Cartesian Products Avoiding Patterns

by

Jacob Denson

BSc. Computing Science, University of Alberta, 2017

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Science

in

THE FACULTY OF SCIENCE
(Mathematics)

The University of British Columbia
(Vancouver)

April 2019

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

Cartesian Products Avoiding Patterns

submitted by **Jacob Denson** in partial fulfillment of the requirements for the degree of **Master of Science** in **Mathematics**.

Examining Committee:

Malabika Pramanik, Mathematics
Supervisor

Joshua Zahl, Mathematics
Supervisor

Abstract

The pattern avoidance problem seeks to construct a set $X \subset \mathbf{R}^d$ with large dimension that avoids a prescribed pattern such as three term arithmetic progressions, or more general patterns such as avoiding points $x_1, \dots, x_n$ such that $f(x_1, \dots, x_n) = 0$ (three term arithmetic progressions are specified by the pattern $x_1 - 2x_2 + x_3 = 0$). Previous work on the subject has considered patterns described by polynomials, or by functions f satisfying certain regularity conditions. We consider the case of 'rough patterns. There are several problems that fit into the framework of rough pattern avoidance. As a first application, if $Y \subset [0,1]$ is a set with Minkowski dimension α, we construct a set $X \subset [0,1]$ with Hausdorff dimension $1-\alpha$ so that $X+X$ is disjoint from Y. As a second application, given a set Y of dimension close to one, we can construct a subset $X \subset Y$ of dimension $1/2$ that avoids isosceles triangles.

Lay Summary

The lay or public summary explains the key goals and contributions of the research/scholarly work in terms that can be understood by the general public. It must not exceed 150 words in length.

Preface

At University of British Columbia (UBC), a preface may be required. Be sure to check the Graduate and Postdoctoral Studies (GPS) guidelines as they may have specific content to be included.

Contents

List of Figures

Glossary

This glossary uses the handy `acroynym` package to automatically maintain the glossary. It uses the package's `printonlyused` option to include only those acronyms explicitly referenced in the LaTeX source.

GPS Graduate and Postdoctoral Studies

Acknowledgments

Thank those people who helped you.

Don't forget your parents or loved ones.

You may wish to acknowledge your funding sources.

Chapter 1

Introduction

In this thesis, we study a simple question from the perspective of geometric measure theory:

How large can Euclidean sets be not containing geometric patterns?.

The patterns manifest as certain configurations of point tuples, often affine invariant. For instance, we might like to find large subsets X of $\mathbf{R}^n$ that contain no three collinear points, or do not contain all of the vertices of any isoceles triangle. Aside from purely geometric interest, these problems provide useful scenarios in which to test the methods of ergodic theory and harmonic analysis.

Geometric measure theory becomes important because sets which avoid the patterns we consider are highly irregular. Any open subset of $\mathbf{R}^n$ must contain a copy of the vertices of any suitably small isoceles triangle, and the Lebesgue density theorem shows that this remains true if we consider any subset of $\mathbf{R}^n$ of positive measure. Any set avoiding any of the patterns we study in this thesis must necessarily have Lebesgue measure zero. This means we cannot use the Lebesgue measure as a measure of the 'maximal size' of a set avoiding patterns. A 'second-order' measure of a set's size has to be introduced, and this is satisfied by the fractional dimension of a set, first introduced by Felix Hausdorff in 1918. We rephrase our problem as determining the maximum Hausdorff dimension a set can have avoiding patterns.

At the time of this writing, many fundamental questions about the relation between the geometric arrangement of a set and it's fractional dimension remain open. It might be expected that sets with sufficiently large Hausdorff dimension contain patterns, but this is not always true. For example, Theorem 6.8 of [2] shows that any set $X \subset \mathbf{R}^d$ with Hausdorff dimension exceeding one must contain three colinear points. On the other hand, [1] constructs a set $X \subset \mathbf{R}^d$ with full Hausdorff dimension such that no four points in X form the vertices of a parallelogram. On the other hand, Theorem 6.8 of [2] shows that any set $X \subset \mathbf{R}^d$ with Hausdorff dimension exceeding one must contain three colinear points. For most geometric configurations, it remains unknown at what threshold patterns are guaranteed, or whether such a threshold exists at all. So finding sets with large Hausdorff dimension avoiding patterns is an important topic to study in the study of contemporary geometric measure theory.

Thus we derive new methods for constructing sets avoiding patterns. Furthermore, rather than studying particular instances of the configuration avoidance problem, we choose to study general methods for finding large subsets of space avoiding any pattern. In particular, we expand on a number of general *pattern dissection methods* which have proven useful in the area, originally developed by Keleti but also studied notably by Mathé, and Pramanik/Fraser.

Our novel contribution to the construction is that in an arbitrary configuration problem, random choices can overcome the lack of presence of structure. We use this to expand the utility of interval dissection methods to a set of *fractal avoidance problems*, that previous methods were completely unavailable to address. Such problems include finding large X such that the angles formed by any three distinct points of X avoid a specified set Y of angles, where Y has a fixed Hausdorff dimension. For instance, Y could be the set of all angles expressable as $q+x$ radians, where q is rational, and x lies in the Cantor set.

Chapter 2

Background

This thesis discusses methods to form large sets avoiding patterns. First, we give a precise definition by what we mean by a pattern, and what it means to avoid a pattern. We consider an ambient set $\mathbf{A}$, often discrete, or equipped with some topological or metric structure. It's *n-point configuration space* is the set of distinct tuples of n points in $\mathbf{A}$, i.e.

$$\mathscr{C}^n(\mathbf{A}) = \{(x_1,\dots,x_n) \in X^n : x_i \neq x_j \text{ if } i \neq j\}.$$

The *configuration space* of $\mathbf{A}$ is $\mathscr{C}(\mathbf{A}) = \bigcup_{n=1}^\infty \mathscr{C}^n(\mathbf{A})$. A *pattern*, or *configuration*, on $\mathbf{A}$ is a subset of $\mathscr{C}(X)$, and we say a subset Y of X avoids a configuration $\mathscr{C}$ if $\mathscr{C}(Y)$ is disjoint from $\mathscr{C}$. An *n point configuration* on X will be a configuration $\mathscr{C}$ which is a subset of $\mathscr{C}^n(X)$.

Example (Isoceles Triangle Configuration)**.** *Consider the problem of finding a set avoiding the vertices of an isoceles triangle in the plane. Set*

$$\mathscr{C} = \{(x_1,x_2,x_3) \in \mathscr{C}^3(\mathbf{R}^2) : |x_1 - x_2| = |x_1 - x_3|\}.$$

Then $\mathscr{C}$ is a 3-point configuration, and a set $X \subset \mathbf{R}^2$ avoids $\mathscr{C}$ if and only if it contains no vertices of an isoceles triangle. Notice that $|x_1 - x_2| = |x_1 - x_3|$ holds if and only if $|x_1 - x_2|^2 = |x_1 - x_3|^2$, which is an algebraic equation in the coordinates of x_1, x_2, and x_3. Thus $\mathscr{C}$ is an algebraic hypersurface in $\mathbf{R}^6$.

Example (General Position Configuration)**.** *Suppose we wish to find a subset of $\mathbf{R}^d$ such that every collection of $k+1$ points in the set, for $k < d$, lies in 'general* *position', i.e. they do not lie in a k dimensional hyperplane. Set*

$$\mathscr{C}^k = \{(x_0, x_1, \dots, x_k) \in \mathscr{C}^{k+1}(\mathbf{R}^d) : x_1 - x_0, \dots, x_k - x_0 \text{ are linearly dependant}\}.$$

and then consider the configuration $\mathscr{C} = \bigcup_{k=1}^{d-1} \mathscr{C}^k$. *A set X avoids* $\mathscr{C}$ *if and only if all of it's points lie in general position. Notice that*

$$\mathscr{C}^k = \bigcup \left\{ span(y_1, \dots, y_k) \times \{y\} : y = (y_1, \dots, y_k) \in \mathscr{C}^k(\mathbf{R}^d) \right\} \cap \mathscr{C}^{k+1}(\mathbf{R}^d).$$

so each $\mathscr{C}^k$ *is essentially a union of k dimensional hyperplanes.*

Even though our problem formulation assumes configurations are formed by distinct sets of points, one can still formulate avoidance problems involving repeated points in our framework, because an instance of a configuration involving n points which may contain repetitions can be seen as an instance of a configuration involving fewer than n distinct points.

Example (Sum Set Configuration)**.** *Let G be an abelian group, and fix* $Y \subset G$. *Set*

$$\mathscr{C}^1 = \{g \in \mathscr{C}^1(G) : 2g \in Y\} \quad \text{and} \quad \mathscr{C}^2 = \{(g_1, g_2) \in \mathscr{C}^2(G) : g_1 + g_2 \in Y\}.$$

Then set $\mathscr{C} = \mathscr{C}^1 \cup \mathscr{C}^2$. *A set* $X \subset G$ *avoids* $\mathscr{C}$ *if and only if* $(X + X) \cap Y = \emptyset$.

Our main focus in this thesis is on the *pattern avoidance problem*: Given a configuration $\mathscr{C}$ on $\mathbf{A}$, how large can $X \subset \mathbf{A}$ be avoiding $\mathscr{C}$. If $\mathbf{A}$ is discrete, i.e. finite, or a discrete limit of finite sets $\mathbf{A}_n$, the goal is to find X with large cardinality, or such that $X \cap \mathbf{A}_n$ has large cardinality asymptotically in n. If $\mathbf{A} = \mathbf{R}^d$, but $\mathscr{C}$ is a sufficiently discrete configuration, then a satisfactory goal is to find X with large Lebesgue measure avoiding $\mathscr{C}$. But in this thesis we establish methods for avoiding non-discrete configurations $\mathscr{C}$, i.e. those for which $X^m \cap \mathscr{C}^m$ are dense in X^m, but taking inspiration from methods in the discrete setting. The next section shows that Lebesgue measure completely fails to measure the degree of success for a solution to the pattern avoidance problems, but provides an alternate measurement which does succeed to establish this point.

2.1 Fractal Dimension

The Lebesgue measure is not the correct measurement for how large a pattern-avoiding set for non-discrete patterns. This is because for most of these patterns, every pattern-avoiding set has measure zero. One intuition as to why this is true is that a set with positive Lebesgue measure behaves in many respects like an open set, and an open set certainly intersects a dense set somewhere. The rigorous instance of this phenomenon we use is the Lebesgue density theorem. It's proof takes us too far afield into differentiation theory, so we merely state the result without proof. The intuitive idea of the result is that a set of positive Lebesgue measure locally contains a large percentage of it's surrounding points.

Theorem 1 (Lebesgue Density Theorem)**.** *Let $E \subset \mathbf{R}^d$ have positive Lebesgue measure. Then for almost every point $x \in E$,*

$$\lim_{r \to 0} \frac{|E \cap B_r(x)|}{|B_r(x)|} = 1.$$

Under mild non-discreteness conditions, which are certainly satisfied by the example configurations given in the last section, no set with positive Lebesgue measure can avoid a configuration.

Theorem 2. *Let $\mathscr{C}$ be an n-point configuration on $\mathbf{R}^d$ such that*

1. Translation Invariance*: For any $b \in \mathbf{R}^d$, $\mathscr{C} + b \subset \mathscr{C}$.*

2. Non-Discrete*: For any $\varepsilon > 0$, there is an instance of the configuration $(c_1, \dots, c_n) \in C$ such that diam$\{c_1, \dots, c_n\} \leq \varepsilon$.*

Then no set with positive Lebesgue measure avoids $\mathscr{C}$.

Proof. Let $X \subset \mathbf{R}^d$ have positive Lebesgue measure. Applying the Lebesgue density theorem, we find a point $x_0 \in X$ such that

$$\lim_{r \to 0} \frac{|X \cap B_r(x_0)|}{|B_r(x_0)|} = 1$$

We fix r_0 such that $|X \cap B_{r_0}(x_0)| \geq (1 - \varepsilon)|B_{r_0}(x)|$. Let $C = (c_1, \dots, c_n) \in \mathscr{C}$ be an instance of the configuration such that $\text{diam}\{c_1, \dots, c_n\} \leq \varepsilon r_0$. For each $p \in B_{r_0}(x_0)$, let $C(p) = (c_1(p), \dots, c_n(p)) \in \mathscr{C}$, where $c_i(p) = p + (c_i - c_1)$. Then a union bound gives

$$\begin{aligned}\left|\bigcup_{i=1}^n \{p \in B_{r_0/2}(x_0) : c_i(p) \not\in X\}\right| &\leq \sum_{i=1}^n |B_{r_0/2}(x_0) \cap (X + (c_1 - c_i))^c|\\ &= \sum_{i=1}^n |B_{r_0/2}(x_0)| - |B_{r_0/2}(x_0 + c_i - c_1) \cap X|\\ &\leq n\left(|B_{r_0/2}(x_0)| - |B_{r_0(1/2-\varepsilon)}(x_0) \cap X|\right)\\ &\leq n\left[(r_0/2)^d - (1-\varepsilon)(1/2-\varepsilon)^d r_0^d\right]\\ &\leq n\left(1 - (1-\varepsilon)(1-2d\varepsilon)\right)(r_0/2)^d\\ &\leq \varepsilon n(2d+1)(r_0/2)^d\end{aligned}$$

If $\varepsilon < 1/n(2d+1)$, we conclude that

$$\begin{aligned}\left|\bigcap_{i=1}^n \{p \in B_{r_0/2}(x_0) : c_i(p) \in X\}\right| &= |B_{r_0/2}(x_0)| - \left|\bigcap_{i=1}^n \{p \in B_{r_0/2}(x_0) : c_i(p) \not\in X\}\right|\\ &\geq (r_0/2)^d - \varepsilon n(2d+1)(r_0/2)^d > 0\end{aligned}$$

Thus there exists p such that $C(p) \in \mathscr{C}(X)$, and so X does not avoid $\mathscr{C}$. □

Fortunately, we have a second order notion of size, which is able to distinguish between sets of measure zero, known as *fractional dimension*. Intuitively, the fractional dimension provides a measure of the local density of a set, and so we view a set which is more dense as ‘larger’, for the purposes of a pattern avoidance problem. There are two definitions of fractional dimension we use in this thesis: Minkowski dimension and Hausdorff dimension. The main difference between the two is that Minkowski dimension measures relative density at a single scale, whereas Hausdorff dimension measures relative density at various small scales.

We begin by discussing the Minkowski dimension, which is the easiest of the two dimension to define. If E is a bounded set in $\mathbf{R}^d$, then we can consider the *delta thickening* $E_\delta = \{x : d(x,E) < \delta\}$. We define the *upper* and *lower* Minkowski dimension as

$$\overline{\dim}_M(E) = \limsup_{\delta \to 0} \left(d - \frac{\log |E_\delta|}{\log \delta} \right) \qquad \underline{\dim}_M(E) = \liminf_{\delta \to 0} \left(d - \frac{\log |E_\delta|}{\log \delta} \right)$$

If $\overline{\dim}_M(E) = \underline{\dim}_M(E)$, then we refer to this common quantity as the *Minkowski* *dimension* of E, denote $\dim_M(E)$. This means that as $\delta \to 0$, if $\dim_M(E) = \alpha$, then $|E_\delta| = \delta^{n-\alpha+o(1)}$. In particular, every set with positive Lebesgue measure has Minkowski dimension d, so Minkowski dimension is really only interesting for sets with Lebesgue measure zero. We can extend the lower and upper Minkowski dimension to unbounded sets E by considering the supremum and infinum of the dimensions of all bounded subsets of E.

Example. *Let*

$$C = \left\{ \sum_{i=1}^{\infty} a_i/4^i : a_i \in \{0,3\} \right\}$$

If $1/4^{N+1} \leq \delta \leq 1/4^N$, *then*

$$\left\{ \sum_{i=1}^{\infty} a_i/4^i : a_1, \dots, a_{N+1} \in \{0,3\} \right\} \subset C_\delta \subset \left\{ \sum_{i=1}^{\infty} a_i/4^i : a_1, \dots, a_N \in \{0,3\} \right\}$$

The former set has volume $2^{N+1}/4^{N+1} = 1/2^{N+1} \geq \delta^{1/2}$, *whereas the latter set has volume* $2^N/4^N = 1/2^N \leq (2\delta)^{1/2}$. *Thus* $\log|C_\delta| = \log\delta/2 + O(1)$, *and so*

$$\dim_{\mathbf{M}}(C) = \lim_{\delta \to 0} \left(1 - \frac{\log|C_\delta|}{\log\delta} \right) = 1 - 1/2 = 1/2$$

So C has Minkowski dimension $1/2$*.*

There are a few things to notice about this calculation. First, we performed an upper and lower bound on powers of $1/4^k$, and then used this to obtain bounds at all scales. And indeed, if we fix M, and consider $1/M^{k+1} \leq \delta \leq 1/M^k$, then as $k \to \infty$, $\log\delta \sim \log(1/M^k)$, and combined with the fact that $C_{1/M^{k+1}} \subset C_\delta \subset C_{1/M^k}$, we find

$$\frac{|C_{1/M_k}|}{\log(1/M_k)} \sim \frac{|C_{1/M^k}|}{\log(\delta)} \leq \frac{|C_\delta|}{\log(\delta)} \leq \frac{|C_{1/M^{k+1}}|}{\log(\delta)} \sim \frac{|C_{1/M^{k+1}}|}{\log(1/M_{k+1})}$$

Thus for any set E and any integer M,

$$\underline{\dim}_{\mathbf{M}}(E) = \liminf_{n \to \infty} \frac{|C_{1/M^k}|}{\log(1/M^k)} \quad \text{and} \quad \overline{\dim}_{\mathbf{M}}(E) = \limsup_{n \to \infty} \frac{|C_{1/M^k}|}{\log(1/M^k)}$$

The second point is that we understood the dimension of C via a ‘Cantor-type’ decomposition. Indeed, the set C can be understood as the limit of a family of sets which are simple unions of intervals. If we set

$$C_k = \left\{ \sum_{i=1}^{\infty} a_i/4^i : a_1, \dots, a_{N+1} \in \{0,3\} \right\}.$$

Then $\{C_k\}$ is a nested family of sets, with $C = \lim C_k$. If for an index k, we set

$$\mathscr{B}_l^d = \{[a_1, a_1+l] \times \cdots \times [a_d, a_d+l] : a_k \in l \cdot \mathbf{Z}\}.$$

then C_k is a union of 2^k cubes in $\mathscr{B}_{4^k}^d$, and

$$\frac{\log(2^k)}{\log(4^k)} = \log_4(2) = 1/2.$$

One can also calculate the Minkowski dimension by counting the number of cubes in $\mathscr{B}_l^d$ intersecting the set. For a set E, set $\mathscr{B}_l^d(E)$ to be the set of all cubes in $\mathscr{B}_l^d$ intersecting E, i.e. $\mathscr{B}_l^d(E) = \{I \in \mathscr{B}_l^d : I \cap E \neq \emptyset\}$.

Lemma 1. *If E is a bounded set in $\mathbf{R}^d$ and M is an integer, then*

$$\underline{\dim}_{\mathbf{M}}(E) = \liminf_{l \to 0} \frac{\# \mathscr{B}_l^d(E)}{\log(1/l)} \quad \textit{and} \quad \overline{\dim}_{\mathbf{M}}(E) = \limsup_{l \to 0} \frac{\# \mathscr{B}_l^d(E)}{\log(1/l)}$$

where the limit is taken over lengths $l = 1/M^k$.

Proof. Let $l = 1/M^k$. For each cube $I \in \mathscr{B}_l^d$, the l thickening I_l is contained in 3^d cubes in $\mathscr{B}_l^d$. Conversely, if $I \in \mathscr{B}_l^d(E)$, then $I \subset E_{d^{1/2}l}$, so $I \subset E_{M^{k_0}l}$ for $M^{k_0} \geq d^{1/2}$. Thus

$$|E_l| \leq 3^d \# \mathscr{B}_l^d(E) l^d \quad \text{and} \quad |E_{M^{k_0}l}| \geq \# \mathscr{B}_l^d(E) l^d$$

So as $l \to 0$,

$$d - \frac{\log |E_l|}{\log l} = \frac{\log(|E_l| l^{-d})}{\log(1/l)} \leq \frac{\log(3^d \# \mathscr{B}_l^d(E))}{\log(1/l)} = \frac{\# \mathscr{B}_l^d(E)}{\log(1/l)} + o_d(1)$$

and

$$d - \frac{\log |E_{M^{k_0}l}|}{\log(M^{k_0}l)} = \frac{\log(|E_{M^{k_0}l}|(d^{1/2}l)^{-d})}{\log(1/l)} \geq o_d(1) + \frac{\# \mathscr{B}_l^d(E)}{\log(1/l)}$$

Taking limits completes the proof. $\square$

We will often use 'Cantor-type' constructions to form pattern avoiding sets. And Lemma 1 will be crucial either for calculating the Minkowski dimension of these constructions. It is especially useful when trying to work at discrete scales, because we can view $\bigcup \mathscr{B}^d_l(E)$ as a 'discretization' of a set E at the scale l.

Example. *We will often consider sets whose dimension behaves differently at various scales. This often occurs when performing multi-scale constructions where the scales in the construction decay inverse superexponentially. A toy example of this phenomenon can be obtained by modifying the last example slightly so that the construction of the Cantor set behaves differently at various scales. We fix an increasing sequence of integers $\{N_k\}$, with $N_0 = 0$, and consider*

$$C = \left\{ \sum_{i=1}^{\infty} a_i/4^i : a_i \in \{0,3\} \textit{ if there is } k \geq 0 \textit{ such that } N_{2k} \leq i \leq N_{2k+1} \right\}$$

Then $C = \lim C_n$, where

$$C_n = \left\{ \sum_{i=1}^{\infty} a_i/4^i : a_i \in \{0,3\} \textit{ if there is } k \leq n \textit{ such that } N_{2k} \leq i \leq N_{2k+1} \right\}$$

Notice that C_n is the union of

$$2^{N_1-N_0}4^{N_2-N_1}2^{N_3-N_2}\dots 4^{N_{2n}-N_{2n-1}} = 2^{-N_1+N_2-N_3+\dots-N_{2n-1}+2N_{2n}} \geq 2^{2N_{2n}-N_{2n-1}}$$

sidelength l_n cubes, where $l_n = 1/4^{N_{2n}}$. Each of these cubes intersects C, so

$$\frac{\log \# \mathscr{B}^1_{l_n}(C)}{\log(1/l_n)} \geq \frac{2N_{2n} - N_{2n-1}}{2N_{2n}} = 1 - \frac{N_{2n-1}}{2N_{2n}}$$

Provided that $N_{2n-1}/N_{2n} = o(1)$, which occurs if the values N_k increase superexponentially, i.e. if $N_k = 2^{k^2}$, we conclude that $\overline{\dim}_{\mathbf{M}}(C) = 1$. On the other hand, C_n is also the union of

$$2^{N_1-N_0}4^{N_2-N_1}2^{N_3-N_2}\dots 4^{N_{2n}-N_{2n-1}}2^{N_{2n+1}-N_{2n}} = 2^{-N_1+N_2-\dots+N_{2n}+N_{2n+1}} \leq 2^{N_{2n}+N_{2n+1}}.$$

sidelength r_n cubes, where $r_n = 1/4^{N_{2n+1}}$. Thus

$$\frac{\log \# \mathscr{B}^1_{l_n}(C)}{\log(1/l_n)} \leq \frac{N_{2n} + N_{2n+1}}{2N_{2n+1}} = 1/2 + \frac{N_{2n}}{2N_{2n+1}}$$

Again, if $N_{2n}/N_{2n+1} = o(1)$, then $\underline{\dim}_{\mathbf{M}}(C) \leq 1/2$. One can fairly easily check that the values r_n are the 'worst case' scales, so that $\underline{\dim}_{\mathbf{M}}(C) = 1/2$. Thus the set C 'looks' half dimension between l_n and r_n, for each n, but 'looks' full dimensional between the scales r_n and l_{n+1}.

Hausdorff dimension is a more stable version of fractal dimension which is obtained by finding a canonical 's dimensional measure' H^s on $\mathbf{R}^d$ for $s \in [0,\infty)$, and then setting the dimension of E to be the supremum of the values s such that $H^s(E) < \infty$. A naive way the Hausdorff measure to construct is to assign a mass r^s to each radius r ball in $\mathbf{R}^n$, and then define

$$H^{s,\infty}(E) = \inf\left\{\sum r_k^s : E \subset \bigcup B(x_k, r_k)\right\}$$

This is an outer measure, and so Caratheodory's extension theorem gives a σ algebra of measurable sets. Unfortunately, most sets are not measurable with respect to this σ algebra. For instance, take $s = 1/2$, and $E = (a,b)$. On one hand, $H^{s,\infty}(E) \leq [(b-a)/2]^{1/2}$. On the other hand, if (a,b) is covered by balls $B(x_k, r_k)$, then $\sum 2r_k \geq b-a$, so applying the concavity of $x \mapsto x^{1/2}$, we conclude

$$\sum r_k^{1/2} \geq \left(\sum r_k\right)^{1/2} \geq \left(\frac{b-a}{2}\right)^{1/2}$$

Thus $H^{s,\infty}(E) = [(b-a)/2]^{1/2}$. But now we see that the additivity property begins to breakdown, since $H^{1/2,\infty}[0,1] = 2^{-1/2}$, whereas $H^{1/2,\infty}[0,1/2] = H^{1/2,\infty}[1/2,1] = 1/2$, and so $H^{1/2,\infty}[0,1] < H^{1/2,\infty}[0,1/2] + H^{1/2,\infty}[1/2,1]$. The reason for this is that $[0,1]$ is most efficiently covered by one large ball, rather than covering $[0,1/2]$ and $[1/2,1]$ separately. This is fixed by limiting the Hausdorff measure to be the value of the most efficient cover by arbitrarily small balls.

For a subset E of Euclidean space, we define

$$H^s_\delta(E) = \inf\left\{\sum_{n=1}^\infty \operatorname{diam}(B_n)^s : E \subset \bigcup_{n=1}^\infty B_n, \operatorname{diam}(B_n) \leq \delta\right\}$$

We then define $H^s(E) = \lim_{\delta \to 0} H^s_\delta(E)$. Then H^s is an exterior measure, and $H^s(E \cup F) = H^s(E) + H^s(F)$ if $d(E,F) > 0$. Thus all Borel sets are measurable with respect to H^s, which is certainly more satisfactory than the last definition.

Example. *Let $s = 0$. Then $H^0_\delta(E) = N^{Ext}_\delta(E)$, which tends to ∞ as $\delta \to 0$ unless E is finite, and then $H^0_\delta(E) \to \#E$. Thus H^0 is just the counting measure.*

Example. *Let $s = n$. If E has Lebesgue measure zero, then for any $\varepsilon > 0$, there exists countable many balls $B(x_k, r_k)$ covering E with $\sum r_k^n < \varepsilon$. Then $r_k < \varepsilon^{1/n}$, so $H^n_{\varepsilon^{1/n}}(E) < \varepsilon$. Letting $\varepsilon \to 0$, we conclude $H^n(E) = 0$. Thus H^n is absolutely continuous with respect to the Lebesgue measure. The measure H^n is translation invariant, so H^n is actually a constant multiple of the Lebesgue measure. We let the constant multiple be defined $1/\omega_n$. The value ω_n can be defined as the volume of a unit ball in $\mathbf{R}^n$, since $H^n(B) = 1$ if B is a unit ball.*

The same argument shows that if V is an m dimensional subspace of $\mathbf{R}^n$, then H^m, restricted to subsets of V, is a constant multiple of the m dimensional Lebesgue measure on V. More generally, H^m measures the m dimensional surface area of smooth, m dimensional submanifolds of $\mathbf{R}^n$.

Theorem 3. *Let U be an open subset of $\mathbf{R}^d$, and let $\phi : U \to \mathbf{R}^n$ be a smooth immersion. Then for any compact set E,*

$$H^d(\phi(E)) \propto \frac{1}{\omega_d} \int_E J(x)\, dx$$

where $J(x)$ is the square root of the sums of squares of the $d \times d$ minors of $D\phi(x)$.

Proof. We may cover E by finitely many open sets $U_1, \dots, U_N$, together with coordinate charts $y_1, \dots, y_N$ such that $(y_k \circ \phi)(x) = (x, f_k(x))$ for some smooth f_k, and fix J_k such that for any $x \in U_k$, $|J(x) - J_k| < \varepsilon$. TODO: PROVE REST OF THEOREM. □

Lemma 2. *If $t < s$ and $H^t(E) < \infty$, $H^s(E) = 0$, and if $H^s(E) = \infty$, $H^t(E) = \infty$.*

Proof. If, for any cover of E by balls $B(x_k, r_k)$, $\sum r_k^t \leq A$, and $r_k \leq \delta$, then $\sum r_k^s \leq \sum r_k^{s-t} r_k^t \leq \delta^{s-t} A$. Thus $H^s_\delta(E) \leq \delta^{s-t} A$, and taking $\delta \to 0$, we conclude $H^s(E) = 0$. The latter point is just proved by taking contrapositives. □

Thus given any Borel set E, there is s such that $H^{s_0}(E) = 0$ for $s_0 < s$, and $H^{s_1}(E) = \infty$ for $s_1 > s$. We refer to s as the Hausdorff dimension of E, denoted $\dim_H(E)$.

Example. *Consider $S = \{(x, \sin(1/x)) : 0 < x \leq 1\}$. Then for each $\delta > 0$, the set $S \cap [\delta, 1] \times \mathbf{R}$ is a smooth curve, and therefore has Hausdorff dimension 1. Thus for any $\varepsilon > 0$, $H^{1+\varepsilon}(S \cap [\delta, 1] \times \mathbf{R}) = 0$. But then taking limits as $\delta \to 0$, we conclude $H^{1+\varepsilon}(S) = 0$. Since $H^1(S) > 0$, this shows S has Hausdorff dimension 1. Compare this to the Minkowski dimension $3/2$ result we obtained previously.*

An easy way to compare the approaches to fractal dimension given by Minkowski and Hausdorff dimension is that Minkowski dimension measures the efficiency of covers of a set at a fixed scale, whereas Hausdorff dimension measures the efficiency of covers of a set at various, small scales.

2.2 Branching Processes

2.3 Rusza: Difference Sets Without Squares

In this section, we describe the work of Ruzsa on the discrete squarefree difference problem, which provides inspiration for our speculated results for the squarefree subset problem in the continuous setting. If X and Y are subsets of integers, we shall let $X \pm Y = \{x \pm y : x \in X, y \in Y, x \pm y > 0\}$ denote the sums and difference of the set. The *differences* of a set X are elements of $X - X$, and so the squarefree difference set problem asks to consider how large a subset of the integers can be, whose differences do not contain the square of any positive integer. We let $D(N)$ denote the maximum number of integers which can be selected from $[1,N]$ whose differences do not contain a square.

Example. *The set* $X = \{1,3,6,8\}$ *is squarefree, because* $X - X = \{2,3,5,7\}$*, and none of these elements are perfect squares. On the other hand,* $\{1,3,5\}$ *is not a squarefree subset, because* $5 - 1 = 4$ *is a perfect square.*

There are a few tricks to constructing large subsets of integers avoiding squares. If p is prime, then $p\mathbf{Z} \cap [1,p^2)$ avoids squares, because the difference of two numbers must be divisible by p, but not by p^2. If $N = p^2$, this gives a set with $N^{1/2}$ elements. However, we can do just as well without using any properties of the set of squares except for their sparsity, by greedily applying a sieve. We start by writing out a large list of integers $1,2,3,4,\dots,N$. Then, while we still have numbers to pick, we greedily select the smallest number x_* we haven't crossed out of the list, add it to our set X of squarefree numbers, and then cross out all integers y such that $y - x_*$ is a positive square. Thus we cross out x_*, $x_* + 1$, $x_* + 4$, and so on, all the way up to $x_* + m^2$, where m is the largest integer with $x_* + m^2 \leq N$. This implies $m \leq \sqrt{N - x_*} \leq \sqrt{N-1}$, hence we cross out at most $\sqrt{N-1} + 1$ integers whenever with add a new element x_* to X. When the algorithm terminates, all integers must be crossed out, and if the algorithm runs n iterations, a union bound gives that we cross out at most $n[\sqrt{N-1}+1]$ integers, hence $n[\sqrt{N-1}+1] \geq N$.

It follows that the set X we end up with contains $\Omega(N^{1/2})$ elements. What's more, this algorithm generates an increasing family of squarefree subsets of the integers as n increases, so we may take the union of these subsets over all N to find an infinite squarefree subset X with $|X \cap [1,N]| = \Omega(\sqrt{N})$ for all N.

In 1978, Sárközy proved an upper bound on the size of squarefree subsets of the integers, showing $D(N) = O(N(\log N)^{-1/3+\varepsilon})$ for every $\varepsilon > 0$. In particular, this proves a conjecture of Lovász that every infinite squarefree subset has density zero, because if X is any infinite squarefree subset, then $|X \cap [1,N]| = o(N)$. Sárközy even conjectured that $D(N) = O(N^{1/2+\varepsilon})$ for all $\varepsilon > 0$. Thus the sieve technique is essentially optimal, an incredibly pessimistic point of view, since the Sieve method doesn't depend on any properties of the set of perfect squares. Ruzsa's results shows we should be more optimistic, taking advantage of the digit expansion of numbers to obtain infinite squarefree subsets X with $|X \cap [1,N]| = \Omega(N^{0.73})$. The method reduces the problem to a finitary problem of maximizing squarefree subsets modulo a squarefree integer m.

Theorem 4. *If m is a squarefree integer, then*

$$D(N) \geq \frac{n^{\gamma_m}}{m} = \Omega_m(n^{\gamma_m})$$

where

$$\gamma_m = \frac{1}{2} + \frac{\log_m |R^*|}{m}$$

and R^* denotes the maximal subset of $[1,m]$ whose differences contain no squares modulo m. Setting $m = 65$ gives

$$\gamma_m = \frac{1}{2}\left(1 + \frac{\log 7}{\log 65}\right) = 0.733077\ldots$$

and therefore $D(N) = \Omega(n^{0.7})$. For $m = 2$, we find $D(N) \geq \sqrt{N}/2$, which is only slightly worse than the sieve result.

Remark. *Let us look at the analysis of the sieve method backwards. Rather than fixing N and trying to find optimal solutions of $[1,N]$, let's fix a particular strategy (to start with, the sieve strategy), and think of varying N and seeing how the size of the solution given by the strategy on $[1,N]$ increases over time. In our analysis, the size of a solution is directly related to the number of iterations the stategy can produce before it runs out of integers to add to a solution set. Because we apply a union bound in our analysis, the cost of each particular new iteration is the same* *as the cost of the other iterations. If the cost of each iteration was independant of N, we could increase the solution size by increasing N by a fixed constant, leading to family of solutions which increases on the order of N. However, as we increase N, the cost of each iteration increases on the order of $\sqrt{N}$, leading to us only being able to perform $N/\sqrt{N} = \sqrt{N}$ iterations for a fixed N. Rusza's method applies the properties of the perfect squares to perform a similar method of expansion. At an exponential cost, Rusza's method increases the solution size exponentially. The advantage of exponentials is that, since Rusza's is based on a particular parameter, a squarefree integer m, we can vary m to improve the iteration numbers more naturally.*

The idea of Rusza's construction is to break the problem into exponentially large intervals, upon which we can solve the problem modulo an integer. More enerally, Rusza constructs a set whose differences are free of d'th powers.

Theorem 5. *Let $R \subset [1,m]$ be a subset of integers such that no difference is a power of d modulo m, where m is a squarefree integer. Construct the set*

$$A = \left\{ \sum_{k=0}^{n} r_k m^k : 0 \le n < \infty, r_k \in \begin{Bmatrix} R & d \text{ divides } N \\ [1,m] & \text{otherwise} \end{Bmatrix} \right\}$$

Then A is squarefree.

Proof. Suppose that we can write $\sum(r_k - r'_k)m^k = N^d$. Let s to be the smallest index with $r_s \neq r'_s$. Then $(r_s - r'_s)m^s + Mm^{s+1} = N^d$ where M is some positive integer. If $s = ds_0$, then $(N/m^{s_0})^d = (r_s - r'_s) + Mm$, and this contradicts the fact that $r_s - r'_s$ cannot be a d'th power modulo m. On the other hand, we know m^s divides N^d, but m^{s+1} does not. This is impossible if s is not divisible by d, because primes in N^d occur in multiples of d, and m is squarefree. $\square$

For any n, we find

$$A \cap [1, m^n - 1] = \left\{ \sum_{k=0}^{n-1} r_k m^k : r_k \in [1,m], r_k \in R \text{ when } d \text{ divides } k \right\}$$

which therefore has cardinality

$$|R|^{1+\lfloor \frac{n-1}{d} \rfloor} m^{n-1-\lfloor \frac{n-1}{d} \rfloor} = m^n \left(\frac{|R|}{m} \right)^{1+\lfloor \frac{n-1}{d} \rfloor} \ge m^n \left(\frac{|R|}{m} \right)^{n/d} = m^{n\gamma(m,d)}$$

where $\gamma(m,d) = 1 - 1/d + \log_m |R|/d$. Therefore, for $m^{n+1} - 1 \geq k \geq m^n - 1$

$$A \cap [1,k] \geq A \cap [1,m^n] \geq m^{n\gamma(m,d)} = \frac{m^{(n+1)\gamma(m,d)}}{m} \geq \frac{k^{\gamma(m,d)}}{m}$$

This completes Rusza's construction. Thus we have proved a more general result than was required.

Theorem 6. *For every d and squarefree integer m, we can construct a set X whose differences contain no dth powers and*

$$|X \cap [1,n]| \geq \frac{n^{\gamma(d,m)}}{m} = \Omega(n^{\gamma(d,m)})$$

where $\gamma(d,m) = 1 - 1/d + \log_m |R^*|/d$*, and* R^* *is the largest subset of* $[1,m]$ *containing no d'th powers modulo m.*

For $m = 65$, the group $\mathbf{Z}_{65}^* \cong \mathbf{Z}_5^* \times \mathbf{Z}_{13}^*$ has a set of squarefree residues of the form $\{(0,0),(0,2),(1,8),(2,1),(2,3),(3,9),(4,7)\}$, which gives the required value for γ_{65}. In 2016, Mikhail Gabdullin proved that if m is squarefree, then in $\mathbf{Z}_m$, any set R such that $R - R$ is squarefree has $|R| \leq me^{-c\log m/\log\log m}$, where n denotes the number of odd prime divisors of m, so that

$$\gamma(d,m) \leq 1 - 1/d + \frac{\log(me^{-c\log m/\log\log m})}{m}$$

Rusza believes that we cannot choose m to construct squarefree subsets of the integers growing better than $\Omega(n^{3/4})$, and he claims to have proved this assuming m is squarefree and consists only of primes congruent to 1 modulo 4. Looking at some sophisticated papers in number theory (Though I forgot to write down the particular references), it seems that using modern estimates this is quite easy to prove. Thus expanding on Rusza's result in the discrete case requires a new strategy, or perhaps Rusza's result is the best possible.

Let $D(N,d)$ denote the largest subset of $[1,N]$ containing no dth powers of positive integer. The last part of Rusza's paper is devoted to lower bounding the polynomial growth of $D(N,d)$ asymptotically.

Theorem 7. *If p is the least prime congruent to one modulo* $2d$*, then*

$$\limsup_{N\to\infty} \frac{\log D(N,d)}{\log N} \geq 1 - \frac{1}{d} + \frac{\log_p d}{d}$$

Proof. The set X we constructed in the last theorem shows that for any m,

$$\frac{\log D(N,d)}{\log n} \geq \gamma(d,m) - \frac{\log m}{\log n} = 1 - \frac{1}{d} + \frac{\log_m |R^*|}{d} - \frac{\log m}{\log n}$$

Hence

$$\limsup_{N\to\infty} \frac{\log D(N,d)}{\log n} \geq 1 - \frac{1}{d} + \frac{\log_m |R^*|}{d}$$

The claim is then completed by the following lemma. □

Lemma 3. *If p is a prime congruent to* 1 *modulo* $2d$*, then we can construct a set $R \subset [1,p]$ whose differences do not contain a dth power modulo p with $|R| \geq d$.*

Proof. Let $Q \subset [1,p]$ be the set of powers $1^k, 2^k, \ldots, p^k$ modulo p. We have

$$|Q| = \frac{p-1}{k} + 1$$

This follows because the nonzero elements of Q are the images of the group homomorphism $x \mapsto x^k$ from $\mathbf{Z}_p^*$ to itself. Since $\mathbf{Z}_p^*$ is cyclic, the equation $x^k = 1$ has the same number of solutions as the equation $kx = 0$ modulo $p-1$, and since $p \equiv 1$ modulo $2k$, there are exactly k solutions to this equation. The sieve method yields a kth power modulo p free subset of size greater than or equal to

$$p/q = \frac{p}{1 + \frac{p-1}{k}} = \frac{pk}{p+k-1} \to k$$

as $p \to \infty$, which is greater than $k-1$ for large enough p. This shows the theorem is essentially trivial for large primes. However, for smaller primes a more robust analysis is required. We shall construct a sequence $b_1, \ldots, b_k \in \mathbf{Z}_p$ such that $b_i - b_j \notin Q$ for any i, j and $|B_j + Q| \leq 1 + j(q-1)$. Given $b_1, \ldots, b_j$, let b_{j+1} be any element of $(B_j + Q + Q) - (B_j + Q)$. Since $b_{j+1} \notin B_j + Q$, $b_{j+1} - b_i \notin Q$ for any i. Since $b_{j+1} \in B_j + Q + Q$, the sets $B_j + Q$ and $b_{j+1} + Q$ are not disjoint (we have used $Q = -Q$, which is implied when $p \equiv 1 \bmod 2k$), and so

$$\begin{aligned}
|B_{j+1} + Q| &= |(B_j + Q) \cup (b_{j+1} + Q)| \\
&\leq |B_j + Q| + |b_{j+1} + Q| - 1 \\
&\leq 1 + j(q-1) + q - 1 \\
&= 1 + (j+1)(q-1)
\end{aligned}$$

This procedure ends when $B_j + Q + Q = B_j + Q$, and this can only happen if $B_j + Q = \mathbf{Z}_p$, because we can obtain all integers by adding elements of Q recursively, so $1 + j(q-1) \geq p$, and thus $j \geq k$. $\square$

Corollary. *In the special case of avoiding squarefree numbers, we find*

$$\limsup \frac{\log D(N)}{\log N} \geq \frac{1}{2} + \frac{\log_5 2}{2} = 0.71533\ldots$$

which is only slightly worse than the bound we obtain with $m = 65$.

Rusza's leaves the ultimate question of whether one can calculate

$$\alpha = \lim_{N\to\infty} \log D(N) / \log N$$

or even whether it exists at all. The consequence of this would essentially solve the squarefree integers problem, since it gives the exact growth of $D(N)$ in terms of a monomial. Because of how conclusive this problem is, we should not expect to find a nontrivial way to calculate this constant.

2.4 Keleti: A Translate Avoiding Set

Keleti's two page paper constructs a full dimensional subset X of $[0,1]$ such that X intersects $t + X$ in at most one place for each nonzero real number t. Malabika has adapted this technique to construct high dimensional subsets avoiding nontrivial solutions to differentiable functions. In this section, and in the sequel, we shall find it is most convenient to avoid certain configurations by expressing them in terms of an equation, whose properties we can then exploit. One feature of translation avoidance is that the problem is specified in terms of a linear equation.

Lemma 4. *A set X avoids translates if and only if there do not exists values $x_1 < x_2 \leq x_3 < x_4$ in X with $x_2 - x_1 = x_4 - x_3$.*

Proof. Suppose $t + X \cap X$ contains two points $a < b$. Without loss of generality, we may assume that $t > 0$. If $a \leq b - t$, then the equation $a - (a-t) = t = b - (b-t)$ satisfies the constraints, since $a - t < a \leq b - t < b$ are all elements of X. We also have $(b-t) - (a-t) = b - a$ which satisfies the constraints if $a - t < b - t \leq a < b$. This covers all possible cases. Conversely, if there are $x_1 < x_2 \leq x_3 < x_4$ in X with $x_2 - x_1 = t = x_4 - x_3$, then $X + t$ contains $x_2 = x_1 + (x_2 - x_1)$ and $x_4 = x_3 + (x_4 - x_3)$. $\square$

The basic, but fundamental idea to Keleti's technique is to introduce memory into Cantor set constructions. Keleti constructs a nested family of discrete sets $X_0 \supset X_1 \supset \dots$ converging to X, with each X_N a union of disjoint intervals in $\mathscr{B}(L_N)$, for a decreasing sequence of lengths L_N converging to zero, to be chosen later, but with $10L_{N+1} \mid L_N$. We initialize $X_0 = [0,1]$, and $L_0 = 1$. Furthermore, we consider a queue of intervals, initially just containining $[0,1]$. To construct $X_1, X_2, \dots$, Keleti iteratively performs the following procedure:

Algorithm 1 Construction of the Sets X_N

Set $N = 0$
Repeat
  Take off an interval I from the front of the queue
  For all $J \in \mathscr{B}(L_N)$ contained in X_N:
    Order the intervals in $\mathscr{B}(L_{N+1})$ contained in J as $J_0, J_1, \dots, J_M$
    If $J \subset I$, add all intervals J_i to X_{N+1} with $i \equiv 0$ modulo 10
    Else add all J_i with $i \equiv 5$ modulo 10
  Add all intervals in $\mathscr{B}(L_{N+1})$ to the end of the queue
  Increase N by 1

After each iteration of the algorithm, we obtain a new set X_{N+1}, and so leaving the algorithm to repeat infinitely produces a sequence of sets $X_1, X_2, \dots$ converging to a set X. We claim that with the appropriate choice of parameters, X is a translate avoiding set.

If X is not translate avoiding, there is $x_1 < x_2 \leq x_3 < x_4$ with $x_2 - x_1 = x_4 - x_3$. Since $L_N \to 0$, there is N such that x_1 is contained in an interval $I \in \mathscr{B}(L_N)$ that x_2, x_3, x_4 are not contained in. At stage N of the algorithm, the interval I is added to the end of the queue, and at a much later stage M, the interval I is retrieved. Find the startpoints $x_1^\circ, x_2^\circ, x_3^\circ, x_4^\circ \in L_M \mathbf{Z}$ to the intervals in $\mathscr{B}(L_M)$ containing x_1, x_2, x_3, and x_4. Then we can find n and m such that $x_4^\circ - x_3^\circ = (10n)L_M$, and $x_2^\circ - x_1^\circ = (10m+5)L_M$. In particular, this means that $|(x_4^\circ - x_3^\circ) - (x_2^\circ - x_1^\circ)| \geq 5L_M$. But

$$\begin{aligned} |(x_4^\circ - x_3^\circ) - (x_2^\circ - x_1^\circ)| &= |[(x_4^\circ - x_3^\circ) - (x_2^\circ - x_1^\circ)] - [(x_4 - x_3) - (x_2 - x_1)]| \\ &\leq |x_1^\circ - x_1| + \dots + |x_4^\circ - x_4| \leq 4L_M \end{aligned}$$

which gives a contradiction.

The algorithm shows that X_N contains $L_{N-1}/10L_N$ times the number of intervals that X_{N-1} has, but they are at a length L_N rather than L_{N-1}. This means that in total, X_N contains $1/10^N L_N$ intervals, of length L_N. Since $L_N/10^N L_N = o(1)$, this shows our set will have Lebesgue measure zero irrespective of our parameters. However, if L_N decays suitably fast, then we might have $L_N^{1-\varepsilon}/10^N L_N \gtrsim_\varepsilon 1$ for all $\varepsilon > 0$, which would imply that X has positive $1-\varepsilon$ dimensional Hausdorff measure for all ε, so X still has Hausdorff dimension one. For this to be true, L_N must decay superexponentially, i.e. the inequalities above are equivalent to $L_N \lesssim_B 1/B^N$ for all choices of B. Choosing an arbitrarily fast decaying sequence, such as $L_N = 1/N! \cdot 10^N$ or $L_N = 1/10^{10^N}$, suffices to obtain a Hausdorff dimension one set.

Lemma 5. *If L_N decays superexponentially, X has Hausdorff dimension one.*

Proof. We use the mass distribution principle, as used in our note on calculating Hausdorff dimensions. It is easy to establish the bounds $\mu_N(I) \lesssim_\varepsilon L(I)^{1-\varepsilon}$ for $I \in \mathscr{B}(L_N)$, and since we can choose L_N suitably slowly decreasing to use the epsilon of room technique, this gives the result. Alternatively, we can use the uniform distribution bounds with $L_N = R_N$, since if $J \in \mathscr{B}(L_{N+1})$, $I \in \mathscr{B}(L_N)$, $\mu(J) = 1/10^{N+1}L_{N+1}$, $\mu(I) = 1/10^N L_N$, and so $\mu(J) \lesssim (L_{N+1}/L_N)\mu(I)$. This gives the result if L_N grows too fast. $\square$

Remark. *Keleti briefly remarks that by replacing the 10 in the algorithm with a slowly increasing set of numbers, one can obtain a Hausdorff dimension one set which is linearly independant over the rational numbers. To see why this works, the condition of linear independence would fail if $a_1 x_1 + \cdots + a_M x_M = 0$, where $x_1 < x_2 < \cdots < x_M$, and $a_1, \dots, a_M$ are integers with no common factor. One can again reduce this by picking intervals with indices congruent to a certain large modulus.*

2.5 Fraser/Pramanik: Extending Keleti Translation to Smooth Configurations

Inspired by Keleti's result, Pramanik and Fraser obtained a generalization of the queue method which allows one to find sets avoiding solutions to *any* smooth function satisfying suitably mild regularity conditions. To do this, rather than making a linear shift in one of the intervals we avoid as in Keleti's approach, one must use the smoothness properties of the function to find large segments of an interval avoiding solutions to another interval.

Theorem 8. *Suppose that $f : \mathbf{R}^{d+1} \to \mathbf{R}$ is a C^1 function, and there are sets $T_0, \dots, T_d \subset [0,1]$, with each T_n a union of almost disjoint closed intervals of length $1/M$ such that $A \leq |\partial_0 f|$ and $|\nabla f| \leq B$ on $T_0 \times \cdots \times T_d$. There there exists a rational constant C and arbitrarily large integers $N \in M\mathbf{Z}$ for which there exist subsets $S_n \subset T_n$ such that*

(i) $f(x) \neq 0$ *for* $x \in S_0 \times \cdots \times S_d$.

(ii) For $n \neq 0$, if we divide each interval T_n into length $1/N$ intervals, then S_n contains an interval of length C/N^d of each of these intervals.

(iii) If T_0 is split into length $1/N$ intervals, then for a fraction $1 - 1/M$ of such intervals, S_0 is a union of length C/N^d intervals with total length C/N.

Proof. We begin by dividing the sets $T_1, \dots, T_d$ into length $1/N$ intervals, and let S_n be defined by including a length C_0/N^d segment, for some constant C_0 to be chosen later. Then once we fix C_0, the S_n will satisfy property (ii) of the theorem. We define

$$\mathbf{A} = \{a \in \mathbf{R}^{d-1} : a_n \text{ is a startpoint of a length } 1/N \text{ interval in } T_n\}$$

Then $|\mathbf{A}| \leq N^d$, since each interval T_n is contained in $[0,1]$, and therefore can only contain at most N almost disjoint intervals of length $1/N$. Hence if we define the set of 'bad points' in T_0 as

$$\mathbf{B} = \{x \in T_0 : \text{there is } a \in \mathbf{A} \text{ such that } f(x,a) = 0\}$$

Then $|\mathbf{B}| \leq MN^d$. This is because for each fixed a, the function $x \mapsto f(x,a)$ is either strictly increasing or decreasing over each interval in the decomposition of T_0, or which there are at most M because $T_0 \subset [0,1]$. If we split T_0 into length $1/N$ intervals, and choose a subcollection of such intervals I such that $|I \cap \mathbf{B}| \leq M^3 N^{d-1}$, then we throw away at most $MN^d/M^3N^{d-1} = N/M^2$ intervals, and so we keep $(N/M)(1-1/M)$ intervals, which is $1 - 1/M$ of the total number of intervals in the decomposition of T_0. The lemma we prove after this theorem implies that there exists a constant C_1 such that if $x \in S_n$, and $f(y,x) = 0$, then $d(y,\mathbf{B}) \leq C_0 C_1/N^d$. If we split each interval I with $|I \cap \mathbf{B}| \leq M^3 N^{d-1}$ into $4M^3N^{d-1}$ length $1/4M^3N^d$ intervals, and we choose C_0 such that $C_0 C_1 < 1/4M^3$, then the set S_0 obtained by discarding each interval that contains or is adjacent to an interval containing an element of $\mathbf{B}$ satisfies $d(S_0, \mathbf{B}) > C_0 C_1/N^d$, and therefore there does not exist any

$x_n \in S_n$ and $y \in S_0$ such that $f(y,x) = 0$. S_0 satisfies property (iii) of the theorem since for the interval I we are considering, we keep at least $M^3 N^{d-1}$ length $1/4M^3N^d$ intervals, which in total has length at least $1/4N$. $\square$

Remark. *The length $1/N$ portion of each interval guaranteed by (iii) is unneccesary to the Hausdorff dimension bound, since the slightly better bounds obtained on scales where an interval is dissected as a $1/N$ are decimated when we eventually divide the further subintervals into $1/N^{d-1}$ intervals. The importance of (iii) is that it implies that the set we will construct has full Minkowski dimension. The reason for this is that Minkowski dimension lacks the ability to look at varying dissection depths at once, and since, at any particular depth, there exists a length $1/N$ dissection, the process appears to Minkowski to be full dimensional, even though at later scales this $1/N$ dissection is dissected into $1/N^{d-1}$ intervals.*

Lemma 6. *Given the f, $T_0, \dots, T_d$, there exists a constant C_1 depending on these quantities, such that for any C_0, and $x \in S_1 \times \cdots \times S_{d-1}$, if $f(y,x) = 0$, then $d(y, \mathbf{B}) \leq C_0 C_1 / N^{d-1}$.*

Proof. Since $T_0 \times \cdots \times T_d$ breaks into finitely many cubes with sidelengths $1/M$, it suffices to prove the theorem for a particular cube J in this decomposition, where we assume the zeroset of f intersects J. If $J = I \times J'$, where I is an interval, we let U be the set of all $x \in J'$ for which there is y in the interior of I such that $f(y,x) = 0$. Then U is open. The implicit function theorem implies that there exists a C^1 function $g : U \to I$ such that $f(x,y) = 0$ if and only if $y = g(x)$. Then the function $h(x) = f(x, g(x))$ vanishes uniformly, so

$$0 = \partial_n h(x) = (\partial_n f)(g(x),x) + (\partial_0 f)(g(x),x)\partial_n g(x)$$

Hence for $x \in U$,

$$|(\nabla g)(x)| = \frac{|(\nabla f)(x)|}{|(\partial_d f)(x, g(x))|} \leq \frac{B}{A}$$

If N is chosen large enough, then for every $x \in U \cap (S_1 \times \cdots \times S_d)$ there is $a \in \mathbf{A} \cap U$ in the same connected component of U as x with $|x - a| \lesssim C_0/N^{d-1}$, and this means that

$$|g(x) - g(a)| \leq \|\nabla g\|_\infty |x - a| \lesssim \frac{BC_0}{AN^{d-1}}$$

and $g(a) \in \mathbf{B}$, completing the proof. $\square$

How do we use this lemma to construct a set avoiding solutions to f? We form an infinite queue which will eventually filter out all the possible zeroes of the equation. Divide the interval $[0,1]$ into d intervals, and consider all orderings of $d-1$ subsets of these intervals, and add them to the queue. Now on each iteration N of the algorithm, we have a set $X_N \subset [0,1]$. We take a particular sequence of intervals $T_1,\dots,T_d$ from the queue, and then use the lemma above to dissect the $X_N \cap T_n$, which are unions of intervals, into sets avoiding solutions to the equation, and describe the remaining points as X_{N+1}. We then add all possible orderings of d intervals created into the end of the queue, and rinse and repeat. The set $X = \lim X_n$ then avoids all solutions to the equation with distinct inputs.

What remains is to bound the Hausdorff dimension of X by constructing a probability measure supported on X with suitable decay. To construct our probability measure, we begin with a uniform measure on the interval, and then, whenever our interval is refined, we uniformly distribute the volume on that particular interval uniformly over the new refinement. Let μ denote the weak limit of this sequence of probability distributions. At each step n of the process, we let $1/M_n$ denote the size of the intervals at the beginning of the n'th subdivision, $1/N_n$ denote the size of the split intervals in the lemma, and C_n the n'th constant. We have the relation $1/M_{n+1} = C_n/N_n^{d-1}$. If K is a length $1/M_{N+1}$ interval, J a length $1/N_N$ interval, and I a length $1/M_N$ interval with $K \subset J \subset I$ and all recieving some mass in μ. To calculate a bound on their mass, we consider the decompositions considered in the algorithm:

- If J is subdivided in the non-specialized manner, then every length $1/N_N$ interval recieves the same mass, which is allocated to a single length $1/M_{N+1}$ interval it contains. Thus $\mu(K) = \mu(J) \leq (M_N/N_N)\mu(I)$.

- In the second case, at least a fraction $1-1/M_N$ of the length $1/N_N$ intervals are assigned mass, so $\mu(J) = (M_N/N_N)(1-1/M_N)^{-1} \leq (2M_N/N_N)\mu(I)$, and more than C_N/N_N of each length $1/N_N$ interval is maintained, so

$$\mu(K) = \frac{N_N}{C_N M_{N+1}}\mu(J) \leq \frac{2M_N}{C_N M_{N+1}}\mu(I)$$

Thus in both cases, we have $\mu(J) \lesssim (N_N/M_N)\mu(I)$, $\mu(K) \lesssim_N |K|$, and $N_N = M_{N+1}^{1/(d-1)}/C_N \lesssim_N M_{N+1}^{1/(d-1)}$. From this, we conclude using the results of the appendix that there exists a family of rapidly decaying parameters which gives a $1/(d-1)$ dimensional set.

Remark. *The set X constructed is precisely a $1/(d-1)$ dimensional set. Recall that $X = \lim X_n$, where X_n is a union of a certain number of length $1/M_n$ intervals*

$I_1,\dots,I_N$. For each n, the interval I_i is inevitably subdivided at a stage J_i into length $C_{J_i}N_{J_i}^{1-d}$ intervals for each length $1/N_{J_i}$ interval that I_i contains. Thus

$$H^{\alpha}_{1/M_n}(X) \leq \sum_{i=1}^{N} \frac{N_{m_i}}{M_n}(C_{m_i}N_{m_i}^{1-d})^{\alpha} = \frac{1}{M_n}\sum_{i=1}^{N} C_{m_i}^{\alpha}N_{m_i}^{1-\alpha(d-1)}$$

We may assume that $C_{m_i} \leq 1$, so if $\alpha > 1/(d-1)$, using the fact that $N \leq M_n$, since X_n is contained in $[0,1]$, we obtain

$$H^{\alpha}_{1/M_n}(X) \leq \frac{1}{M_n}\sum_{i=1}^{N} N_{m_i}^{1-\alpha(d-1)} \leq N_{\max(m_i)}^{1-\alpha(d-1)} \leq 1$$

Thus, taking $n \to \infty$, we conclude $H^{\alpha}(X) \leq 1 < \infty$, so as $\alpha \downarrow 1/(d-1)$, we conclude that X has Hausdorff dimension bounded above by $1/(d-1)$.

2.6 A Set Avoiding All Functions With A Common Derivative

In the latter part's of their paper, Pramanik and Fraser apply an iterative technique to construct, for each α with $\sum \alpha_n = 0$ and $K > 0$, a set E of positive Hausdorff dimension avoiding solutions to any function $f : \mathbf{R}^d \to \mathbf{R}$ satisfying wth $(\partial_n f)(0) = \alpha_n$,

$$\left| f(x) - \sum \alpha_n x_n \right| \leq K \sum_{n \neq 1} (x_n - x_1)^2$$

The set of such f is an uncountable family, which makes this situation interesting. The technique to create such a set relies on another iterative procedure.

Lemma 7. *Let $I \subsetneq [1,d]$ be a strict subset of indices, and $\delta_0 > 0$. Then there exists $\varepsilon > 0$ such that for any $\lambda > 0$ and two disjoint intervals J_1 and J_2, with J_1 occuring before J_2, and if we set*

$$[a_n, b_n] = \begin{cases} J_1 & n \in I \\ J_2 & n \notin I \end{cases}$$

then for $\delta < \delta_0$, either for all $x_n \in [a_n, a_n + \varepsilon\lambda]$ or for all $x_n \in [b_n - \varepsilon\lambda, b_n]$,

$$\left|\sum \alpha_n x_n\right| \geq \delta\lambda$$

Proof. If $C^* = \sum |\alpha_n|$, then for $|x_n - a_n| \leq \varepsilon\lambda$,

$$|\sum \alpha_n (x_n - a_n)| \leq C^* \varepsilon \lambda$$

Thus if $|\sum \alpha_n a_n| > (\delta + \varepsilon C^*)\lambda$, then $|\sum \alpha_n x_n| \geq \delta\lambda$. If this does not occur $\square$

2.7 Equidistribution Results

The classical Weyl equidistribution theorem says that if α is an irrational number, then the decimal parts of the numbers $n\alpha$ are equidistributed in $\mathbf{T} = \mathbf{R}/\mathbf{Z}$, in the sense that for any continuous function $f : \mathbf{T} \to \mathbf{R}$,

$$\frac{1}{N} \sum_{n \leq N} f(n\alpha) \to \int_{\mathbf{T}} f(x)\, dx$$

This is equivalent to prove that for any interval $[a,b] \in [0,1)$,

$$\frac{\#\{1 \leq n \leq N : x_n \in [a,b]\}}{N} \to b - a$$

as $N \to \infty$. By approximating a continuous function f by a Fourier series, to prove this is true for a particular sequence, it suffices to prove it for $f(x) = e(nx) = e^{2\pi inx}$, for each nonzero integer n. Certain techniques we are developing in the theory of cantor decompositions require a higher dimensional variant of such a result, so this section details some information which might help us in the future. We will encounter sequences that are not equidistributed over the entire space, so if G is any closed subgroup of $\mathbf{T}$, we say a sequence x_n is equidistributed over G if $x_n \in G$ for all n, and for any continuous function $f : G \to \mathbf{R}$,

$$\frac{1}{N} \sum_{n \leq N} f(n\alpha) \to \int_{G} f(x)\, dx$$

or alternatively, if for any closed set K in G,

$$\frac{\#\{1 \leq n \leq N : x_n \in K\}}{N} \to |K|$$

where $|K|$ is taken with respect to the Haar probability measure on G.

Theorem 9. *A sequence $x_1, x_2, \dots$ is equidistributed in $\mathbf{T}^d$ if and only if for every* *nonzero $\xi \in \mathbf{Z}^n$,*

$$\frac{1}{N}\sum_{n\leq N} e(\xi \cdot x_n) \to 0$$

Proof. We prove that the exponential sum condition implies the general result, noting that the other direction is clear. Clearly the exponential sum condition implies the result for all functions f which are trigonometric polynomials. But then, by the Stone Weirstrass theorem and basic Abelian Harmonic analysis, the multivariate trigonometric polynomials are dense in $C(\mathbf{T}^d)$, and we may apply a standard limiting argument. □

Example. *If $x_n = n\alpha + \beta$, then*

$$\frac{1}{N}\sum_{n\leq N} e(\xi \cdot x_n) = \frac{e(\xi \cdot \beta)}{N}\sum_{n\leq N} e(n\xi \cdot \alpha) = \begin{cases} \frac{1}{N}\frac{e(\xi\cdot\beta)(e((n+1)\xi\cdot\alpha)-1)}{e(\xi\cdot\alpha)-1} & : \xi\cdot\alpha \notin \mathbf{Z} \\ e(\xi\cdot\beta) & : \xi\cdot\alpha \in \mathbf{Z} \end{cases}$$

Weyl's exponential sum theorem implies that x_n is equidistributed on $\mathbf{T}^n$ precisely when $\xi\cdot\alpha \notin \mathbf{Z}$ for all $\xi \in \mathbf{Z}^n$. What's more, it is simple to see from this that x_n is still equidistributed for any subsequence whose indices form an arithmetic progression.

Thus in one dimension, an arithmetic sequence is either equidistributed over the entire torus, or over a discrete set of points forming a *discrete subgroup* of the torus. We can think of this discrete subgroup as a zero dimension torus, which leads us to suspect that in higher dimensions, an arithmetic sequence is always equidistributed, but not necessarily over the whole torus, but instead over a lower dimensional subtorus.

Theorem 10 (Ratner). *If $x_n = n\alpha + \beta$, then we can write $\alpha = \alpha_0 + \alpha_1$, where the sequence $n\alpha_0 + \beta$ is periodic in $\mathbf{T}^d$, and $n\alpha_1 + \beta$ is equidistributed over a subtorus of $\mathbf{T}^d$. In particular, if $\beta = 0$, then Ratner's theorem says that x_n has an evenly spaced subsequence equidistributed over a subtorus of the space containing the origin.*

Proof. We induct on the dimension. For $d = 1$, the theorem is obvious, since either α is rational, and therefore $n\alpha + \beta$ is periodic, or α is irrational, and $n\alpha + \beta$ is equidistributed on $\mathbf{T}$. So now let us consider a sequence on the torus $\mathbf{T}^{d+1}$. If α is irrational, then x_n is equidistributed, and the theorem is obvious. Otherwise, there exists $\xi \in \mathbf{Z}^n$ such that $\xi\cdot\alpha \in \mathbf{Z}$. We may write $\alpha = \alpha_0 + \alpha_1$, where $\alpha_0 \in \mathbf{Q}^d$, and $\xi\cdot\alpha_1 = 0$. The sequence $n\alpha_0 + \beta$ is periodic, whereas $n\alpha_1$ takes values in the subtorus T of points x with $\xi \cdot x = 0$. Since $\xi \neq 0$, T is a d dimensional compact subgroup of $\mathbf{T}^{d+1}$ isomorphic to $\mathbf{T}^d$. To see this, we assume for simplicity that $\xi_d \neq 0$. Then $\xi^\perp$ is generated by the basis of d vectors $v_n = \xi_d e_n - \xi_n e_d$, for $1 \leq n \leq d$. Thus we get a homomorphism between $\mathbf{R}^d$ and T given by the map $x \mapsto \sum x_n v_n$. If $\sum x_n v_n \in \mathbf{Z}$, so that $\sum x_n v_n = 0$ on T, then this means that $x_n \in \mathbf{Z}/\xi_d$ for each n, implying that the kernel of this homomorphism is discrete. Lattice theory implies that we can write the kernel as $\bigoplus \mathbf{Z}\langle w_n \rangle$, for d generating vectors $w_1, \dots, w_d$. But then the map $f(x) = \sum x_n w_n$ gives an isomorphism between $\mathbf{T}^d$ and T. If we set $\beta = f^{-1}(\alpha_0)$, then $f^{-1}(n\alpha_0) = n\beta$, and so by induction, we can write $\beta = \beta_0 + \beta_1$, where $n\beta_0$ is periodic, and $n\beta_1$ is equidistributed over a subtorus of $\mathbf{T}^d$. But then $nf(\beta_0) = f(n\beta_0)$ is periodic on T, and $nf(\beta_1) = f(n\beta_1)$ is equidistributed over a subtorus of T. Since the sum of two periodic sequences is periodic, $n\alpha_0 + nf(\beta_0)$ is periodic, and $nf(\beta_1)$ is equidistributed over a subtorus. We have $\alpha_0 + f(\beta_0) + f(\beta_1) = \alpha_0 + f(\beta) = \alpha_0 + \alpha_1$, completing the proof. □

Corollary. *Any linear sequence in $\mathbf{T}^d$ is equidistributed in a finite union of cosets of a subtorus of $\mathbf{T}^d$.*

In general, ergodic theory results do not give rates on how long it takes for a sequence to equidistribute over a set. This is not a problem in the constructions we perform, since the rates that our intervals shrink can be arbitrarily fast. However, it is important to note that the convergence rates are uniform across all intervals.

Theorem 11. *If x_n is equidistributed over a torus $\mathbf{T}^d$, and*

$$A_N = \sup_I \left| \frac{\#\{1 \leq n \leq N : x_n \in I\}}{N} - |I| \right|$$

where I ranges over all boxes in $\mathbf{T}^d$, then $A_N \to 0$ as $N \to \infty$.

Proof. For notational simplicity, we let

$$\#(I,N) = \#\{1 \leq n \leq N : x_n \in I\}$$

For each n, we can partition $\mathbf{T}^d$ into finitely many disjoint cubes $\{I_n\}$ with sidelengths $1/n$. Since there are only finitely many such cubes, there is N_n such that for $M \geq N_n$, $|\#(I_n,M)/M - |I_n|| \leq 1/n$. Now given any box J, we can find sets J_1 and J_2, each unions of the cubes I_n, with $J_1 \subset J \subset J_2$ and $|J - J_1|, |J_2 - J| \lesssim_d 1/n$. Thus

$$|J| - \frac{2}{n} \leq |J_1| - \frac{1}{n} \leq \frac{\#(J_1,M)}{M} \leq \frac{\#(J,M)}{M} \leq \frac{\#(J_2,M)}{M} \leq |J_2| + \frac{1}{n} \leq |J| + \frac{2}{n}$$

which completes the proof, since N_n is independent of J. □

2.8 Results about Hypergraphs

Lemma 8 (Turán). *For any k uniform hypergraph $H = (V,E)$ with $|E| \leq |V|^{\alpha}$, V contains an independant set of size $\Omega(|V|^{(k-\alpha)/(k-1)})$.*

Proof. We create an independant set I by the following procedure. First, select a subset S of vertices, including each independantly with probability p. Delete a single vertex from each edge in each hypergraph entirely contained in S, obtaining an independant set I. We find that each edge in V is entirely included in S with probability p^k, and S has expected size $p|V|$, so $\mathbf{E}|I| = p|V| - p^k|E|$. If $|E| = |V|^{\alpha}$ for $\alpha \geq 1$, then setting $p = (1/2)|V|^{(1-\alpha)/(k-1)}$ induces a set I with size

$$|V|^{(k-\alpha)/(k-1)}(1/2 - 1/2^k)$$

We create an independant set I by the following procedure. First, select a subset S of vertices, including each vertex independantly with probability p. Delete a single vertex from each edge in each hypergraph which is entirely contained in S. Then I is an independant set with respect to each hypergraph, and we shall show that for an appropriate choice of p, $\mathbf{E}|I| \geq h$.

Trivially, we find $\mathbf{E}|S| = p|V|$. For any $i \geq 2$, the expected number of edges of H_i falling entirely in S is

$$p^i|E_i| \leq \frac{p^i|V|^i}{c_k h^{i-1}}$$

therefore

$$\mathbf{E}|I| = p|V| - \sum_{i=2}^{k} \frac{p^i|V|^i}{c_k h^{i-1}}$$

Setting $p = 2h/|V|$ and $c_k = 2^{k+1}$ gives

$$\mathbf{E}|I| = h\left(2 - \sum_{i=2}^{k} \frac{1}{2^{k+1-i}}\right) > h$$

which completes the proof. □

2.9 Hyperdyadic Covers

Recall the definition of the Hausdorff measure $H^\alpha(E) = \lim_{\delta\to 0} H^\alpha_\delta(E)$, where $H^\alpha_\delta(E)$ is the greatest lower bound of $\sum r_n^\alpha$, over all choices of covers of E by cubes $I_1, I_2, \dots$, where I_n has sidelengths r_n. We then define the Hausdorff dimension of E to be the least upper bound of the scalars α such that $H^\alpha(E) = 0$, or alternatively, the greatest lower bound of α such that $H^\alpha(E) = \infty$.

To determine the Hausdorff dimension of E, it suffices to consider only dyadic cubes in the cover of E. Define $H^\alpha(E) = \lim_{\delta\to 0} H^\alpha_{D,\varepsilon}(E)$, where $H^\alpha_{D,\varepsilon}(E)$ is the greatest lower bound of $\sum r_n^\alpha$ over *dyadic* covers $I_1, I_2, \dots$, with $I_n \in \mathscr{B}(r_n)$. Then H^α_D is comparable with H^α.

Theorem 12. *For any set E, $H^\alpha(E) \leq H^\alpha_D(E) \leq 2^{d+\alpha} H^\alpha(E)$.*

Proof. Given any not necessarily dyadic cover $I_1, I_2, \dots$, we can replace each sidelength r_n cube I_n with at most 2^d dyadic cubes with radius at most $2r_n$, which gives $H^\alpha_{D,\varepsilon}(E) \leq 2^{1+\alpha} H^\alpha_\varepsilon(E)$, and taking the limit as $\varepsilon \to 0$ then gives the required upper bound for H^α_D. □

If we are restricting ourselves to cubes lying at a series of discrete scales, it seems as if the dyadic sequence is about as fast as we can use so that the resultant Hausdorff measure is comparable to the usual Hausdorff measure. Nonetheless, using a weak type bound we can get results for a faster decreasing family of scales. This is necessary for our calculations. We fix a positive δ, and consider a sequence of **hyperdyadic scales** $H_N = 2^{-\lfloor (1+\delta)^N \rfloor}$. A **hyperdyadic cube** is then a cube in $\mathscr{B}(H_N)$ for some N.

Proof. For any sidelength L cube, we can cover the cube by at most 2^d hyperdyadic cubes with sidelength at most $2L^{1-\delta} \geq 2L^{(1+\delta)^{-1}}$. This is because

$$2H_{N+1}^{(1+\delta)^{-1}} = 2^{1-(1+\delta)^{-1}\lfloor (1+\delta)^{N+1} \rfloor} \geq 2^{1-(1+\delta)^N} \geq 2^{\lfloor (1+\delta)^N \rfloor} = H_N$$

If E has Hausdorff dimension α, for every ε and N we can find a collection of dyadic cubes $I_1, I_2, \dots$ covering E with I_k sidelength $L_k \leq H_N$, and $\sum L_{N,i}^{\alpha+\varepsilon} \lesssim_\varepsilon 1$. A weak type bound implies the number of cubes I_k with $H_{N+1} \leq L_k \leq H_N$ is $O_\varepsilon(1/H_{N+1}^{\alpha+\varepsilon})$. But

$$1/H_{N+1}^{\alpha+\varepsilon} \leq (H_N/H_{N+1})^{\alpha+\varepsilon} 1/H_N^{\alpha+\varepsilon} \lesssim 1/H_N^{\alpha+\varepsilon+\delta}$$

and so the cover of E by hyperdyadic cubes contains $O_\varepsilon(1/H_N^{\alpha+\varepsilon+\delta})$ length H_N cubes for each N.

If we swap each cube $I_{N,i}$ with 2^d hyperdyadic cubes of length at most $2L^{1-\delta}$, we obtain

$$\sum 2^d (2L_{N,i}^{1-\delta})^{\alpha+\varepsilon} = 2^{d+\alpha+\varepsilon} \sum L_{N,i}^{(1-\delta)(\alpha+\varepsilon)} \lesssim_\varepsilon 1$$

Thus $H_{HD}^{(1-\delta)\alpha+\varepsilon}(E) \lesssim_\varepsilon 1$.

We can swap each cube I_i with 2^d hyperdyadic cubes of length at most $2L^{(1+\delta)^{-1}}$, without effecting the estimate too much.

Then for every hyperdyadic number H_N, we can find a collection of cubes $I_{N,1}, I_{N,2}, \dots$ covering E with $I_{N,i}$ sidelength $r_{N,i} \leq H_N$, and $\sum r_{N,i}^{\alpha+\varepsilon} \lesssim_\varepsilon 1$. Covering each cube by 2^d cubes with hyperdyadic sidelengths, which magnifies $r_{N,i}$ by at most

$$2 \cdot 2^{(1+\delta)^{N+1} - (1+\delta)^N} = 2 \cdot 2^{\delta(1+\delta)^N} \lesssim 2 \cdot r_{N,i}^{-\delta}$$

We conclude that

$$2^{d+\alpha+\varepsilon} 2^{(\alpha+\varepsilon)\delta(1+\delta)^N} C_\varepsilon$$

□

We assume δ and ε are some fixed parameters. If $A(\varepsilon,\delta)$ and $B(\varepsilon,\delta)$ are two quantities depending on ε and δ, we write $A \preccurlyeq B$ mean $A \lesssim_\varepsilon \delta^{-C\varepsilon} B$ for some C, and for every ε. We let $A \approx B$ mean $A \preccurlyeq B$ and $B \preccurlyeq A$ hold simultaneously. We say a union of balls is δ discretized if it is the union of balls with radius $\approx \delta$. Thus there exists C_ε and C such that for each ball B_r of radius r, $|r - \delta| \leq C_\varepsilon \delta^{1-C\varepsilon}$. Thus

$$\delta(1 - C_\varepsilon \delta^{-C\varepsilon}) \leq r \leq \delta(1 + C_\varepsilon \delta^{-C\varepsilon})$$

In particular, the dyadic scales $2^{-\lfloor (1+\varepsilon)^k \rfloor}$ are allowed in a discretization of a hyperdyadic scale $2^{-(1+\varepsilon)^k}$, since we can choose C_ε and C such that

$$1 - C_\varepsilon 2^{C(1+\varepsilon)^k \varepsilon} \leq 1 \leq 2^{(1+\varepsilon)^k - \lfloor (1+\varepsilon)^k \rfloor} \leq 2 \leq 1 + C_\varepsilon 2^{(1+\varepsilon)^k C\varepsilon}$$

Theorem 13. *Let E be a compact subset of $\mathbf{R}^n$. If $0 < \alpha < n$, and $\dim(E) \leq \alpha$, then for each hyperdyadic number δ, we can associate a δ discretized set X_δ with $|X_\delta \cap B(x,r)| \preccurlyeq \delta^n (r/\delta)^\alpha$ for all $\delta \leq r \leq 1$ and $x \in \mathbf{R}^n$, and every element of E is contained in infinitely many of the X_δ.*

Proof. Fix E. For every hyperdyadic δ, we can find a cover of E by balls $B(x_{\delta n}, r_{\delta n})$

such that $r_{\delta n} < \delta$, and
$$\sum_n r_{\delta n}^{\alpha + C\varepsilon} \lesssim 1 \tag{2.1}$$

Choose $m_{\delta n}$ such that $2^{-(1+\varepsilon)^{m_{\delta n}+1}} \leq r_{\delta n} \leq 2^{-(1+\varepsilon)^{m_{\delta n}}}$. We calculate

$$\begin{aligned}
\frac{2^{-(\alpha+C'\varepsilon)(1+\varepsilon)^{m_{\delta n}}}}{r_{\delta n}^{\alpha+C\varepsilon}} &\leq \frac{2^{-(\alpha+C'\varepsilon)(1+\varepsilon)^{m_{\delta n}}}}{2^{-(\alpha+C\varepsilon)(1+\varepsilon)^{m_{\delta n}+1}}} \\
&= \left(2^{\varepsilon(1+\varepsilon)^{m_{\delta n}}}\right)^{\alpha+(C(1+\varepsilon)-C')}
\end{aligned}$$

Provided that $C' > \alpha + C(1+\varepsilon)$, the quantity on the left is ≤ 1, which is independant of ε provided that ε is bounded from above, and so we conclude

$$\sum_n \left(2^{-(1+\varepsilon)^{m_{\delta n}}}\right)^{\alpha+C'\varepsilon} \leq \sum_n r_{\delta n}^{\alpha+C\varepsilon}$$

Thus we may assume by changing the value of C that the quantities $r_{\delta n}$ are hyperdyadic from the outset. This means that at each hyperdyadic scale δ, the number of hyperdyadic balls at the scale δ in each cover is $\lesssim (1/\delta)^{\alpha+C\varepsilon}$. STOP IS THIS ALL WE NEED, THEN COME BACK TO THE PROOF.

For a pair of hyperdyadic numbers δ and γ we set

$$Y_{\delta\gamma} = \bigcup_{r_{\delta n}=\gamma} B(x_{\delta n}, r_{\delta n})$$

Every element of X is in infinitely many of the $Y_{\delta n}$. For each δ and γ, we let $Q_{\delta\gamma}$ be the collection of hyperdyadic cubes with sidelength at least γ covering $Y_{\delta\gamma}$ and minimizing $\sum_{Q \in Q_{\delta\gamma}} l(Q)^\alpha$. From condition (1.1) we obtain that $Y_{\delta\gamma}$ can be covered by at most $r^{-\alpha-\varepsilon}$ sidelength r cubes, so

$$\sum_{Q \in Q_{\delta\gamma}} l(Q)^\alpha \leq C r^{-\varepsilon}$$

and so $l(Q) \leq C r^{-\varepsilon/\alpha}$ for all $Q \in Q_{\delta\gamma}$. From the construction of $Q_{\delta\gamma}$, we see that the Q are all disjoint, and for any hyperdyadic cube I,

$$\sum_{\substack{Q \in Q_{\delta\gamma} \\ Q \subset I}} l(Q)^\alpha \leq l(I)^\alpha$$

since otherwise we could replace such elements of Q in $Q_{\delta\gamma}$ by I itself. $\square$

Bibliography

[1] P. Maga. Full dimensional sets without given patterns. *Real Anal. Exchange*, 36:79–90, 2010. → page 2

[2] P. Mattila. *Fourier analysis and Hausdorff dimension*. Cambridge University Press, 2015. → pages 1, 2