given by $\mathbf{f}(H) = (f_1(H), \ldots, f_n(H))$, where $f_i : A^+ \to A_i$ is called node $i$'s *reaction function*.

We now describe the asynchronous dynamics of our model, i.e., the ways that a system's state can evolve due to interactions between nodes. Informally, there is some initial state, and, in each discrete time step $1, 2, 3, \ldots$, a subset of the nodes are *activated* according to a *schedule*. The nodes that are activated in a given timestep react simultaneously; each applies its reaction function to the current state to choose a new action. This updated action is *immediately observable* to all other nodes.[1]

**Definition.** Let $S \subseteq [n]$ be a set of nodes. Define the function $\mathbf{f}_S : A^+ \to A$ by $\mathbf{f}_S(H) = (\hat{f}_1(H), \ldots, \hat{f}_n(H))$, where each function $\hat{f}_i : A^+ \to A_i$ is given by

$$\hat{f}_i(\mathbf{a}^0, \ldots, \mathbf{a}^\ell) = \begin{cases} f_i(\mathbf{a}^0, \ldots, \mathbf{a}^\ell) & \text{if } i \in S \\ a_i^\ell & \text{otherwise}. \end{cases}$$

A *schedule* is a function $\sigma : \mathbb{Z}_+ \to 2^{[n]}$ that maps each $t$ to a (possibly empty) subset of the computational nodes.[2] If $i \in \sigma(t)$, then we say that node $i$ is *activated* at time $t$.

Since the reaction functions are deterministic, an initial history and a schedule completely determine the resulting infinite state sequence; we call these state sequences *trajectories*.

[1] This model has "perfect monitoring." While this is clearly unrealistic in some important real-life contexts, this restriction only strengthens our non-convergence result.

[2] $[n]$ denotes $\{1, \ldots, n\}$, and for any set $S$, $2^S$ is the set of all subsets of $S$.

**Definition.** Let $H = (\mathbf{a}^0, \ldots, \mathbf{a}^\ell) \in A^+$ be a history, and let $\sigma$ be a schedule. The $(H, \sigma)$-*trajectory* of the system is the infinite sequence $\mathbf{a}^0, \mathbf{a}^1, \mathbf{a}^2, \ldots$ extending $H$ such that for every $t > \ell$,

$$\mathbf{a}^t = \mathbf{f}_{\sigma(t)}(\mathbf{a}^0, \ldots, \mathbf{a}^{t-1})$$

The history $(\mathbf{a}^0, \ldots, \mathbf{a}^{t-1})$ is the length-$t$ *prefix* of the $(H, \sigma)$-trajectory.

### 7.1.1 Fairness and Convergence

The main theorem of this chapter is an impossibility result, in which we show that an adversarially chosen initial history and schedule can prevent desirable system behavior. Notice that an arbitrary schedule might never allow some or all nodes to react, or might stop activating them after some time. Hence, we limit this adversarial power (thereby strengthening our impossibility result) by restricting our attention to *fair* schedules, which never permanently stop activating any node.

**Definition.** A *fair schedule* is a schedule $\sigma$ that activates each node infinitely many times, i.e., for each $i \in [n]$, the set $\{t \in \mathbb{Z}_+ : i \in \sigma(t)\}$ is infinite. A *fair trajectory* is one that is the $(H, \sigma)$-trajectory for some history $H$ and some fair schedule $\sigma$.

We are especially interested in whether a system's fair trajectories *converge*, eventually remaining at a single state forever.

**Definition.** A trajectory $\mathbf{a}^0, \mathbf{a}^1, \mathbf{a}^2, \ldots$ *converges* to a state $\mathbf{b}$ if there exists some $T \in \mathbb{Z}_+$ such that, for all $t > T$, $\mathbf{a}^t = \mathbf{b}$. The system is *convergent* if every fair trajectory converges. A state $\mathbf{b}$ is a *limit state* of the system if some fair trajectory converges to $\mathbf{b}$.

Note that it is possible for a trajectory to visit a limit state without converging to that state, meaning that limit states are not necessarily "stable" or "absorbing." They may, however, have basins of attraction, in the sense that reaching certain histories might guarantee convergence to a given limit state.

**Definition.** A history $H$ is *committed* to a limit state $\mathbf{b}$ if, for every fair schedule $\sigma$, the $(H, \sigma)$-trajectory converges to $\mathbf{b}$. An *uncommitted* history is one that is not committed to any state.

### 7.1.2 Informational Restrictions on Reaction Functions

This framework allows for very powerful reaction functions. We now present several possible restrictions on the information they may use. These are illustrated in Fig. 7.1.

Our non-convergence theorem concerns systems in which the reaction functions are *self-independent*, meaning that each node ignores its own past and present actions when reacting to the system's state. In discussing self independence, we use the notation

$$A_{-i} = A_1 \times \ldots \times A_{i-1} \times A_{i+1} \times \ldots \times A_n \,,$$

the state space of the system when $i$ is ignored. Similarly, for a state $\mathbf{a}$, $\mathbf{a}_{-i} \in A_{-i}$ denotes $(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n)$, and given a history $H = (\mathbf{a}^0, \ldots, \mathbf{a}^{\ell-1})$, we write $H_{-i}$ for $(\mathbf{a}^0_{-i}, \ldots, \mathbf{a}^{\ell-1}_{-i})$. Using this notation, we formally define self independence.

**Definition.** A reaction function $f_i$ is *self-independent* if there exists a function $g_i : A^+_{-i} \to A_i$ such that $f_i(H) = g_i(H_{-i})$ for every history $H \in A^+$.

A reaction function has *bounded recall* if it only depends on recent states.

**Definition.** Given $k \in \mathbb{Z}_+$ and a history $H = (\mathbf{a}^0, \ldots, \mathbf{a}^{t-1}) \in A^t$ with $t \geq k$, the *$k$-history* at $H$ is $H_{|k} := (\mathbf{a}^{t-k}, \ldots, \mathbf{a}^{t-1})$, the $k$-tuple of most recent states. A reaction function $f_i$ has *$k$-recall* if it only depends on the $k$-history and the time counter, i.e., there exists a function $g_i : A^k \times \mathbb{Z}_+ \to A_i$ such that $f_i(H) = g_i(H_{|k}, t)$ for every time $t \geq k$ and history $H \in A^t$.

We sometimes slightly abuse notation by referring to the restricted-domain function $g_i$, rather than $f_i$, as the node's reaction function.

A bounded-recall reaction function is *stationary* if it also ignores the time counter.

**Definition.** We say that a $k$-recall reaction function is *stationary* if the time counter $t$ is of no importance. That is, if there exists a function $g_i : A^k \to A_i$ such that $f_i(H) = g_i(H_{|k})$ for every time $t \geq k$ and history $H \in A^t$. A reaction function $f_i$ is *historyless* if $f_i$ is both 1-recall and stationary. That is, if $f_i$ only depends on the nodes' most recent actions.

While seemingly very restricted, historyless dynamics capture the prominent and extensively studied best-response dynamics from game theory (as we discuss in Section 7.3). Historyless dynamics also encompass a host of other applications of interest, ranging from Internet protocols to the adoption of technologies in social network.

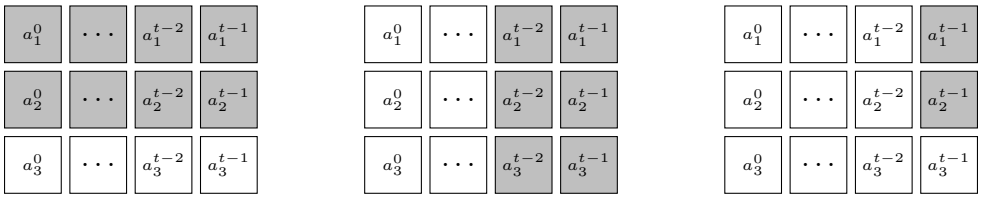


Figure 7.1: Shading shows the information about past and current actions available to node 3 at time $t$ given different reaction function restrictions. Left: self-independent. Node 3 can see the entire record of other nodes' past actions, but not its own. The length of this record gives the current timestamp $t$. Center: 2-recall. Node 3 can see only the two most recent states. Unless the reaction function is stationary, it may also use the value of the current timestamp. Right: self-independent and historyless. Node 3 can only see other nodes' most recent actions and cannot even see the value of the current timestamp.

## 7.2 General Non-convergence Result

We now present a general impossibility result for convergence of nodes' actions under bounded-recall dynamics in asynchronous, distributed computational environments.

**Theorem 43.** *In an interaction system where every reaction function is self-independent and has bounded recall, the existence of multiple limit states implies that the system is not convergent.*

We prove this theorem below, using a valency argument. We then show that the hypotheses of Theorem 43 are necessary and discuss connections of this work to the famous result of Fischer, et al. [36] on the impossibility of resilient consensus.

Note that system convergence is closely related to *self stabilization*, which is a guarantee that the system will reach and remaining within a set of *legitimate states*. For a set $L \subseteq A$, we say that a system *self-stabilizes to* $L$ if for every fair trajectory $\mathbf{a}^0, \mathbf{a}^1, \mathbf{a}^2, \ldots$, there is some $T \in \mathbb{Z}_+$ such that, for every $t > T$, $\mathbf{a}^t \in L$. Theorem 43 precludes self stabilization to any set containing only committed states.

In proving our non-convergence result, we use the following sequence of lemmas. We first show in Lemma 44 that it is sufficient to consider systems with 1-recall reaction functions. Then in Lemma 45, we argue that such a system can be convergent only if every fair trajectory has no committed prefix. To show the existence of a fair trajectory with no committed prefix, we show that in any such system, uncommitted histories exist (Lemma 47), and can be extended to a longer uncommitted histories in a way that activates any given node (Lemma 48). This means that committed prefixes can be avoided forever on a trajectory that activates every node infinitely many times, i.e., a fair trajectory.

**Lemma 44.** *If there exists a convergent interaction system with multiple limit states and bounded-recall, self-independent reaction functions, then there is also a convergent interaction system with multiple limit states and* 1*-recall, self-independent reaction functions.*

*Proof.* Assume that $\Gamma = (n, A, \mathbf{f})$ is a convergent system with self-independent, $k$-recall reaction functions and multiple limit states, for some $k \in \mathbb{Z}_+$. Consider a 1-recall system $\Gamma' = (n, A', \mathbf{f}')$, where $A' = A_1^k \times \ldots \times A_n^k$ and $\mathbf{f}' : A' \times \mathbb{Z}_+ \to A'$ is given by

$$f_i'((a_1^1, \ldots, a_1^k), \ldots, (a_n^1, \ldots, a_n^k), t) = \left[f_i((a_1^1, \ldots, a_n^1), \ldots, (a_1^k, \ldots, a_n^k), kt)\right]^k.$$

Informally, a state in $\Gamma'$ is the transpose of a $k$-history for $\Gamma$. The reaction function $f_i'$ applies $f_i$ to this transpose and repeats the output $k$ times. Notice that $\Gamma'$ has self-independent reaction functions. Furthermore, if $(a_1, \ldots, a_n)$ is a limit state of $\Gamma$, then $((a_1, \ldots, a_1), \ldots, (a_n, \ldots, a_n))$ is a limit state of $\Gamma'$, so $\Gamma'$ also has multiple limit states.

Let $\sigma : \mathbb{Z}_+ \to 2^{[n]}$ be a fair schedule, and let $H \in (A^k)^\ell$ be a history of $\Gamma'$ for some

$\ell \in \mathbb{Z}_+$. Define the schedule $\sigma' : \mathbb{Z}_+ \to 2^{[n]}$ by

$$\sigma'(t) = \begin{cases} \sigma(t/k) & t/k \in \mathbb{Z}_+ \\ \emptyset & \text{otherwise}\,. \end{cases}$$

Notice that $\sigma'$ is also fair. Let $H' \in A^{k\ell}$ be the history for $\Gamma$ formed by concatenating the $k$-tuples in $H$. It is easy to see that the $(H, \sigma)$-trajectory of $\Gamma'$ converges if and only if the $(H', \sigma')$-trajectory of $\Gamma$ converges. Since we assumed that $\Gamma$ is convergent, it follows that $\Gamma'$ is also. □

**Lemma 45.** *Let $\Gamma$ be a convergent system with self-independent, 1-recall reaction functions. Then every fair trajectory in $\Gamma$ has a committed finite prefix.*

*Proof.* Assume there exist some history $H$ and fair schedule $\sigma$ for $\Gamma$ such that the $(H, \sigma)$-trajectory converges to a state $\mathbf{a} = (a_1, \ldots, a_n)$ but has no committed finite prefix. We will construct a fair schedule $\sigma'$ such that the $(H, \sigma')$-trajectory does not converge, giving a contradiction.

Let $\mathbf{u}^0, \mathbf{u}^1, \mathbf{u}^2, \ldots$ be the $(H, \sigma)$-trajectory. Then there is some $t_0 \in \mathbb{Z}_+$ such that $\mathbf{u}^t = \mathbf{a}$ for all $t \geq t_0$. The fairness of $\sigma$ implies that there is some $t_1 > t_0$ such that every node is activated by $\sigma$ between $t_0$ and $t_1$, i.e., $\bigcup_{t_0 < t < t_1} \sigma(t) = [n]$. By assumption, $(\mathbf{u}^0, \ldots, \mathbf{u}^{t_1})$ is not committed to $\mathbf{a}$, which means there is some time $t_2 \geq t_1$ and node $i \in [n]$ such that $f_i(\mathbf{a}, t_2) \neq a_i$. The fairness of $\sigma$ also implies that there is some $t_3 > t_2$ such that $i \in \sigma(t_3)$. Since $t_3 \geq t_0$, we must have $f_i(\mathbf{a}, t_3) = a_i$. By self-independence, then, $f_i(\mathbf{a}', t_3) = a_i$ for all $\mathbf{a}'$ such that $\mathbf{a}'_{-i} = \mathbf{a}_{-i}$.

We use these facts to iteratively build our fair schedule $\sigma'$. In the $(H, \sigma')$-trajectory $\mathbf{v}^0, \mathbf{v}^1, \mathbf{v}^2, \ldots$, the system will repeatedly enter and exit the state $\mathbf{a}$. First, let $\sigma'(t) = \sigma(t)$ for all $1 \leq t \leq t_0$, so that $\mathbf{v}^{t_0} = \mathbf{a}$. Define $\sigma'(t)$ on values $t_0 < t \leq t_2$ as follows.

$$\sigma'(t) = \begin{cases} \sigma(t) & t_0 < t < t_2 \\ \{i\} & t_2 \leq t \leq t_3\,. \end{cases}$$

By our choices of $t_0$, $t_2$, and $t_3$, this partial schedule activates every node and induces a segment

$$(\mathbf{v}^{t_0+1}, \ldots, \mathbf{v}^{t_3})$$

of the $(H, \sigma')$-trajectory such that $\mathbf{v}^t = \mathbf{a}$ whenever $t_0 < t < t_2$ or $t = t_3$, but $\mathbf{v}^{t_2} \neq \mathbf{a}$.

Now set $t_0 = t_3$, select new $t_1$, $t_2$, $i$, and $t_3$ relative to this $t_0$, and iterate this process to define $\sigma'(t)$ for all values $t \in \mathbb{Z}_+$. Notice that $\sigma'$ is fair and that the $(H, \sigma')$-trajectory $\mathbf{v}^0, \mathbf{v}^1, \mathbf{v}^2, \ldots$ does not converge, which contradicts the assumption that the system is convergent. Therefore every fair trajectory in the system must have a committed finite prefix. □

We will use the following consequence of self independence in the course of proving Lemmas 47 and 48.

**Observation 46.** *Let $H' = (\mathbf{a}^0, \ldots, \mathbf{a}^\ell)$ and $H' = (\mathbf{b}^0, \ldots, \mathbf{b}^\ell)$ be committed histories in an system with self-independent, 1-recall reaction functions. If $\mathbf{a}^\ell_{-i} = \mathbf{b}^\ell_{-i}$ for some $i \in [n]$, then $H$ and $H'$ are committed to the same limit state.*

*Proof.* Let $H = (\mathbf{a}^0, \ldots, \mathbf{a}^\ell)$ and $H' = (\mathbf{b}^0, \ldots, \mathbf{b}^\ell)$ be committed histories such that, for some $i \in [n]$, $\mathbf{a}^\ell_{-i} = \mathbf{b}^\ell_{-i}$, as in Fig. 7.2. Let $\sigma$ be any fair schedule such that $\sigma(\ell + 1) = \{i\}$, and consider the $(H, \sigma)$- and $(H', \sigma)$-trajectories. When node $i$ is activated, it will choose the same action regardless of whether the history is $H$ or $H'$, by self independence. As the reaction functions have 1-recall, this means that both these trajectories are identical after time $\ell + 1$. Thus, since $H$ and $H'$ are both committed, they must be committed to the same limit state. □

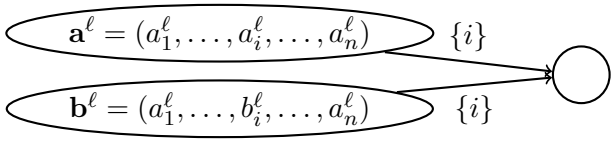


Figure 7.2: Activating $\{i\}$ from $H = (\mathbf{a}^0, \ldots, \mathbf{a}^\ell)$ or $H' = (\mathbf{b}^0, \ldots, \mathbf{b}^\ell)$ will have the same outcome.

**Lemma 47.** *Every interaction system with 1-recall, self-independent reaction functions and more than one limit state has at least one uncommitted history.*

*Proof.* Suppose that every history of length one is committed, and consider two such histories $(\mathbf{a}) = ((a_1, \ldots, a_n))$ and $(\mathbf{b}) = ((b_1, \ldots, b_n))$. Observation 46 implies that, for all $1 \leq i < n$, the histories $((a_1, \ldots, a_{i-1}, b_i, \ldots, b_n))$ and $((a_1, \ldots, a_i, b_{i+1}, \ldots, b_n))$ are committed to the same limit state, and therefore that $\mathbf{a}$ and $\mathbf{b}$ are committed to the same limit state. Thus, all histories of length one must be committed to the same limit state, and it follows that all histories must be committed to the same limit state. This contradicts the system having more than one limit state. □

**Lemma 48.** *Let $(n, A, \mathbf{f})$ be an interaction system with self-independent, 1-recall reaction functions and more than one limit state, let $H = (\mathbf{a}^0, \ldots, \mathbf{a}^{\ell-1}) \in A^\ell$ be an uncommitted history, for some $\ell \in \mathbb{Z}_+$, and let $i \in [n]$ be a node. Then there exist some $t \geq \ell$ and schedule $\sigma$ such that $i \in \sigma(t)$ and the length-$(t+1)$ prefix of the $(H, \sigma)$-trajectory is uncommitted.*

*Proof.* Assume for contradiction that no such $t$ and $\sigma$ exist. Consider all histories that result from activating a set containing $i$ at history $H$. By assumption, each of these histories is committed. Notice that for all $S \subseteq [n]$ and $j \in [n]$, the states $\mathbf{f}_S(H)$ and $\mathbf{f}_{S \cup \{j\}}(H)$ can only differ at coordinate $j$. Hence, we can iteratively apply Observation 46, much as in the proof of Lemma 47, to see that all these histories must be committed to the same limit state, which we call $\mathbf{b}$.

Let $\sigma$ be any fair schedule, and let $\mathbf{a}^0, \mathbf{a}^1, \ldots$ be the $(H, \sigma)$-trajectory. For each $t \in \mathbb{Z}_+$, let $H^t = (\mathbf{a}^0, \ldots, \mathbf{a}^t)$, and notice that $H^{\ell-1} = H$. For each $t \geq \ell$, let $\mathbf{v}^t = \mathbf{f}_{\sigma(t) \cup \{i\}}(H^{t-1})$, and let $I^t = (H^{t-1}, \mathbf{v}^t)$. Since $i \in \sigma(t) \cup \{i\}$, our assumption implies that each history $I^t$ is committed. Let $\mathbf{w}^t = \mathbf{f}_{\{i\}}(H^{t-1})$, and note that by self independence, $\mathbf{f}_{\{i\}}(I^{t-1}) = \mathbf{w}^t$ also, as illustrated in Fig. 7.3. Let $J^t = (I^{t-1}, \mathbf{w}^t)$.

We now show by induction on $t$ that, for every $t \geq \ell$, the history $I^t$ is committed to $\mathbf{b}$. This holds for $t = \ell$ by our definition of $\mathbf{b}$. Fix $t > \ell$, and suppose that $I^{t-1}$ is committed to

**b**. Then $J^t$ is also committed to **b**. Consider all histories that result from activating a set containing $i$ at history $H^{t-1}$. As before, our assumption implies that all these histories are committed, and iterative application of Observation 46 shows that they are all committed to the same limit state. In particular, $I^t$ must be committed to the same limit state as $J^t$, namely **b**.

Since $\sigma$ is a fair schedule, there is some time $t$ for which $i \in \sigma(t)$. For this $t$, we have $H^t = I^t$, so $H^t$ is committed to **b**. Thus for every fair schedule $\sigma$, the $(H, \sigma)$-trajectory converges to **b**, contradicting the assumption that $H$ is uncommitted. We conclude that our assumption was false and that the lemma holds. □

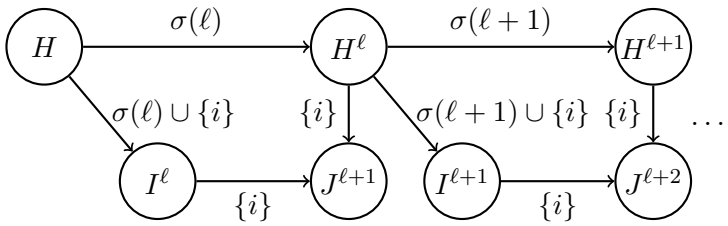


Figure 7.3: All histories in the bottom row are committed to the same limit state **b**.

*Proof of Theorem 43.* It follows from Lemmas 47 and 48 that, in every system with 1-recall, self-independent reaction functions and multiple limit states, it is be possible to activate each node infinitely many times without ever reaching a committed history. This means that every such system has a fair trajectory with no committed prefix. By Lemma 45, this implies that no such system may be convergent. The theorem follows immediately by Lemma 44.[3] □

[3] Although our primary focus is on discrete state spaces, we note that in continuous metric spaces, the standard notion of convergence only requires indefinite approach; the limit point might never be reached. Accordingly, given a metric $d$ on an infinite state space $A$, one could modify Definition 7.1.1 to say that a trajectory $\mathbf{a}^0, \mathbf{a}^1, \mathbf{a}^2, \ldots$ converges to a state **b** if for every $\varepsilon > 0$ there exists some $T \in \mathbb{Z}_+$ such that, for all $t > T$, $d(\mathbf{a}^t, \mathbf{b}) < \varepsilon$. If we require every limit state to have a committed neighborhood, then our proof of Theorem 43 still holds in this setting.

### 7.2.1 Tightness of Theorem 43

The following two examples demonstrate that the statement of Theorem 43 does not hold if either the self-independence restriction or the bounded-recall restriction is removed.

*Example. The self-independence restriction cannot be removed.*

Consider a system with one node, with action space $\{\alpha, \beta\}$. When activated, the node always re-selects its own current action. Observe that the system is convergent despite having two limit states.

*Example. The bounded-recall restriction cannot be removed.*

Consider a system with two nodes, 1 and 2, each with the action space $\{\alpha, \beta\}$. The self-independent reaction functions of the nodes are as follows: node 2 always chooses node 1's action; node 1 will choose $\beta$ if node 2's action changed from $\alpha$ to $\beta$ in the past, and $\alpha$ otherwise. Observe that node 1's reaction function has unbounded recall: it depends on the entire history of interaction. We make the observations that the system is convergent and has two limit states. Observe that if node 1 chooses $\beta$ at some point in time due to the fact that node 2's action changed from $\alpha$ to $\beta$, then it will continue to do so thereafter; if, on the other hand, 1 never does so, then from some point in time onward, node 1's action is constantly $\alpha$. In both cases, node 2 will have the same action as node 1 eventually, and thus convergence to one of the two limit states, $(\alpha, \alpha)$ and $(\beta, \beta)$, is guaranteed. Hence, two limit states exist and the system is convergent nonetheless. Notice also that node 1's reaction functions requires only two states, so the bounded-recall restriction cannot be replaced by a memory restriction.

### 7.2.2 Connection to Consensus Protocols

We now discuss the relationship of our non-convergence result (Theorem 43) to the seminal result of Fischer et al. on the impossibility of *fault-resilient consensus protocols* [36]. The consensus problem is fundamental to distributed computing research. We give a brief

description of it here, and we refer the reader to [36] for a detailed explanation of the model.

Fischer et al. studied an environment in which a group of *processes*, each with an initial value in $\{0, 1\}$, communicate with each other via *messages*. The objective is for all *non-faulty* processes to eventually agree on some *consensus* value $x \in \{0, 1\}$, where $x$ must match the initial value of some process. Fischer et al. established that no consensus protocol is resilient to even a single failure. Their proof of this breakthrough non-termination result introduced the idea of a *valency* argument. They showed that there exists some initial configuration that is *bivalent*, meaning that the resulting consensus could be either 0 and 1 (the outcome depends on the asynchronous schedule of message transmission), and that this bivalence can be maintained. Our proof of Theorem 43 also uses a valency argument, where uncommitted histories play the role of bivalent configurations.

Intuitively, the risk of protocol non-termination in the environment studied by Fischer et al. stems from the possibility of failures; a computational node cannot tell whether another node is silent due to a failure or is simply taking a long time to react. Our non-convergence result concerns environments in which nodes/communication channels do not fail. Thus, each node is guaranteed that all other nodes will eventually react. Observe that in such an environment reaching a consensus is easy; one pre-specified node $i$ (the "dictator") waits until it learns all other nodes' inputs (this is guaranteed to happen as failures are impossible) and then selects a value $v_i$ and informs all other nodes; then, all other nodes select $v_i$. By contrast, the possibility of non-convergence shown in Theorem 43 stems from limitations on nodes' behaviors. Hence, there is no immediate translation from the result of Fischer et al. to ours (and vice versa).

## 7.3 Asynchronous Game Dynamics

Traditionally, work in game theory on game dynamics (e.g., best-response dynamics) relies on the explicit or implicit premise that players' behavior is somehow synchronized (in some contexts play is sequential, while in others it is simultaneous). Here, we consider

the realistic scenario that there is no computational center than can synchronize players' selection of strategies. We describe these dynamics in the setting of this work and exhibit an impossibility result for best-response, and more general, dynamics.

A *game* is characterized by a triple $(n, S, \mathbf{u})$. There are $n$ *players*, $1, \ldots, n$. Each player $i$ has a *strategy set* $S_i$. $S = S_1 \times \ldots \times S_n$ is the space of *strategy profiles* $\mathbf{s} = (s_1, \ldots, s_n)$. Each player $i$ has a *utility function* $u_i : S \to \mathbb{R}$, where $\mathbf{u} = (u_1 \ldots u_n)$. Intuitively, player $i$ "prefers" states for which $u_i$ is higher. Informally, a player is *best responding* when it has no incentive to unilaterally change its strategy.

**Definition.** In a game $U = (n, S, \mathbf{u})$, player $i$ is *best responding* at $\mathbf{s} \in S$ if $u_i(\mathbf{s}) \geq u_i(\mathbf{s}')$ for every $\mathbf{s}' \in S$ such that $s_{-i} = s'_{-i}$. We write $s_i \in BR_i^U(\mathbf{s})$. A strategy profile $\mathbf{s} \in S$ is a *pure Nash equilibrium* (*PNE*) if every player is best responding at $\mathbf{s}$.

There is a natural relationship between games and the interaction systems described in Section 7.1. A player with a strategy set corresponds directly to a node with an action space, and a strategy profile may be viewed as a state. These correspondences are so immediate that we often use these terms interchangeably.

Consider the case of *best-response dynamics* for a game in which best responses are unique (a *generic* game): starting from some arbitrary strategy profile, each player chooses its unique best response to other players' strategies when activated. Convergence to pure Nash equilibria under best-response dynamics is the subject of extensive research in game theory and economics, and both positive [94, 82] and negative [45, 46] results are known. If we view each player $i$ in a game $(n, S, \mathbf{u})$ as a node in an interaction system, then under best-response dynamics its utility function $u_i$ induces a self-independent historyless reaction function $f_i : S_{-i} \to S_i$, as long as best responses are unique. Formally,

$$f_i(\mathbf{a}_{-i}) = \arg\max_{\alpha \in S_i} u_i(a_1, \ldots, \alpha, \ldots, a_n)\,.$$

Conversely, any system with historyless and self-independent reaction functions can be described as following best-response dynamics for a game with unique best responses. Given

reaction functions $f_1, \ldots, f_n$, consider the game where each player $i$'s utility function is given by

$$u_i(\mathbf{a}) = \begin{cases} 1 & \text{if } f_i(\mathbf{a}) = a_i \\ 0 & \text{otherwise}. \end{cases}$$

Best-response dynamics on this game replicate the dynamics induced by those reaction functions. Thus historyless and self-independent dynamics are exactly equivalent to best-response dynamics. Since pure Nash equilibria are fixed points of these dynamics, the historyless case of Theorem 43 may be restated in the following form.

**Theorem 49.** *If there are two or more pure Nash equilibria in a game with unique best responses, then asynchronous best-response dynamics can potentially oscillate indefinitely.*

In fact, best-response dynamics are just one way to derive reaction functions from utility functions, i.e., to translate preferences into behaviors. In general, a *game dynamics protocol* is a mapping from games to systems that makes this translation. Given a game $(n, S, \mathbf{u})$ as input, the protocol selects reaction functions $\mathbf{f} = (f_1, \ldots, f_n)$, and returns an interaction system $(n, S, \mathbf{f})$. The above non-convergence result holds for a large class of these protocols. In particular, it holds for bounded-recall and self-independent game dynamics, whenever pure Nash equilibria are limit states. When cast into game-theoretic terminology, Theorem 43 says that if players' choices of strategies are not synchronized, then the existence of two (or more) pure Nash equilibria implies that this broad class of game dynamics are not guaranteed to reach a pure Nash equilibrium. This result should be contrasted with positive results for such dynamics in the traditional synchronous game-theoretic environments. In particular, this result applies to best-response dynamics with bounded recall and consistent tie-breaking rules (studied by Zapechelnyuk [112]).

**Theorem 50.** *If there are two or more pure Nash equilibria in a game with unique best responses, then all bounded-recall self-independent dynamics for which those equilibria are fixed points can fail to converge in asynchronous environments.*

# Chapter 8

# Self-Stabilizing Uncoupled Dynamics

We now consider dynamic behavior when nodes still have limited knowledge of system history, as in Chapter 7 but the self-independence restriction of Theorem 43 is replaced by a restriction on each node's knowledge of the other nodes' preferences. We use a model of interaction in which the nodes do not have reaction functions *a priori*. Instead, each node is given an input before interaction begins, and the nodes choose reaction functions based on these inputs, thereby defining an interaction system. We say that a protocol for choosing these reaction functions is *uncoupled* if the inputs are private, i.e., if each node must independently choose its reaction function based only on its own input, without knowledge of the other nodes' inputs [47].

In this chapter we ask, in a synchronous environment, when an uncoupled protocol can reliably produce a *self-stabilizing* system. Self stabilization is a failure-resilience property that captures the ability of a system to recover from temporary errors by returning to a designated set of "legitimate" states. An uncoupled protocol that always yields self-stabilizing interaction systems, then, captures the distributed computing concepts of both distribution and resilience to temporary failures. We investigate the ability of game dynamics to reach strategic equilibrium in this distributed setting.

Uncoupled dynamics were introduced by Hart and Mas-Colell [45], who introduced the concept of uncoupled game dynamics and gave a non-convergence result in a continuous-time setting. In a later paper [46], the same authors showed that this result held in a discrete-time setting, even in the presence of randomness, and addressed convergence to mixed Nash

equilibria by bounded-recall uncoupled dynamics. Babichenko [3, 4] investigated the situation when the uncoupled nodes are finite-state automata, as well as *completely uncoupled dynamics*, in which each node can see only the history of its own actions and payoffs. Hart and Mansour [44] analyzed the time to convergence for uncoupled dynamics. In a related model, Young [110] and Pradelski and Young [90] gave completely uncoupled dynamics that achieve an equilibrium in a high proportion of steps but do not necessarily converge.

In Section 8.1, we incorporate the discrete-time game dynamics setting of Hart and Mas-Colell [46] into the distributed computing framework of Chapter 7 and state a separation result showing that randomized uncoupled dynamics are more powerful than deterministic uncoupled dynamics for some games; the proof of this result is in Section 8.4. In Section 8.2, we give an uncoupled protocol for deterministic bounded-recall dynamics that self-stabilizes on all games. We give another such protocol that self-stabilizes on all games in which each node has at least four available actions, and we prove that the recall used by this protocol is optimal for that class of games. Section 8.3 is concerned with randomized uncoupled dynamics; we describe earlier work on this topic and characterize the state spaces over which historyless randomized uncoupled game dynamics can be self-stabilizing. Table 8 summarizes the results of this section. For the sake of brevity, we only discuss systems in which every node is *non-trivial*, having an action space of size at least 2.

| | Deterministic | Randomized |
|---|---|---|
| 1-recall | *never* (Theorem 52) | *iff* $n = 2$ *and some* $\lvert A_i\rvert = 2$ (Theorem 59) |
| 2-recall | *if every* $\lvert A_i\rvert \geq 4$ (Theorem 54) | always (Theorem 56 [46]) |
| 3-recall | *always* (Theorem 53) | always |

Table 8.1: For $n \geq 2$ non-trivial nodes, is the given amount of recall sufficient for a stationary uncoupled protocol to be self-stabilizing on all games over state space $A = A_1 \times \ldots \times A_n$? Entries in *italics* are contributions of this work.

## 8.1 Historyless Uncoupled Game Dynamics

In uncoupled game dynamics, the nodes' private inputs are real-valued utility functions over the state space, and the legitimate states for self stabilization are the pure Nash equilibria (PNE) of the game induced by these utility functions. Intuitively, each node acts as a strategic agent with private preferences about the system's state, and the global objective is for the system to be in a state that is compatible with all those preferences, in the sense that every node is best-responding at that state. Game dynamics were discussed more generally in Section 7.3. In this context, a self-stabilizing system is one that, given a completely synchronous schedule, is guaranteed to reach and remain within the set of PNE, regardless of the initial state.

**Definition.** A synchronous system with game dynamics is *self-stabilizing* if, for every 1-fair trajectory $\mathbf{a}^0, \mathbf{a}^1, \mathbf{a}^2, \ldots$, there is some $T \in \mathbb{Z}_+$ such that, for every $t > T$, $\mathbf{a}^t$ is a PNE.

Traditional study of convergence to equilibria in game dynamics makes various assumptions about the "rationality" of nodes' behavior in each timestep, restricting them to always play the game in ways that are somehow consistent with their self-interest given their current knowledge. In contrast to these behavioral restrictions on the nodes, uncoupledness is an informational restriction, in that the nodes have no knowledge of each other's payoffs. In this situation, no individual node can recognize a PNE, so finding an equilibrium is a truly distributed task. We formally define an uncoupled protocol in this setting.

**Definition.** An *uncoupled protocol* for game dynamics takes a game $(n, A, \mathbf{u})$ and returns an interaction system $(n, A, \mathbf{f})$. For each node $i$, the protocol provides a mapping that takes the utility function $u_i$ as input (which implicitly reveals $n$ and $A$), and returns a reaction function $f_i$.

Given an $n$-node game, the protocol describes an $n$-node interaction system whose state space is the game's strategy profile space. Informally, when each node chooses its reaction

function for this system, it knows the number of nodes, the state space, and its own preferences, but it does not know anything about the other nodes' preferences. We say this protocol is *self-stabilizing* if it produces a self-stabilizing system whenever a PNE exists.

**Definition.** An uncoupled protocol for game dynamics is *self-stabilizing* on a given class of games if the protocol yields a self-stabilizing system for every game in that class that has at least one PNE.

If uncoupledness is the only restriction on the dynamics, then the nodes can find a PNE through a straightforward exhaustive search. However, this changes when nodes' ability to remember past actions is restricted. Hart and Mas-Colell [46] showed that when the reaction functions are required to be historyless, no uncoupled protocol can guarantee convergence to a PNE, even in games with a unique PNE. (They also showed this in a continuous time setting in an earlier paper [45].) They proved this theorem by two counterexamples. The first was a $3 \times 3$ game, and the second was a $3 \times 3 \times 3$ game in which best responses are unique. We restate this result using the terminology of the present work.

**Theorem 51** (Hart and Mas-Colell [46])**.** *No historyless uncoupled protocol is self-stabilizing for all games. In particular, any such protocol fails to self-stabilize on some* $3 \times 3$ *game and on some* $3 \times 3 \times 3$ *game with unique best responses.*

To obtain positive results for self stabilization in bounded-recall uncoupled game dynamics, we might place restrictions on the state space or the inputs, or we might make the nodes more powerful, allowing them to use more sophisticated reaction functions. Restricting the state space $A$ can be done by limiting the number of nodes $n$ or the number of actions each node has available. The inputs, which are utility functions, may be restricted so that we only require self stabilization on some natural class of games, e.g., zero-sum games or games with unique best responses. To make the nodes more powerful, we might increase the number of steps of recall they may access.

Another possibility is to permit the reaction functions to be randomized, a relaxation

we discuss in detail in Section 8.3. Hart and Mas-Colell [46] showed that the analogue of Theorem 51 for nodes with randomized reaction functions holds, but that self stabilization is possible for historyless uncoupled protocols when the nodes are allowed to use both randomness and an additional step of recall, or when both the state space and the input space are restricted (i.e., in 2-node games with unique best responses). We show in Lemma 60 that using randomness also enables historyless uncoupled protocols to self-stabilize over sufficiently restricted state spaces (e.g., for all $2 \times 2$ games). In contrast, there is no nontrivial state space over which historyless deterministic uncoupled protocols can self-stabilize. We defer the proof of this theorem until Section 8.4.

**Theorem 52.** *For every $n \geq 2$ and every $n$-node state space $A$, there is no historyless deterministic uncoupled protocol that self-stabilizes on all games over $A$.*

## 8.2 Self-stabilizing Deterministic Uncoupled Dynamics

The challenge for deterministic nodes in searching the state space for a PNE is keeping track of their progress in that search. Here we show that by using repeated states to coordinate, the nodes can take advantage of two additional steps of recall to perform an exhaustive search, thereby executing a self-stabilizing uncoupled protocol.

**Theorem 53.** *There is a stationary 3-recall deterministic uncoupled protocol that is self-stabilizing on all games.*

*Proof.* Let $n, k_1, \ldots, k_n \geq 2$, and $A = [k_1] \times \ldots \times [k_n]$. Let $\pi : A \to A$ be a cyclic permutation on the states. We describe, for any game $U = (n, A, \mathbf{u})$, how our protocol chooses reaction functions. We write $\pi_i(\mathbf{a})$ for the action of node $i$ in $\pi(\mathbf{a})$. For every node $i$, the 3-recall

reaction function $f_i : A^3 \to A_i$ is given by

$$f_i(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \begin{cases} c_i & \text{if } \mathbf{b} = \mathbf{c} \text{ and } c_i \in BR_i(\mathbf{c}) \\ \min BR_i(\mathbf{c}) & \text{if } \mathbf{b} = \mathbf{c} \text{ and } c_i \notin BR_i(\mathbf{c}) \\ \pi_i(\mathbf{a}) & \text{if } \mathbf{a} = \mathbf{b} \neq \mathbf{c} \\ c_i & \text{otherwise}. \end{cases}$$

Notice that this protocol is uncoupled: $f_i$ does not depend on any other node's utility function. Informally, the nodes use repetition to keep track of which state is the current "PNE candidate" in each stage. If a state has just been repeated, then it is the current candidate, and each node plays a best response to it, with a preference against moving. If the nodes look back and see that some state $\mathbf{a}$ was repeated in the past but then followed by a different state, they infer that $\mathbf{a}$ was rejected as a candidate and move on by playing $\mathbf{a}$'s successor, $\pi(\mathbf{a})$. Otherwise the nodes repeat the most recent state, establishing it as the new candidate. We call these three types of histories *query*, *move-on*, and *repeat* histories, respectively. Here "query" refers to asking each node for one of its best responses to $\mathbf{b}$.

We now show that every 1-fair trajectory in the system $(n, A, \mathbf{f})$ converges to a PNE if one exists. Assume that $U$ has at least one PNE. Let $X = (\mathbf{a}, \mathbf{b}, \mathbf{c}) \in A^3$, and let $Y$ be the next history $(\mathbf{b}, \mathbf{c}, \mathbf{f}(X))$. If $X$ is a repeat history, then $Y = (\mathbf{b}, \mathbf{c}, \mathbf{c})$, which is a query history. If $X$ is a move-on history, then $\mathbf{b} \neq \mathbf{c}$, and $Y = (\mathbf{b}, \mathbf{c}, \pi(\mathbf{a}))$. If $\mathbf{c} = \pi(\mathbf{a})$, then this is a query history; otherwise, it is a repeat history, which will be followed by the query history $(\mathbf{c}, \pi(\mathbf{a}), \pi(\mathbf{a}))$. Thus, every non-query history will be followed within two stages by a query history.

Now let $X = (\mathbf{a}, \mathbf{b}, \mathbf{b}) \in A^3$ be a query history, and let $Y$ and $Z$ be the next two histories. If $\mathbf{b}$ is a PNE, then $Y = (\mathbf{b}, \mathbf{b}, \mathbf{b})$, and the system has converged. Otherwise, $Y = (\mathbf{b}, \mathbf{b}, \mathbf{c})$ for some $\mathbf{c} \neq \mathbf{b}$, so $Y$ is a move-on history, which will be followed by a query history $(\mathbf{b}, \pi(\mathbf{b}), \pi(\mathbf{b}))$ or $(\mathbf{c}, \pi(\mathbf{b}), \pi(\mathbf{b}))$ within two stages. Let $\mathbf{p}$ be a PNE for $U$. Since $\pi$ is cyclic, $\mathbf{p} = \pi^r(\mathbf{b})$ for some $r \in \mathbb{Z}_+$. So $(\mathbf{p}, \mathbf{p}, \mathbf{p})$ is reachable from $X$ unless $\pi^s(\mathbf{b})$ is a PNE for some $s < r$. It follows that the system always converges to a PNE, so it is

self-stabilizing for $U$. It follows that the protocol is self-stabilizing on all games. □

We can improve on this recall bound when every node has at least four actions in its action space. It is perhaps surprising that larger action spaces are an advantage here, since they correspond to a larger state space that must be searched. This essentially gives the nodes a larger alphabet, though, with which to encode information about the current status of the search. Notice that 2-recall is optimal for this task; by Theorem 52, historyless deterministic uncoupled dynamics could not self-stabilize on this class of games.

**Theorem 54.** *There is a stationary 2-recall deterministic uncoupled protocol that is self-stabilizing on all games in which every node has at least four available actions.*

*Proof.* Let $n \geq 2$, $k_1, \ldots, k_n \geq 4$ and $A = [k_1] \times \ldots \times [k_n]$. Define a permutation $\pi : A \to A$ such that for every $\mathbf{a} \in A$, $\pi(\mathbf{a})$ is $\mathbf{a}$'s lexicographic successor. Formally,

$$\pi(\mathbf{a}) = (\pi_1(\mathbf{a}), \ldots, \pi_n(\mathbf{a}))\,,$$

where for $i = 1, \ldots, n-1$,

$$\pi_i(\mathbf{a}) = \begin{cases} a_i + 1 \bmod k_i & \text{if } a_j = k_j \text{ for every } j \in \{i+1, \ldots, n\} \\ a_i & \text{otherwise}\,, \end{cases}$$

and $\pi_n(\mathbf{a}) = a_n + 1 \bmod k_n$. Observe then that $\pi$ is cyclic, and for each node $i$ and $\mathbf{a} \in A$, we have $\pi_i(\mathbf{a}) - a_i \bmod k_i \in \{0, 1\}$.

Let $U = (n, A, u)$ be a game. We now describe how our protocol chooses the $n$-tuple of reaction functions $\mathbf{f} = (f_1, \ldots, f_n)$ for $U$. The reaction functions will differentiate between three types of histories, each named according to the event it prompts. Let $(\mathbf{a}, \mathbf{b}) \in A^2$.

- *move-on*: If $\mathbf{a} \neq \mathbf{b}$ and $a_j - b_j \bmod k_j \in \{0, 1\}$ for every $j \in [n]$, then the nodes "move on" from $\mathbf{a}$, in the sense that each node $i$ plays $\pi_i(\mathbf{a})$, giving $f(\mathbf{a}, \mathbf{b}) = \pi(\mathbf{a})$.

- *query*: If $b_j - a_j \bmod k_j \in \{0, 1, 2\}$ for every $j \in [n]$, then we "query" each node's utility function to check whether it is best-responding at $\mathbf{b}$. Each node $i$ answers by

playing $b_i$ if it is best-responding and $b_i - 1 \bmod k_i$ if it is not. So at a query history $(\mathbf{a}, \mathbf{b})$, for $i \in [n]$,

$$f_i(\mathbf{a}, \mathbf{b}) = \begin{cases} b_i & \text{if } b_i \in BR_i^U(\mathbf{b}) \\ b_i - 1 \bmod k_i & \text{otherwise}. \end{cases}$$

- *repeat*: Otherwise, each node $i$ "repeats" by playing $b_i$, giving $\mathbf{f}(\mathbf{a}, \mathbf{b}) = \mathbf{b}$.

Notice that because $k_1, \ldots, k_n \geq 4$, it is never the case that both $a_j - b_j \bmod k_j \in \{0, 1\}$ and $b_j - a_j \bmod k_j \in \{0, 1, 2\}$. Thus, the conditions for the *move-on* and *query* types are mutually exclusive, and the three history types are all disjoint.

The history following $(\mathbf{a}, \mathbf{b})$ is $(\mathbf{b}, \mathbf{f}(\mathbf{a}, \mathbf{b}))$. There are three cases:

1. $(\mathbf{a}, \mathbf{b})$ is a move-on history. Then this is $(\mathbf{b}, \pi(\mathbf{a}))$. Since for every node $i$, $a_i - b_i \bmod k_i \in \{0, 1\}$ and $\pi(\mathbf{a})_i - a_i \bmod k_i \in \{0, 1\}$, we have $\pi_i(\mathbf{a}) - b_i \bmod k_i \in \{0, 1, 2\}$, so $(\mathbf{b}, \pi(\mathbf{a}))$ is a query history.

2. $(\mathbf{a}, \mathbf{b})$ is a query history. Then $b_i - f_i(\mathbf{a}, \mathbf{b}) \bmod k_i \in \{0, 1\}$ for every node $i$, so $(\mathbf{b}, \mathbf{f}(\mathbf{a}, \mathbf{b}))$ is a move-on history unless $\mathbf{b} = \mathbf{f}(\mathbf{a}, \mathbf{b})$, in which case it is a query history. But if $\mathbf{b} = \mathbf{f}(\mathbf{a}, \mathbf{b})$ and $(\mathbf{a}, \mathbf{b})$ was a query history, then $b_i \in BR_i^U(\mathbf{b})$ for every node $i$, i.e., $\mathbf{b}$ is a PNE.

3. $(\mathbf{a}, \mathbf{b})$ is a repeat history. Then it is followed by the repeat history $(\mathbf{b}, \mathbf{b})$.

Thus, histories of move-on or repeat type are always followed by histories of query type, and query histories are never followed by repeat histories. We conclude that with the possible exception of the initial history, every history will be of move-on or query type, and no two consecutive histories will be of move-on type. In particular, some query history is reachable from every initial history.

Any query history $(\mathbf{a}, \mathbf{b})$ will be followed by $(\mathbf{b}, \mathbf{b})$ if and only if $\mathbf{b}$ is a PNE, in which case the system will have converged. If $\mathbf{b}$ is not a PNE, then $(\mathbf{a}, \mathbf{b})$ will be followed by a move-on history $(\mathbf{b}, \mathbf{c})$, for some $\mathbf{c} \in A$. This will be followed by the query history $(\mathbf{c}, \pi(\mathbf{b}))$.

Continuing inductively, since $\pi$ is cyclic, unless the system converges to a PNE, the nodes will examine every state $\mathbf{a} \in A$ with a query state of the form $(\mathbf{a}, \mathbf{b})$. Thus, if $U$ has at least one PNE, then the system will converges to a PNE, meaning that this protocol is self-stabilizing on the given class of games. □

We have shown that 3-recall is sufficient for a stationary deterministic uncoupled protocol to be self-stabilizing on all games, and that 2-recall is necessary. It remains open whether 2-recall is sufficient. We conjecture that it is not.

**Conjecture 55.** *There is no stationary 2-recall deterministic uncoupled protocol that is self-stabilizing on all games.*

## 8.3 Uncoupled Game Dynamics with Randomness

In this section we characterize the state spaces for which historyless uncoupled protocols can be self-stabilizing if the nodes are permitted to use randomized reaction functions. In this case each reaction function $f_i$ has $\Delta(A_i)$, the set of probability distributions over $A_i$, as its range. In every stage, $i$'s action is drawn from the distribution given by its reaction function: $a_i^t \sim f_i(\mathbf{a}^0, \ldots, \mathbf{a}^{t-1})$. We say that the system is self-stabilizing if, regardless of the initial state, the probability that $\mathbf{a}^t$ is a PNE approaches 1 as $t \to \infty$.

Hart and Mas-Colell [45, 46] showed that historyless uncoupled protocols cannot be self-stabilizing on all games (Theorem 51 above). They also showed, however, that there is a historyless randomized uncoupled protocol that self-stabilizes on two-node games in which every node has a unique best response to every state, and that adding recall admits a protocol that self-stabilizes on all games.

**Theorem 56** (Hart and Mas-Colell [46])**.** *There is a historyless randomized uncoupled protocol that is self-stabilizing on all two-node games with unique best responses. There is a stationary 2-recall randomized uncoupled protocol that is self-stabilizing on all games.*

Both of these protocols are simple, especially the historyless protocol, which we refer to as the *stay-or-roll protocol*; each node "stays" if it is best-responding and "rolls the dice" otherwise. More formally, in every stage, each node checks whether it is currently best-responding. If it is, then it continues with the same action. If it is not, then it selects an action uniformly at random from its action space [46].

Lemma 57 states that any historyless uncoupled protocol that self-stabilizes on all games over a given state space must be similar to this protocol: each node must continue playing the same action when best-responding and otherwise move to a new action with positive probability (*w.p.p.*). This lemma expands on the main idea in Hart and Mas-Colell's proof of Theorem 51, that convergence to a unique PNE can be guaranteed only if best-responding nodes never change their actions. Intuitively, a historyless node that is best-responding doesn't know whether the system is at the PNE, so this node changing its action w.p.p. would mean that the system could not guarantee permanent convergence to that state. The only exception is when a node is certain that there is more than one PNE, which happens when the game has two players and that node has a multiple actions that are always best responses.

**Definition.** An action $\alpha \in A_i$ is *weakly dominant* if it is a best response for node $i$ at every state. It is *strongly dominant* if it is a unique best response for $i$ at every state.

**Lemma 57.** *Suppose that a historyless uncoupled protocol* $\Pi$ *is self-stabilizing on all games over an* $n$*-node state space* $A$*. Let* $U$ *be a game over* $A$*. Under the reaction functions given by* $\Pi(U)$*, the following hold.*

1. *Any node that is not best-responding will change its action w.p.p.*
2. *Any node that is best-responding will change its action with probability 0, unless* $n = 2$ *and that node has more than one weakly dominant action in* $U$*.*

*Proof.* Assume that $\Pi$ is a historyless uncoupled protocol that is self-stabilizing on all games over $A$. Let $u_i$ be a utility function for node $i$, and suppose that $i$ is best-responding for $u_i$

at state $\mathbf{a}$. Let $f_i$ be $i$'s reaction function given by $\Pi$ for $u_i$. If $n > 2$ or $i$ has at most one weakly dominant action in $u_i$, then there is some game over $A$ in which $\mathbf{a}$ is the unique PNE and $i$'s utility function is $u_i$, meaning that $\mathbf{a}$ must be a fixed point of $f_i$, since otherwise the system for that game would not converge to $\mathbf{a}$. Thus, each node changes its action with probability 0 whenever it is best-responding.

Let $\mathbf{b} \in A$ be a state such that $i$ is not best-responding for $u_i$ at $\mathbf{b}$. There is some game in which $\mathbf{b}$ is a PNE, $i$'s utility function is $u_i$, and no node other than $i$ has more than one weakly dominant action. Then as shown above, it is certain that no node $j \neq i$ will change its action from $b_j$ in the system given by $\Pi$ for this game. This means that unless $i$ changes its action w.p.p., the system will converge to $\mathbf{b}$, which is not a PNE. This would contradict the self stabilization of $\Pi$ on games over $A$, so we conclude that any node that is not best-responding must change its action w.p.p. □

An important consequence of Lemma 57 is that the stay-or-roll protocol is essentially the best historyless randomized uncoupled protocol, in the sense that it is self-stabilizing whenever any such protocol is. Like Lemma 57, the following lemma expands on an idea that is implicit in the work of Hart and Mas-Colell [46].

**Lemma 58.** *For any state space $A$, there is a historyless randomized uncoupled protocol that self-stabilizes on all games over $A$ if and only if the stay-or-roll protocol self-stabilizes on all games over $A$.*

*Proof.* The stay-or-roll protocol is a historyless randomized uncoupled protocol, so one direction is trivial. For the other direction, assume that $\Pi$ is a historyless randomized uncoupled protocol that is self-stabilizing on all games over an $n$-node state space $A$. Let $U$ be a game over $A$ that has at least one PNE. Assume that $n \geq 3$ or no node has multiple weakly dominant actions, so that Lemma 57 applies. For any state $\mathbf{a} \in A$, some PNE $\mathbf{p}$ is reachable in the system $\Pi(U)$. In $\Pi(U)$ a node can change its action only when it is not best-responding, so for every action change on the path from $\mathbf{a}$ to $\mathbf{p}$ in $\Pi(U)$, the reaction

function given by the stay-or-roll protocol would have made the same change w.p.p. This means that $\mathbf{p}$ is also reachable from $\mathbf{a}$ in the stay-or-roll system for $U$, so the stay-or-roll protocol self-stabilizes on $U$.

To prove the lemma, then, it suffices to show that the stay-or-roll protocol self-stabilizes on all two-node games in which some node has a weakly dominant action. Let $A$ be a two-node state space, and let $U$ be a game over $A$ in which node 1 has at least one weakly dominant action. Consider the system given by the stay-or-roll protocol for $U$. Let $\mathbf{a}$ be any state in $A$. If $a_1$ is weakly dominant, then either node 2 is best-responding at $\mathbf{a}$ or, w.p.p., node 2 will play a best response to $a_1$ in the next stage. Since node 1 will continue playing $a_1$, which is a best response to all of node 2's actions, a PNE is reachable from every state in which node 1 plays a weakly dominant action. If node 1 is not best-responding at $\mathbf{a}$, then w.p.p. $f_1(\mathbf{a})$ is a weakly dominant action, so a PNE is also reachable from every state in which node 1 is not best-responding. If node 1 is best-responding at $\mathbf{a}$ but $a_1$ is not weakly dominant, then either $\mathbf{a}$ is a PNE or node 2 is not best-responding at $\mathbf{a}$. In the latter case, node 2 will change w.p.p. to an action that $a_1$ is not a best response to. Then w.p.p. node 1 will play a weakly dominant action in the next stage. Thus, a PNE is reachable from every state $\mathbf{a} \in A$. We conclude that the stay-or-roll protocol is self-stabilizing on all such games. □

With Lemma 58 in hand, we can determine the recall needed for a stationary randomized uncoupled protocol to self-stabilize on all games over any state space: a self-stabilizing historyless uncoupled protocol exists if and only if the stay-or-roll protocol is self-stabilizing. Otherwise, two steps of recall is necessary, and by Theorem 56 is also sufficient.

Recall from Theorem 51 and Theorem 56 that Hart and Mas-Colell showed that a self-stabilizing historyless randomized uncoupled protocol exists for all two-node games with unique best responses, but not for all $3 \times 3$ games or all $3 \times 3 \times 3$ games with unique best responses. Theorem 59 essentially says that such protocols also exist for all $2 \times k$ games, but that further expansion of the state space makes self stabilization over all games impossible

for these protocols.

**Theorem 59.** *Let $n \geq 2$, and let $A$ be an $n$-node state space. If $n = 2$ and $|A_1|$ or $|A_2| = 2$, then there is a historyless randomized uncoupled protocol that self-stabilizes on all games over $A$. Otherwise, there is not.*

To prove this theorem, we use four lemmas. The first, Lemma 60, shows the positive part of the theorem, that the stay-or-roll protocol self-stabilizes on games over two-node state spaces where one node has only two actions. The second and third are reduction lemmas, showing that enlarging a state space—either by adding an action to some node's action space (Lemma 61) or by adding a node (Lemma 62)—never improves the stay-or-roll protocol's ability to self-stabilize on all games over that state space. Finally, Lemma 63 says that the stay-or-roll protocol fails to self-stabilize on all $2 \times 2 \times 2$ games.

**Lemma 60.** *For $k \geq 2$, the stay-or-roll protocol self-stabilizes on all $2 \times k$ games.*

*Proof.* Let $A = \{1, 2\} \times \{1, \ldots k\}$, and let $u_1$ and $u_2$ be the nodes' respective utility functions over this state space. Suppose that the resulting game has at least one PNE, and let $\mathbf{a} = (a_1, a_2) \in A$ be the initial state. Let $\mathbf{f} = (f_1, f_2)$ be the pair of reaction functions given by the stay-or-roll protocol for this game. Notice that if the nodes reach a PNE, they will remain in that state. Consider four cases:

1. Node 1 is best-responding at $\mathbf{a}$ and $a_1 = p_1$ for some PNE $\mathbf{p} = (p_1, p_2)$. Then either node 2 is also best-responding and $\mathbf{a}$ is a PNE, or node 2 is not best-responding and will change its action to $p_2$ in the next stage w.p.p.

2. Node 1 is not best-responding at $\mathbf{a}$ and there is no PNE $\mathbf{p}$ such that $a_1 = p_1$. Then w.p.p. $f_1(\mathbf{a}) \neq a_1$. Since the game has a PNE, $\mathbf{f}(\mathbf{a})$ is then an instance of case 1.

3. Node 1 is best-responding at $\mathbf{a}$ and there is no PNE $\mathbf{p}$ such that $a_1 = p_1$. Then node 2 is not best-responding at $\mathbf{a}$, so w.p.p. $f_2(\mathbf{a}) \in BR_2(\mathbf{a})$, but $\mathbf{f}(\mathbf{a})$ cannot be a PNE

since $f_1(\mathbf{a}) = a_1$. Then node 1 is not best-responding at $\mathbf{f}(\mathbf{a})$, so $\mathbf{f}(\mathbf{a})$ is an instance of case (2).

4. Node 1 is not best-responding at $\mathbf{a}$ and $a_1 = p_1$ for some PNE $\mathbf{p}$. Then w.p.p. $f_1(\mathbf{a}) \neq a_1$ and $f_2(\mathbf{a}) = a_2$, in which case node 1 is best-responding at $\mathbf{f}(\mathbf{a}) = (f_1(\mathbf{a}), a_2)$, since node 1 has only two actions. Then $\mathbf{f}(\mathbf{a})$ is an instance of (1) or (3).

Thus, from every state $\mathbf{a} \in A$, some PNE for is reachable from $\mathbf{a}$. We conclude that the protocol is self-stabilizing on all such games. □

Informally, Lemma 61 says that expanding a node's action space does not make a state space any "easier" for self-stabilizing game dynamics. The proof relies on a reduction in which the nodes take advantage of a protocol for a larger game by "pretending" to play the larger game. Whenever node $i$ plays its highest-indexed action $k_i$, all nodes guess randomly whether $i$ would have played $k_i$ or $k_i + 1$ in the larger game.

**Lemma 61.** *Let $A$ and $A'$ be state spaces for $n \geq 2$ nodes such that $|A'_i| = |A_i| + 1$ for some $i$ and $|A'_j| = |A_j|$ for every $j \neq i$. If the stay-or-roll protocol is self-stabilizing on all games over $A'$, then it is self-stabilizing on all games over $A$.*

*Proof.* For $n, k_1, \ldots, k_n \geq 2$, let

$$A = [k_1] \times \ldots \times [k_n] .$$

Since nodes can be reordered, assume $i = 1$ and

$$A' = [k_1 + 1] \times [k_2] \times \ldots \times [k_n] .$$

Suppose that the stay-or-roll protocol is self-stabilizing on all games over $A'$. Let $U = (n, A, \mathbf{u})$ be a game that has at least one PNE, and define another game $U' = (n, A', \mathbf{u}')$ such that for every node $j \in [n]$ and state $\mathbf{a} \in A$, $u'_j(\mathbf{a}) = u_j(\mathbf{a})$ and

$$u'_j(k_1 + 1, a_2, \ldots, a_n) = u_j(k_1, a_2, \ldots, a_n) .$$

Thus, in $U'$, every node is always indifferent to whether node $i$ plays $k_i$ or $k_i + 1$.

We now define a new protocol by its randomized reaction functions for $U$. Let $\mathbf{f} = (f_1, \ldots, f_n)$ be the reaction functions given by the stay-or-roll protocol for $U'$. In our new protocol, the reaction functions $\mathbf{g} = (g_1, \ldots, g_n)$ for $U$ are defined for each node $j$ by $g_j(\mathbf{a}) = f_j(\mathbf{a})$ when $a_1 \neq k_1$. When $a_1 = k_1$,

$$g_j(\mathbf{a}) = \begin{cases} \min\{k_j, f_j(\mathbf{a})\} & \text{with probability } 1/2 \\ \min\{k_j, f_j(k_1 + 1, a_2, \ldots, a_n)\} & \text{otherwise}\,. \end{cases}$$

That is, whenever the nodes see that node 1 has played $k_1$, each decides independently at random to interpret that action either as $k_1$ or as $k_1 + 1$, then plays the action prescribed by the stay-or-roll protocol. The minimum operator ensures that node 1 is never instructed to play the action $k_1 + 1$, which is not in its action space for $U$.

Let $\mathbf{p}$ be a PNE for $U$. To see that $\mathbf{p}$ is an absorbing state in the system $(n, A, \mathbf{g})$, notice that $\mathbf{p}$ is also a PNE for $U'$, hence $f(p_j) = p_j$ for every node $j$, and

$$g_j(\mathbf{p}) = \min\{k_j, f_j(\mathbf{p})\} = p_j\,.$$

It remains to show that the system always reaches a PNE. Let $\mathbf{a} \in A \subseteq A'$. Since $U'$ has a PNE and the stay-or-roll protocol is self-stabilizing on $U'$, there is some PNE

$$\mathbf{q} = (q_1, \ldots, q_n) \in A'$$

for $U'$ such that $\mathbf{q}$ is reachable from $\mathbf{a}$ in the system $(n, A', \mathbf{f})$. So $q$ is reachable from $\mathbf{a}$ in that system, i.e., there is some $T \in \mathbb{Z}_+$ and 1-fair path $\mathbf{a}^0, \ldots, \mathbf{a}^T$ in such that $\mathbf{a}^0 = \mathbf{a}$ and $\mathbf{a}^T = \mathbf{q}$. Since $\mathbf{q}$ is a PNE for $U'$,

$$\mathbf{q}' = (\min\{q_1, k_1\}, q_2, \ldots, q_n)$$

is a PNE for both $U$ and $U'$.

Now let $\mathbf{b}^0, \ldots, \mathbf{b}^T$ be a 1-fair path in the system $(n, A, \mathbf{g})$ such that $\mathbf{b}^0 = \mathbf{a}$. Suppose, for some $0 \leq t < T$, that

$$\mathbf{b}^t = (\min\{a_1^t, k_1\}, a_2^t, \ldots, a_n^t)\,.$$

Then w.p.p.,

$$\mathbf{b}^{t+1} = \mathbf{g}(\mathbf{b}^t) = (\min\{a_1^{t+1}, k_1\}, a_2^{t+1}, \ldots, a_n^{t+1})\,.$$

Hence, by induction, $\mathbf{b}^T = \mathbf{q}'$ w.p.p., meaning that a PNE is reachable from $\mathbf{a}$ in this system. We conclude that the protocol giving $\mathbf{f}$ is self-stabilizing on all games over $A$, and by Lemma 58, it follows that the stay-or-roll protocol is also self-stabilizing on this class of games. □

Lemma 62 says that adding another node also doesn't help, and its proof uses another reduction. Here, the nodes use a protocol for an $(n+1)$-node game to play a $n$-node game by pretending there is an additional node that never wants to change its action.

**Lemma 62.** *Let $A$ be a state space for $n \geq 2$ nodes, and let $A'$ be a state space for $n+1$ nodes such that $|A_i| = |A'_i|$ for every $i \in [n]$. If the stay-or-roll protocol is self-stabilizing on all games over $A'$, then it is self-stabilizing on all games over $A$.*

*Proof.* For $n, k_1, \ldots, k_{n+1} \geq 2$, let $A = [k_1] \times \ldots \times [k_n]$ and $A' = [k_1] \times \ldots \times [k_{n+1}]$. Let $U = (n, A, \mathbf{u})$ be a game that has at least one PNE. Define another game $U' = (n, A', \mathbf{u}')$ such that for every state $\mathbf{a}' = (a_1, \ldots, a_{n+1}) \in A'$ and every node $i \in [n]$,

$$u'_i(\mathbf{a}) = u_i(a_1, \ldots, a_n)\,,$$

and

$$u'_{n+1}(\mathbf{a}) = \begin{cases} 1 & \text{if } a_{n+1} = 1 \\ 0 & \text{otherwise}\,. \end{cases}$$

Informally, the first $n$ nodes are apathetic about node $n+1$'s action, and node $n+1$ always prefers to play 1. Notice that $(a_1, \ldots, a_n) \in A$ is a PNE for $U$ if and only if $(a_1, \ldots, a_n, 1)$ is a PNE for $U'$. Assume that the stay-or-roll protocol is self-stabilizing on all games over $A'$.

Let $\mathbf{f} = (f_1, \ldots, f_n)$ be the reaction functions given by the stay-or-roll protocol for $U'$. Our new protocol's reaction functions $\mathbf{g} = (g_1, \ldots, g_n)$ for $U$ are defined very simply: for

every node $i \in [n]$ and $(a_1, \ldots, a_n) \in A$,

$$g_i(a_1, \ldots, a_n) = f_i(a_1, \ldots, a_n, 1)\,.$$

A PNE is an fixed point in the system $(n, A, \mathbf{g})$.

Since $U$ has a PNE, $U'$ does also, and node $n+1$'s action is 1 in every PNE for $U'$. Let $\mathbf{a} = (a_1, \ldots, a_n)$ be any state in $A$. By the self stabilization of the stay-or-roll protocol on $U'$, there is a 1-fair path in the system $(n, A', \mathbf{f})$ from $\mathbf{a}' = (a_1, \ldots, a_n, 1)$ to some PNE $\mathbf{p}' = (p_1, \ldots, p_n, 1)$ for $U'$. We now show that there is also a path from $(a_1, \ldots, a_n)$ to $\mathbf{p} = (p_1, \ldots, p_n)$, which is a PNE for $U$.

For some $T \in \mathbb{Z}_+$, let $\mathbf{a}^0, \ldots, \mathbf{a}^T$ be a 1-fair path in $(n, A', \mathbf{f})$ such that $\mathbf{a}^0 = \mathbf{a}'$ and $\mathbf{a}^T = \mathbf{p}'$. Node $n+1$ will never change its action from 1 in the stay-or-roll protocol, so its action is 1 in every state of this path. Let $\mathbf{b}^0, \ldots, \mathbf{b}^T$ be a 1-fair path in the system $(n, A, \mathbf{g})$. For $0 \leq t < T$ and $i \in [n]$, if

$$\mathbf{a}^t = (b_1^t, \ldots, b_n^t, 1)\,,$$

then $g_i(\mathbf{b}^t) = f_i(\mathbf{a}^t)$. Thus, w.p.p.,

$$\mathbf{a}^{t+1} = (b_1^{t+1}, \ldots, b_n^{t+1}, 1)\,,$$

and it follows by induction that w.p.p. $\mathbf{b}^T = \mathbf{p}$. Hence every 1-fair trajectory in $(n, A, \mathbf{g})$ converges to a PNE, so this protocol is self-stabilizing on all games over $A$. By Lemma 58, the stay-or-roll protocol is also self-stabilizing on this class of games. □

Finally, we give an example of a $2 \times 2 \times 2$ game on which the stay-or-roll protocol fails to self-stabilize.

**Lemma 63.** *The stay-or-roll protocol is not self-stabilizing on all $2 \times 2 \times 2$ games.*

*Proof.* Let $A = [2] \times [2] \times [2]$. Consider the game $U = (3, A, \mathbf{u})$ where $u_i(\alpha, \beta, \gamma)$ is the $i^{\text{th}}$ coordinate of $M_\alpha[\beta, \gamma]$, for

$$M_1 = \begin{bmatrix} 1,1,1 & 1,0,1 \\ 1,0,0 & 0,1,1 \end{bmatrix} \qquad M_2 = \begin{bmatrix} 0,1,0 & 0,1,1 \\ 0,0,0 & 1,0,1 \end{bmatrix}.$$

The unique PNE of $U$ is $\mathbf{p} = (1, 1, 1)$. Let $\mathbf{a} \in A$ with $a_3 = 2$. Then node 3 will continue playing 2 under the stay-or-roll protocol, since 2 is always a best response. It follows that under the stay-or-roll protocol, if the third node initially plays 2, then it will never play 1, so $\mathbf{p}$ will never be reached from, for example, $(1, 1, 2)$. □

Given these last three lemmas, along with Theorem 51, proving the negative part of Theorem 59 is straightforward. Informally, there are no self-stabilizing historyless uncoupled protocols for $2 \times 2 \times 2$ games or $3 \times 3$ games, and adding actions or nodes doesn't change that.

*Proof of Theorem 59.* The positive part of the theorem statement follows immediately from Lemma 60. For the negative part, let $A$ be an $n$-node state space, for some $n \geq 2$, such that some historyless uncoupled protocol is self-stabilizing on all games over a $A$. Then by Lemma 58, the stay-or-roll protocol is self-stabilizing on all games over $A$. Assume first that $n \geq 3$. Then by repeated application of Lemma 62 and Lemma 61, the stay-or-roll protocol must also be self-stabilizing over $2 \times 2 \times 2$ state spaces, which contradicts Lemma 63. If instead we assume that $n = 2$ and $|A_1|, |A_2| \geq 3$, then repeated application of Lemma 61 shows that the stay-or-roll protocol must be self-stabilizing over $3 \times 3$ state spaces, which is false by Theorem 51. □

## 8.4 Proof of Theorem 52

We now have the tools to prove our negative result for deterministic uncoupled protocols, which we restate here.

**Theorem 52.** *For every $n \geq 2$ and every $n$-node state space $A$, there is no historyless deterministic uncoupled protocol that self-stabilizes on all games over $A$.*

*Proof.* Except when $n = 2$ and either $|A_1|$ or $|A_2|$ is 2, this follows directly from Theorem 59. So let $k \geq 2$ and $A = [2] \times [k]$, and assume for contradiction there is some historyless deterministic uncoupled protocol $\Pi$ that self-stabilizes on all games over $A$.

Consider the game $U = (n, A, \mathbf{u})$, where

$$u_1(\mathbf{a}) = \begin{cases} 1 & \text{if } a_1 = 1 \\ 0 & \text{if } a_1 = 2 \end{cases}$$

$$u_2(\mathbf{a}) = \begin{cases} 1 & \text{if } a_2 = a_1 = 1 \text{ or } a_2 \geq a_1 = 2 \\ 0 & \text{otherwise}\,, \end{cases}$$

for every $\mathbf{a} = (a_1, a_2) \in A$. The unique PNE of this game is $\mathbf{p} = (1, 1)$. Let $f_1$ and $f_2$ be the reaction functions given for this game by the protocol. Since it is self-stabilizing, every 1-fair trajectory in the resulting interaction system must converge to $\mathbf{p}$.

Define a new game $U' = (n, A, \mathbf{u}')$, where

$$u'_1(\mathbf{a}) = \begin{cases} 2 & \text{if } a_2 \geq a_1 = 2 \text{ and } f_2(1, a_2) = 1 \\ u_1(\mathbf{a}) & \text{otherwise} \end{cases}$$

$$u'_2(\mathbf{a}) = u_2(\mathbf{a})\,,$$

for every $\mathbf{a} = (a_1, a_2) \in A$. Informally, each node's preferences are exactly the same as in $U$, except that node 1 now prefers to play 2 wherever $f_2$ would instruct node 2 to change its action to 1. Let $f'_1$ and $f'_2$ be the reaction functions given by $\Pi$ for this game. Notice that this game also has $\mathbf{p} = (1, 1)$ as its unique PNE, and that by uncoupledness, $f'_2(\mathbf{a}) = f_2(\mathbf{a})$ for every $\mathbf{a} \in A$.

Let $\mathbf{a} = (1, \alpha) \in A$, for some action $\alpha \neq 1$. Notice that $u'_1(\mathbf{a}) = u_1(\mathbf{a}) = 1$, and that no node in either game has multiple weakly dominant actions. Consider two cases:

1. $f_2(\mathbf{a}) = 1$. Then $u'_1(2, \alpha) = 2$, so node 1 is not $U'$-best-responding at $\mathbf{a}$. Thus, by Lemma 57, $f'_1(\mathbf{a}) \neq 1$. Since $f'_2(\mathbf{a}) = f_2(\mathbf{a}) = 1$, we then have $\mathbf{f}'(\mathbf{a}) = (2, 1)$.

2. $f_2(\mathbf{a}) \neq 1$. Then $u'_1(2, \alpha) = u_1(2, \alpha) = 0$, so node 1 is $U'$-best-responding at $\mathbf{a}$, and by Lemma 57, $f'_1(\mathbf{a}) = 1$. Since $f'_2(\mathbf{a}) = f_2(\mathbf{a}) \neq 1$, we have $\mathbf{f}'(\mathbf{a}) = (1, \beta)$ for some $\beta \neq 1$.

Now let $\mathbf{b} = (2, 1)$. Then $u'_1(\mathbf{b}) = u_1(\mathbf{b}) = 0$, and $u'_2(\mathbf{b}) = u_2(\mathbf{b}) = 0$, so neither node is best-responding. Then by Lemma 57 $f'_1(\mathbf{b}) \neq 2$ and $f'_2(\mathbf{b}) \neq 1$, i.e., $\mathbf{f}'(\mathbf{a}) = (1, \beta)$ for some

$\beta \neq 1$. It follows that $\mathbf{p} = (1, 1)$ is not reachable from $(2, 1)$ in the system $(2, A, \mathbf{f}')$, so $\Pi$ is not self-stabilizing on $U'$. $\square$

# Chapter 9

# Conclusion

By examining fractal geometry and dynamics through a computational lens, this dissertation made progress in several areas. Part I both advanced the theory of algorithmic dimension and showed a new way in which theoretical computer science can be applied to answer questions in pure mathematics that may appear unrelated to computation. Our point-to-set principles for fractal dimensions (Theorems 2 and 3) constitute a strong and useful connection between seemingly unrelated areas, allowing us to viewing classical fractal dimensions as pointwise, algorithmic information theoretic quantities. With conditional dimensions, we developed a major component of the information theoretic apparatus for studying dimension. The utility of these dimensions is immediately apparent in the correspondence between our chain rule for effective dimensions (Corollary 19) and the product inequalities for classical dimensions (Theorem 42).

We further demonstrated the value of the point-to-set principles and conditional dimension by employing them in the first two uses of algorithmic fractal dimensions to prove new theorems in classical fractal geometry. The dissimilarity between these results, in content and in proof technique, is evidence of the power and versatility of algorithmic information theoretic methods for classical dimension bounds.

The first, our lower bound on the Hausdorff dimension of generalized Furstenberg sets (Theorem 37), required nuanced technical arguments about Kolmogorov complexity. In the process, we proved a general lower bound on the dimension of points on lines in $\mathbb{R}^2$ and resolved an open problem on the structure of algorithmic information in Euclidean spaces by proving the non-existence of a line in $\mathbb{R}^2$ with singleton dimension spectrum.

The second (Theorem 38) is a fundamental bound on the Hausdorff dimension of intersections of fractals. With the algorithmic information theoretic tools of Chapters 2 and 3 in hand, the proof is simple and intuitive and does not directly invoke Kolmogorov complexity. Beyond its direct mathematical value, this theorem demonstrates that our approach can be used to strengthen the foundations of fractal geometry.

These successes motivate further investigation of algorithmic fractal geometry in general and of effective Hausdorff dimension on lines in particular; improvements on our lower bound or extensions to higher dimensions would have implications for important questions about Furstenberg or Kakeya sets. Our results also motivate a broader search for potential applications of algorithmic dimension to problems in classical fractal geometry.

In Part II, we applied distributed computing and algorithmic game theoretic perspectives to dynamical systems. Our general non-convergence result for asynchronous dynamics (Theorem 43) shows that any non-singleton invariant set rules out guaranteed convergence to a point for a large and natural class of dynamics. For uncoupled dynamics, which capture the existence of private information in multi-agent systems, we conducted a detailed investigation of the recall demands of self-stabilizing dynamics. We described the first deterministic uncoupled protocol that self-stabilizes to a Nash equilibrium using bounded recall (Theorem 53). We also gave a lower bound showing that this protocol is nearly optimal (Theorem 59), and we gave an optimal protocol for the case that the agents have sufficient expressive power (Theorem 54). Together, these results constitute significant progress toward an understanding of discrete dynamics, particularly strategic dynamics, in distributed settings.

We conclude by outlining several promising directions for future research in this area, some of which is already in progress.

**More refined algorithmic information tools for geometric measure theory**

In particular, can sets with finite or non-zero $s$-dimensional Hausdorff or packing measures be characterized using Kolmogorov complexity? More quantitative versions of Theorems 2 and 3 would allow us to study a broader range of questions in fractal geometry using algorithmic information theoretic techniques.

**Results on dimension spectra of lines in $\mathbb{R}^n$, for $n > 2$**

An extension of Theorem 24 to any higher-dimensional Euclidean space (with $a, b \in \mathbb{R}^{n-1}$) would be a major breakthrough, and more modest results in that direction should also be pursued. It is immediate from the work of J. H. Lutz and Weihrauch [68] that the dimension spectrum of any line in $\mathbb{R}^n$ has nonempty intersection with the interval $[1, n-1]$, but little else it known about these spectra.

**Complete descriptions of dimension spectra**

As we showed in Chapter 5 the dimension spectrum of any line $\mathcal{L}_{a,b} \subseteq \mathbb{R}^2$ has cardinality at least 2. It is conjectured that this spectrum must furthermore contain a unit interval. Some progress has already been made: in a recent paper with D. M. Stull [71], we show that the conjecture holds when $\dim(a, b) = \mathrm{Dim}(a, b)$. Completely describing the dimension spectra of lines and circles would represent a significant step toward understanding effective Hausdorff dimension in the plane.

**Fractal geometry on schedule spaces**

Consider, for a given asynchronous, discrete dynamical system, the class of asynchronous schedules that do not lead to convergence. For stationary and bounded-recall dynamics, this class may be characterized by some set of forbidden activation sequences, which induces a self-similar fractal structure. In general, what can the fractal geometry of this class tell us about the dynamics?

**Topological and knowledge-based approaches to asynchronous game dynamics**

Topological [13, 49, 95] and knowledge-based [42] approaches have been very successful in addressing fundamental questions in distributed computing. Can these approaches shed new light on the implications of asynchrony for strategic dynamics?

**Impossibility results for stronger dynamics and weaker asynchrony**

Finding limitations on less restricted dynamics operating in less hostile distributed environments is an important research direction. In joint work with Danny Dolev, Michael Erdmann, Michael Schapira, and Adva Zair [25], we investigate a broad class of *stateless protocols*, which include historyless dynamics but also permit nodes to send messages that are distinct from their outputs. That work gives a non-convergence result that holds even for a "fairer" variant of asynchrony.

# References


[1] Ittai Abraham, Danny Dolev, Rica Gonen, and Joe Halpern. Distributed computing meets game theory: robust mechanisms for rational secret sharing and multiparty computation. In *Proceedings of the 25th annual ACM Symposium on Principles of Distributed Computing (PODC '06)*, pages 53–62, New York, NY, 2006. ACM.

[2] Krishna B. Athreya, John M. Hitchcock, Jack H. Lutz, and Elvira Mayordomo. Effective strong dimension in algorithmic information and computational complexity. *SIAM J. Comput.*, 37(3):671–705, 2007.

[3] Yakov Babichenko. Uncoupled automata and pure Nash equilibria. *Int. J. Game Theory*, 39(3):483–502, 2010.

[4] Yakov Babichenko. Completely uncoupled dynamics and Nash equilibria. *Games and Economic Behavior*, 76(1):1–14, 2012.

[5] Yakov Babichenko, Yuval Peres, Ron Peretz, Perla Sousi, and Peter Winkler. Hunter, Cauchy rabbit, and optimal Kakeya sets. *Trans. Amer. Math. Soc.*, 366:5567–5586, 2014).

[6] Maria-Florina Balcan, Florin Constantin, and Ruta Mehta. The weighted majority algorithm does not converge in nearly zero-sum games. *ICML Workshop on Markets, Mechanisms and Multi-Agent Models*, 2012.

[7] Michael Ben-Or. Randomized agreement protocols. In *Fault-Tolerant Distributed Computing*, pages 72–83, 1986.

[8] A. S. Besicovitch. Sur deux questions d'intégrabilité des fonctions. *Journal de la Société de physique et de mathematique de l'Universite de Perm*, 2:105–123, 1919.

[9] A. S. Besicovitch. On Kakeya's problem and a similar one. *Mathematische Zeitschrift*, 27:312–320, 1928.

[10] P. Billingsley. Hausdorff dimension in probability theory II. *Illinois Journal of Mathematics*, 5(2):291–298, 1961.

[11] Christopher J. Bishop. Personal communication, April 27, 2017.

[12] Christopher J. Bishop and Yuval Peres. *Fractals in Probability and Analysis*. Cambridge University Press, 2017.

[13] Elizabeth Borowsky and Eli Gafni. Generalized FLP impossibility result for t-resilient asynchronous computations. In *Proceedings of the 25th annual ACM Symposium on Theory of Computing (STOC '93)*, pages 91–100, 1993.

[14] James E. Burns. Self-stabilizing rings without demons. Technical Report GIT-ICS-87/36, School of Information and Computer Science, Georgia Institute of Technology, November 1987.

[15] Adam Case and Jack H. Lutz. Mutual dimension. *ACM Transactions on Computation Theory*, 7(3):12, 2015.

[16] Adam Case and Jack H. Lutz. Mutual dimension and random sequences. In *Mathematical Foundations of Computer Science 2015 - 40th International Symposium, MFCS 2015, Milan, Italy, August 24-28, 2015, Proceedings, Part II*, pages 199–210, 2015.

[17] Gregory J. Chaitin. On the length of programs for computing finite binary sequences. *J. ACM*, 13(4):547–569, 1966.

[18] Gregory J. Chaitin. On the length of programs for computing finite binary sequences: statistical considerations. *J. ACM*, 16(1):145–159, 1969.

[19] Thomas R. Cover and Joy A. Thomas. *Elements of Information Theory.* Wiley, second edition, 2006.

[20] Colleen D. Cutler. Strong and weak duality principles for fractal dimension in Euclidean space. *Math. Proc. Camb. Phil. Soc.*, 118:393–410, 1995.

[21] Constantinos Daskalakis, Rafael Frongillo, Christos H. Papadimitriou, George Pierrakos, and Gregory Valiant. On learning algorithms for Nash equilibria. In Spyros Kontogiannis, Elias Koutsoupias, and Paul G. Spirakis, editors, *Algorithmic Game Theory: Third International Symposium, SAGT 2010, Athens, Greece, October 18–20, 2010. Proceedings*, pages 114–125. Springer Berlin Heidelberg, 2010.

[22] Roy O. Davies. Some remarks on the Kakeya problem. *Proc. Cambridge Phil. Soc.*, 69:417–421, 1971.

[23] Edsger W. Dijkstra. Self-stabilizing systems in spite of distributed control. *Commun. ACM*, 17(11):643–644, 1974.

[24] Danny Dolev, Cynthia Dwork, and Larry Stockmeyer. On the minimal synchronism needed for distributed consensus. *J. ACM*, 34(1):77–97, 1987.

[25] Danny Dolev, Michael Erdmann, Neil Lutz, Michael Schapira, and Adva Zair. Brief announcement: Stateless computation. In *Proceedings of the 2017 ACM Symposium on Principles of Distributed Computing (PODC 2017)*, to appear.

[26] Shlomi Dolev. *Self-Stabilization.* MIT Press, 2000.

[27] Randall Dougherty, Jack H. Lutz, R. Daniel Mauldin, and Jason Teutsch. Translating the Cantor set by a random real. *Transactions of the American Mathematical Society*, 366:3027–3041, 2014.

[28] Rod Downey and Denis Hirschfeldt. *Algorithmic Randomness and Complexity.* Springer-Verlag, 2010.

[29] Zeev Dvir. On the size of Kakeya sets in finite fields. *J. Amer. Math. Soc.*, 22:1093–1097, 2009.

[30] K. J. Falconer. *Fractal Geometry: Mathematical Foundations and Applications*. Wiley, 1990.

[31] K. J. Falconer. *Techniques in Fractal Geometry*. Wiley, 1997.

[32] Kenneth Falconer. *Fractal Geometry: Mathematical Foundations and Applications*. Wiley, third edition, 2014.

[33] Kenneth J. Falconer. *The Geometry of Fractal Sets*. Cambridge University Press, 1985.

[34] Kenneth J. Falconer. Sets with large intersection properties. *Journal of the London Mathematical Society*, 49(2):267–280, 1994.

[35] Faith Fich and Eric Ruppert. Hundreds of impossibility results for distributed computing. *Distributed Computing*, 16(2–3):121–163, 2003.

[36] Michael J. Fischer, Nancy A. Lynch, and Michael S. Paterson. Impossibility of distributed consensus with one faulty process. *J. ACM*, 32(2):374–382, 1985.

[37] O. Frostman. Potential d'equilibre et capacité des ensembles avec quelques applications à la theórie des fonctions. *Meddel. Lunds Univ. Math. Sen.*, 3:1–118, 1935.

[38] Timothy G. Griffin, F. Bruce Shepherd, and Gordon Wilfong. The stable paths problem and interdomain routing. *IEEE/ACM Trans. Netw.*, 10(2):232–243, 2002.

[39] Xiaoyang Gu and Jack H. Lutz. Dimension characterizations of complexity classes. *Computational Complexity*, 17(4):459–474, 2008.

[40] Xiaoyang Gu, Jack H. Lutz, and Elvira Mayordomo. Points on computable curves. In *FOCS*, pages 469–474, 2006.

[41] Xiaoyang Gu, Jack H. Lutz, Elvira Mayordomo, and Philippe Moser. Dimension spectra of random subfractals of self-similar fractals. *Ann. Pure Appl. Logic*, 165(11):1707–1726, 2014.

[42] Joseph Y. Halpern and Yoram Moses. Knowledge and common knowledge in a distributed environment. *J. ACM*, 37(3):549–587, July 1990.

[43] Sergiu Hart. Adaptive heuristics. *Econometrica*, 73(5):1401–1430, 2005.

[44] Sergiu Hart and Yishay Mansour. How long to equilibrium? The communication complexity of uncoupled equilibrium procedures. *Games and Economic Behavior*, 69(1):107–126, 2010.

[45] Sergiu Hart and Andreu Mas-Colell. Uncoupled dynamics do not lead to Nash equilibrium. *American Economic Review*, 93(5):1830–1836, 2003.

[46] Sergiu Hart and Andreu Mas-Colell. Stochastic uncoupled dynamics and Nash equilibrium. *Games and Economic Behavior*, 57(2):286–303, 2006.

[47] Sergiu Hart and Andreu Mas-Colell. *Simple Adaptive Strategies: From Regret-matching to Uncoupled Dynamics*. World Scientific Publishing Co., Inc., River Edge, NJ, 2012.

[48] Felix Hausdorff. Dimension und äusseres Mass. *Mathematische Annalen*, 79:157–179, 1919.

[49] Maurice Herlihy and Nir Shavit. The topological structure of asynchronous computability. *J. ACM*, 46(6):858–923, 1999.

[50] Michael Hochman. Lectures on dynamics, fractal geometry, and metric number theory. *Journal of Modern Dynamics*, 8(3–4):437–497, 2014.

[51] Aaron D. Jaggard, Neil Lutz, Michael Schapira, and Rebecca N. Wright. Self-stabilizing uncoupled dynamics. In *Proceedings of the 7th International Symposium on Algorithmic Game Theory (SAGT '14)*, pages 74–85, 2014.

[52] Aaron D. Jaggard, Neil Lutz, Michael Schapira, and Rebecca N. Wright. Dynamics at the boundary of game theory and distributed computing. *ACM Transactions on Economics and Computation*, to appear.

[53] Aaron D. Jaggard, Michael Schapira, and Rebecca N. Wright. Distributed computing with adaptive heuristics. In *Proceedings of the Second Symposium on Innovations in Computer Science (ICS 2011)*, pages 417–443, 2011.

[54] Jean-Pierre Kahane. Sur la dimension des intersections. In Jorge Alberto Barroso, editor, *Aspects of mathematics and its applications*, North-Holland Mathematical Library, 34, pages 419–430. Elsevier, 1986.

[55] Nets Hawk Katz and Terence Tao. Some connections between Falconers distance set conjecture and sets of Furstenburg type. *New York J. Math*, 7:149–187, 2001.

[56] Robert D. Kleinberg, Katrina Ligett, Georgios Piliouras, and Éva Tardos. Beyond the Nash equilibrium barrier. In *Innovations in Computer Science - ICS 2010, Tsinghua University, Beijing, China, January 7–9, 2011. Proceedings*, pages 125–140, 2011.

[57] Andrei N. Kolmogorov. On the Shannon theory of information transmission in the case of continuous signals. *IRE Transactions on Information Theory*, 2(4):102–108, 1956.

[58] Elias Koutsoupias and Christos H. Papadimitriou. Worst-case equilibria. In *STACS 99, 16th Annual Symposium on Theoretical Aspects of Computer Science, Trier, Germany, March 4–6, 1999, Proceedings*, pages 404–413, 1999.

[59] Roger Lagunoff and Akihiko Matsui. Asynchronous choice in repeated coordination games. *Econometrica*, 65(6):1467–1477, 1997.

[60] Leonid A. Levin. On the notion of a random sequence. *Soviet Math Dokl.*, 14(5):1413–1416, 1973.

[61] Leonid A. Levin. Laws of information conservation (nongrowth) and aspects of the foundation of probability theory. *Problemy Peredachi Informatsii*, 10(3):30–35, 1974.

[62] Ming Li and Paul M.B. Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications*. Springer, third edition, 2008.

[63] Jack H. Lutz. Dimension in complexity classes. *SIAM J. Comput.*, 32(5):1236–1259, 2003.

[64] Jack H. Lutz. The dimensions of individual strings and sequences. *Inf. Comput.*, 187(1):49–79, 2003.

[65] Jack H. Lutz and Neil Lutz. Lines missing every random point. *Computability*, 4(2):85–102, 2015.

[66] Jack H. Lutz and Neil Lutz. Algorithmic information, plane kakeya sets, and conditional dimension. In *Proceedings of the 34th Symposium on Theoretical Aspects of Computer Science, STACS 2017, March 8–11, 2017, Hannover, Germany*, pages 53:1–53:13, 2017.

[67] Jack H. Lutz and Elvira Mayordomo. Dimensions of points in self-similar fractals. *SIAM Journal of Computing*, 38(3):1080–1112, 2008.

[68] Jack H. Lutz and Klaus Weihrauch. Connectivity properties of dimension level sets. *Mathematical Logic Quarterly*, 54:483–491, 2008.

[69] Neil Lutz. Fractal intersections and products via algorithmic dimension, 2016.

[70] Neil Lutz and D. M. Stull. Bounding the dimension of points on a line. In TV Gopal, Gerhard Jaeger, and Silvia Steila, editors, *Theory and Applications of Models of Computation: 14th Annual Conference, TAMC 2017, Bern, Switzerland, April 20-22, 2017, Proceedings*, pages 425–439, 2017.

[71] Neil Lutz and D. M. Stull. Dimension spectra of lines. In *Computability in Europe 2017: Unveiling Dynamics and Complexity (CiE 2017)*, to appear.

[72] Nancy Lynch. A hundred impossibility proofs for distributed computing. In *Proceedings of the Eighth annual ACM Symposium on Principles of Distributed Computing (PODC '89)*, pages 1–28, 1989.

[73] Jason R Marden and Jeff S Shamma. Revisiting log-linear learning: Asynchrony, completeness and payoff-based implementation. *Games and Economic Behavior*, 75(2):788–808, 2012.

[74] John M. Marstrand. Some fundamental geometrical properties of plane sets of fractional dimensions. *Proceedings of the London Mathematical Society*, 4(3):257–302, 1954.

[75] Per Martin-Löf. The definition of random sequences. *Information and Control*, 9(6):602–619, 1966.

[76] Pertti Mattila. Hausdorff dimension and capacities of intersections of sets in $n$-space. *Acta Mathematica*, 152:77–105, 1984.

[77] Pertti Mattila. On the Hausdorff dimension and capacities of intersections. *Mathematika*, 32:213–217, 1985.

[78] Pertti Mattila. *Geometry of sets and measures in Euclidean spaces: fractals and rectifiability*. Cambridge University Press, 1995.

[79] Elvira Mayordomo. A Kolmogorov complexity characterization of constructive Hausdorff dimension. *Inf. Process. Lett.*, 84(1):1–3, 2002.

[80] Joseph S. Miller. "Open Questions" section of personal webpage. Retrieved October 30, 2016. https://web.archive.org/web/20161021183028/http://www.math.wisc.edu/~jmiller/open.html.

[81] Ursula Molter and Ezequiel Rela. Furstenberg sets for a fractal set of directions. *Proc. Amer. Math. Soc.*, 140:2753–2765, 2012.

[82] Dov Monderer and Lloyd S. Shapley. Potential games. *Games and Economic Behavior*, 14:124–143, 1996.

[83] Andre Nies. *Computability and Randomness*. Oxford University Press, Inc., New York, NY, USA, 2009.

[84] Noam Nisan and Amir Ronen. Algorithmic mechanism design. *Games and Economic Behavior*, 35(1–2):166–196, 2001.

[85] Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V. Vazirani, editors. *Algorithmic Game Theory*. Cambridge University Press, New York, NY, 2007.

[86] Tuomas Orponen. An improved bound on the packing dimension of Furstenberg sets in the plane, 2016.

[87] Christos H. Papadimitriou. On the complexity of the parity argument and other inefficient proofs of existence. *J. Comput. Syst. Sci.*, 48(3):498–532, 1994.

[88] Christos H. Papadimitriou and Georgios Piliouras. From Nash equilibria to chain recurrent sets: Solution concepts and topology. In *Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science, Cambridge, MA, USA, January 14–16, 2016*, pages 227–235, 2016.

[89] Yakov Pesin. *Dimension Theory in Dynamical Systems: Contemporary Views and Applications*. University of Chicago Press, 1998.

[90] Bary S. R. Pradelski and H. Peyton Young. Learning efficient Nash equilibria in distributed systems. *Games and Economic Behavior*, 75(2):882–897, 2012.

[91] Jan Reimann. Effective multifractal spectra. http://www.personal.psu.edu/jsr25/Lectures/CFTM_2014_talk_Reimann.pdf. Lecture slides.

[92] Jan Reimann. *Computability and fractal dimension*. PhD thesis, Heidelberg University, 2004.

[93] Ezequiel Rela. Refined size estimates for Furstenberg sets via Hausdorff measures: a survey of some recent results. In M. Cepedello Boiso, H. Hedenmalm, M.A. Kaashoek, A. Montes Rodrguez, and S. Treil, editors, *Concrete Operators, Spectral Theory, Operators in Harmonic Analysis and Approximation*, pages 421–454. Springer, 2014.

[94] Robert W. Rosenthal. A class of games possessing pure-strategy Nash equilibria. *Int. J. Game Theory*, 2:65–67, 1973. 10.1007/BF01737559.

[95] Michael E. Saks and Fotios Zaharoglou. Wait-free k-set agreement is impossible: The topology of public knowledge. *SIAM J. Comput.*, 29(5):1449–1483, 2000.

[96] Claude E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3–4):379–423, 623–656, 1948.

[97] Alexander Shen and Nikolai K. Vereshchagin. Logical operations and Kolmogorov complexity. *Theoretical Computer Science*, 271(1–2):125–129, 2002.

[98] Pablo Shmerkin. On Furstenberg's intersection conjecture, self-similar measures, and the $l^q$ norms of convolutions, 2016.

[99] Ray J. Solomonoff. A formal theory of inductive inference. *Information and Control*, 7(1-2):1–22, 224–254, 1964.

[100] Elias M. Stein and Rami Shakarchi. *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*. Princeton Lectures in Analysis. Princeton University Press, 2005.

[101] Dennis Sullivan. Entropy, Hausdorff measures old and new, and limit sets of geometrically finite Kleinian groups. *Acta Mathematica*, 153(1):259–277, 1984.

[102] Terence Tao. From rotating needles to stability of waves: emerging connections between combinatorics, analysis, and PDE. *Notices Amer. Math. Soc*, 48:294–303, 2000.

[103] Claude Tricot. Two definitions of fractional dimension. *Mathematical Proceedings of the Cambridge Philosophical Society*, 91(1):57–74, 1982.

[104] Daniel Turetsky. Connectedness properties of dimension level sets. *Theoretical Computer Science*, 412(29):3598–3603, 2011.

[105] Alan M. Turing. On computable numbers, with an application to the Entscheidungsproblem. A correction. *Proceedings of the London Mathematical Society*, 43(2):544–546, 1937.

[106] Klaus Weihrauch. *Computable Analysis: An Introduction*. Springer, 2000.

[107] T. Wolff. Recent work connected with the Kakeya problem. *Prospects in Mathematics*, pages 129–162, 1999.

[108] Meng Wu. A proof of Furstenberg's conjecture on the intersections of $\times p$ and $\times q$-invariant sets, 2016.

[109] Kiho Yoon. The effective minimax value of asynchronously repeated games. *Int. J. Game Theory*, 32(4):431–442, 2004.

[110] H. Peyton Young. Learning by trial and error. *Games and Economic Behavior*, 65(2):626–643, 2009.

[111] Lai-Sang Young. Dimension, entropy and Lyapunov exponents. *Ergodic Theory and Dynamical Systems*, 2:109–124, 1982.

[112] Andriy Zapechelnyuk. Better-reply dynamics with bounded recall. *Math. Oper. Res.*, 33(4):869–879, 2008.