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.

Table of 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:

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 $\mathbf{R}^d$ that contain no three collinear points, or do not contain 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.

The Lebesgue density theory shows any subset of $\mathbf{R}^d$ with positive measure must contain a copy of the vertices of any suitably small isosceles triangle. Similar techniques show sets avoiding any of the patterns considered in this thesis must have measure zero. Thus non-trivial pattern avoiding sets must be highly irregular, and require techniques from geometric measure theory to be understood.

In particular, the Lebesgue measure cannot quantify the 'maximal size' of a pattern avoiding set. Thus a 'second-order' quantification of the size of a measure zero must be introduced, and this is satisfied by the fractional dimension of a set, first introduced by Felix Hausdorff in 1918. We shall take the Hausdorff dimension he introduced as a primary measure of a set's size.

At the time of the writing of this thesis, many fundamental questions about the relation between the geometric arrangement of a set and it's fractional dimension remain unsolved. 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 [12] shows that any set $X \subset \mathbf{R}^d$ with Hausdorff dimension exceeding one must contain three colinear points. On the other hand, [10] 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. 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 determine for which classes of patterns our intuition remains true.

Our goal in this thesis is to derive new methods for constructing sets avoiding patterns. And 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.

The main contribution we give to pattern dissection methods is using random dissections to avoid patterns without much structual information. Thus we can expand the utility of interval dissection methods from regular patterns 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

2.1 Configuration Avoidance

This thesis discusses methods to form large sets avoiding patterns. First, we introduce notation in a general context. We consider an ambient set $\mathbf{A}$. 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 general *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 $\mathscr{C}$. We say $\mathscr{C}$ is an *n point configuration* if it is a subset of $\mathscr{C}^n(X)$.

Example (Isoceles Triangle Configuration)**.** *Set*

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

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.

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*

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

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 $\mathbf{A}$, what is the maximal size of a set $X \subset \mathbf{A}$ avoiding $\mathscr{C}$? Depending on the structure of the ambient space $\mathbf{A}$ and the configuration $\mathscr{C}$, there are various ways of measuring the size of the set X:

- If $\mathbf{A}$ is finite, the goal is to find a set X with large cardinality.
- If $\{\mathbf{A}_n\}$ is an increasing family of finite sets with $\mathbf{A} = \lim \mathbf{A}_n$, the goal is to find a set X such that $X \cap \mathbf{A}_n$ has large cardinality asymptotically in n.
- If $\mathbf{A} = \mathbf{R}^d$, but $\mathscr{C}$ is sufficiently discrete, then a satisfactory goal is to find a set X with large Lebesgue measure avoiding $\mathscr{C}$.

In this thesis, inspired by results in these three settings, we establish methods for avoiding non-discrete configurations $\mathscr{C}$ in $\mathbf{R}^d$. Here, Lebesgue measure completely fails to measure the size of pattern avoiding solutions, as the next theorem shows, under the often true assumption that the configuration is *translation invariant*, i.e. that if $(a_1, \dots, a_n) \in \mathscr{C}$ and $b \in \mathbf{R}^d$, $(a_1 + b, \dots, a_n + b) \in \mathscr{C}$.

Theorem 1. *Let $\mathscr{C}$ be a translation-invariant n-point configuration on $\mathbf{R}^d$. Suppose*

(A) *For any $\varepsilon > 0$, there is $(a_1, \dots, a_n) \in \mathscr{C}$ with $\operatorname{diam}\{a_1, \dots, a_n\} \leq \varepsilon$.*

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

Proof. Let $X \subset \mathbf{R}^d$ have positive Lebesgue measure. For a point $x \in \mathbf{R}^d$, we let

$$I(x,r) = [x_1 - r/2, x_1 + r/2] \times \cdots \times [x_n - r/2, x_n + r/2].$$

denote the coordinate-axis oriented cube centered at x with sidelength r. The Lebesgue density theorem shows that there exists a point $x \in X$ such that

$$\lim_{r \to 0} \frac{|X \cap I(x,r)|}{|I(x,r)|} = 1.$$

We fix $\varepsilon > 0$, to be chosen later, and pick r_0 small enough such that

$$|X \cap I(x,r_0)| \geq (1-\varepsilon)|I(x,r_0)|.$$

Applying Property (A), we find $C = (a_1, \dots, a_n) \in \mathscr{C}$ with $\operatorname{diam}\{a_1, \dots, a_n\} \leq \varepsilon r_0$. For each $p \in I(x,r_0)$, let $C(p) = (a_1(p), \dots, a_n(p))$, where $a_i(p) = p + (a_i - a_1)$. By translation invariance, $C(p) \in \mathscr{C}$ for each $p \in \mathbf{R}^d$. A union bound gives

$$\begin{aligned}
|\{p \in I(x,r_0/2) : C(p) \notin \mathscr{C}(X)\}| &= \left| \bigcup_{i=1}^n \{p \in I(x,r_0/2) : a_i(p) \notin X\} \right| \\
&\leq \sum_{i=1}^n |I(x,r_0/2) \cap (X + (a_1 - a_i))^c| \\
&= \sum_{i=1}^n |I(x,r_0/2)| - |I(x + a_i - a_1, r_0/2) \cap X| \\
&\leq n\big[|I(x,r_0/2)| - |I(x,r_0(1/2-\varepsilon)) \cap X|\big] \\
&\leq n r_0^d \left[1/2^d - (1/2-\varepsilon)^d\right] \\
&\leq \varepsilon \cdot n(2d+1) \cdot (r_0/2)^d.
\end{aligned}$$

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

$$\begin{aligned}
|\{p \in I(x,r_0/2) : C(p) \in \mathscr{C}(X)\}| &= |I(x,r_0/2)| - |\{p \in I(x,r_0/2) : C(p) \notin \mathscr{C}(X)\}| \\
&\geq \varepsilon \cdot n(2d+1)(r_0/2)^d - (r_0/2)^d > 0.
\end{aligned}$$

Thus there is $p \in I(x, r_0/2)$ such that $C(p) \in \mathscr{C}(X)$, so X does not avoid $\mathscr{C}$. $\square$

Since no set of positive Lebesgue measure can avoid non-discrete configurations, we cannot use the Lebesgue measure to quantify the size of pattern avoiding sets. Fortunately, there is a quantity which can distinguish between the size of sets of measure zero. This is the *fractional dimension* of a set, introduced in the next section.

2.2 Fractal Dimension

Intuitively, fractional dimension provides a measure of the local density of a set, and so we view a set which is more dense as 'larger'. There are two main definitions of fractional dimension we use in this thesis: Minkowski dimension and Hausdorff dimension. The main difference between the first two is that Minkowski dimension measures relative density at a single scale, whereas Hausdorff dimension measures relative density at various scales. In this section, we indicate the subtle differences between the two dimensional quantities which are relavent to this thesis.

We begin by discussing the Minkowski dimension, which is the easiest of the two dimensions to define. Given a length l, and a bounded set $E \subset \mathbf{R}^d$, we let $N(l,E)$ denote the length l covering number of E, i.e. the minimum number of sidelength l cubes required to cover E. We define the *lower* and *upper* Minkowski dimension as

$$\underline{\dim}_{\mathbf{M}}(E) = \liminf_{l \to 0} \left(\frac{\log(N(l,E))}{\log(1/l)} \right) \quad \overline{\dim}_{\mathbf{M}}(E) = \limsup_{l \to 0} \left(\frac{\log(N(l,E))}{\log(1/l)} \right).$$

If $\overline{\dim}_{\mathbf{M}}(E) = \underline{\dim}_{\mathbf{M}}(E)$, then we refer to this common quantity as the *Minkowski dimension* of E, denoted $\dim_M(E)$. Intuitively, $\underline{\dim}_{\mathbf{M}}(E) < s$ if there *exists* a sequence of lengths $\{l_k\}$ converging to zero with $N(l_k,E) \leq (1/l_k)^s$, and $\overline{\dim}_{\mathbf{M}}(E) < s$ if $N(l,E) \leq (1/l)^s$ for *all* sufficiently small lengths l.

Later on, it will be helpful to further discretize the covering number of sets so that all the cubes in the cover lie on a grids. We let

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

Given any $E \subset \mathbf{R}^d$, we let

$$\mathscr{B}(l,E) = \left\{ I \in \mathscr{B}(l,\mathbf{R}^d) : I \cap E \neq \emptyset \right\}.$$

Up to a constant factor, $\mathscr{B}(l,E)$ is an optimal covering of E by cubes.

Lemma 1. *For any bounded set E,*

$$N(l,E) \leq \#(\mathscr{B}(l,E)) \leq 2^d N(l,E)$$

Proof. Since $\mathscr{B}(l,E)$ is a cover of E, and $N(l,E)$ is the size of a minimal cover, $N(l,E) \leq \#(\mathscr{B}(l,E))$. But if I is an arbitrary sidelength l cube, then it intersects at most 2^d cubes in $\mathscr{B}(l,E)$, so $\#\mathscr{B}(l,E) \leq 2^d N(l,E)$. □

It is also useful to work at a discrete sequence of 'dyadic' scales. If we fix $M > 1$, and consider $1/M^{k+1} \leq l \leq 1/M^k$, then as $k \to \infty$, $\log(1/l) \sim \log(M^k)$. To determine the Minkowski dimension, it suffices to look at the scales $\{1/M^k\}$, which we can combine with the discretization of cubes.

Corollary. *For any set E,*

$$\underline{\dim}_{\mathbf{M}}(E) = \liminf_{k\to\infty} \frac{\log_M(\#\mathscr{B}(1/M^k,E))}{k} \quad \overline{\dim}_{\mathbf{M}}(E) = \limsup_{k\to\infty} \frac{\log_M(\#\mathscr{B}(1/M^k,E))}{k}.$$

Proof. Suppose $1/M^{k+1} \leq l \leq 1/M^k$. Then the upper bound of Lemma 1 implies

$$\frac{\log(N(l,E))}{\log(1/l)} \sim \frac{\log_M(N(l,E))}{k+1} \leq \frac{\log_M(N(1/M^{k+1},E))}{k+1} \leq \frac{\log_M(\#\mathscr{B}(1/M^{k+1},E))}{k+1}.$$

Conversely, the lower bound of Lemma 1 implies that

$$\begin{aligned}\frac{\log(N(l,E))}{\log(1/l)} \sim \frac{\log_M(N(l,E))}{k} &\geq \frac{\log_M(N(1/M^k,E))}{k} \\ &\geq \frac{\log_M(2^d \# \mathscr{B}(1/M^k,E))}{k} = \frac{\log_M(\# \mathscr{B}(1/M^k,E))}{k} + o(1).\end{aligned}$$

Taking $l \to 0$, and correspondingly, $k \to \infty$, completes the proof. □

Hausdorff dimension is a version of fractal dimension which is more stable under analytical operations, but is less easy to calculate. It is obtained by finding a canonical 's dimensional measure' H^s on $\mathbf{R}^d$ for $s \in [0,\infty)$, which for $s = 1$, measures the 'length' of a set, for $s = 2$, measure the 'area', and so on and so forth. We then set the dimension of E to be the supremum of the values s such that $H^s(E) < \infty$. Given a cube I, we let $l(I)$ denote it's sidelength. For $E \subset \mathbf{R}^d$, we define the *Hausdorff content*

$$H^s_\delta(E) = \inf \left\{ \sum_{k=1}^\infty l(I_k)^s : E \subset \bigcup_{k=1}^\infty I_k, l(I_k) \leq \delta \right\}.$$

We then define $H^s(E) = \lim_{\delta \to 0} H^s_\delta(E)$. It is easy to see H^s is an exterior measure, and $H^s(E \cup F) = H^s(E) + H^s(F)$ if the Hausdorff distance $d(E,F)$ between E and F is positive. So H^s is actually a metric exterior measure, and the Caratheodory extension theorem shows all Borel sets are measurable with respect to H^s.

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

Proof. If, for any cover of E by intervals I_k, $\sum l(I_k)^t \leq A$, and $l(I_k) \leq \delta$, then

$$\sum l(I_k)^s \leq \sum l(I_k)^{s-t} l(I_k)^t \leq \delta^{s-t} A.$$

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. □

Example. *We note $H^1[-N,N]^d = (2N)^d$. If $s > d$, then Lemma 2 shows $H^s[-N,N]^d = 0$, and so by countable additivity, $H^s(\mathbf{R}^d) = 0$. Thus H^s is trivial if $s > d$.*

Given any Borel set E, the last example, combined with Lemma 2, implies there is a unique value $s_0 \in [0,d]$ such that $H^s(E) = 0$ for $s > s_0$, and $H^s(E) = \infty$ for $0 \leq s < s_0$, though it is possible for $H^{s_0}(E)$ to take any value in $[0,\infty]$. We refer to s_0 as the *Hausdorff dimension* of E, denoted $\dim_{\mathbf{H}}(E)$.

Theorem 2. *For any bounded set E, $\dim_{\mathbf{H}}(E) \leq \underline{\dim}_{\mathbf{M}}(E) \leq \overline{\dim}_{\mathbf{M}}(E)$.*

Proof. We consider the simple bound $H^s_l(E) \leq \#\mathscr{B}(l,E) \cdot l^s$. Taking $l \to 0$, we conclude that $H^s(E) \leq \liminf_{l \to 0} \#\mathscr{B}(l,E) \cdot l^s$. Fix $\varepsilon > 0$. If $s \geq \underline{\dim}_{\mathbf{M}}(E) + 2\varepsilon$, consider a sequence of lengths $\{l_k\}$ converging to zero with

$$\frac{\log(\#\mathscr{B}(l_k,E))}{\log(1/l_k)} \leq \underline{\dim}_{\mathbf{M}} + \varepsilon \leq s - \varepsilon.$$

Thus $\#\mathscr{B}(l_k,E) \leq (1/l_k)^{s-\varepsilon}$, so $H^s(E) \leq \lim \#\mathscr{B}(l_k,E) l_k^s = \lim l_k^\varepsilon = 0$. Taking $\varepsilon \to 0$ proves the claim. $\square$

A simple intuition behind the two approaches to fractal dimension we have considered here 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 a simultaneous set of infinitely many scales, which explains both why in certain cases the Hausdorff dimension is smaller than the Minkowski dimension, and also why the Hausdorff dimension is more stable under analytical operations. For instance, given any sequence $\{E_k\}$,

$$\dim_{\mathbf{H}}\Big(\bigcup_{k=1}^{\infty} E_k\Big) = \sup \dim_{\mathbf{H}}(E_k).$$

This need not be true for the Minkowski dimension; a single point has Minkowski dimension zero, but the rational numbers, which are a countable union of points, have Minkowski dimension one.

Just like with Minkowski dimension, we can also discretize Hausdorff dimension to cubes lying on a grid, and also lying on a set of dyadic scales.

Lemma 3. *For a fixed M, let*

$$H_N^{s,M}(E) = \inf \left\{ \sum_{k=1}^{\infty} 1/M^{si_k} : E \subset \bigcup_{k=1}^{\infty} I_k,\ I_k \in \mathscr{B}(1/M^{i_k}, \mathbf{R}^d),\ \textit{and } i_k \geq N \right\},$$

and let $H^{s,M}(E) = \lim_{N \to \infty} H_N^{s,M}(E)$. Then

$$H^s(E) \leq H^{s,M}(E) \leq 3^d M^s H^s(E),$$

for any Borel set E.

Proof. Let $E \subset \bigcup_{k=1}^{\infty} I_k$, where $l(I_k) \leq 1/M^N$ for each k. If $1/M^{i_k}$ is the largest power of i_k such that $1/M^{i_k} \geq l(I_k)$, then $i_k \geq N$, and $\#\mathscr{B}(1/M^{i_k}, I_k) \leq 3^d$. If we now consider the infinite family of cubes $\bigcup_{k=1}^{\infty} \mathscr{B}(1/M^{i_k}, I_k)$, then

$$H_N^{s,M}(E) \leq \sum_{k=1}^{\infty} \#\mathscr{B}(1/M^{i_k}, I_k) \cdot 1/M^{si_k} \leq 3^d M^s \sum_{k=1}^{\infty} l(I_k)^s$$

Since the cover was arbitrary, we conclude that $H_N^{s,M}(E) \leq 3^d M^s H^s_{1/M^N}(E)$, and taking $N \to \infty$ gives $H^{s,M}(E) \leq 3^d M^s H^s(E)$. The lower bound $H^{s,M}(E) \geq H^s(E)$ is trivial. □

It is often easy to upper bound Hausdorff dimension, but non-trivial to *lower bound* the Hausdorff dimension of a given set. A key technique to finding a lower bound is *Frostman's lemma*, which says that a set has large Hausdorff dimension if and only if it supports a probability measure which is suitably sparse. We say a measure μ is a *Frostman measure* of dimension s if it is non-zero, compactly supported, and for any length l, if I is a cube with sidelength l, then $\mu(I) \lesssim l^s$. The proof of Frostman's lemma will indicate an important technique, known as the *mass distribution principle*.

Lemma 4 (Mass Distribution Principle)**.** *Let $\{l_k\}$ be a decreasing sequence of lengths, and let $\mathscr{B} = \bigcup_{k=1}^{\infty} \mathscr{B}(l_k, \mathbf{R}^d)$. Suppose μ is a function from $\mathscr{B}$ to $[0,\infty)$ such that for each k, and for any $I \in \mathscr{B}(l_k, \mathbf{R}^d)$,*

$$\sum \{\mu(J) : J \in \mathscr{B}(l_{k+1}, I)\} = \mu(I)$$

Then μ extends uniquely to a regular Borel measure on $\mathbf{R}^d$.

Proof. We can define a sequence of regular Borel measures μ_k by setting for each $f \in C_c(\mathbf{R}^d)$,

$$\int f d\mu_k = \sum \left\{ \mu(I) \int_I f\, dx : I \in \mathscr{B}(1/M^k, \mathbf{R}^d) \right\}$$

Let E_k be the operator on regular Borel measures given by the formula

$$\int f(x) dE_k(\nu) = \sum \left\{ \nu(I) \int_I f(x) : I \in \mathscr{B}(l_k, \mathbf{R}^d) \right\}.$$

The main condition of the theorem then says that $E_j(\mu_k) = \mu_j$ if $j \leq k$. Note that the operators E_k are each continuous with respect to the weak topology; If $\nu_i \to \nu$ weakly, then $\nu_i(I) \to \nu(I)$ for each fixed $I \in \mathscr{B}(l_k, \mathbf{R}^d)$. Thus if $f \in C_c(\mathbf{R}^d)$, then the support of f intersects only finitely many intervals in $\mathscr{B}(l_k, \mathbf{R}^d)$, so

$$\int f(x) dE_k(\nu_i) = \sum \nu_i(I) \int_I f\, dx \to \sum \nu(I) \int_I f\, dx = \int f dE_k(\nu).$$

Now fix an interval $I \in \mathscr{B}(l_0, \mathbf{R}^d)$. The measures $\{\mu_k|_I\}$, restricted to I, are uniformly bounded by $\mu(I)$. Thus the Banach-Alaoglu theorem implies that is a subsequence μ_{k_i} converging weakly on I to some measure μ_I. By continuity,

$$E_j(\mu_I) = \lim E_j(\mu_{k_i}|_I) = \lim E_j(\mu_{k_i})|_I = \mu_j|_I.$$

This means precisely that μ_I is the extension of μ to a Borel measure on I. If we patch together the measures μ_I over all I, we obtain a measure extending μ on all of $\mathbf{R}^d$. The uniqueness of the extension is guaranteed by the fact that the intervals upon which μ are defined generate the entire Borel sigma algebra. $\square$

In the next lemma, it will help to notice that the definition of the Frostman measure is also robust to working over dyadic-type lengths. Suppose we can establish a result $\mu(I) \lesssim 1/M^{ks}$ for all $I \in \mathscr{B}(1/M^k, \mathbf{R}^d)$ and all indices k. Given any length l, there is a value of k such that $1/M^{k+1} \leq l \leq 1/M^k$. For any $I \in \mathscr{B}(l, \mathbf{R}^d)$, $\#\mathscr{B}(1/M^k, I) \leq 3^d$, so

$$\mu(I) \leq \mu\left(\bigcup \mathscr{B}(1/M^k, I)\right) \lesssim 3^d/M^{ks} \leq 3^d (l/M)^s \lesssim l^s.$$

Thus μ is a Frostman measure.

Lemma 5 (Frostman's Lemma)**.** *If E is Borel, $H^s(E) > 0$ if and only if there exists an s dimensional Frostman measure supported on E.*

Proof. Suppose that μ is s dimensional and supported on E. If $H^s(F) = 0$, then for $\varepsilon > 0$ there is a sequence of cubes $\{I_k\}$ with $\sum_{k=1}^\infty l(I_k)^s \leq \varepsilon$. But then

$$\mu(F) \leq \mu\left(\bigcup_{k=1}^\infty I_k\right) \leq \sum_{k=1}^\infty \mu(I_k) \lesssim \sum_{k=1}^\infty l(I_k)^s \leq \varepsilon.$$

Taking $\varepsilon \to 0$, we conclude $\mu(F) = 0$, so μ is absolutely continuous with respect to H^s. But since $\mu(E) > 0$, this means that $H^s(E) > 0$.

Conversely, suppose $H^s(E) > 0$. Then by translating, we may assume that $H^s(E \cap [0,1)^d) > 0$, and so without loss of generality we may assume that $E \subset [0,1)^d$. Fix M, and for each $I \in \mathscr{B}(1/M^k, \mathbf{R}^d)$, define

$$\mu^+(I) = H^{s,M}_{1/M^k}(E \cap I)$$

Then $\mu^+(I) \leq 1/M^{ks}$, and μ^+ is subadditive. We use it to recursively define a measure μ to which we can apply Lemma 4, such that $\mu(I) \leq \mu^+(I)$ for each $I \in \mathscr{B}(1/M^k)$. We initially define μ by setting $\mu([0,1)^d) = \mu^+([0,1)^d)$. Given

$I \in \mathscr{B}(1/M^k, [0,1)^d)$, we enumerate all the children $J_1, \dots, J_M \in \mathscr{B}(1/M^{k+1}, I)$. We then consider any values of $A_1, \dots, A_M$ such that

$$A_1 + \cdots + A_M = \mu(I) \quad \text{and} \quad A_i \leq \mu^+(J_i)$$

This is feasible to do because $\sum_{i=1}^M \mu^+(J_i) \geq \mu^+(I) \geq \mu(I)$. We then define $\mu(J_i) = A_i$. The recursive constraint is satisfied, so μ is well defined. The mass distribution principle then implies that μ extends to a full measure, which satisfies $\mu(I) \leq \mu^+(I) \leq 1/M^{ks}$ for each $I \in \mathscr{B}(1/M^k, \mathbf{R}^d)$, so it is a Frostman measure with dimension s. $\square$

2.3 Dimensions of Cantor-Type Sets

A *Cantor-type* decomposition of a set E is a decreasing sequence of sets $\{E_k\}$, together with a decreasing sequence of lengths $\{l_k\}$, such that each E_k is a finite union of cubes in $\mathscr{B}(l_k, \mathbf{R}^d)$, and $E = \lim_{k \to \infty} E_k$. For simplicity, we assume $E \subset [0,1]^d$, and $l_0 = 1$. If a sequence of lengths $\{l_k\}$ is fixed, then given a cube $I \in \mathscr{B}(l_{k+1}, \mathbf{R}^d)$, we let $I^* \in \mathscr{B}(l_k, \mathbf{R}^k)$ denote the unique *parent cube* of I, i.e. the unique cube with $I \subset I^*$.

In this thesis, Cantor-type decompositions naturally arise from the constructions of sets we produce. One advantage of the decomposition is the ability to construct Frostman measures. For instance, a natural choice is obtained by recursively defining a function μ by setting $\mu[0,1]^d = 1$, and setting, for each $I \in \mathscr{B}(l_{k+1}, \mathbf{R}^d)$,

$$\mu(I) = \frac{\mu(I^*)}{\#\mathscr{B}(l_k, E_k \cap I^*)}$$

if $I \in \mathscr{B}(l_{k+1}, E_k)$, and $\mu(I) = 0$ otherwise. The mass distribution principle then extends μ to a probability measure. We will refer to this as the *canonical construction*. Simple combinatorial reasoning related to the construction of the set E often enables one to establish bounds of the form $\mu(I) \lesssim l_k^s$ for $I \in \mathscr{B}(l_k, \mathbf{R}^k)$. If the lengths $\{l_k\}$ are dyadic, we have already shown this is sufficient to show μ is a Frostman measure of dimension s. But if we weaken the condition to being a Frostman measure of dimension $s-\varepsilon$ for all $\varepsilon > 0$, we can allow the lengths to decrease at a much faster rate, even up to where the lengths are *hyperdyadic*.

Theorem 3. *Let $l_k = 2^{-a_k}$, and suppose $a_{k+1} - a_k = o(a_k)$. If μ is a Borel measure such that $\mu(I) \lesssim l_k^s$ for each $I \in \mathscr{B}(l_k, \mathbf{R}^d)$, then μ is a Frostman measure of dimension $s-\varepsilon$ for each $\varepsilon > 0$.*

Proof. The idea of the argument, which has been implicit in most of our other discretization approaches, is a covering argument. Given a cube I with sidelength l, find k such that $l_{k+1} \leq l \leq l_k$. Then we can cover I by $O(1)$ cubes in $\mathscr{B}(l_k, \mathbf{R}^d)$, which shows

$$\mu(I) \lesssim l_k^s = [(l_k/l)^s l^\varepsilon] l^{s-\varepsilon}.$$

The proof is completed by noticing

$$(l_k/l)^s l^\varepsilon \leq (l_k/l_{k+1})^s l^\varepsilon \leq \exp(s(a_{k+1} - a_k) - \varepsilon a_k) \lesssim_\varepsilon 1,$$

so $\mu(I) \lesssim_\varepsilon l^{s-\varepsilon}$. $\square$

This result applies, for instance, in the case where $l_k = 2^{-k^n}$ for some fixed integer n, since then $(k+1)^n - k^n = O(k^{n-1}) = o(k^n)$. We can even get 'just' over polynomial powers of n. Let $l_k = 1/2^{2^{\psi(k)k}}$, where ψ is a decreasing function of k, which tends to zero as $k \to \infty$, but for which $\psi(k)k$ is an increasing function which tends to ∞. Such a sequence is often called *hyperdyadic*. It is simple to calculate that

$$2^{\psi(k+1)(k+1)} - 2^{\psi(k)k} \leq 2^{\psi(k)k}(2^{\psi(k)} - 1) = o(2^{\psi(k)k}).$$

thus it suffices to analyze Frostman measures at a hyperdyadic set of lengths. But it is of course easy to construct superexponential examples of l_k where we cannot use this technique. And the next example shows we shouldn't be too hopeful in this scenario.

Example. *Let $l_k = 1/2^{a_k}$, where $a_{k+1} \geq 2a_k$ for each k. Form a set $E = \lim_{k\to\infty} E_k$, where $E_0 = [0,1]$, each E_k is a union of sidelength l_k^2 cubes, and E_{k+1} is obtained* *by selecting a single cube in $\mathscr{B}(l_{k+1}^2, I)$ from each element $I \in \mathscr{B}(l_{k+1}, E_k)$. For each k, E is covered by at most $1/l_k$ sidelength l_k^2 cubes, so E has lower Minkowski dimension at most $1/2$, and thus Hausdorff dimension at most $1/2$. Nonetheless, if we define the canonical measure μ from the Cantor-type decomposition $\{E_k\}$, then one verifies that μ assigns the same amount of mass to each cube in $\mathscr{B}(l_k, E)$, and*

$$\#\mathscr{B}(l_k, E) = \#\mathscr{B}(l_{k-1}, E)(l_{k-1}^2/l_k).$$

Thus as $k \to \infty$,

$$\#\mathscr{B}(l_k, E) = \frac{l_{k-1}^2 \dots l_1^2}{l_k \dots l_1} = \frac{l_{k-1} \dots l_1}{l_k}.$$

Provided $a_1 + \dots + a_{k-1} = o(a_k)$, this implies that $\#\mathscr{B}(l_k, E) \gtrsim l_k^{1-o(1)}$, so if $I \in \mathscr{B}(l_k, E)$, $\mu(I) \lesssim l_k^{1-o(1)}$. Thus at the scales $\{l_k\}$, μ looks like a Frostman measure of full dimension, which cannot be the case. This example works if $a_k = 2^{k^{1+\varepsilon}}$, for which the lengths $\{l_k\}$ are just *past the point of being hyperdyadic.*

Thus, if we are to use a faster decreasing sequence of lengths $\{l_k\}$ than hyperdyadic, we must have to exploit some extra property of the decomposition which is not always present. Here, we rely on some kind of *uniformity* between scales. Given the uniformity assumption, the lengths can decrease as fast as desired.

Theorem 4. *Let μ be a measure supported on a set E, and $\{l_k\}$ and $\{r_k\}$ two decreasing sequences of lengths, with $l_{k+1} \leq r_{k+1} \leq l_k$ for each k. Write $l_k = 2^{-a_k}$ and $r_k = 2^{-b_k}$. Suppose that*

(A) (Discrete Bound)*: For any $I \in \mathscr{B}(l_k, \mathbf{R}^d)$, $\mu(I) \lesssim l_k^s$.*

(B) (Hyperdyadic Scale Change)*: $b_k - a_k = o(a_k)$.*

(C) (Uniform Mass Distribution)*: For any $I \in \mathscr{B}(l_k, \mathbf{R}^d)$, and $J \in \mathscr{B}(r_{k+1}, I)$,*

$$\mu(J) \lesssim (r_{k+1}/l_k)^d \mu(I).$$

Then μ is a Frostman measure of dimension $s - \varepsilon$ for each $\varepsilon > 0$.

Proof. Suppose an interval I has length l, and we can find k with $r_{k+1} \leq l \leq l_k$. Then we can cover I by at most $(l/r_{k+1})^d$ cubes in $\mathscr{B}(r_{k+1}, \mathbf{R}^d)$. By Properties (A) and (C), each of these cubes has measure at most $(r_{k+1}/l_k)^d l_k^s$, so we obtain that for any $\varepsilon > 0$,

$$\mu(I) \lesssim (l/r_{k+1})^d (r_{k+1}/l_k)^d l_k^s = l^d / l_k^{d-s} \lesssim l^s.$$

In particular, this implies $\mu(I) \lesssim r_k^s$ if $I \in \mathscr{B}(r_k, \mathbf{R}^d)$. On the other hand, suppose there exists k with $l_k \leq l \leq r_k$. Then we can cover I by $O(1)$ cubes in $\mathscr{B}(r_k, \mathbf{R}^d)$, so we find

$$\mu(I) \lesssim r_k^s \leq [(r_k/l_k)^s l^\varepsilon] l^{s-\varepsilon}.$$

Property (B) shows $(r_k/l_k)^s l^\varepsilon \lesssim_\varepsilon 1$, so $\mu(I) \lesssim_\varepsilon l^{s-\varepsilon}$ for each $\varepsilon > 0$. This covers all possible cases, so μ is a Frostman measure of dimension $s - \varepsilon$ for each $\varepsilon > 0$. $\square$

Chapter 3

Related Work

3.1 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. If this is true, we say that X *avoids translates*. 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 6. *Let X be a set. Then 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_k a union of disjoint intervals in $\mathscr{B}(l_k,\mathbf{R}^d)$, for a decreasing sequence of lengths $\{l_k\}$ converging to zero, to be chosen later, but with $10l_{k+1}\mid l_k$. We initialize $X_0=[0,1]$, and $l_0=1$. Furthermore, we consider a queue of intervals, initially just containining $[0,1]$. To construct the sequence $\{X_k\}$, Keleti iteratively performs the following procedure:

Algorithm 1 Construction of the Sets X_N

Set $k=0$
Repeat
  Take off an interval I from the front of the queue
  For all $J\in\mathscr{B}(l_k,\mathbf{R}^d)$ contained in X_k:
    Order the intervals in $\mathscr{B}(l_{k+1},\mathbf{R}^d)$ contained in J as $J_0,J_1,\dots,J_N$
    If $J\subset I$, add all intervals J_i to X_{k+1} with $i\equiv 0$ modulo 10
    Else add all J_i with $i\equiv 5$ modulo 10
  Add all intervals in $\mathscr{B}(l_{k+1})$ to the end of the queue
  Increase k by 1

Each iteration of the algorithm produces a new set X_k, and so leaving the algorithm to repeat infinitely produces a sequence $\{X_k\}$ converging to a set X, which is translate avoiding.

Lemma 7. *The set X is translate avoiding.*

Proof. 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 $l_k \to 0$, there is a suitably large integer N such that x_1 is contained in an interval $I \in \mathscr{B}(l_N, \mathbf{R}^d)$ not containing x_2, x_3, or x_4. 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°, x_2°, $x_3^\circ, x_4^\circ \in l_M \mathbf{Z}$ to the intervals in $\mathscr{B}(l_M, \mathbf{R}^d)$ 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

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

which gives a contradiction. $\square$

The algorithm shows that

$$\mathscr{B}(l_k, X_k) = (l_{k-1}/10 l_k)\, \mathscr{B}(l_{k-1}, X_{k-1}).$$

Thus closing the recursive definition shows

$$\mathscr{B}(l_k, X_k) = \frac{1}{10^k l_k}.$$

In particular, this means $|X_k| = 1/10^k$, so X has measure zero irrespective of our parameters. Nonetheless, the canonical measure μ on X defined with respect to the decomposition $\{X_k\}$ satisfies $\mu(I) = 10^k l_k$ for all $I \in \mathscr{B}(l_k, X)$. If $10^k l_k^\varepsilon \lesssim_\varepsilon 1$ for all ε, then we can establish the bounds $\mu(I) \lesssim_\varepsilon l_k^{1-\varepsilon}$ for all ε. In particular this is true if we set $l_k = 1/10^{\lfloor k \log k \rfloor}$. And because this sequence is not too fast increasing, we can apply Theorem 3 to show μ is a Frostman measure of dimension $1 - \varepsilon$ for each $\varepsilon > 0$, so X has full Hausdorff dimension.

3.2 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 5. *Let $f : \mathbf{R}^{d+1} \to \mathbf{R}$ be a C^1 function, and consider sets $T_0, \dots, T_d \subset [0,1]$, with each T_n a union of intervals in $\mathscr{B}(1/M, [0,1])$, and such that $\partial_0 f$ is non-vanishing on $T_0 \times \cdots \times T_d$. Then there exists a rational number C, and arbitrarily large integers $N \in M\mathbf{Z}$ for which there are $S_k \subset T_k$, each a union of cubes in $\mathscr{B}(C/N^d, T_k)$ with*

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

(B) *If $k \neq 0$, S_k contains an interval in $\mathscr{B}(C/N^d, I)$ for each $I \in \mathscr{B}(1/N, T_k)$.*

(C) *For at least a fraction $1 - 1/M$ of the cubes $I \in \mathscr{B}(1/N, T_0)$, $|S_0 \cap I| \geq C/N$.*

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 $d(S_0,\mathbf{B}) > C_0C_1/N^d$, 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^3N^{d-1} length $1/4M^3N^d$ intervals, which in total has length at least $1/4N$. □

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 8. *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 $d(y,\mathbf{B}) \leq C_0C_1/N^{d-1}$.*

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 $1/(d-1)$.

3.3 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 9. *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 $\square$

Chapter 4

Avoiding Rough Sets

A major question in modern geometric measure theory is whether sets with sufficiently large Hausdorff dimension necessarily contain copies of certain patterns. For example, Theorem 6.8 of [12] shows that each set $X \subset \mathbf{R}^d$ with Hausdorff dimension exceeding one must contain three colinear points. On the other hand, the main result of [10] constructs a set $X \subset \mathbf{R}^2$ with full Hausdorff dimension such that no four points in X form the vertices of a parallelogram. At the time of this writing, many fundamental questions about the relation between the geometric patterns of a set and it's fractional dimension remain open. The threshold dimension above which patterns are guaranteed can vary for different patterns, or not even exist at all. In this paper, we provide lower bounds for the threshold by solving the pattern avoidance problem; constructing sets with large dimension avoiding patterns.

One natural way to approach the pattern avoidance problem is to find general methods for constructing pattern avoiding sets by exploiting a particular geometric feature of the pattern. In [11], Máthé shows that, given a pattern specified by a countable union of rational coefficient bounded degree polynomials $f_1, f_2, \dots$ in nd variables, one can construct a set $X \subset \mathbf{R}^d$ such that for every collection of distinct points $x_1, \dots, x_n$, and each index k, $f_k(x_1, \dots, x_n) \neq 0$. In [6], Fraser and the second author consider a related problem where the polynomials f_k are replaced with full rank C^1 functions.

In this paper, we also find general methods for constructing pattern avoid sets. But rather than avoiding the zeroes of a function, we fix $Z \subset \mathbf{R}^{dn}$, and construct sets X such that for any distinct $x_1, \dots, x_n \in X$, $(x_1, \dots, x_n) \notin Z$. We then say X avoids the pattern specified by Z. For instance, if $Z = \{(x_1, x_2, x_3) \in (\mathbf{R}^d)^3 : (x_1, x_2, x_3)$ are colinear$\}$, then $X \subset \mathbf{R}^d$ avoids the pattern specified by Z precisely when X does not contain three colinear points. Our problem generalizes the problem statements considered in [11] and [6] if we set $Z = \bigcup_{k=1}^{\infty} f_k^{-1}(0)$. From this perspective, Máthé constructs sets avoiding a set Z formed from a countable union of algebraic varieties, while Fraser and the second author construct sets avoiding a set Z formed from the countable union of C^1 manifolds.

The advantage of the formulation of pattern avoidance we consider is that it is now natural to consider 'rough' sets $Z \subset \mathbf{R}^{dn}$ which are not naturally specified as the zero set of a function. In particular, in this paper we consider Z formed from the countable union of sets, each with lower Minkowski dimension bounded above by some α. In contrast with previous work, no further assumptions are made on Z. The dimension of the avoiding set X we will eventually construct depends only on the codimension $nd - \alpha$ of the set Z, and the number of variables n.

Theorem 6. *Let $Z \subset \mathbf{R}^{nd}$ be a countable union of compact sets, each with lower Minkowski dimension at most α, with $d \leq \alpha < dn$. Then there exists a set $X \subset [0,1)^d$ with Hausdorff dimension at least $(nd - \alpha)/(n-1)$ such that whenever $x_1, \dots, x_n \in X$ are distinct, then $(x_1, \dots, x_n) \notin Z$.*

Remark. *When $\alpha < d$, the pattern avoidance problem is trivial, since $X = [0,1)^d - \pi(Z)$ is full dimensional and solves the pattern avoidance problem, where $\pi(x_1, \dots, x_n) = x_1$ is a projection map from $\mathbf{R}^{dn}$ to $\mathbf{R}^d$. The case $\alpha = dn$ is trivial as well, since we can set $X = \emptyset$.*

When Z is a countable union of smooth manifolds, Theorem 6 generalizes Theorem 1.1 and 1.2 from [6]. Section 4.6 is devoted to a comparison of our methods with [6], as well as other general pattern avoidance methods. But our result is more interesting when it can be applied to truly ‘rough’ sets. One surprising application to our method is obtained by considering a ‘rough’ set Y in addition to a set Z, and finding a set $X \subset Y$ of large dimension avoiding Z. Previous methods in the literature fundamentally exploit the ability to select points on the entirety of Euclidean space, and so it remains unlikely that we can obtain any result of this form using their methods, unless Y is suitably flat, i.e. a smooth surface. To contrast this, our result even applies for certain sets Y which are of Cantor type. We discuss applications of our result in Section 5.

Theorem 6 is proved using a Cantor-type construction, a common theme in the surrounding literature, described explicitly in Section 4.3. For a sequence of decreasing lengths $\{l_n\}$ with $l_n \to 0$, the construction specifies a selection mechanism for a nested family of sets $X_n \to X$, with X_n a union of sidelength l_n cubes and avoiding Z at scales close to l_n. Our *key* contribution to this process is using the probabilistic method to guarantee the existence of efficient selections at each scale, described in Section 2. This selection also assigns mass uniformly on intermediate scales, akin to [6], ensuring that the selection procedure at a single scale is the sole reason for the Hausdorff dimension we calculate in Section 4. Furthermore, the randomization allows us to avoid the complicated queueing techniques in [9] and [6], which has the additional benefit that we can confront the entire behaviour of Z at scales close to l_n simultaneously during the nth step of the construction.

4.1 Frequently Used Notation and Terminology

Our argument heavily depends upon discretizing sets into unions of cubes. Throughout our argument, we use the following notation:

(A) A *dyadic length* is a length l equal to 2^{-k} for some non-negative integer k.

(B) Given a length $l > 0$, we let $\mathscr{B}_l^d$ denote the family of all half open cubes in

$\mathbf{R}^d$ with sidelength l and corners on the lattice $(l \cdot \mathbf{Z})^d$. That is,

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

If $E \subset \mathbf{R}^d$, $\mathscr{B}_l^d(E)$ is the family of cubes in $\mathscr{B}_l^d$ intersecting E, i.e.

$$\mathscr{B}_l^d(E) = \{I \in \mathscr{B}_l^d : I \cap E = \emptyset\}.$$

(C) The *lower* and *upper Minkowski dimension* of a compact set $Z \subset \mathbf{R}^d$ are defined as

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

If $\underline{\dim}_{\mathbf{M}}(Z) = \overline{\dim}_{\mathbf{M}}(Z)$, then we define $\dim_{\mathbf{M}}(Z)$ equal to the common value.

(D) If $0 \leq \alpha$ and $\delta > 0$, we define the dyadic Hausdorff content of a set $E \subset \mathbf{R}^d$ as

$$H_\delta^\alpha(E) = \inf \left\{ \sum_{k=1}^m l_k^\alpha : E \subset \bigcup_{k=1}^m I_k \text{ and } I_k \in \mathscr{B}_{l_k}^d, l_k \leq \delta \text{ for all } k \right\}.$$

The α-dimensional dyadic Hausdorff measure H^α on $\mathbf{R}^d$ is $H^\alpha(E) = \lim_{\delta \to 0} H_\delta^\alpha(E)$, and the *Hausdorff dimension* of a set E is $\dim_{\mathbf{H}}(E) = \inf\{\alpha \geq 0 : H^\alpha(E) = 0\}$.

(E) Given $I \in \mathscr{B}_l^{dn}$, we can decompose I as $I_1 \times \cdots \times I_n$ for unique cubes $I_1, \dots, I_n \in \mathscr{B}_l^d$. We say I is *strongly non-diagonal* if the cubes $I_1, \dots, I_n$ are distinct. Strongly non-diagonal cubes will play an important role in Section 4.2, when we solve a discrete version of Theorem 6.

(F) Adopting the terminology of [8], we say a collection of sets $\{U_k\}$ is a *strong cover* of a set E if $E \subset \limsup U_k$, which means every element of E is contained in infinitely many of the sets U_k. This idea will be useful in Section 4.3.

(G) A *Frostman measure* of dimension α is a non-zero compactly supported probability measure μ on $\mathbf{R}^d$ such that for every cube I of sidelength l, $\mu(I) \lesssim l^\alpha$. Note that a measure μ satisfies this inequality for every cube I if and only if it satisfies the inequality for cubes whose sidelengths are dyadic lengths. *Frostman's lemma* says that

$$\dim_{\mathbf{H}}(E) = \sup \left\{ \alpha : \begin{array}{c} \text{there is an } \alpha \text{ dimensional Frostman} \\ \text{measure supported on } E \end{array} \right\}.$$

4.2 Avoidance at Discrete Scales

In this section we describe a method for avoiding Z at a single scale. We apply this technique in Section 4.3 at many scales to construct a set X avoiding Z at all scales. This single scale avoidance technique is the core part of our construction, and the efficiency with which we can avoid Z at a single scale has direct consequences on the Hausdorff dimension of the set X obtained in Theorem 6.

At a single scale, we solve a discretized version of the problem, where all sets are unions of cubes at two dyadic lengths $l \geq s$. In the discrete setting, Z is replaced by a discretization version of itself, i.e. a union of cubes in $\mathscr{B}_s^{dn}$, denoted by Z_s. Given a set E, which is a union of cubes in $\mathscr{B}_l^d$, our goal is to construct a a union of cubes in $\mathscr{B}_s^d$, denoted F, such that $F \subset E$ and F^n is disjoint from the strongly non-diagonal cubes of Z_s (see Definition (E)).

In order to ensure the final set X obtained in Theorem 6 has large Hausdorff dimension regardless of the rapid decay of scales used in the construction of X, it is crucial that F is spread uniformly over E. We achieve this by decomposing E into sub-cubes in $\mathscr{B}_r^d$ for some intermediate scale $r \in [s,l]$, and distributing F evenly among as many of these intermediate sub-cubes as possible. Assuming a mild regularity condition on the volume of Z_s, this is possible.

Lemma 10. *Fix two dyadic lengths $l \geq s$. Let E be a nonempty union of cubes in $\mathscr{B}_l^d$, and let Z_s be a union of cubes in $\mathscr{B}_s^d$ such that $|Z_s| \leq l^{dn}/2$. Then there exists a dyadic length $r \in [s,l]$ such that*

$$A_l |Z_s|^{1/d(n-1)} \leq r \leq \max(s, 2A_l|Z_s|^{1/d(n-1)}) \quad \textit{where} \quad A_l = (2^{1/d}/l)^{1/(n-1)}. \tag{4.1}$$

For this scale, there exists a set $F \subset E$, which is a nonempty union of cubes in $\mathscr{B}_s^d$, satisfying the following three properties:

(A) Avoidance*: For any distinct $J_1, \dots, J_n \in \mathscr{B}_s^d(F)$, $J_1 \times \cdots \times J_n \notin \mathscr{B}_s^{dn}(Z_s)$.*

(B) Non-Concentration*: For any $I \in \mathscr{B}_r^d(E)$, there is at most one $J \in \mathscr{B}_s^d(F)$ with $J \subset I$.*

(C) Large Size*: For any $I \in \mathscr{B}_l^d(E)$, $\#\mathscr{B}_s^d(F \cap I) \geq \#\mathscr{B}_r^d(I)/2 = (l/r)^d/2$.*

In other words, F avoids strongly non-diagonal cubes in Z_s, and contains a single sidelength s portion of more than half of the sidelength r cubes contained in any sidelength l cube in E.

Proof. Let r be the smallest dyadic length larger than than s and $A_l|Z_s|^{1/d(n-1)}$, so that (4.1) is satisfied. The assumption that $|Z_s| \leq l^{dn}/2$ implies

$$A_l|Z_s|^{1/d(n-1)} \leq A_l l^{n/(n-1)}/2^{1/d(n-1)} = l.$$

Thus we have gauranteed that $r \in [s,l]$. For each $I \in B_r^d(E)$, let J_I be a random element of $\mathscr{B}_s^d(I)$ chosen uniformly at random, independantly from the other random variables $J_{I'}$ with $I \neq I'$. Define a random set

$$U = \bigcup \{J_I : I \in \mathscr{B}_s^d(I)\},$$

and for each U, set

$$\mathscr{K}(U) = \{K \in \mathscr{B}_s^{dn}(Z_s) : K \in U^n, K \text{ strongly non-diagonal}\}$$

Then set

$$F_U = U - \{\pi(K) : K \in \mathscr{K}(U), K \text{ is strongly diagonal}\}, \tag{4.2}$$

where $\pi : \mathbf{R}^{dn} \to \mathbf{R}^d$ maps $K_1 \times \cdots \times K_n$ to K_1 for each $K_i \in \mathbf{R}^d$. Given any strongly non-diagonal cube $J_1 \times \cdots \times J_n \in \mathscr{B}_s^{dn}(Z_s)$, either $J_1 \times \cdots \times J_n \notin \mathscr{B}_s^{dn}(U^n)$, or $J_1 \times \cdots \times J_n \in \mathscr{B}_s^{dn}(U^n)$. If the former occurs then $J_1 \times \cdots \times J_n \notin \mathscr{B}_s^{dn}(F_U^n)$ since $F_U \subset U$, while if the latter occurs then $K \in \mathscr{K}(U)$, so $J_1 \notin \mathscr{B}_s^d(F_U)$. In either case, $J_1 \times \cdots \times J_n \notin \mathscr{B}_s^{dn}(F_U^n)$, so F_U satisfies Property (A). By construction, U contains at most one subcube $J \in \mathscr{B}_s^{dn}$ for each $I \in \mathscr{B}_l^{dn}(E)$. Since $F_U \subset U$, F_U satisfies Property (B). These properties are satisfied for any instance of U, but it is not true that Property (C) holds for each F_U. The remainder of this proof is devoted to showing this property holds for F_U with non-zero probability, guaranteeing the existence of the required set F satisfying the properties of the lemma.

For each cube $J \in \mathscr{B}_s^d(E)$, there is a unique 'parent' cube $I \in \mathscr{B}_r^d(E)$ such that $J \subset I$. Since I contains $(r/s)^d$ elements of $\mathscr{B}_s^d(E)$, and J_I is chosen uniformly at random from $\mathscr{B}_s^d(I)$,

$$\mathbf{P}(J \subset U) = \mathbf{P}(J_I = J) = (s/r)^d.$$

The cubes J_I are chosen independantly, so if $J_1, \dots, J_k$ are distinct cubes in $\mathscr{B}_s^d(E)$, then the last calculation combined with Property (B) shows that

$$\mathbf{P}(J_1, \dots, J_k \in U) = \begin{cases} (s/r)^{dk} & : \text{if } J_1, \dots, J_k \text{ have distinct parents} \\ 0 & : \text{otherwise} \end{cases}$$

Let $K = J_1 \times \cdots \times J_n \in \mathscr{B}_s^{dn}(Z_s)$ be a strongly non-diagonal cube. Then cubes $J_1, \dots, J_n$ are distinct, so the calculation we just performed implies

$$\mathbf{P}(K \subset U^n) = \mathbf{P}(J_1, \dots, J_k \in U) \leq (s/r)^{dn}.$$

Together with linearity of expectation, and (4.1), if K ranges over the strongly non-diagonal cubes of $\mathscr{B}_s^{dn}(Z_s)$, we find

$$\begin{aligned}\mathbf{E}(\#\mathscr{K}(U)) &= \sum_K \mathbf{P}(K \subset U^n) \leq \#\mathscr{B}_s^{dn}(Z_s) \cdot (s/r)^{dn} = |Z_s| r^{-dn} \\ &= \left[|Z_s| r^{-d(n-1)}\right] r^{-d} \leq \left[|Z_s| (A_l |Z_s|^{1/d(n-1)})^{-d(n-1)}\right] r^{-d} = \left[l^d/2\right] r^{-d} = (l/r)^d/2.\end{aligned}$$

Thus there exists at least one (non-random) set U_0 such that

$$\#\mathscr{K}(U_0) \leq \mathbf{E}(\#\mathscr{K}(U)) \leq (l/r)^d/2. \tag{4.3}$$

This means that the resultant set F_{U_0} is obtained by removing at most $(l/r)^d/2$ cubes in $\mathscr{B}_s^d$ from U_0. Since U_0 contains $(l/r)^d$ cubes in $\mathscr{B}_s^d(I)$ for each $I \in \mathscr{B}_l^d(E)$, F_{U_0} contains at least $(l/r)^d/2$ cubes in $\mathscr{B}_s^d(I)$ for each $I \in \mathscr{B}_l^d(E)$, so F_{U_0} satisfies Property (C). Setting $F = F_{U_0}$ completes the proof. □

Remark. *While the existence of the set F in Lemma 10 was obtained by probabilistic techniques, we emphasize that it's existence is a purely deterministic statement. One can find a candidate F constructively by checking all of the finitely many possible choice of U to find one particular choice U_0 which satisfies* (4.3)*, and then defining F by* (4.2)*. Thus the set we obtain in Theorem 6 exists by purely constructive means.*

Our inability to select almost every cube in Lemma 10 means that repeated applications of the result will lead to a loss in Hausdorff dimension. In fact, in the worst case, applying the lemma causes us to lose as much Hausdorff dimension as is permitted by Theorem 1. Note that the value r specified in Theorem 10 is always bounded below by $|Z_s|$. We will later see that if Z is the countable union of sets with lower Minkowski dimension α, the scale s discretization Z_s satisfies $|Z_s| \leq s^{dn-\alpha-\varepsilon}$, for some small positive ε converging to zero as $s \to 0$. But we can certainly find sets Z with lower Minkowski dimension α satisfying $|Z_s| \geq s^{dn-\alpha}$

for all s. In this situation, (4.1) shows

$$r \geq A_l |Z_s|^{1/d(n-1)} \geq A_l s^{(dn-\alpha)/d(n-1)} \geq s^{(dn-\alpha)/d(n-1)} \tag{4.4}$$

If we combine this inequality with Property (B) of Lemma 10, we conclude

$$\frac{\log \# \mathscr{B}_s^d(F)}{\log(1/s)} \leq \frac{\log \# \mathscr{B}_r^d(E)}{\log(1/s)} = \frac{\log |E| r^{-d}}{\log(1/s)} = \frac{d \log(1/r) - \log |E|^{-1}}{\log(1/s)} \leq \frac{dn-\alpha}{n-1}.$$

Given a set $F = \bigcap F_k$, where $\{F_k\}$ are infinitely many sets obtained from an application of Lemma 10 at a sequence of scales $\{s_k\}$ with $s_k \to 0$ as $k \to \infty$, and where $|Z_{s_k}| \geq s_k^{dn-\alpha}$, then the last computation shows

$$\dim_{\mathbf{H}}(F) \leq \overline{\dim}_{\mathbf{M}}(F) \leq \lim_{s_k \to 0} \frac{\log \# \mathscr{B}_{s_k}^d(F_k)}{\log(1/s_k)} \leq \frac{dn-\alpha}{n-1}.$$

This is why we can only lower the Hausdorff dimension of the set X obtained in Theorem 6 by $(dn-\alpha)/(n-1)$. Furthermore, we see that we must be very careful to ensure applications of the discrete lemma are the only place in our proof where dimension is lost.

Remark. *Lemma 10 is the core method in our avoidance technique. The remaining argument is fairly modular. If, for a special case of Z, one can improve the result of Lemma 10 so that r is chosen on the order of $s^{\beta/d}$, then the remaining parts of our paper can be applied near verbatim to yield a set X with Hausdorff dimension β, as in Theorem 6. The last paragraph shows that when $|Z_s| \geq s^{dn-\alpha}$, which is certainly possible given the hypothesis of Theorem 10, the length r is chosen on the order of $s^{(dn-\alpha)/d(n-1)}$, as we saw in* (4.4)*, which is why we obtain a Hausdorff dimension $(dn-\alpha)/(n-1)$ set.*

4.3 Fractal Discretization

In this section we will construct the set X by applying Lemma 10 at many scales. Since Z is a countable union of compact sets with Minkowski dimension at most

α, there exists a strong cover (see Definition (F)) of Z by cubes restricted to a sequence of dyadic lengths $\{l_k\}$. We will select this strong cover so that the scales l_k converge to 0 very quickly.

Lemma 11. *Let $Z \subset \mathbf{R}^{dn}$ be a countable union of compact sets, each with lower Minkowski dimension at most α. Let $\{\varepsilon_k\}$ be a sequence of positive numbers and let $\{f_k\}$ be a sequence of functions such that $f_k\colon (0,\infty) \to (0,\infty)$. Then there exists a sequence of dyadic lengths $\{l_k\}$ and compact sets $\{Z_k\}$ such that*

(A) *For each index $k \geq 2$, $l_k \leq f_{k-1}(l_{k-1})$.*

(B) *For each index k, Z_k is a union of cubes in $\mathscr{B}^{dn}_{l_k}$.*

(C) *Z is strongly covered by the sets $\{Z_k\}$.*

(D) *For each index k, $\#\mathscr{B}^{dn}_{l_k}(Z_k) \leq (1/l_k)^{\alpha+\varepsilon_k}$.*

Proof. Let Z be the union of sets $\{Y_k\}$ with $\underline{\dim}_{\mathbf{M}}(Y_k) \leq \alpha$ for any k. Let $m_1, m_2, \ldots$ be a sequence of integers that repeats each integer infinitely often. For each positive integer k, since $\underline{\dim}_{\mathbf{M}}(Y_{m_k}) \leq \alpha$, definition (C) implies that there exists arbitrarily small lengths l which satisfy $\#\mathscr{B}^{dn}_{l}(Y_{m_k}) \leq 1/l^{\alpha+(\varepsilon_k/2)}$. Replacing l with a dyadic length at most twice the size of l, there are infinitely many *dyadic* scales l with

$$\#\mathscr{B}^{dn}_{l}(Y_{m_k}) \leq \frac{1}{(l/2)^{\alpha+\varepsilon_k}} \leq \frac{2^{dn}}{l^{\alpha+(\varepsilon_k/2)}} = \frac{\left(2^{dn} l^{\varepsilon_k/2}\right)}{l^{\alpha+\varepsilon_k}}.$$

In particular, we may select a dyadic length l such that $2^{dn} l^{\varepsilon_k/2} \leq 1$, and, if $k \geq 2$, also satisfying $l \leq f_{k-1}(l_{k-1})$. The first constraint together with the last calculation implies $\#\mathscr{B}^{dn}_{l}(Y_{m_k}) \leq 1/l^{\alpha+\varepsilon_k}$. We then set $l_k = l$, and define Z_k to be the union of all cubes in $\mathscr{B}^{dn}_{l_k}(Y_{m_k})$. $\square$

We now construct X by avoiding the various discretizations of Z at each scale. The aim is to find a nested decreasing family of discretized sets $\{X_k\}$ with $X = \bigcap X_k$. One condition guaranteeing that X avoids Z is that X_k^n is disjoint from *strongly non-diagonal* cubes in Z_k.

Lemma 12. *Let $Z \subset \mathbf{R}^{dn}$ and let $\{l_k\}$ be a sequence of lengths converging to zero. For each index k, let Z_k be a union of cubes in $\mathscr{B}^{dn}_{l_k}$, and suppose the sets $\{Z_k\}$ strongly cover Z. For each index k, let X_k be a union of cubes in $\mathscr{B}^{d}_{l_k}$. Suppose that for any k, X_k^n avoids strongly non-diagonal cubes in Z_k. If $X = \bigcap X_k$, then $(x_1,\dots,x_n) \notin Z$ for any distinct $x_1,\dots,x_n \in X$.*

Proof. Let $z \in Z$ be a point with distinct coordinates $z_1,\dots,z_n$. Set

$$\Delta = \{(w_1,\dots,w_n) \in \mathbf{R}^{dn} : \text{there exists } i \neq j \text{ such that } w_i = w_j\}.$$

Then $d(\Delta,z) > 0$, where d is the Hausdorff distance between Δ and z. Since $\{Z_k\}$ strongly covers Z, there is a subsequence $\{k_m\}$ such that $z \in Z_{k_m}$ for any index m. For suitably large m, the sidelength l_k cube I in Z_{k_m} containing z is disjoint from Δ. But this means I is strongly non-diagonal, and so $z \notin X^n_{k_m}$. In particular, z is not an element of X^n. $\square$

We are now ready to construct the set X in Theorem 6. Let $l_0 = 1$ and $X_0 = [0,1)^d$. For each $k \geq 1$, define $\varepsilon_k = c_0/k$, where we view $c_0 = (dn-\alpha)/4$ as a irrelevant constant small enough that $dn - \alpha - 2\varepsilon_k > 0$ for any k. We set

$$f_k(x) = \min\left(x^{k^2}, (x^{dn}/2)^{1/(dn-\alpha-\varepsilon_{k+1})}, (1/2A_x)^{d(n-1)/\varepsilon_{k+1}}\right).$$

Apply Lemma 11 to Z with this choice of $\{\varepsilon_k\}$ and $\{f_k\}$; let $\{l_k\}$ be the resulting sequence of dyadic lengths and let $\{Z_k\}$ be the resulting strong cover of Z. Observe that the definition of f_k implies that for any index k,

$$l_{k+1} \leq l_k^{k^2}, \quad l_{k+1}^{dn-\alpha-\varepsilon_{k+1}} \leq l_k^{dn}/2, \quad \text{and} \quad 2A_{l_k} l_{k+1}^{\varepsilon_{k+1}/d(n-1)} \leq 1. \tag{4.5}$$

For each index $k \geq 1$, define $l = l_k$ and $s = l_{k+1}$. Observe that X_k is a non-empty union of cubes in $\mathscr{B}^d_l$; that Z_{k+1} is a union of cubes in $\mathscr{B}^{dn}_s$, and that (4.5) implies

$$|Z_{k+1}| \leq l_{k+1}^{dn-\alpha-\varepsilon_{k+1}} \leq l_k^{dn}/2 = l^{dn}/2.$$

Setting $Z_s = Z_{k+1}$, we are therefore justified in applying Lemma 10. This produces a dyadic length r, which we denote by r_{k+1}. Assuming $\alpha \geq d$, by (4.1) and (4.5), we find

$$\begin{aligned} r_{k+1} &\leq \max\left(l_{k+1}, 2A_{l_k}|Z_{k+1}|^{1/d(n-1)}\right) \leq \max\left(l_{k+1}, 2A_{l_k} l_{k+1}^{(dn-\alpha-\varepsilon_{k+1})/d(n-1)}\right) \\ &\leq \max\left(l_{k+1}, l_{k+1}^{(dn-\alpha-2\varepsilon_{k+1})/d(n-1)}\right) = l_{k+1}^{(dn-\alpha-2\varepsilon_{k+1})/d(n-1)}. \end{aligned} \tag{4.6}$$

For this choice of r, we obtain a set $F \subset X_k$ which is a union of cubes in $\mathscr{B}_s^d(E)$ satisfying Properties (A), (B), and (C) from the lemma, and we define $X_{k+1} = F$. Property (A) implies X_{k+1} avoids strongly non-diagonal cubes in Z_{k+1}, so if we define $X = \bigcap X_k$, then Lemma 12 implies $(x_1, \dots, x_n) \notin Z$ for any distinct $x_1, \dots, x_n \in X$.

Lemma 13.

$$\overline{\dim}_{\mathbf{M}}(X) \geq \frac{dn-\alpha}{n-1}$$

Proof. Consider X as the limit of the sequence $\{X_k\}$. Since every cube in $\mathscr{B}_{l_{k+1}}^d(X_{k+1})$ intersects X, $\mathscr{B}_{l_{k+1}}^d(X_{k+1}) = \mathscr{B}_{l_{k+1}}^d(X)$. And so by Property (C) of Lemma 10, (4.5), and (4.6), we conclude

$$\begin{aligned} \frac{\log(\#\mathscr{B}_{l_{k+1}}^d(X))}{\log(1/l_{k+1})} &= \frac{\log(\#\mathscr{B}_{l_{k+1}}^d(X_{k+1}))}{\log(1/l_{k+1})} \geq \frac{\log((l_k/r_{k+1})^d \cdot \#\mathscr{B}_{l_k}^d(X_k))}{\log(1/l_{k+1})} \\ &= \frac{d\log(1/r_{k+1})}{\log(1/l_{k+1})} - \frac{d\log(1/l_k)}{\log(1/l_{k+1})} \\ &\geq \frac{d\log\left(1/l_{k+1}^{(dn-\alpha-2\varepsilon_{k+1})/d(n-1)}\right)}{\log(1/l_{k+1})} - \frac{d\log(1/l_k)}{\log(1/l_k^{k^2})} \\ &= \frac{dn-\alpha-2\varepsilon_{k+1}}{n-1} - \frac{d}{k^2} \\ &= \frac{dn-\alpha}{n-1} - o(1) \end{aligned}$$

Taking $k \to \infty$, we conclude $\overline{\dim}_{\mathbf{M}}(X) \geq (dn-\alpha)/(n-1)$. $\square$

4.4 Dimension Bounds

To complete the proof of Theorem 6, we must show that $\dim_{\mathbf{H}}(X) \geq \beta$, where

$$\beta = \frac{dn - \alpha}{n-1}.$$

We begin with a rough outline of our proof strategy. Recall that from the previous section, we have a decreasing sequence of lengths $\{l_k\}$. The most convenient way to examine the dimension of X at various scales is to use Frostman's lemma (see Definition (G)). We construct a probability measure μ supported on X such that for all $\varepsilon > 0$, for all dyadic lengths l, and for all $I \in \mathscr{B}^d_l$, $\mu(I) \lesssim_\varepsilon l^{\beta-\varepsilon}$. We begin by proving the bound $\mu(I) \lesssim l_k^{\beta - O(1/k)}$ when $I \in \mathscr{B}^d_{l_k}$, which we view as saying X looks like a set with dimension $\beta - O(1/k)$ at the lengths $\{l_k\}$. To obtain the complete dimension bound, it then suffices to interpolate to get an acceptable bound at all intermediate scales. In this construction, as in [6], the rapid decay of the lengths $\{l_k\}$ forced on us in the construction means interpolation poses a significant difficulty. We avoid this difficulty because of the uniform way that we have selected cubes in consecutive scales. This will imply that between the scales l_k and r_{k+1}, the mass of μ distributes with similar properties to the full dimensional Lebesgue measure, which makes interpolation easy.

We now define the measure μ, by first defining it as a premeasure on $\bigcup_{i=1}^\infty \mathscr{B}^d_{l_i}[0,1)^d$. We initially set $\mu([0,1)^d) = 1$. Given $I \in \mathscr{B}^d_{l_k}$, we find a parent cube $I' \in \mathscr{B}^d_{l_{k-1}}$ with $I \subset I'$. If $I \subset X_k$, then we set

$$\mu(I) = \frac{\mu(I')}{\#\mathscr{B}^d_{l_k}(X_k \cap I')}. \tag{4.7}$$

Otherwise, if $I \not\subset X_k$, we set $\mu(I) = 0$. Notice

$$\mu(I') = \sum_{I \in \mathscr{B}^d_{l_k}(X_k \cap I')} \frac{\mu(I')}{\#\mathscr{B}^d_{l_k}(X_k \cap I')} = \sum_{I \in \mathscr{B}^d_{l_k}(I')} \mu(I).$$

Thus mass is maintained at each stage of the construction, and then Proposition 1.7 of [5] implies μ is a premeasure, and thus extends to a measure on the entire Borel sigma algebra. The remainder of this section is devoted to showing that μ is a Frostman measure of dimension $\beta - \varepsilon$ for any $\varepsilon > 0$.

Lemma 14. *If $I \in \mathscr{B}^d_{l_k}$, then*

$$\mu(I) \leq 2^k \left[\frac{r_k \dots r_1}{l_{k-1} \dots l_1} \right]^d. \tag{4.8}$$

Proof. We prove the theorem inductively on k. For $k = 0$, the theorem is obvious, because $\mu(I) \leq 1$. For the purposes of induction, let $I \in \mathscr{B}^d_{l_k}$, together with a parent cube $I' \in \mathscr{B}^d_{l_{k-1}}$ with $I \subset I'$. If $\mu(I) > 0$, $I \subset X_k$, so $I' \subset X_{k-1}$. Because X_k was obtained from X_{k-1} via an application of Lemma 1, Property (C) of that lemma states that $\#\mathscr{B}^d_{l_k}(X_k \cap I) \geq (l_{k-1}/r_k)^d/2$, so together with the inductive hypothesis and (4.7), we conclude

$$\mu(I) = \frac{\mu(I')}{\#\mathscr{B}^d_{l_k}(X_k \cap I)} \leq \frac{\mu(I')}{(l_{k-1}/r_k)^d/2} \leq \frac{2^{k-1} \cdot \left[\frac{r_{k-1} \dots r_1}{l_{k-2} \dots l_1}\right]^d}{(l_{k-1}/r_k)^d/2} = 2^k \left[\frac{r_k \dots r_1}{l_{k-1} \dots l_1} \right]^d. \quad \square$$

Treating all parameters in (4.8) which depend on indices smaller than k as essentially constant, and using (4.6), we 'conclude' that

$$\mu(I) \lesssim r_k^d \lesssim l_k^{\beta - 2\cdot\varepsilon_k/(n-1)} = l_k^{\beta - O(1/k)}.$$

The bounds in (4.5) imply l_k decays very rapidly, which enables us to ignore quantities depending on previous indices, and obtain a true inequality.

Corollary. *There exists $c > 0$ such that for all $I \in \mathscr{B}^d_{l_k}$, $\mu(I) \lesssim l_k^{\beta - c/k}$.*

Proof. Given $\varepsilon > 0$, Lemma 14, Equation (4.6), the inequality $l_k \leq l_{k-1}^{(k-1)^2}$ from

(4.5), and the fact that there exists a constant c_0 such that $\varepsilon_{k+1} = c_0/k$, we find

$$\mu(I) \leq 2^k \left[\frac{r_k \dots r_1}{l_{k-1} \dots l_1}\right]^d \leq \left(\frac{2^k}{l_{k-1}^d \dots l_1^d}\right) l_k^{\beta - \varepsilon_k/(n-1)} = \left(\frac{2^k l_k^{2d/k}}{l_{k-1}^{d(k-1)}}\right) l_k^{\beta - \varepsilon_k/(n-1) - 2d/k}$$
$$\leq \left(2^k l_{k-1}^{(2d/k)(k-1)^2 - d(k-1)}\right) l_k^{\beta - (c_0/(n-1)+2d)/k} = o\left(l_k^{\beta - (c_0/(n-1)+2d)/k}\right),$$

so we can then set $c = c_0/(n-1) + 2d$. □

Corollary 3 gives a clean expression of the β dimensional behaviour of μ at discrete scales. To obtain a Frostman measure bound at *all* scales, we need to apply a covering argument. This is where the uniform mass assignment technique comes into play. Because μ behaves like a full dimensional set between the scales l_k and r_{k+1}, we won't be penalized for making the gap between l_k and r_{k+1} arbitrarily large. This is essential to our argument, because l_k decays faster than 2^{-k^m} for any $m > 0$.

Lemma 15. *If l is dyadic and $I \in \mathscr{B}_l^d$, then $\mu(I) \lesssim_k l^{\beta - c/k}$ for each integer k.*

Proof. We begin by assuming $l \leq l_k$. To bound $\mu(I)$, we apply a covering argument, which breaks into cases depending on the size of l in proportion to the scales l_k and r_k:

- If $r_{k+1} \leq l \leq l_k$, we can cover I by $(l/r_{k+1})^d$ cubes in $\mathscr{B}_{r_{k+1}}^d$. Because of Property (B) and (C) of Lemma 10, we know that the mass of each cube in $\mathscr{B}_{r_{k+1}}^d$ is bounded by at most $2(r_{k+1}/l_{k+1})^d$ times the mass of a cube in $\mathscr{B}_{l_k}^d$. Thus

  $$\mu(I) \lesssim (l/r_{k+1})^d (2(r_{k+1}/l_k)^d) l_k^{\beta - c/k} \leq 2l^d / l_k^{d-\beta+c/k} \leq 2l^{\beta - c/k},$$

  where we used the fact that $d - \beta + c/k \geq 0$, so $l_k^{d-\beta+c/k} \geq l^{d-\beta+c/k}$.

- If $l_{k+1} \leq l \leq r_{k+1}$, we can cover I by a single cube in $\mathscr{B}_{r_{k+1}}^d$. Because of Property (B) of Lemma 10, each cube in $\mathscr{B}_{r_{k+1}}^d$ contains at most one cube of $\mathscr{B}^d_{l_{k+1}}(X_{k+1})$, so

$$\mu(I) \lesssim l_{k+1}^{\beta - c/k} \leq l^{\beta - c/k}.$$

- If $l \leq l_{k+1}$, there certainly exists m such that $l_{m+1} \leq l \leq l_m$, and one of the previous cases yields that $\mu(I) \lesssim l^{\beta - c/m} \leq l^{\beta - c/k}$.

If $l \geq l_k$, then $\mu(I) \leq 1 \lesssim_k l_k^{\beta - c/k} \leq l^{\beta - c/k}$, so $\mu(I) \lesssim_k l^{\beta - c/k}$ for arbitrary dyadic l. □

Applying Frostman's lemma to Lemma 15 gives $\dim_{\mathbf{H}}(X) \geq \beta - c/k$ for each $k > 0$. Taking $k \to \infty$ proves the needed dimension bound. Since we have already shown in Section 4.3 that $(x_1, \dots, x_n) \notin Z$ for any distinct $x_1, \dots, x_n \in X$, this concludes the proof of Theorem 6.

4.5 Applications

As discussed in the introduction, Theorem 1 generalizes Theorems 1.1 and 1.2 from [6]. In this section, we present two applications of Theorem 6 in settings where previous methods cannot obtain any results.

Theorem 7 (Sum-sets avoiding specified sets)**.** *Let $Y \subset \mathbf{R}^d$ be a countable union of sets of Minkowski dimension at most α. Then there exists a set $X \subset \mathbf{R}^d$ with Hausdorff dimension at least $1-\alpha$ such that $X+X$ is disjoint from Y.*

Proof. Define $Z = Z_1 \cup Z_2$, where

$$Z_1 = \{(x,y) : x+y \in Y\} \quad \text{and} \quad Z_2 = \{(x,y) : y \in Y/2\}.$$

Since Y is a countable union of sets of Minkowski dimension at most α, Z is a countable union of sets with lower Minkowski dimension at most $1+\alpha$. Applying Theorem 6 with $d=1$, $n=2$, giving a set $X \subset \mathbf{R}^d$ with Hausdorff dimension $1-\alpha$ avoiding Z. Since X avoids Z_1, whenever $x, y \in X$ are distinct, $x+y \notin Y$. Since X avoids Z_2, $X \cap (Y/2) = \emptyset$, and thus for any $x \in X$, $x+x \notin Y$. Thus $X+X$ is disjoint from Y. □

Remark. *One weakness of our result is that as the number of variables n increases, the dimension of X tends to zero. If we try and make the n-fold sum $X + \cdots + X$ disjoint from Y, current techniques only yield a set of dimension* $(1-\alpha)/(n-1)$. *We have ideas on how to improve our main result when Z is 'flat', in addition to being low dimension, which will enable us to remove the dependence of* $\dim_{\mathbf{H}}(X)$ *on n. In particular, we expect to be able to construct a set X of dimension $1-\alpha$, such that X is disjoint from Y, and X is closed under addition, and multiplication by rational numbers. In particular, given a* $\mathbf{Q}$ *subspace V of* $\mathbf{R}^d$ *with dimension α, we can always find a 'complementary'* $\mathbf{Q}$ *vector space W with complementary fractional dimension $d-\alpha$ such that* $V \cap W = (0)$.

One of the most interesting uses of our method is to construct subsets of fractals avoiding patterns. In [6], Fraser and the second author show that if γ is a C^2 curve with non-vanishing curvature, then there exists a set $E \subset \gamma$ of Hausdorff dimension $1/2$ that does not contain isoceles triangles. Our method can extend this result from the case of curves to more general sets, which gives an example of the flexibility of our method. For simplicity, we stick to an analysis of planar sets.

Theorem 8 (Restricted sets avoiding isoceles triangles)**.** *Let $Y \subset$* $\mathbf{R}^2$ *and let π :* $\mathbf{R}^2 \to \mathbf{R}$ *be an orthogonal projection such that $\pi(Y)$ has non-empty interior. Let d be an arbitrary metric on* $\mathbf{R}^2$. *Suppose that*

$$Z_0 = \{(y_1,y_2,y_3) \in Y^3 : d(y_1,y_2) = d(y_1,y_3)\}$$

is the countable union of sets with lower Minkowski dimension at most α, for $\varepsilon \geq 0$. *Then there exists a set $X \subset Y$ with dimension at least* $(3-\alpha)/2$ *so that no triple of points* $(x_1,x_2,x_3) \in X^3$ *form the vertices of an isoceles triangle.*

Proof. Without loss of generality, by translation and rescaling, we may assume $\pi(Y)$ contains $[0,1)$. Form the set

$$Z = \pi(Z_0) = \{(\pi(y_1),\pi(y_2),\pi(y_3)) : y \in Z_0\}$$

Then Z is the projection of an α dimensional set, and therefore has dimension at most α. Applying Theorem 6 with $d = 1$ and $n = 3$, we construct a set $X_0 \subset [0,1)$ with Hausdorff dimension at least $(3-\alpha)/2$ such that for any distinct $x_1, x_2, x_3 \in X_0$, $(x_1, x_2, x_3) \notin Z$. Thus if we form a set X by picking, from each $x \in X_0$, a single element of $\pi^{-1}(x)$, then X avoids isoceles triangles, and has Hausdorff dimension at least as large as X_0. $\square$

To see that Theorem 3 indeed generalizes the result of Fraser and the second author, observe that if d is the Euclidean metric, then for every pair of points $x, y \in \mathbf{R}^2$, the set

$$\{z \in \mathbf{R}^2 : d(x,z) = d(y,z)\}$$

is the perpendicular bisector B_{xy} of x and y. If γ is a compact portion of a smooth curve with non-vanishing curvature, then the number of points in $\gamma \cap B_{xy}$ is bounded independantly of x and y. Thus the set Z_0 in the statement of Theorem 3 has Minkowski dimension at most 2, and we can find a set with Hausdorff dimension $(3-2)/2 = 1/2$ on the curve avoiding isoceles triangles.

Results about slice of measures, such as those detailed in Chapter 6 of [12], show that for any one dimensional set Y, for almost every line L, $L \cap Y$ consists of a finite collection of points. This suggests that if Y is any set with fractional dimension one, then Z_0 has dimension at most 2. This implies that we can find a subset of Y with dimension $1/2$ avoiding curves. We are unsure if this is true for every set with dimension one, but we provide two examples suggesting this is true for a generic set. The first result shows that for any $\varepsilon > 0$, there is an infinite family of Cantor-type sets with dimension $1+\varepsilon$ such that Z_0 has dimension at most $2+\varepsilon$. The second shows that for any rectifiable curve, Z_0 has dimension at most 2. Thus Theorem 3 can be applied in settings where Y is incredibly 'rough', i.e. totally disconnected.

To study the first example, we consider a probabilistic model for a Cantor-type set which almost surely has the required properties. This model is obtained by considering a nested decreasing family of discretized random sets $\{C_k\}$, with each C_k a union of sidelength $1/2^k$ squares. First, we fix $p \in [0,1]$. Then we set $C_0 = [0,1]^2$. To construct C_{k+1}, we split each sidelength $1/2^k$ cube in C_k into four sidelength $1/2^{k+1}$ squares, and keep each square in C_{k+1} with probability p. To study this model, we employ some results about tail bounds and asymptotics for branching processes.

Lemma 16. *If $p > 1/4$, then with non-zero probability,* $\dim_{\mathbf{M}}(C) = 2 - \log_2(1/p)$.

Proof. Let $p > 1/4$. For each k, let Z_k denote the number of sidelength $1/2^k$ cubes in $[0,1]^2$. Then the sequence $\{Z_k\}$ is a branching process, where each cube can produce between zero and four subcubes, with each of these four cubes kept with probability p. Thus $\mathbf{E}[Z_{k+1}|Z_k] = (4p)Z_k$. Since $4p > 1$, A simple calculation, summarized in Theorem 8.1 of [7], shows that the process $W_k = Z_k/(4p)^k$ is an L^2 bounded martingale, and so there exists a random variable W such that $W_k \to W$ almost surely, and $\mathbf{E}(W|W_k) = W_k$ for all k. Whenever W is non-zero,

$$\dim_{\mathbf{M}}(C) = \lim_{k\to\infty} \frac{\log Z_k}{k\log 2} = \lim_{k\to\infty} \frac{\log(W_k/W) + \log(W(4p)^k)}{k\log 2} = 2 - \log_2(1/p).$$

Since $\mathbf{E}(W) = \mathbf{E}(\mathbf{E}(W|W_0)) = \mathbf{E}(W_0) = 1$, W is non-zero with positive probability. $\square$

Similar asymptotics for branching processes show that the set Z_0 associated with C as in Theorem 3 almost surely has Minkowski dimension $3 - \alpha$. We do this first by proving a supplementary result.

Lemma 17. *Let $\{Z_k\}$ be a supercritical branching process with extinction probability q. If we set $\tilde{M} = q + M$, then there exists small positive constants λ and ε, depending only on the offspring law of $\{Z_k\}$, such that* $\mathbf{P}(Z_k \geq k\tilde{M}^k) \lesssim \exp\left(-\lambda k^{1+\varepsilon}\right)$.

Proof. Let $q_0, q_1, \dots, q_M$ denote the offspring law for the branching process, so $q = q_0$. Then there exists a grid of i.i.d discrete random variables X_{ij} with $\mathbf{P}(X_{ij} =$

$k) = q_k$ such that

$$Z_{k+1} = \sum_{j=1}^{Z_k} X_{ij}.$$

Now consider the branching process $\{\tilde{Z}_k\}$ defined by setting $\tilde{Z}_0 = 1$, and

$$\tilde{Z}_{k+1} = \sum_{j=1}^{\tilde{Z}_k} \max(X_{ij}, 1).$$

We find $Z_k \leq \tilde{Z}_k$, and if $\tilde{M} = q + M$, then

$$\mathbf{E}(\tilde{Z}_{k+1} | \tilde{Z}_k) = \left(q + \sum_{k=1}^{N} k q_k \right) \tilde{Z}_k = (q + M)\tilde{Z}_k = \tilde{M}\tilde{Z}_k.$$

Most importantly for our purposes, $\{\tilde{Z}_k\}$ has zero chance of extinction. Theorem 5 of [1] implies that for such a supercritical branching process, there exists $\lambda > 0$ depending only on the offspring distribution of $\tilde{Z}$, such that if $\tilde{W}_k = \tilde{Z}_k / \tilde{M}^k$, and $\tilde{W} = \lim \tilde{W}_k$, then

$$\mathbf{P}(|\tilde{W} - \tilde{W}_k| \geq t) \lesssim \exp\left(-\lambda t^{2/3} \tilde{M}^{k/3} \right)$$

In particular, this means

$$\begin{aligned} \mathbf{P}\left(Z_k \geq (1 + \tilde{W})\tilde{M}^k \right) &\leq \mathbf{P}(\tilde{Z}_k \geq (1 + \tilde{W})\tilde{M}^k) \\ &\leq \mathbf{P}\left(|\tilde{W}\tilde{M}^k - \tilde{Z}_k| \geq \tilde{M}^k \right) \\ &= \mathbf{P}\left(|\tilde{W} - \tilde{W}_k| \geq 1 \right) \lesssim \exp\left(-\lambda \tilde{M}^{k/3} \right), \end{aligned}$$

Theorem 2 of [3] implies that as $t \to \infty$, there are constants C and $\varepsilon > 0$ depending only on the offspring distribution of $\tilde{Z}$ such that

$$-\log \mathbf{P}(\tilde{W} \geq t) \geq (C + o(1)) t^{1+\varepsilon}$$

In particular, there exists a small constant λ such that

$$\mathbf{P}(1+\tilde{W} \geq k) \leq \exp\left(-\lambda k^{1+\varepsilon}\right)$$

Applying a union bound gives

$$\begin{aligned}\mathbf{P}(Z_k \geq k\tilde{M}^k) &\leq \mathbf{P}(Z_k \geq (1+\tilde{W})\tilde{M}^k) + \mathbf{P}(1+\tilde{W} \geq k) \\ &\lesssim \exp\left(-\lambda \tilde{M}^{k/3}\right) + \exp\left(-\lambda k^{1+\varepsilon}\right) \lesssim \exp\left(-\lambda k^{1+\varepsilon}\right). \qquad \square\end{aligned}$$

For a set E, let $E_\delta = \{x : d(x,E) \leq \delta\}$ denote the δ thickened version of E.

Lemma 18. *There exists a constant B such that for any k, we can find lines $L_{k,1},\dots,L_{k,M}$ with $M \leq B \cdot 8^{kN}$ such that for any line L, there exists i such that $[0,1]^2 \cap L_{1/2^{kN+1}} \subset (L_{k,i})_{2^{-kN}}$.*

Proof. For each $(x,\theta) \in [0,1]^2 \times [0,1]$, let

$$L^{\theta,x} = \{x + te^{2\pi i\theta} : t \in \mathbf{R}\}$$

Note that

$$L^{\theta,x} \cap [0,1]^2 \subset \{x + te^{2\pi i\theta} : t \in [-2,2]\}$$

Any line intersecting $[0,1]^2$ is equal to $L^{\theta,x}$ for some θ and some x. For any k, we consider the set of $8^5 \cdot 8^{kN}$ points $E = (\mathbf{Z}/2^{kN+5})^3 \cap [0,1]^3$. For any (x_1,θ_1) and (x_2,θ_2), and $t \in [-2,2]$, we calculate that

$$|(x_1 + te^{2\pi i\theta_1}) - (x_2 + te^{2\pi i\theta_2})| \leq |x_1 - x_2| + 2\pi t|\theta_1 - \theta_2| \leq |x_1 - x_2| + 4\pi|\theta_1 - \theta_2|$$

Thus $L^{\theta_1,x_1} \cap [0,1]^2$ is contained in the $|x_1 - x_2| + 4\pi|\theta_1 - \theta_2|$ thickening of L^{θ_2,x_2}. For any line $L^{\theta,x}$, there exists $(\theta_0,x_0) \in E$ such that $|\theta - \theta_0| \leq 1/2^{kN+2}$ and $|x - x_0| \leq \sqrt{2}/2^{kN+5}$, and so $L^{\theta,x} \cap [0,1]^2$ is contained in the $\sqrt{2}/2^{kN+5} + 4\pi/2^{kN+5} \leq 14/2^{kN+5} \leq 1/2^{kN+1}$, and thus $L^{\theta,x}_{1/2^{kN+1}} \cap [0,1]^2 \subset L_{1/2^{kN}}$. Thus we may set $B = 8^5$, completing the proof. $\square$

Lemma 19. *Fix N. If $p > 1/2$, then almost surely, there exists a value k_0 such that if $k \geq k_0$, and L is any line, then*

$$\#\mathscr{B}^2_{2^{-kN}}(L_{2^{-kN}}) \leq k \cdot 10^k (2p)^{kN}$$

Proof. Fix N, and for any index k let $\delta_k = 1/2^{Nk}$. Given a line L, let I be a sidelength δ_k square intersecting the δ_k thickened line L_{δ_k}. Then L passes from one edge of I to another edge, and so if we split I into 4^N sidelength δ_{k+1} squares, then $L_{\delta_{k+1}}$ intersects at most $5 \cdot 2^N$ of these boxes. Conditioned on I being contained in C_{kN}, each of these subboxes occurs in $C_{(k+1)N}$ with probability p^N. We let $Z_k(L)$ denote the number of sidelength $1/2^{Nk}$ boxes in C_{Nk} intersecting L_{δ_k}. We now find $\tilde{Z}_k(L) \geq Z_k(L)$ by adding *exactly* $5 \cdot 2^N$ potential subboxes of each box at each subsequent stage, then $\tilde{Z}_k(L)$ is a branching process, whose offspring distribution is independant of L. Since each subbox is added with probability p^N, the extinction probability of $\tilde{Z}_k$ is $(1-p^N)^{5 \cdot 2^N}$. Also, $\mathbf{E}(\tilde{Z}_{k+1}(L) | \tilde{Z}_k(L)) = 5 \cdot (2p)^N \tilde{Z}_k$. Setting $M = 5(2p)^N$ and $q = (1-p^N)^{5 \cdot 2^N}$, Lemma 17 shows that there exists positive constants λ and ε, independant of L, such that if $\tilde{M} = M + q$,

$$\mathbf{P}\left(\tilde{Z}_k(L) \geq k\tilde{M}^k\right) \lesssim \exp(-\lambda k^{1+\varepsilon}).$$

Elementary bounds on the logarithm show that

$$5 \cdot 2^N \log(1-p^N) \leq \frac{-5(2p)^N}{2-p^N} \leq -5(2p)^N$$

so if $p > 1/2$,

$$q = (1-p^N)^{5 \cdot 2^N} \leq e^{-5(2p)^N} \leq 1 \leq 5(2p)^N$$

and this implies $\tilde{M} \leq 10(2p)^N$, so we have shown

$$\mathbf{P}(\tilde{Z}_k(L) \geq k \cdot 10^k (2p)^{kN}) \lesssim \exp(-\lambda k^{1+\varepsilon}).$$

This gives an upper bound on $\tilde{Z}_k(L)$ up to a superexponentially decaying term in

k.

For each k, Lemma 18 shows we can find lines $L_{k,1},\dots,L_{k,M}$, with $M \leq B \cdot 8^{kN}$ such that for any line L, there exists i such that $[0,1]^2 \cap L_{\delta_k/2} \subset (L_{k,i})_{\delta_k}$. Applying a union bound, we find that

$$\mathbf{P}\left(\text{there is } i \text{ such that } Z_k(L_{k,i}) \geq k \cdot 10^k (2p)^{kN}\right) \lesssim B \cdot 8^{kN} \exp(-ck^\lambda)$$

But since $\lambda > 1$, we therefore find

$$\sum_{k=1}^{\infty} \mathbf{P}\left(\text{there is } i \text{ such that } Z_k(L_{k,i}) \geq k \cdot 10^k (2p)^{kN}\right) < \infty$$

Applying the Borel-Cantelli lemma, we conclude that almost surely, it is eventually true for sufficiently large k that $Z_k(L_{k,i}) \leq k \cdot 10^k (2p)^{kN}$ for all i. In particular, the number of cubes intersecting $L_{\delta_k/2}$ for any line L is upper bounded by $k \cdot 10^k (2p)^{kN}$. $\square$

Lemma 20. *There exists a constant A such that if $I,J \in \mathscr{B}^2_{1/2^{kN}}[0,1]^2$, with $d(I,J) \geq A/2^{kN}$, then*

$$\bigcup \{L_{xy} : x \in I, y \in J\} \subset L_{1/2^{kN+1}}$$

where we recall that L_{xy} is the bisector of the points x and y.

Proof. Denote the bottom left vertices of the boxes I and J by $(N_I,M_I) \cdot 1/2^{kN}$ and $(N_J,M_J) \cdot 1/2^{kN}$, with $0 \leq N_I,M_I,N_J,M_J \leq 2^{kN}$. Let $\Delta_N = |N_I - N_J|$ and $\Delta_M = |M_I - M_J|$. We may assume for simplicity that $N_I \leq N_J$ and $M_I \leq M_J$. Since $d(I,J) \geq A/2^{kN}$,

$$(\Delta_N^2 + \Delta_M^2)^{1/2} \geq A$$

If we let θ_1 and θ_2 denote the minimal and maximal angle between the lines connecting pairs of points in I and J and the horizontal line, then

$$\begin{aligned}\sin(\theta_2-\theta_1) &= \sin(\theta_2)\cos(\theta_1)-\sin(\theta_1)\cos(\theta_2)\\ &= \frac{(\Delta_M+1)(\Delta_N+1)}{\Delta_N^2+\Delta_M^2}-\frac{\Delta_N\Delta_M}{\Delta_N^2+\Delta_M^2}\\ &\le \frac{\Delta_N+\Delta_M+1}{\Delta_N^2+\Delta_M^2}\le 1/A\end{aligned}$$

Thus $\theta_2-\theta_1 \le 2/A$. Thus all the bisectors are contained in a

$$\sqrt{2}/2^{kN}+4\pi(2/A)\le 1/2^{kN}$$

□

Theorem 9. *If $p>1/2$, then almost surely, the set Z_0 associated with C has lower Minkowski dimension at most $5-3\log_2(1/p)$.*

Proof. Almost surely, Lemma 16 shows that for sufficiently large k we can cover C_{kN} by $O(2^{\alpha kN})$ boxes $I_1,\dots,I_M \in \mathscr{B}^2_{2^{-kN}}$, with $\alpha = 2-\log_2(1/p)$. Lemma 19 shows that if k is sufficiently large, then the $1/2^{kN+1}$ thickened line L intersects at most $k\cdot 10^k(2p)^{kN}$ squares in $\mathscr{B}^2_{2^{-kN}}$. There exists a constant A such that for any i and j, if $d(I_i,I_j)>A2^{-kN}$, then there exists a line L_{ij} such that as x ranges over all points in I_i, and y over all points in I_j,

$$\bigcup_{x,y} B_{xy}\cap[0,1]^2\subset (L_{ij})_{\delta_k/2}$$

This means that $Z_0\cap(I_i\times I_j\times[0,1]^2)\subset I_i\times I_j\times(L_{ij})_{\delta_k/2}$, and the last paragraph implies that $(L_{ij})_{\delta_k/2}$ is covered by $k\cdot 10^k(2p)^{kN}$ cubes in $\mathscr{B}^2_{l_k}[0,1]^2$. On the other hand, if $d(I_i,I_j)<A\delta_k$, we can apply the obvious bound that $Z_0\cap(I_i\times I_j\times[0,1]^2)\subset I_i\times I_j\times C_{kN}$, and C_{kN} is coverable by $O(2^{\alpha kN})$ boxes. Thus we conclude

that almost surely, for sufficiently large k,

$$\begin{aligned}
\#\mathscr{B}^6_{\delta_k}(Z_0) &= \sum_{i,j} \mathscr{B}^6_{\delta_k}(Z_0 \cap (I_i \times I_j \times [0,1]^2)) \\
&\leq \sum_i \left(\sum_{d(I_i,I_j) \leq A\delta_k} \#\mathscr{B}^6_{\delta_k}(Z_0 \cap (I_i \times I_j \times (L_{ij})_{\delta_k/2})) \right) \\
&\quad + \left(\sum_{d(I_i,I_j) > A\delta_k} \#\mathscr{B}^6_{\delta_k}(Z_0 \cap (I_i \times I_j \times C_{kN})) \right) \\
&\leq \sum_i \left(\sum_{d(I_i,I_j) \leq A\delta_k} \#\mathscr{B}^2_{\delta_k}((L_{ij})_{\delta_k/2}) \right) + \left(\sum_{d(I_i,I_j) > A\delta_k} \#\mathscr{B}^2_{\delta_k}(C_{kN}) \right) \\
&\lesssim \sum_i 2^{\alpha kN} \cdot k10^k (2p)^{kN} + 2^{\alpha kN} \lesssim 2^{2\alpha kN} k10^k (2p)^{kN}
\end{aligned}$$

Thus we conclude that almost surely,

$$\begin{aligned}
\dim_{\mathbf{M}}(Z_0) &= \lim_{k\to\infty} \frac{\log\left(\#\mathscr{B}^6_{\delta_k}(Z_0)\right)}{\log(1/\delta_k)} \\
&\leq \lim_{k\to\infty} \frac{\log(2^{2\alpha kN} k10^k (2p)^{kN}) + O(1)}{\log(2^{kN})} \\
&= \lim_{k\to\infty} \frac{2\alpha kN \log(2) + k\log 10 + kN\log(2p)}{kN\log(2)} \\
&= 2\alpha + 3/N + 1 - \log_2(1/p) \\
&= 5 - 3\log_2(1/p) + 3/N
\end{aligned}$$

Taking $N \to \infty$, we obtain the required result. $\square$

TODO: does a projection of C have non-empty interior almost surely?

Corollary. *If $p = 1/2^{1-\varepsilon}$, then conditioned on C having Minkowski dimension $1+\varepsilon$, we can find a subset X of C avoiding isoceles triangles with* $\dim_{\mathbf{H}}(X) \geq 1/2 - (3/2)\varepsilon$.

4.6 Relation to Literature, and Future Work

Our result is part of a growing body of work finding general methods to find sets avoiding patterns. The main focus of this section is comparing our method to the two other major results in the literature. [6] constructs sets with dimension $k/(n-1)$ avoiding the zero sets of rank k C^1 functions. In [11], sets of dimension d/l are constructed avoiding a degree l algebraic hypersurface specified by a polynomial with rational coefficients.

We can view our result as a robust version of Pramanik and Fraser's result. Indeed, if we try and avoid the zero set of a C^1 rank k function, then we are really avoiding a dimension $dn-k$ dimensional manifold. Our method gives a dimension

$$\frac{dn-(dn-k)}{n-1} = \frac{k}{n-1}$$

set, which is exactly the result obtained in [6].

That our result generalizes [6] should be expected because the technical skeleton of our construction is heavily modeled after their construction technique. Their result also reduces the problem to a discrete avoidance problem. But they *deterministically* select a particular side length S cube in every side length R cube. For arbitrary Z, this selection procedure can easily be exploited for a particularly nasty Z, so their method must rely on smoothness in order to ensure some cubes are selected at each stage. Our discrete avoidance technique was motivated by other combinatorial optimization problems, where adding a random quantity prevents inefficient selections from being made in expectation. This allows us to rely purely on combinatorial calculations, rather than employing smoothness, and greatly increases the applicability of the sets Z we can apply our method to. Furthermore, it shows that the underlying problem is robust to changes in dimension; slightly 'thickening' Z only slightly perturbs the dimension of X.

One useful technique in [6], and its predecessor [9], is the use of a Cantor set construction 'with memory'; a queue in the construction algorithm for their sets allows storage of particular discrete versions of the problem to be stored, and then retrieved at a much later stage of the construction process. This enables them to 'separate' variables in the discrete version of the problem, i.e. instead of forming a single set F from a set E, they from n sets $F_1, \dots, F_n$ from disjoint sets $E_1, \dots, E_n$. The fact that our result is more general, yet does not rely on this technique is an interesting anomaly. An obvious advantage is that the description of the technique is much more simple. But an additional advantage is that we can attack 'one scale' of the problem at a time, rather than having to rely on stored memory from a vast number of steps before the current one. We believe that we can exploit the single scale approach to the problem to generalize our theorem to a much wider family of 'dimension α' sets Z, which we plan to discuss in a later paper.

As a generalization of the result in [6], our result has the same issues when compared to the result of [11]. When the parameter n is large, the dimension of our result suffers greatly, as with the n fold sum application in the last section. Furthermore, our result can't even beat trivial results if Z is almost full dimensional, as the next example shows.

Example. *Consider an α dimensional set of angles Y, and try and find $X \subset \mathbf{R}^2$ such that the angle formed from any collection of three points in X avoids Y. If we form the set*

$$Z = \left\{ (x,y,z) : \textit{There is } \theta \in Y \textit{ such that } \frac{(x-y)\cdot(x-z)}{|x-y||x-z|} = \cos\theta \right\}$$

Then we can find X avoiding Z. But one calculates that Z has dimension $3d+\alpha-1$, which means X has dimension $(1-\alpha)/2$. Provided the set of angles does not contain π, the trivial example of a straight line beats our result.

Nonetheless, we still believe our method is a useful inspiration for new techniques in the 'high dimensional' setting. Most prior literature studies sets Z only implicitly as zero sets of some regular function f. The features of the function f imply geometric features of Z, which are exploited by these results. But some geometric features are not obvious from the functional perspective; in particular, the fractional dimension of the zero set of f is not an obvious property to study.

We believe obtaining methods by looking at the explicit geometric structure of Z should lead to new techniques in the field, and we already have several ideas in mind when Z has geometric structure in addition to a dimension bound, which we plan to publish in a later paper.

We can compare our randomized selection technique to a discrete phenomenon that has been recently noticed, for instance in [2]. There, certain combinatorial problems can be rephrased as abstract problems on hypergraphs, and one can then generalize the solutions of these problems using some strategy to improve the result where the hypergraph is sparse. Our result is a continuous analogue of this phenomenon, where sparsity is represented by the dimension of the set Z we are trying to avoid. One can even view Lemma 1 as a solution to a problem about independent sets in hypergraphs. In particular, we can form a hypergraph by taking the cubes $\mathscr{B}_s^d(E)$ as vertices, and adding an edge $(I_1, \dots, I_n)$ between n distinct cubes $I_k \in \mathscr{B}_s^d(E)$ if $I_1 \times \dots \times I_n$ intersects Z_s. An independent set of cubes in this hypergraph corresponds precisely to a set F with F^n disjoint except on a discretization of the diagonal. And so Lemma 1 really just finds a ‘uniformly chosen’ independent set in a sparse graph. Thus we really just applied the discrete phenomenon at many scales to obtain a continuous version of the phenomenon.

Chapter 5

Extensions to Low Rank Configurations

5.1 Boosting the Dimension of Pattern Avoiding Sets by Low Rank Coordinate Changes

We now consider finding subsets of $[0,1]$ avoiding solutions to the equation $y = f(Tx)$, where T is a rank k linear transformation with integer coefficients with respect to standard coordinates, and f is real-valued and Lipschitz continuous. Fix a constant A bounding the operator norm of T, in the sense that $|Tx| \leq A|x|$ for all $x \in \mathbf{R}^n$, and a constant B such that $|f(x+y) - f(x)| \leq B|y|$ for all x and y for which the equation makes sense (if f is C^1, this is equivalent to a bound $\|\nabla f\|_\infty \leq B$). Consider sets $J_0, J_1, \dots, J_n, \subset [0,1]$, which are unions of intervals of length $1/M$, with startpoints lying on integer multiples of $1/M$. The next theorem works as a 'building block lemma' used in our algorithm for constructing a set avoiding solutions to the equation with Hausdorff dimension k and full Minkowski dimension.

Theorem 10. *For infinitely many integers N, there exists $S_i \subset J_i$ avoiding solutions to $y = f(Tx)$ with $y \in S_0$ and $x_n \in S_n$, such that*

- *For $n \neq 0$, if we decompose each J_i into length $1/N$ consecutive intervals, S_i contains an initial portion $\Omega(1/N^k)$ of each length $1/N$ interval. This part of the decomposition gives the Hausdorff dimension $1/k$ bound for the set we will construct.*

- *If we decompose J_i into length $1/N$ intervals, and then subdivide these intervals into length $\Omega(1/N^k)$ intervals, then S_0 contains a subcollection of these $1/N^k$ intervals which contains a total length $\Omega(1/N)$ of a fraction $1-1/M$ of the length $1/N$ intervals. This property gives that our resultant set will have full Minkowski dimension.*

The implicit constants in these bounds depend only on A, B, n, and k.

Proof. Split each interval of J_a into length $1/N$ intervals, and then set

$$\mathbf{A} = \{x : x_a \text{ is a startpoint of a } 1/N \text{ interval in } J_a\}$$

Since the startpoints of the intervals are integer multiples of $1/N$, $T(\mathbf{A})$ is contained with a rank k sublattice of $(\mathbf{Z}/N)^m$. The operator norm also guarantees $T(\mathbf{A})$ is contained within the ball B_A of radius A in $\mathbf{R}^m$. Because of the lattice structure of the image, $|x-y| \gtrsim_n 1/N$ for each distinct pair $x,y \in T(\mathbf{A})$. For any R, we can cover $\Sigma \cap B_A$ by $O_{n,k}((A/R)^k)$ balls of radius R. If $R \gtrsim_n 1/N$, then each ball can contain only a single element of $T(\mathbf{A})$, so we conclude that $|T(\mathbf{A})| \lesssim_{n,k} (AN)^k$. If we define the set of 'bad points' to be

$$\mathbf{B} = \{y \in [0,1] : \text{there is } x \in \mathbf{A} \text{ such that } y = f(T(x))\}$$

Then

$$|\mathbf{B}| = |f(T(\mathbf{A}))| \leq |T(\mathbf{A})| = O_{A,n,k}(N^k)$$

For simplicity, we now introduce an integer constant $C_0 = C_0(A,n,k,M)$ such that $|\mathbf{B}| \leq (C_0/M^2)N^k$. We now split each length $1/M$ interval in J_0 into length $1/N$ intervals, and filter out those intervals containing more than $C_0 N^{k-1}$ elements of

$\mathbf{B}$. Because of the cardinality bound we have on $\mathbf{B}$ there can be at most N/M^2 such intervals, so we discard at most a fraction $1/M$ of any particular length $1/M$ interval in J_0. If we now dissect the remaining intervals into $4C_0N^{k-1}$ intervals of length $1/4C_0N^k$, and discard any intervals containing an element of $\mathbf{B}$, or adjacent to such an interval, then the remaining such intervals I satisfy $d(I,\mathbf{B}) \geq 1/4C_0N^k \gtrsim_{A,n,M,k} (1/N^k)$, and because of our bound on the number of elements of $\mathbf{B}$ in these intervals, there are at least C_0N^{k-1} intervals remaining, with total length exceeding $C_0N^{k-1}/4C_0N^k = \Omega(1/N)$. If f is C^1 with $\|\nabla f\|_\infty \leq B$, or more generally, if f is Lipschitz continuous of magnitude B, then

$$|f(Tx) - f(Tx')| \leq AB|x - x'|$$

and so we may choose $S_i \subset J_i$ by thickening each startpoint $x \in J_i$ to a length $O(1/N^k)$ interval while still avoiding solutions to the equation $y = f(T(x))$. $\square$

Remark. *If T is a rank k linear transformation with rational coefficients, then there is some number a such that aT has integer coefficients, and then the equation $y = f(Tx)$ is the same as the equation $y = f_0((aT)(x))$, where $f_0(x) = f(x)/a$. Since f_0 is also Lipschitz continuous, we conclude that we still get the dimension $1/k$ bound if T has rational rather than integral coefficients. More generally, this trick shows the result applies unperturbed if all coefficients of T are integer multiples of some fixed real number. More generally, by varying the lengths of our length $1/N$ decomposition by a constant amount, we can further generalize this to the case where each column of T are integers multiples of some fixed real number.*

Remark. *To form $\mathbf{A}$, we take startpoints lying at equal spaced $1/N$ points. However, by instead taking startpoints at varying points in the length $1/N$ intervals, we might be able to make points cluster more than in the original algorithm. Maybe the probabalistic method would be able to guarantee the existence of a choice of startpoints whose images are tightly clustered together.*

Remark. *Since the condition $y = f(Tx)$ automatically assumes a kind of 'non-vanishing derivative' condition on our solutions, we do not need to assume the* *regularity of f, and so the theorem extends naturally to a more general class of functions than Rob's result, i.e. the Lipschitz continuous functions.*

Using essentially the same approach as the last argument shows that we can avoid solutions to $y = f(Tx)$, where y and x are now vectors in some $\mathbf{R}^m$, and T has rank k. If we consider unions of $1/M$ cubes $J_0, \dots, J_n$. If we fix startpoints of each x_k forming lattice spaced apart by $\Omega(1/N)$, and consider the space $\mathbf{A}$ of products, then there are $O(N^k)$ points in $T(\mathbf{A})$, and so there are $O(N^k)$ elements in $\mathbf{B}$. We now split each $1/M$ cube in J_0 into length $1/N$ cubes, and discard those cubes which contain more than $O(N^{k-m})$ bad points, then we discard at most $1 - 1/M$ of all such cubes. We can dissect the remaining length $1/N$ cubes into $O(N^{k-m})$ length $\Omega(1/N^{k/m})$ cubes, and as in the previous argument, the cubes not containing elements of $\mathbf{B}$ nor adjacent to an element have total volume $\Omega(1/N^m)$, which we keep. The startpoints in the other intervals T_i may then be thickened to a length $\Omega(1/N^{m/k})$ portion while still avoiding solutions. This gives a set with full Minkowski dimension and Hausdorff dimension m/k avoiding solutions to $y = f(Tx)$. (I don't yet understand Minkowski dimension enough to understand this, but the techniques of the appendix make proving the Hausdorff dimension $1/k$ bound easy)

5.2 Extension to Well Approximable Numbers

If the coefficients of the linear transformation T in the equation $y = f(Tx)$ are non-rational, then the images of startpoints under the action of T do not form a lattice, and so points may not overlap so easily when avoiding solutions to the equations $y = f(Tx)$. However, if T is 'very close' to a family of rational coefficient linear transformations, then we can show the images of the startpoints are 'very close' to a lattice, which will still enable us to find points avoiding solutions by replacing the direct combinatorial approach in the argument for integer matrices with a covering argument.

Suppose that T is a real-coefficient linear transformation with the property that for each coefficient x there are infinitely many rational numbers p/q with

$|x - p/q| \leq 1/q^{\alpha}$, for some fixed α. For infinitely many K, we can therefore find a linear transformation S with coefficients in $\mathbf{Z}/K$ with each coefficient of T differing from the corresponding coefficient in S by at most $1/K^{\alpha}$. Then for each x, we find

$$\|(T-S)(x)\|_{\infty} \leq (n/K^{\alpha})\|x\|_{\infty}$$

If we now consider $T_0, \dots, T_n$, splitting $T_1, \dots, T_n$ into length $1/N$ intervals, and considering $\mathbf{A}$ as in the last section, then $S(\mathbf{A})$ lie in a k dimensional sublattice of $(\mathbf{Z}/KN)^m$, hence containing at most $(2A)^k(KN)^k = O_{T,n}((KN)^k)$ points. By our error term calculation of $T-S$, the elements of $T(\mathbf{A})$ are contained in cubes centered at these lattice points with side-lengths $2n/K^{\alpha}$, or balls centered at these points with radius $n^{3/2}/K^{\alpha}$. If $\|\nabla f\| \leq B$, then the images of the radius $n^{3/2}/K^{\alpha}$ balls under the action of f are contained in length $Bn^{3/2}/K^{\alpha}$ intervals. Thus the total length of the image of all these balls under f is $(Bn^{3/2}/K^{\alpha})(2A)^k(KN)^k = (2A)^k Bn^{3/2}K^{k-\alpha}N^k$. If $k < \alpha$, then we can take K arbitrarily large, so that there exists intervals with $\text{dist}(I, \mathbf{B}) = \Omega_{A,k,M}(1/N^k)$. But I believe that, after adding the explicit constants in, we cannot let $k = \alpha$.

Remark. *One problem is that, if T has rank k, we might not be able to choose S to be rank k as well. Is this a problem? If T has full rank, then the set of all such matrices is open so if T and S are close enough, S also has rank k, but this need not be true if T does not have full rank.*

Example. *If T has rank 1, then Dirichlet's theorem says that every irrational number x can be approximated by infinitely many p/q with $|x - p/q| < 1/q^2$, so every real-valued rank 1 linear transformation can be avoided with a dimension one bound.*

5.3 Equidistribution and Real Valued Matrices

If T is a non-invertible matrix containing irrational coefficients, then the values Tx, for $x \in \mathbf{Z}^n$, do not form a lattice, and therefore we cannot use the direct combinatorial arguments of the past section to obtain the decomposition lemma. However, without loss of generality, we can write $T(x) = S(x_1) + U(x_2)$, where $x = (x_1, x_2)$, $x_1 \in \mathbf{R}^k$, $x_2 \in \mathbf{R}^{n-k}$, and S has full rank k. Then S is an embedding of $\mathbf{R}^k$ into $\mathbf{R}^n$, so $\Gamma = S((\mathbf{Z}/N)^k)$ forms a lattice with points spaced apart by a distance on the order of $\Omega(1/N)$. Since T has rank k, the image of U is contained within the image of S. We let $\mathbf{T}$ denote the torus obtained by quotienting the k dimensional subplane forming the image of T by Γ. Then the image of S in $\mathbf{T}$ is contained within an α dimensional subtorus of $\mathbf{T}$. Note that $\alpha = 0$ precisely when T still has rank k over the rational numbers, so that in a suitable basis T is an integer valued matrix. If $\alpha = 1$, then by an appropriate scaling in the values x_2 we can still make the values of S lie at lattice points, which should give a Hausdorff dimension one set. When $\alpha = 2$, we run into problems.

If $\pi : \mathbf{R}^k \to T$ is the homomorphism obtained by composing the quotient map onto the torus with the linear map T, then $\pi(x_1, 0) = 0$ for all $x_1 \in (\mathbf{Z}/N)^k$. On the other hand, Ratner's theorem implies that for each $x_2 \in \mathbf{R}^{n-k}$, there is some M such that the sequence $\pi(0, nMx_2) = Mn\pi(0, x_2)$ is equidistributed on a subtorus of T. Equidistribution may be useful in extending the rational matrix result to all real matrices with some dimension loss, since a matrix is rational if and only if $nM\pi(0, x_2)$ is equidistributed on a zero dimensional lattice – it may ensure that points are closely clustered to lattice points.

Lets consider the simplest case, where $n = k+1$, so $x_2 \in \mathbf{R}$. Consider our setup, with intervals $J_0, \dots, J_n$, and an equation $y = f(Tx)$, where f is Lipschitz with Lipschitz norm bounded by B. Now if $T = S + U$, then $S(\mathbf{Z}^k)$ forms a rank k lattice, and $U(\mathbf{Z})$ equidistributes over an α dimensional subtorus of the torus generated over the lattice. In particular, since the set $\mathbf{Z} \cap NJ_n$ contains $\Omega(N|J_n|) = \Omega(N)$ consecutive points, for any ε and suitably large N, $\mathbf{Z} \cap NJ_n$ contains $\Omega(Nr^\alpha)$ points x such that $S(x)$ is within a distance r from a lattice point, for any r. Dividing by N tells us that we have $O(Nr^\alpha)$ points x in $\mathbf{Z}/N \cap J_n$ such that $U(x)$ is at a distance r/N from a lattice point in $S((\mathbf{Z}/N)^k)$. If $r = 1/N^\beta$, then we have $O(N^{1-\alpha\beta})$ points at a distance $1/N^{\beta+1}$ from a lattice point. There are

$O(N^k)$ points in the lattice, and so provided that $k < 1+\beta$, we can find a large subset avoiding the images of these startpoints, and we should be able to thicken the startpoints to lengt h $\Omega(1/N^k)$ intervals, hence we should expect the set we construct to have Hausdorff dimension $1/k(k-1)$ if this process is repeated to construct our solution avoiding set.

5.4 Applications of Low Rank Coordinate Changes

Example. *Our initial exploration of low rank coordinate changes was inspired by trying to find solutions to the equation*

$$y - x = (u-w)^2$$

Our algorithm gives a Hausdorff dimension $1/2$ set avoiding solutions to this equation. This equals Mathé's result. But this dimension for us now depends on the shifts involved in the equation, not on the exponent, so we can actually avoid solutions to the equation

$$y - x = (u-w)^n$$

for any n, in a set of Hausdorff dimension $1/2$. More generally, if X is a set, then given a smooth function f of n variables, we can find a set X of Hausdorff dimension $1/n$ such that there is no $x \in X$, and $y_1,\dots,y_n \in X - X$ such that $x = f(y_1,\dots,y_n)$. This is better than the $1/2n$ bound that is obtained by Malabika and Fraser's result.

Example. *For any fixed m, we can find a set $X \subset \mathbf{R}^n$ of full Hausdorff dimension which contains no solutions to*

$$a_1x_1 + \cdots + a_nx_n = 0$$

for any rational numbers a_n which are not all zero. Since Malabika/Fraser's technique's solutions are bounded by the number of variables, they cannot let $n \to \infty$ to obtain a linearly independant set over the rational numbers. But since the Haus-

dorff dimension of our sets now only depends on the rank of T, rather than the total number of variables in T, we can let $n \to \infty$ to obtain full sets linearly independant over the rationals. More generally, for any Lipschitz continuous function $f : \mathbf{R} \to \mathbf{R}$, we can find a full Hausdorff dimensional set such that there are no solutions

$$f(a_1 x_1 + \cdots + a_n x_n, y)$$

for any n, and for any rational numbers a_n that are not all zero.

Example. *The easiest applications of the low rank coordinate change method are probably involving configuration problems involving pairwise distances between m points in $\mathbf{R}^n$, where $m \ll n$, since this can best take advantage of our rank condition. Perhaps one way to encompass this is to avoid m vertex polyhedra in n dimensional space, where $m \ll n$. In order to distinguish this problem from something that can be solved from Mathé's approach, we can probably find a high dimensional set avoiding m vertex polyhedra on a parameterized n dimensional manifold, where $m \ll n$. There is a result in projective geometry which says that every projectively invariant property of m points in $\mathbf{RP}^d$ is expressible as a function in the $\binom{m}{d}$ bracket polynomials with respect to these m points. In particular, our result says that we can avoid a countable collection of such invariants in a dimension $1/\binom{m}{d}$ set. This is a better choice of coordinates than Euclidean coordinates if $\binom{m}{d} \leq m$. Update: I don't think this is ever the case.*

Remark. *Because of how we construct our set X, we can find a dimension $1/k$ set avoiding solutions to $y = f(Tx)$ for all rank k rational matrices T, without losing any Hausdorff dimension. Maybe this will help us avoid solutions to more general problems?*

Example. *Given a smooth curve Γ in $\mathbf{R}^n$, can we find a subset E with high Hausdorff dimension avoiding isoceles triangles. That is, if the curve is parameterized by $\gamma : [0,1] \to \mathbf{R}^n$, can we find $E \subset [0,1]$ such that for any t_1, t_2, t_3, $\gamma(t_1)$, $\gamma(t_2)$, and $\gamma(t_3)$ do not form the vertices of an isoceles triangle. This is, in a sense, a non-linear generalization of sets avoiding arithmetic progressions, since if Γ is a* *line, an isoceles triangle is given by arithmetic progressions. Assuming our curve is simple, we must avoid zeroes of the function*

$$|\gamma(t_1) - \gamma(t_2)|^2 = |\gamma(t_2) - \gamma(t_3)|^2$$

If we take a sufficiently small segment of this curve, and we assume the curve has non-zero curvature on this curve, we can assume that $t_1 < t_2 < t_3$ in our dissection method.

If the coordinates of γ are given by polynomials with maximum degree d, then the equation

$$|\gamma(t_1) - \gamma(t_2)|^2 - |\gamma(t_2) - \gamma(t_3)|^2$$

is a polynomial of degree $2d$, and so Mathé's result gives a set of dimension $1/2d$ avoiding isoceles triangles. In the case where Γ is a line, then the function $f(t_1,t_2,t_3) = \gamma(t_1) + \gamma(t_3) - 2\gamma(t_2)$ avoids arithmetic progresions, and Mathé's result gives a dimension one set avoiding such progressions. Rob and Malabika's algorithm easily gives a set with dimension $1/2$ for any curve Γ. Our algorithm doesn't seem to be able to do much better here.

Example. *What is the largest dimension of a set in Euclidean space such that for any value λ, there is at most one pair of points x,y in the set such that $|x-y| = \lambda$.*

Example. *What is the largest dimension of a set which avoids certain angles, i.e. for which a triplet x,y,z avoids certain planar configurations.*

Example. *A set of points $x_0,\dots,x_d \in \mathbf{R}^d$ lie in a hyperplane if and only if the determinant formed by the vectors $x_n - x_0$, for $n \in \{1,\dots,d\}$, is zero. This is a degree d polynomial, hence Mathé's result gives a dimension one set with no set of $d+1$ points lying in a hyperplane. On the other hand, a theorem of Mattila shows that every analytic set E with dimension exceeding one contains $d+1$ points in a hyperplane. Can we generalize this to a more general example avoiding points on a rotational, translation invariant family of manifolds using our results?*

Example. *Given a set F not containing the origin, what is the largest Hausdorff dimension of a set E such that for any for any distinct rational $a_1, \dots, a_N$, the sum $a_1 E + \cdots + a_N E$ does not contain any elements of F. Thus the vector space over the rationals generated by E does not contain any elements of F. We can also take the non-linear values $f(a_1 E + \cdots + a_N E)$ avoiding elements of F. F must have non-empty interior for the problem to be interesting. Then can we find a smooth function f with non-nanishing derivative which vanishes over F, or a family of smooth functions with non-vanishing derivative around F.*

5.5 Idea: Generalizing This Problem to low rank smooth functions

Suppose we are able to find dimension $1/k$ sets avoiding configurations $y = g(f(x))$, where f is a smooth function from $\mathbf{R}^n \to \mathbf{R}^m$ of rank k. Then given any function $g(f(x))$, where f has rank k, if $g(f(x)) = 0$, then the implicit function theorem guarantees that there is a cover U_α and functions $h_\alpha : U^k \to \mathbf{R}^{n-k}$ such that for each if $g(f(x)) = 0$, for $x \in U_\alpha$, then there is a subset of k indices I such that $x_{I^c} = h_\alpha(x_I)$.

Then given any function $f(x)$ with rank k, we can use the implicit function theorem to find sets U_α, indices n_α, and functions g_α such that if $f(x) = 0$, for $x \in U_\alpha$, then $x_{n_\alpha} = g_\alpha(x_1, \dots, \widehat{x_{n_\alpha}}, \dots, x_n)$. Thus we need only avoid this type of configuration to avoid configurations of a general low rank function. If we don't believe that we are able to get dimension $1/(k-1)$ sets for rank k configurations, then we shouldn't be able to find sets of Hausdorff dimension $1/k$ avoiding configurations of the form $y = f(x)$, where f has rank k.

Remark. *If the functions g_α are only partially defined, this makes the problem easier than if the functions were globally defined, because the constraint condition is now smaller than the original constraint.*

Remark. *This would solve our problem of avoiding $y - x = (u - v)^2$, since if*

$f(x,y,u,v) = y - x - (u-v)^2$, *then*

$$\nabla f$$

5.6 Idea: Algebraic Number Fields

If $\mathbf{Q}(\omega)$ is a quadratic extension of the rational numbers, then the ring of integers in this field form a lattice. Perhaps we can use this to generalize our approach to avoiding configurations $y = f(Tx)$, where all coefficients of the matrix T lie in some common quadratic extension of the rational numbers.

5.7 A Scheme for Avoiding Configurations

Mathé's result can be reconfigured in terms of a building block strategy for implementation in our algorithm.

Theorem 11. *Let f be a polynomial of degree m, and consider unions of length $1/M$ intervals $T_0, \dots, T_d \subset [0,1]$, with rational start-points. If $\partial_0 f$ is non-vanishing on $T_0 \times \cdots \times T_d$, then there exists arbitrarily large integers N and a constant C not depending on N and sets $S_n \subset T_n$ such that*

- *$f(x) \neq 0$ for $x \in S_0 \times \cdots \times S_d$.*
- *If $T_0, \dots, T_d$ are split into length $1/N$ intervals, then S_n contains a length C/N^d region of each interval.*

Proof. Without loss of generality (by subdividing the initial intervals), let M be the greatest common divisor of all of the startpoints of the intervals in T_n. Divide each interval T_n into length $1/N$ intervals, and let $\mathbf{A} \subset (\mathbf{Z}/N)^d$ be the cartesian product of all startpoints of these length $1/N$ intervals. Since f has degree m, $f(\mathbf{A}) \subset \mathbf{Z}/N^m$. If $A_0 \leq |\partial_0 f| \leq A_1$ on $T_0 \times \cdots \times T_d$, then for any $a \in \mathbf{A}$, and δ_0, there exists δ_1 between 0 and δ_0 for which

$$|f(a + \delta_0 e_0)| - f(a)| = \delta_0 |(\partial_0 f)(a + \delta_1)|$$

If K is fixed such that $A_1 \leq (K-1)A_0$, so that we can choose

$$\frac{1/K}{A_0 N^m} \leq \delta_0 \leq \frac{(1-1/K)}{A_1 N^m}$$

Then

$$\frac{1/K}{N^m} \leq |f(a+\delta_0 e_0) - f(a)| \leq \frac{1-1/K}{N^m}$$

Thus $d(f(\mathbf{A}+\delta_0 e_0), \mathbf{Z}/N^m) \geq 1/KN^m$. Thus if we thicken the coordinates of $\mathbf{A}+\delta_0$ to intervals of length $O(1/N^m)$, then we obtain sets $S_0, \dots, S_n$ avoiding solutions. $\square$

TODO: CAN WE USE THE COMBINATORIAL NULLSTELLENSATZ TO COME UP WITH AN ALTERNATE BUILDING BLOCK LEMMA FOR ARBITRARY FIELDS?

5.8 Square Free Sets

We now look at avoiding solutions to the equation $x-y=(u-v)^2$. We consider two sets I and J. Suppose that we can select a subset $\mathbf{S}$ from $\mathbf{Z} \cap N^2 I$ such that if $x, y \in \mathbf{S}$ are distinct, $x-y$ is not a perfect square, and $|\mathbf{S}| \gtrsim |\mathbf{Z} \cap N^2 I|^\alpha$. Then for any distinct $u, v \in \mathbf{Z} \cap NJ$, $(u-v)^2 \not\in \mathbf{S} - \mathbf{S}$. But this means that if we thicken the points in $\mathbf{S}/N^2$ to length $O(1/N^2)$ intervals, and the points in $\mathbf{Z}/N \cap J$ into length $O(1/N)$ intervals, then the resultant set will avoid solutions to $x-y=(u-v)^2$. This should give a dimension α set.

Chapter 6

Fourier Dimension and Fractal Avoidance

In the last few chapters, we have discovered that the presence of many configurations is not guaranteed by large Hausdorff dimension. We now turn to an analysis of a different dimensional quantity which in many cases gives much more structure than Hausdorff dimension.

in many problems involving arithmetic structure,

6.1 Schmerkin's Method

6.2 Tail Bounds on Cube Selection

The aim of this section is to obtain tail bounds on the total number of intersections between the random set we select in the discrete selection schema, with the goal of obtaining a Fourier dimension bound on the set we construct. Consider the random process generating our interval selection scheme. We have three fixed lengths l, r, and s. We start with a set E, which is a non-empty union of cubes in $\mathscr{B}_l^d$, and a set G, which is a union of cubes in $\mathscr{B}_s^{dn}$. For each cube $J \in \mathscr{B}_r^d(E)$, we select a single random element of $\mathscr{B}_s^d(J)$, denoted J_I. We let $U = \bigcup J_I$. The aim of this section is to obtain tail bounds on the size of the random set

$$\mathscr{K}(U) = \{K \in \mathscr{B}_s^{dn}(G) : K \in U^n, K \text{ strongly non-diagonal}\}.$$

Here, this is obtained by applying the statistical decoupling techniques of Bourgain and Tzafriri.

To understand the distribution of this random variable, it will be convenient to simplify notation. We write the random variable $\mathbf{I}(I_J = I_0)$ as X_{J,I_0}. We then set

$$Z = \#\mathscr{K}(U) = \sum_{J_1,\dots,J_n} \sum_K X_{J_i,K_i},$$

where $J_1,\dots,J_n$ ranges over all distinct choices of n elements from $\mathscr{B}_l^d(E)$, and $K = K_1 \times \cdots \times K_n$ ranges over all elements of $\mathscr{B}_l^{dn}(G) \cap \mathscr{B}_s^{dn}(J_1 \times \cdots \times J_n)$. The random cubes $\{I_J\}$ form an independant family, which should make the distribution easier to bound. But Z is a multiplicative linear combination of this independant family. Random variables of this form as known as a *chaos*.

The idea behind statistical decoupling is to obtain tail bounds on the chaos Z by studying the tail bounds on the modified random variable

$$W = \sum_{J_1,\dots,J_n} \sum_K \left(\prod_{i=1}^n X_{J_i,K_i}^i \right),$$

where $X^1,\dots,X^n$ are independant and identically distributed copies of X.

Lemma 21. *Let $F : [0,\infty) \to [0,\infty)$ be a monotone, convex function. Then*

$$\mathbf{E}(F(Z)) \leq \mathbf{E}(F(n^n W)).$$

Proof. Consider a partition $\mathscr{J}_1,\dots,\mathscr{J}_n$ selected at random from all possible partitions of $\mathscr{B}_s^d(E)$ into n classes. Then, for any distinct set of intervals $J_1,\dots,J_n \in$

$\mathscr{B}_s^d(E)$,

$$\mathbf{P}(J_1 \in \mathscr{J}_1, \ldots, J_n \in \mathscr{J}_n) = \mathbf{P}(J_1 \in \mathscr{J}_1) \ldots \mathbf{P}(J_n \in \mathscr{J}_n) = 1/n^n. \tag{6.1}$$

Thus if we consider

$$Z' = \sum_{J_1 \in \mathscr{J}_1, \ldots, J_n \in \mathscr{J}_n} \sum_K \left(\prod_{i=1}^n X_{J_i, K_i} \right),$$

where J_i ranges over elements of $\mathscr{J}_i$ for each $i \in [n]$. If we let $\mathbf{E}_{\mathscr{J}}$ denote conditioning with respect to all random variables except those depending on $\mathscr{J}_1, \ldots, \mathscr{J}_n$, then (6.1) implies

$$\mathbf{E}_{\mathscr{J}}(Z') = \sum_{J_1, \ldots, J_n} \sum_K \mathbf{P}(J_1 \in \mathscr{J}_1, \ldots, J_n \in \mathscr{J}_n) \left(\prod_{i=1}^n X_{J_i, K_i} \right) = Z/n^n.$$

Thus Jensen's inequality implies that

$$F(Z) = F(n^n \mathbf{E}_{\mathscr{J}}(Z')) \leq \mathbf{E}_{\mathscr{J}} F(n^n Z').$$

In particular, this means we can select a single, non-random partition $\mathscr{J}_1, \ldots, \mathscr{J}_n$ of $\mathscr{B}_s^d(E)$ such that $F(Z) \leq F(n^n Z')$. Because $\mathscr{J}_1, \ldots, \mathscr{J}_n$ are disjoint sets, Z' is identically distributed to

$$W' = \sum_{J_1 \in \mathscr{J}_1, \ldots, J_n \in \mathscr{J}_n} \sum_K \left(\prod_{i=1}^n X^i_{J_i, K_i} \right).$$

But $W' \leq W$, and so by monotonicity,

$$\mathbf{E}(F(Z)) \leq \mathbf{E}(F(n^n W')) \leq \mathbf{E}(F(n^n W)). \qquad \square$$

Lemma 22. *MOMENT BOUND THEOREM.*

Proof. We now bound the moments of W, which by Lemma 21 bounds the moments of Z. We write

$$W^m = \sum_{J_{j,i}} \sum_{K_1,\ldots,K_m} \left(\prod_{i=1}^{n} \prod_{j=1}^{m} X^i_{J_{j,i},K_{j,i}} \right).$$

Thus

$$\mathbf{E}(W^m) = \sum_{J_{j,i}} \sum_{K_1,\ldots,K_m} \prod_{i=1}^{n} \mathbf{E} \left(\prod_{j=1}^{m} X^i_{J_{j,i},K_{j,i}} \right)$$

For any collection of cubes $K_1,\ldots,K_m$, how many distinct parent cubes do they generate? At most nm, but as few as n. But how many are as few as n? $\square$

6.3 Singular Value Decomposition of Tensors

An *n tensor* with respect to $\mathbf{R}^{k_1},\ldots,\mathbf{R}^{k_n}$, is a multi-linear function

$$A : \mathbf{R}^{k_1} \times \cdots \times \mathbf{R}^{k_n} \to \mathbf{R}.$$

The family of all n tensors forms a vector space, denoted $\mathbf{R}^{k_1} \otimes \cdots \otimes \mathbf{R}^{k_n}$, or $\bigotimes \mathbf{R}^{k_i}$. Given a family of vectors v_i, with $v_i \in \mathbf{R}^{k_i}$, we can consider a tensor $v_1 \otimes \cdots \otimes v_n$, such that

$$(v_1 \otimes \cdots \otimes v_n)(w_1,\ldots,w_n) = (v_1 \cdot w_1)\ldots(v_n \cdot w_n).$$

It is easily verified that a basis for $\mathbf{R}^{k_1} \otimes \cdots \otimes \mathbf{R}^{k_n}$ is given by

$$\{e_{j_1} \otimes \cdots \otimes e_{j_n} : j_i \in [k_i]\}.$$

Thus a tensor A can be identified with a family of coefficients

$$\{A_{j_1\ldots j_n} : 1 \le j_i \le k_i\}$$

such that

$$A(v_1,\ldots,v_n) = \sum A_{j_1\ldots j_n} \cdot e_{j_1} \otimes \cdots \otimes e_{j_n}.$$

From this perspective, a tensor is a multi-dimensional array given by n indices. This basis induces a natural inner product for which the basis is orthonormal. The norm with respect to this inner product is called the *Frobenius norm*, denoted $\|A\|_F$.

Recall that if A is a rank r, $m \times n$ matrix, then the *singular value decomposition* says that there are orthonormal families of vectors $\{u_1, \dots, u_r\}$ and $\{v_1, \dots, v_r\}$ in $\mathbf{R}^m$ and $\mathbf{R}^n$ respectively, and values $s_1, \dots, s_r > 0$ such that $A = \sum s_i (u_i^T v_i)$. We may view A as a 2 tensor in $\mathbf{R}^m \otimes \mathbf{R}^n$ by setting $A(v,w) = v^T A w$. Then the singular value decomposition says that $A = \sum s_i (u_i \otimes v_i)$. Thus, viewing A as a bilinear function, for any vectors $x \in \mathbf{R}^m$ and y in $\mathbf{R}^n$, $A(x,y) = \sum s_i (u_i \cdot x)(v_i \cdot y)$. We want to try and generalize this result to n tensors in $\mathbf{R}^{k_1} \otimes \cdots \otimes \mathbf{R}^{k_n}$.

One can reduce certain quantities over tensors to quantities defined with respect to linear maps. For instance, for each index $i_0 \in [n]$, we can define a linear map

$$A_i : \bigotimes_{i' \neq i} \mathbf{R}^{k_{i'}} \to \mathbf{R}^{k_i}$$

by letting $A_{i_0}(v_1, \dots, \widehat{v_i}, \dots, v_n)$ to be the unique vector $w \in \mathbf{R}^{k_i}$ such that for any vector $v_i \in \mathbf{R}^{k_i}$,

$$v_i \cdot w = A(v_1, \dots, v_n).$$

The matrix A_i is known as the *i'th unfolding of* A. We define the i rank of A to be $\text{rank}_i(A) = \text{rank}(A_i)$. This generalizes the column and row rank of matrices. Unlike in the case of matrices, these quantities can differ for different indices i. One can also obtain the i rank

The natural definition of the rank of a tensor in the study of the higher order singular value decomposition is defined in analogy with the fact that a rank r matrix can be written as the sum of r rank one matrices. We say a tensor is rank one if it can be written in the form $v_1 \otimes \cdots \otimes v_n$ for some vectors $v_i \in \mathbf{R}^{k_i}$. The *rank* of a tensor A, denoted $\text{rank}(A)$, is the minimal number of rank one tensors which can be written in a linear combination to form A. We have $\text{rank}(A_i) \leq \text{rank}(A)$ for each i, but there may be a gap in these ranks.

To obtain an SVD decomposition, fix an index i. Then we can perform a singular value decomposition on A_i, obtaining orthonormal basis

6.4 Decoupling Concentration Bounds

We require some methods in non-asymptotic probability theory in order to understand the concentration structure of the random quantities we study. Most novel of these is relying on a probabilistic decoupling bound of Bourgain.

In this section, we fix a single underlying probability space, assumed large enough for us to produce any additional random quantities used in this section. Given a scalar valued random variable X, we define its *Orlicz norm* as

$$\|X\|_{\psi_2} = \inf\left\{\lambda \geq 0 : \mathbf{E}\left(\exp\left(|X|^2/\lambda^2\right)\right) \leq 2\right\}.$$

The family of random variables with finite Orlicz norm forms a Banach space, known as the space of *subgaussian random variables*.

Theorem 12. *Let X be a sub-gaussian random variable.*

- *For all $t \geq 0$,*
$$\mathbf{P}\left(|X| \geq t\right) \leq 2\exp\left(\frac{-t^2}{\|X\|_{\psi_2}^2}\right).$$
- *Conversely, if there is a constant $K > 0$ such that for all $t \geq 0$,*
$$\mathbf{P}\left(|X| \geq t\right) \leq 2\exp\left(\frac{-t^2}{K^2}\right),$$
then $\|X\|_{\psi_2} \leq 2K$.
- *For any bounded random variable X, $\|X\|_{\psi_2} \lesssim \|X\|_\infty$.*
- $\|X - \mathbf{E}X\|_{\psi_2} \lesssim \|X\|$.

- *If $X_1, \dots, X_n$ are independent, then*

$$\|X_1 + \cdots + X_n\|_{\psi_2} \lesssim \left(\|X_1\|_{\psi_2}^2 + \cdots + \|X_n\|_{\psi_2}^2 \right)^{1/2}.$$

Proof. See [16]. □

Given a random vector X, taking values in $\mathbf{R}^n$, we say it is *uniformly subgaussian* if there is a constant A such that for any $x \in \mathbf{R}^n$, $\|X \cdot x\|_{\psi_2} \leq A \cdot |x|$. We define

$$\|X\|_{\psi_2} = \sup_{|x|=1} \|X \cdot x\|_{\psi_2},$$

so that $\|X \cdot x\|_{\psi_2} \leq \|X\|_{\psi_2} \cdot |x|$ for any random vector X.

The classic Hanson-Wright inequality gives a concentration bound for the magnitude of a quadratic form $X^T A X$ in terms of the matrix norms of A, assuming X is a random vector with independant, mean zero, subgaussian coordinates.

Theorem 13 (Hanson-Wright)**.** *Let $X = (X_1, \dots, X_k)$ be a random vector with independant, mean zero, subgaussian coordinates, with $\|X_i\|_{\psi_2} \leq K$ for each i, and let A be a diagonal-free $k \times k$ matrix. Then there is a universal constant c such that for every $t \geq 0$,*

$$\mathbf{P}\left(|X^T A X| \geq t\right) \leq 2 \exp\left(-c \cdot \min\left(\frac{t^2}{K^4 \|A\|_F^2}, \frac{t}{K^2 \|A\|} \right) \right).$$

Here $\|A\|$ is the operator norm of A, and $\|A\|_F$ it's Frobenius norm.

In our work, we require a modified version of the Hanson Wright inequality where we replace the independant random variables $X_1, \dots, X_k$ with independant random vectors taking values in $\mathbf{R}^d$, and where A is replaced with a n tensor on $\mathbf{R}^{k \times d}$. To make things more tractable, we assume that $A_I = 0$ for each multi-index $I = ((i_1, j_1), \dots, (i_n, j_n))$ if there is $r \neq r'$ with $i_r = i_{r'}$. If this is satisfied, we say A is a *non-diagonal tensor*.

Theorem 14. *Let $X = (X_1, \ldots, X_k)$, where $X_1, \ldots, X_k$ are independant, mean zero random vectors taking values in $\mathbf{R}^d$, with $\|X_i\|_{\psi_2} \leq K$ for each i. Let A be a non-diagonal n-tensor in $\mathbf{R}^{k \times d}$. Then for $t \geq 0$,*

$$\mathbf{P}(|A(X, \ldots, X)| \geq t) \leq 2 \cdot \exp\left(-c(n) \cdot \min\left(\frac{t^2}{K^4\|A\|_F^2}, \ldots, \frac{t^{2/n}}{K^{4/n}\|A\|_F^{2/n}}\right)\right).$$

The main technique involved is a statistical decoupling result due to J. Bourgain and L. Tzafriri. We follow the proof of the Hanson-Wright inequality from [16], altering various calculations to incorporate the vector-valued coordinates $X_1, \ldots, X_n$, and incorporating more robust calculations of the moments of higher order Gaussian chaos.

Lemma 23 (Decoupling)**.** *If $F : \mathbf{R} \to \mathbf{R}$ is convex, and $X = (X_1, \ldots, X_k)$, where $X_1, \ldots, X_k$ are independant, mean zero random vectors taking values in $\mathbf{R}^d$. If A is a non-diagonal n-tensor in $\mathbf{R}^{d \times k}$, then*

$$\mathbf{E}(F(A(X, \ldots, X))) \leq \mathbf{E}(F(n^n \cdot A(X^1, X^2, \ldots, X^n))),$$

where $X^1, \ldots, X^n$ are independant copies of X.

Proof. Given a set $I \subset [k]$, let $\pi_I : \mathbf{R}^{k \times d} \to \mathbf{R}^{k \times d}$ be defined by

$$\pi_I(X)_{ij} = \begin{cases} X_{ij} & : \text{if } i \in I, \\ 0 & : \text{if } i \notin I \end{cases}.$$

Thus given $X = (X_1, \ldots, X_k)$, the vector $\pi_I(X)$ is obtained by zeroeing out all vectors X_i with $i \notin I$. Consider a random partition of $[k]$ into n classes $I_1, \ldots, I_n$, chosen uniformly at random. Thus given any $x \in \mathbf{R}^{d \times k}$, we have

$$\mathbf{E}(A(\pi_{I_1}(x), \ldots, \pi_{I_n}(x))) = \sum_{I = ((i_1, j_1), \ldots, (i_n, j_n))} A_I \cdot x_{i_1 j_1} \ldots x_{i_n j_n} \mathbf{P}(i_1 \in I_1, \ldots, i_n \in I_n).$$

If $A_I \neq 0$, the indices $\{i_r\}$ are distinct from one another. Since $(I_1, \ldots, I_n)$ is selected uniformly at random, this means the positions of the indices i_r are independent of one another. Thus

$$\mathbf{P}(i_1 \in I_1, \ldots, i_n \in I_n) = \mathbf{P}(i_1 \in I_1) \ldots \mathbf{P}(i_n \in I_n) = 1/n^n.$$

Thus we have shown

$$A(x, \ldots, x) = n^n \cdot \mathbf{E}(A(\pi_{I_1}(x), \ldots, \pi_{I_n}(x))).$$

In particular, substituting X for x, we have

$$A(X, \ldots, X) = n^n \cdot \mathbf{E}_I(A(\pi_{I_1}(X), \ldots, \pi_{I_n}(X)))$$

Applying Jensen's inequality, we find

$$\begin{aligned}
\mathbf{E}_X(F(A(X, \ldots, X))) &= \mathbf{E}_X(F(n^n \cdot \mathbf{E}_I(A(\pi_{I_1}(X), \ldots, \pi_{I_n}(X))))) \\
&\leq \mathbf{E}_X \, \mathbf{E}_I \, F(n^n \cdot A(\pi_{I_1}(X), \ldots, \pi_{I_n}(X))) \\
&= \mathbf{E}_I \, \mathbf{E}_X \, F(n^n \cdot A(\pi_{I_1}(X), \ldots, \pi_{I_n}(X))).
\end{aligned}$$

Thus we may select a *deterministic* partition I for which

$$\mathbf{E}(F(A(X, \ldots, X))) \leq \mathbf{E}(F(n^n \cdot A(\pi_{I_1}(X), \ldots, \pi_{I_n}(X)))).$$

Since the classes $I_1, \ldots, I_n$ are disjoint from one another, $(\pi_{I_1}(X), \ldots, \pi_{I_n}(X))$ is identically distributed to $(\pi_{I_1}(X^1), \ldots, \pi_{I_n}(X^n))$, so

$$\mathbf{E}(F(n^n \cdot A(\pi_{I_1}(X), \ldots, \pi_{I_n}(X)))) = \mathbf{E}(F(n^n \cdot A(\pi_{I_1}(X^1), \ldots, \pi_{I_n}(X^n)))).$$

Let $Z = A(X^1, \ldots, X^n) - A(\pi_{I_1}(X^1), \ldots, \pi_{I_n}(X^n))$. If $\mathbf{E}'$ denotes conditioning with respect to $\pi_{I_1}(X^1), \ldots, \pi_{I_n}(X^n)$, then this expectation fixes $A(\pi_{I_1}(X_1), \ldots, \pi_{I_n}(X_n))$, but $\mathbf{E}'(Z) = 0$. Thus we can apply Jensen's inequality again to conclude

$$\begin{aligned}
\mathbf{E}(F(n^n \cdot A(\pi_{I_1}(X^1),\dots,\pi_{I_n}(X^n)))) &= \mathbf{E}(F(n^n \cdot \mathbf{E}'(A(\pi_{I_1}(X^1),\dots,\pi_{I_n}(X^n)) + Z))) \\
&= \mathbf{E}(F(n^n \cdot \mathbf{E}'(A(X^1,\dots,X^n)))) \\
&\leq \mathbf{E}(\mathbf{E}'(F(n^n A(X^1,\dots,X^n)))) \\
&= \mathbf{E}(F(n^n A(X^1,\dots,X^n))). \qquad \square
\end{aligned}$$

Remark. *Note this theorem also implies that for each $\lambda \in \mathbf{R}$,*

$$\mathbf{E}(F((\lambda/n^n) \cdot A(X,\dots,X))) \leq \mathbf{E}(F(\lambda \cdot A(X^1,\dots,X^n))).$$

Lemma 24 (Comparison)**.** *Let $X^1,\dots,X^n$ be independant random vectors taking values in $\mathbf{R}^{k \times d}$*

To make things easy for ourselves, we assume that A vanishes along it's diagonal. Reinterpreting the quantities in the Hanson-Wright inequality in this setting gives the result. For each i and j, we define A_{ij} be an $d \times d$ matrix, then we set $X^T A X = \sum X_i^T A_{ij} X_j$,

$$\|A\|_F = \sqrt{\sum \|A_{ij}\|_F^2}, \quad \text{and} \quad \|A\| = \sup\left\{ |\sum x_i^T A_{ij} y_j| : \sum |x_i|^2 = \sum |y_j|^2 = 1 \right\}.$$

These are precisely the definitions of the operator norm and Frobenius norm if we interpret A as a $dn \times dn$ matrix by expanding out the entries of each A_{ij}, and X as a random vector taking values in $\mathbf{R}^{dn}$ by expanding out the entries of each X_i.

Chapter 7

Constructing Squarefree Sets

7.1 Ideas For New Work

A continuous formulation of the squarefree difference problem is not so clear to formulate, because every positive real number has a square root. Instead, we consider a problem which introduces a similar structure to avoid in the continuous domain rather than the discrete. Unfortunately, there is no direct continuous anology to the squarefree subset problem on the interval $[0,1]$, because there is no canonical subset of $[0,1]$ which can be identified as 'perfect squares', unlike in $\mathbf{Z}$. If we only restrict ourselves to perfect squares of a countable set, like perfect squares of rational numbers, a result of Keleti gives us a set of full Hausdorff dimension avoiding this set. Thus, instead, we say a set $X \subset [0,1]$ is (continuously) **squarefree** if there are no nontrivial solutions to the equation $x - y = (u - v)^2$, in the sense that there are no $x, y, u, v \in X$ satisfying the equation for $x \neq y, u \neq v$. In this section we consider some blue sky ideas that might give us what we need.

How do we adopt Rusza's power series method to this continuous formulation of the problem? We want to scale up the problem exponentially in a way we can vary to give a better control of the exponentials. Note that for a fixed m, every elements $x \in [0,1]$ has an essentially unique m-ary expansion

$$x = \sum_{n=1}^{\infty} \frac{x_n}{m^n}$$

and the pullback to the Haar measure on $\mathbf{F}_m^{\infty}$ is measure preserving (with respect to the natural Haar measure on $\mathbf{F}_m^{\infty}$), so perhaps there is a way to reformulate the problem natural as finding nice subsets of $\mathbf{F}_m^{\infty}$ avoiding squares. In terms of this expansion, the equation $x - y = (u-v)^2$ can be rewritten as

$$\sum_{n=1}^{\infty} \frac{x_n - y_n}{m^n} = \left(\sum_{k=1}^{\infty} \frac{u_n - v_n}{m^n} \right)^2 = \sum_{n=1}^{\infty} \left(\sum_{k=1}^{n-1} (u_k - v_k)(u_{n-k} - v_{n-k}) \right) \frac{1}{m^n}$$

One problem with this expansion is that the sums of the differences of each element do not remain in $\{0, \dots, m-1\}$, so the sum on the right cannot be considered an equivalent formal expansion to the expansion on the left. Perhaps $\mathbf{F}_m^{\infty}$ might be a simpler domain to explore the properties of squarefree subsets, in relation to Ruzsa's discrete strategy. What if we now consider the problem of finding the largest subset X of $\mathbf{F}_m^{\infty}$ such that there do not exist $x, y, u, v \in \mathbf{F}_m^{\infty}$ such that if $x, y, u, v \in X$, $x \neq y$, $u \neq v$, then for any n

$$x_n - y_n \neq \sum_{k=1}^{n-1} (u_k - v_k)(u_{n-k} - v_{n-k})$$

What if we consider the problem modulo m, so that the convolution is considered modulo m, and we want to avoid such differences modulo m. So in particular, we do not find any solutions to the equation

$$\begin{aligned}
x_2 - y_2 &= (u_1 - v_1)^2 \\
x_3 - y_3 &= 2(u_1 - v_1)(u_2 - v_2) \\
x_4 - y_4 &= (u_1 - v_1)(u_3 - v_3) + (u_2 - v_2)^2 \\
&\vdots
\end{aligned}$$

which are considered modulo m. The topology of the p-adic numbers induces a power series relationship which 'goes up' and might be useful to our analysis, if the measure theory of the p-adic numbers agrees with the measure theory of normal numbers in some way, or as an alternate domain to analyze the squarefree problem as with $\mathbf{F}_m^\infty$.

The problem with the squarefree subset problem is that we are trying to optimize over two quantities. We want to choose a set X such that the number of distinct differences $x-y$ as small as possible, while keeping the set as large as possible. This double optimization is distinctly different from the problem of finding squarefree difference subsets of the integers. Perhaps a more natural analogy is to fix a set V, and to find the largest subset X of $[0,1]$ such that $x-y=(u-v)^2$, where $x \neq y \in X$, and $u \neq v \in V$. Then we are just avoiding subsets of $[0,1]$ which avoid a particular set of differences, and I imagine this subset has a large theory. But now we can solve the general subset problem by finding large subsets X such that $(X-X)^2 \subset V$ and X containing no differences in V. Does Rusza's method utilize the fact that the problem is a single optimization? Can we adapt Rusza's method work to give better results about finding subsets X of the integers such that $X-X$ is disjoint from $(X-X)^2$?

7.2 Squarefree Sets Using Modulus Techniques

We now try to adapt Ruzsa's idea of applying congruences modulo m to avoid squarefree differences on the integers to finding high dimensional subsets of $[0,1]$ which satisfy a continuous analogy of the integer constraint. One problem with the squarefree problem is that solutions are non-scalable, in the sense that if $X \subset [N]$ is squarefree, αX may not be squarefree. This makes sense, since avoiding solutions to $\alpha(x-y)=\alpha^2(u-v)^2$ is clearly not equivalent to the equation $x-y=(u-v)^2$. As an example, $X=\{0,1/2\}$ is squarefree, but $2X=\{0,1\}$ isn't. On the other hand, if X avoids squarefree differences modulo N, it *is* scalable by a number congruent to 1 modulo N. More generally, if α is a rational number of the form p/q, then αX will avoid nontrivial solutions to $q(x-y)=p(u-v)^2$, and if p and

q are both congruent to 1 modulo N, then X is squarefree, so modulo arithmetic enables us to scale down. Since the set of rational numbers with numerator and denominator congruent to 1 is dense in $\mathbf{R}$, *essentially* all scales of X are continuously squarefree. Since X is discrete, it has Hausdorff dimension zero, but we can 'fatten' the scales of X to obtain a high dimension continuously squarefree set. To initially simplify the situation, we now choose to avoid nontrivial solutions to $y - x = (z - x)^2$, removing a single degree of freedom from the domain of the equation.

So we now fix a subset X of $\{0, \dots, m-1\}$ avoiding squares modulo m. We now ask how large can we make ε such that nontrivial solutions to $x - y = (x - z)^2$ in the set

$$E = \bigcup_{x \in X} [\alpha x, \alpha x + \varepsilon)$$

occur in a common interval, if α is just short of $1/m^n$. This will allow us to recursively place a scaled, 'fattened' version of X in every interval, and then consider a limiting process to obtain a high dimensional continuously squarefree set. If we have a nontrivial solution triple, we can write it as $\alpha x + \delta_1, \alpha y + \delta_2$, and $\alpha z + \delta_3$, with $\delta_1, \delta_2, \delta_3 < \varepsilon$. Expanding the solution leads to

$$\alpha(x-y) + (\delta_1 - \delta_2) = \alpha^2 (x-z)^2 + 2\alpha(x-z)(\delta_1 - \delta_3) + (\delta_1 - \delta_3)^2$$

If x, y, and z are all distinct, then, as we have discussed, we cannot have $\alpha(x-y) = \alpha^2(x-z)^2$. if α is chosen close enough to $1/m^n$, then we obtain an approximate inequality

$$|\alpha(x-y) - \alpha^2(x-z)^2| \geq \alpha^2$$

(we require α to be close enough to $1/n$ for some n to guarantee this). Thus we can guarantee at least two of x, y, and z are equal to one another if

$$|2\alpha(x-z)(\delta_1 - \delta_3) + (\delta_1 - \delta_3)^2 - (\delta_1 - \delta_2)| < \frac{1}{m^{2n}}$$

We calculate that

$$2\alpha(x-z)(\delta_1-\delta_3)+(\delta_1-\delta_3)^2-(\delta_1-\delta_2)<2\alpha(m-1)\varepsilon+\varepsilon^2+\varepsilon$$

$$(\delta_1-\delta_2)-2\alpha(x-z)(\delta_1-\delta_3)-(\delta_1-\delta_3)^2\leq\varepsilon+2\alpha(m-1)\varepsilon$$

So it suffices to choose ε such that

$$\varepsilon^2+[2\alpha(m-1)+1]\varepsilon\leq\alpha^2$$

This is equivalent to picking

$$\varepsilon\leq\sqrt{\left(\frac{2\alpha(m-1)+1}{2}\right)^2+\alpha^2}-\frac{2\alpha(m-1)+1}{2}\approx\frac{\alpha^2}{2\alpha(m-1)+1}$$

We split the remaining discussion of the bound we must place on ε into the three cases where two of x, y, and z are equal, but one is distinct, to determine how small ε must be to prevent this from happening. Now

- If $y=z$, but x is distinct, then because we know $\alpha(x-y)=\alpha^2(x-y)^2$ has no solution in X, we obtain that (provided α is close enough to $1/m^n$),

  $$|\alpha(x-y)-\alpha^2(x-y)^2|\geq\alpha^2$$

  and the same inequality that worked for the case where the three equations are distinct now applies for this case.

- If $x=y$, but z is distinct, we are left with the equation

  $$\delta_1-\delta_2=\alpha^2(x-z)^2+2\alpha(x-z)(\delta_1-\delta_3)+(\delta_1-\delta_3)^2$$

  Now $\alpha^2(x-z)^2\geq\alpha^2$, and

  $$\delta_1-\delta_2-2\alpha(x-z)(\delta_1-\delta_3)-(\delta_1-\delta_3)^2<\varepsilon+2\alpha(m-1)\varepsilon$$

so we need the additional constraint $\varepsilon + 2\alpha(m-1)\varepsilon \leq \alpha^2$, which is equivalent to saying

$$\varepsilon \leq \frac{\alpha^2}{1 + 2\alpha(m-1)}$$

- If $x = z$, but y is distinct, we are left with the equation

  $$\alpha(x-y) + (\delta_1 - \delta_2) = (\delta_1 - \delta_3)^2$$

  Now $|\alpha(x-y)| \geq \alpha$, and

  $$(\delta_1 - \delta_3)^2 - (\delta_1 - \delta_2) < \varepsilon^2 + \varepsilon$$

  $$(\delta_1 - \delta_2) - (\delta_1 - \delta_3)^2 < \varepsilon$$

  so to avoid this case, we need $\varepsilon^2 + \varepsilon \leq \alpha$, or

  $$\varepsilon \leq \frac{\sqrt{1+4\alpha} - 1}{2} \approx \alpha$$

Provided ε is chosen as above, all solutions in E must occur in a common interval. Thus, if we now replace the intervals with a recursive fattened scaling of X, all solutions must occur in smaller and smaller intervals. If we choose the size of these scalings to go to zero, these solutions are required to lie in a common interval of length zero, and thus the three values must be equal to one another. Rigorously, we set $\varepsilon \approx 1/m^2$, and $\alpha \approx 1/m$, we can define a recursive construction by setting

$$E_1 = \bigcup_{x \in X} [\alpha x, \alpha x + \varepsilon_1)$$

and if we then set X_n to be the set of startpoints of the intervals in E_n, then

$$E_{n+1} = \bigcup_{x \in X_n} (x + \alpha^2 E_n)$$

Then $\bigcap E_n$ is a continuously squarefree subset. But what is it's dimension?

7.3 Idea; Delaying Swaps

By delaying the removing in the pattern removal queue, we may assume in our dissection methods that we are working with sets with certain properties, i.e. we can swap an interval with a dimension one set avoiding translates.

7.4 Squarefree Subsets Using Interval Dissection Methods

The main idea of Keleti's proof was that, for a function f, given a method that takes a sequence of disjoint unions of sets $J_1, \dots, J_N$, each a union of almost disjoint closed intervals of the same length, and gives large subsets $J_n' \subset J_n$, each a union of almost disjoint intervals of a much smaller length, such that $f(x_1, \dots, x_n) \neq 0$ for $x_n \in J_n'$. Then one can find high dimensional subsets K of the real line such that $f(x_1, \dots, x_n) \neq 0$ for a sequence of distinct $x_1, \dots, x_n \in K$. The larger the subsets J_n' are compared to J_n, the higher the Hausdorff dimension of K. We now try and apply this method to construct large subsets avoiding solutions to the equation $f(x,y,z) = (x-y)-(x-z)^2$. In this case, since solutions to the equation above satisfy $y = x-(x-z)^2$, given J_1, J_2, J_3, finding J_1', J_2', J_3' as in the method above is the same as choosing J_1' and J_3' such that the image of $J_1' \times J_3'$ under the map $g(x,z) = x-(x-z)^2$ is small in J_2. We begin by discretizing the problem, splitting J_1 and J_3 into unions of smaller intervals, and then choosing large subsets of these intervals, and finding large intervals of J_2 avoiding the images of the startpoints to these intervals.

So suppose that J_1, J_2, and J_3 are unions of intervals of length $1/M$, for which we may find subsets $A, B \subset [M]$ of the integers such that

$$J_1 = \bigcup_{a \in A} \left[\frac{a}{M}, \frac{a+1}{M} \right] \qquad J_3 = \bigcup_{b \in B} \left[\frac{b}{M}, \frac{b+1}{M} \right]$$

If we split J_1 and J_3 into intervals of length $1/NM$, for some $N \gg M$ to be specified later (though we will assume it is a perfect square), then

$$J_1 = \bigcup_{\substack{a \in A \\ 0 \le k < N}} \left[\frac{Na+k}{NM}, \frac{Na+k}{NM} + \frac{1}{NM} \right] \quad J_3 = \bigcup_{\substack{b \in A \\ 0 \le l < N}} \left[\frac{Nb+l}{NM}, \frac{Nb+l}{NM} + \frac{1}{NM} \right]$$

We now calculate g over the startpoints of these intervals, writing

$$\begin{aligned} g\left(\frac{Na+k}{NM}, \frac{Nb+l}{NM}\right) &= \frac{Na+k}{NM} - \left(\frac{N(a-b)+(k-l)}{NM}\right)^2 \\ &= \frac{a}{M} - \frac{(a-b)^2}{M^2} + \frac{k}{NM} - \frac{2(a-b)(k-l)}{NM^2} + \frac{(k-l)^2}{(NM)^2} \end{aligned}$$

which splits the terms into their various scales. If we write $m = k - l$, then m can range on the integers in $(-N, N)$, and so, ignoring the first scale of the equation, we are motivated to consider the distribution of the set of points of the form

$$\frac{k}{NM} - \frac{2(a-b)m}{NM^2} + \frac{m^2}{(NM)^2}$$

where k is an integer in $[0, N)$, and m an integer in $(-N, N)$. To do this, fix $\varepsilon > 0$. Suppose that we find some value $\alpha \in [0, 1]$ such that S intersects

$$\left[\alpha, \alpha + \frac{1}{N^{1+\varepsilon}}\right]$$

Then there is k and m such that

$$0 \le \frac{kNM - 2N(a-b)m + m^2}{(NM)^2} - \alpha \le \frac{1}{N^{1+\varepsilon}}$$

Write $m = q\sqrt{N} + r$ (remember that we chose N so it's square root is an integer), with $0 \le r < \sqrt{N}$. Then $m^2 = qN + 2qr\sqrt{N} + r^2$, and if $2qr = Q\sqrt{N} + R$, where

$0 \leq R < \sqrt{N}$, then we find

$$-\frac{R}{M^2N^{3/2}} - \frac{r^2}{(NM)^2} \leq \frac{kM - 2(a-b)m + q + Q}{NM^2} - \alpha \leq \frac{1}{N^{1+\varepsilon}} - \frac{R}{M^2N^{3/2}} - \frac{r^2}{(NM)^2}$$

Thus

$$d(\alpha, \mathbf{Z}/NM^2) \leq \max\left(\frac{1}{N^{1+\varepsilon}} - \frac{R}{\sqrt{N}} - \frac{r^2}{N}, \frac{R}{M^2N^{3/2}} + \frac{r^2}{(NM)^2}\right)$$

If we now restrict our attention to the set S consisting of the expressions we are studying where $R \leq (\delta_0/2)\sqrt{N}$, $r \leq \sqrt{\delta_0 N/2}$, then if the interval corresponding to α intersects S, then

$$d(\alpha, \mathbf{Z}/NM^2) \leq \max\left(\frac{1}{N^{1+\varepsilon}}, \frac{\delta_0}{NM^2}\right)$$

If $N^\varepsilon \geq M^2/\delta_0$, then we can force $d(\alpha, \mathbf{Z}/NM^2) \leq \delta_0/NM^2$ for all α intersecting S. Thus, if we split J_2 into intervals starting at points of the form

$$\frac{k+1/2}{NM^2}$$

each of length $1/N^{1+\varepsilon}$, then provided $\delta_0 < 1/2$, we conclude that these intervals do not contain any points in S, since

$$d\left(\frac{k+1/2}{NM^2}, \mathbf{Z}/NM^2\right) = \frac{1}{2NM^2} > \frac{\delta_0}{NM^2}$$

So we're well on our way to using Pramanik and Fraser's recursive result, since this argument shows that, provided points in J_1 and J_3 are chosen carefully, we can keep $O_M(1/N^{1+\varepsilon})$ of each interval in J_2, which should lead to a dimension bound arbitrarily close to one.

7.5 Finding Many Startpoints of Small Modulus

To ensure a high dimension corresponding to the recursive construction, it now suffices to show J_1 and J_3 contain many startpoints corresponding to points in S, so that the refinements can be chosen to obtain $O_M(1/N)$ of each of the original intervals. Define T to be the set of all integers $m \in (-N,N)$ with $m = q\sqrt{N} + r$ and $r \leq \sqrt{\delta_0 N/2}$ and $2qr = Q\sqrt{N} + R$ with $R \leq (\delta_0/2)\sqrt{N}$. Because of the uniqueness of the division decomposition, we find T is in one to one correspondence with the set T' of all pairs of integers (q,r), with $q \in (-\sqrt{N}, \sqrt{N})$ and $r \in [0, \sqrt{N})$, with $r \leq \sqrt{\delta_0 N/2}$, $2qr = Q\sqrt{N} + R$, and $R \leq (\delta_0/2)\sqrt{N}$. Thus we require some more refined techniques to better upper bound the size of this set.

Let's simplify notation, generalizing the situation. Given a fixed ε, We want to find a large number of integers $n \in (-N,N)$ with a decomposition $n = qr$, where $r \leq \varepsilon\sqrt{N}$, and $q \leq \sqrt{N}$. The following result reduces our problem to understanding the distribution of the smooth integers.

Lemma 25. *Fix constants A,B, and let $n \leq AN$ be an integer. If all prime factors of n are $\leq BN^{1-\delta}$, then n can be decomposed as qr with $r \leq \varepsilon\sqrt{N}$ and $q \leq \sqrt{N}$.*

Proof. Order the prime factors of n in increasing order as $p_1 \leq p_2 \leq \cdots \leq p_K$. Let $r = p_1 \dots p_m$ denote the largest product of the first prime factors such that $r \leq \varepsilon\sqrt{N}$. If $r = n$, we can set $q = 1$, and we're finished. Otherwise, we know $rp_{m+1} > \varepsilon\sqrt{N}$, hence

$$r > \frac{\varepsilon\sqrt{N}}{p_{m+1}} \geq \frac{\varepsilon\sqrt{N}}{BN^{1-\delta}} = \frac{\varepsilon}{B}N^{\delta-1/2}$$

And if we set $q = n/r$, the inequality above implies

$$q < \frac{nB}{\varepsilon}N^{1/2-\delta} \leq \frac{AB}{\varepsilon}N^{3/2-\delta}$$

But now we run into a problem, because the only way we can set $q < \sqrt{N}$ while keeping A, B, and ε fixed constants is to set $\delta = 1$, and $AB/\varepsilon \leq 1$. $\square$

Remark. *Should we expect this method to work? Unless there's a particular reason why values of (q,r) should accumulate near $Q=0$, we should expect to lose all but $N^{-1/2}$ of the N values we started with, so how can we expect to get $\Omega(N)$ values in our analysis. On the other hand, if a number n is suitably smooth, in a linear amount of cases we should be able to divide up primes into two numbers q and r such that r is small and q fits into a suitable value of Q, so maybe this method will still work.*

Regardless of whether the lemma above actually holds through, we describe an asymptotic formula for perfect numbers which might come in handy. If $\Psi(N,M)$ denotes the number of integers $n \leq N$ with no prime factor exceeding M, then Karl Dickman showed

$$\Psi(N,N^{1/u}) = N\rho(u) + O\left(\frac{uN}{\log N}\right)$$

This is essentially linear for a fixed u, which could show the set of (q,r) is $\Omega_\varepsilon(N)$, which is what we want. Additional information can be obtained from Hildebrand and Tenenbaum's survey paper "Integers Without Large Prime Factors".

7.6 A Better Approach

Remember that we can write a general value in our set as

$$x = \frac{-2(a-b)m}{NM^2} + \frac{q}{NM^2} + \frac{Q}{NM^2} + \frac{R}{N^{3/2}M^2} + \frac{r^2}{N^2M^2}$$

with the hope of guaranteeing the existence of many points, rather than forcing R to be small, we now force R to be close to some scaled value of $\sqrt{N}$,

$$|R - n\varepsilon\sqrt{N}| = \delta\sqrt{N} \leq \varepsilon\sqrt{N}$$

Then

$$x = \frac{-2(a-b)m + q + Q + n\varepsilon + \delta}{NM^2} + \frac{r^2}{N^2M^2}$$

So

$$d\left(x, \mathbf{Z}/NM^2 + n\varepsilon/NM^2\right) \leq \frac{\varepsilon}{NM^2} + \frac{1}{4NM^2} = \frac{\varepsilon + 1/4}{NM^2}$$

By the pidgeonhole principle, since $R < \sqrt{N}$, there are $1/\varepsilon$ choices for n, whereas there are

The choice has the benefit of automatically possessing a lot of points by the pidgeonhole principle,

Chapter 8

Conclusions

Bibliography

[1] K. Athreya. Large deviation rates for branching processes–i. single type case. *Ann. Appl. Probab.*, 4:779–790, 1994. → page 45

[2] J. Balogh, R. Morris, and W. Samotij. Independent sets in hypergraphs. *J. Amer. Math. Soc.*, 28:669–709, 2015. → page 53

[3] J. Biggins and N. Bingham. Large deviations in the supercritical branching process. *Adv. in Appl. Probab.*, 25:757–772. → page 45

[4] K. Dickman. On the frequency of numbers containing prime factors of a certain relative magnitude.

[5] K. Falconer. *Fractal Geometry: Mathematical Foundations and Applications*. Wiley, 2003. → page 39

[6] R. Fraser and M. Pramanik. Large sets avoiding patterns. *Anal. PDE*, 11: 1083–1111, 2018. → pages 26, 27, 28, 38, 41, 42, 51, 52

[7] T. Harris. *The Theory of Branching Processes*. Springer-Verlag, 1963. → page 44

[8] N. H. Katz and T. Tao. Some connections between Falconer's distance set conjecture, and sets of Furstenburg type. *New York J. Math.*, 7:149–187, 2001. → page 29

[9] T. Keleti. A 1-dimensional subset of the reals that intersects each of its translates in at most a single point. *Real Anal. Exchange*, 24:843–845, 1999. → pages 28, 51

[10] P. Maga. Full dimensional sets without given patterns. *Real Anal. Exchange*, 36:79–90, 2010. → pages 2, 26

[11] A. Máthé. Sets of large dimension not containing polynomial configurations. *Adv. Math.*, 316:691–709, 2017. → pages 26, 27, 51, 52

[12] P. Mattila. *Fourier analysis and Hausdorff dimension*. Cambridge University Press, 2015. → pages 2, 26, 43

[13] I. Z. Ruzsa. Difference sets without squares.

[14] B. Sudakov, E. Szemerédi, and V. Vu. On a question of erdös and moser.

[15] T. Tao. 245 C, Notes 5: Hausdorff dimension, 2009. URL https://terrytao.wordpress.com/2009/05/19/245c-notes-5-hausdorff-dimension-optional/.

[16] Vershynin. → pages 72, 73