

Squarefree Sets

Jacob Denson

December 30, 2018

Contents

| | | |
|----------|--|-----------|
| 0.1 | Interval Dissection | 4 |
| 0.2 | Fractal Avoidance and Random Selection | 5 |
| 0.3 | Low Rank Functions and Shifting? | 7 |
| 0.4 | Future Work? | 7 |
| 1 | Background | 8 |
| 1.1 | Rusza: Difference Sets Without Squares | 8 |
| 1.2 | Cantor Sets and Discrete Configurations | 13 |
| 1.3 | Keleti's Translate Avoiding Set | 13 |
| 1.4 | Fraser/Pramanik: Extending Keleti Translation to Smooth Configurations | 16 |
| 1.5 | A Set Avoiding All Functions With A Common Derivative | 19 |
| 1.6 | Equidistribution Results | 20 |
| 1.7 | Results about Hypergraphs | 22 |
| 1.8 | Hyperdyadic Covers | 23 |
| 2 | Ideas on Squarefree Subsets | 26 |
| 2.1 | Ideas For New Work | 26 |
| 2.2 | Squarefree Sets Using Modulus Techniques | 27 |
| 2.3 | Idea; Delaying Swaps | 30 |
| 2.4 | Squarefree Subsets Using Interval Dissection Methods | 30 |
| 2.5 | Finding Many Startpoints of Small Modulus | 32 |
| 2.6 | A Better Approach | 33 |
| 3 | Low Rank Coordinate Changes | 34 |
| 3.1 | Boosting the Dimension of Pattern Avoiding Sets by Low Rank Coordinate Changes | 34 |
| 3.2 | Extension to Well Approximable Numbers | 36 |
| 3.3 | Equidistribution and Real Valued Matrices | 37 |
| 3.4 | Applications of Low Rank Coordinate Changes | 38 |
| 3.5 | Idea: Generalizing This Problem to low rank smooth functions | 40 |
| 3.6 | Idea: Algebraic Number Fields | 41 |
| 3.7 | A Scheme for Avoiding Configurations | 41 |
| 3.8 | Square Free Sets | 42 |

| | | |
|----------|--|-----------|
| 4 | Extension to Arbitrary Cartesian Avoidance Problems | 43 |
| 4.1 | From Minkowski to Hausdorff dimension | 46 |
| 5 | Appendix: Dimensions of Cantor Like Sets | 48 |
| 5.1 | Basic Covering | 49 |
| 5.2 | An Epsilon of Room | 49 |
| 5.3 | Bounds on Uniform Constructions | 50 |

We study a simple question:

How large can objects with a fixed irregularity be?

The objects we study are subsets of Euclidean space, and the irregularities manifest as avoiding a function's locus. For instance, the problem

How large can a subset X of \mathbf{R}^n be such that the angle formed by any three distinct points of X are distinct for distinct triples.

can be rephrased as finding $X \subset \mathbf{R}^n$ such that for any *distinct* pair of triples $(x_1, y_1, z_1), (x_2, y_2, z_2) \in X^3$,

$$\frac{(x_1 - z_1) \cdot (y_1 - z_1)}{|x_1 - z_1||y_1 - z_1|} - \frac{(x_2 - z_2) \cdot (y_2 - z_2)}{|x_2 - z_2||y_2 - z_2|} \neq 0$$

Thus we are avoiding the locus of zeroes of the function defined on a subset of X^6 by the equation on the left. We call the problem of finding X such that $f(x) \neq 0$ for any distinct $x_1, \dots, x_n \in X$ a *configuration avoidance problem*. The reasons for studying such problems emerge from scenarios in ergodic theory, harmonic analysis, and the analysis of partial differential equations. Here the boundedness of operators can be disproved by taking the characteristic functions of sets possessing irregularities.

Rather than studying particular instances of the configuration avoidance problem, we choose to study general methods to find large subsets of space avoiding configurations. In particular, we expand on a number of general *interval dissection methods* which have proven useful in the area, originally developed by Keleti but also studied notably by Mathé, and Pramanik/Fraser. Our novel contribution to the construction is that in an arbitrary configuration problem, random choices can overcome the lack of presence of structure. We use this to expand the utility of interval dissection methods to a set of *fractal avoidance problems*, that previous methods were completely unavailable to address. Such problems include:

How large can a subset X of \mathbf{R}^n be such that the angle 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 of the form $\cos(q + x)$, where q is a rational number, and x is an element of the Cantor set.

Since we study subsets of Euclidean space, a natural way to quantify size seems to be the Lebesgue measure of an object. However, objects with positive Lebesgue measure will not possess the irregularities we wish to construct. This is because sets of positive measure 'almost' contain an open set, and open sets contain will essentially contain all the configurations we wish to avoid in interesting problems. All the sets avoiding irregularities we study will have to have Lebesgue measure zero. Thus a 'second-order' measure of a set's size has to be introduced, and this is satisfied by the notion of Hausdorff and Minkowski dimension. These are the notions of size we will use in this thesis.

If you haven't encountered the notion of fractal dimensions before, the following intuitions should help you through this talk. A set in \mathbf{R}^n has Minkowski dimension α if it is covered by approximately N^α sidelength $1/N$ cubes. Hausdorff dimension is a refinement to Minkowski dimension which is more stable under infinite operations, such as countable unions. It lower bounds the Minkowski dimension.

0.1 Interval Dissection

It is easy to avoid configurations that occur at a particular scale. For instance, if $f(x) = 0$ implies that $x_i \geq \delta$ for some fixed δ , then any length δ interval will solve the configuration avoidance problem. Thus the interesting configuration avoidance problems occur when the zero sets of functions need to be analyzed at infinitely many suitably fine scales. Interval dissection is a way of turning the infinitely many fine scales of configuration avoidance problems into an infinite series of discrete problems involving cutting a family of intervals into smaller pieces to avoid configurations at a particular scale. The idea is simple.

The idea of interval dissection is rather simple. You've encountered it, in a suitable guise, if you've ever seen the construction of the Cantor set. Recall that the Cantor set is constructed by an infinite series of iterations, where at each scale, we have a union of intervals, and we obtain the next scale by dividing each interval into three, and removing the middle third. One multi scale property of the Cantor set is that every number in the Cantor set has a trinary expansion not containing any twos. This property is self evident when you see that at the n 'th iteration of the algorithm, the middle third sets we remove are precisely the elements of the previous iteration whose trinary expansion has a two or a three in the n th place. Since the Cantor set is obtained from applying infinitely many iterations, all possible numbers which have a two in their trinary expansion are eventually removed in the iteration process.

In general, the interval dissection techniques we consider are slightly more technical, because the property of avoiding zeroes of a real valued function on \mathbf{R}^n is much more general, and therefore not as easy to break down into discrete scale arguments like as we saw for the Cantor set. But using a queueing system, we can break our fine scale argument into discrete scales. Each piece of data in the queue will consist of a sequence of disjoint regions of the domain I_1, \dots, I_n . At each stage of the process, we have a union X_n of intervals. The idea is that at this step of the process, we take the first set of regions I_1, \dots, I_n off the front of the queue, and dissect the intervals defining X_n to form a set X_{n+1} such that for any $x_i \in I_i \cap X_{n+1}$, $f(x) \neq 0$. The set X_{n+1} is also a region of intervals, and for each choice of n intervals in the set, we add this interval to the back of the queue.

Every choice of intervals will be processed eventually, and the set $X = \lim X_n$ will avoid the required configurations, because for any distinct pair of $x_1, \dots, x_n \in X$, there is N such that each x_i occurs in a interval I_i of X_N , such that the intervals I_i are pairwise disjoint from one another and so at some

iteration X_{N+M} , any x_i with $f(x) = 0$ will have been removed, so $f(x_1, \dots, x_n)$ must be nonzero. Thus we have reduced the problem of finding configuration avoiding sets to dissection a disjoint family of unions of intervals into small unions. As should be intuitive, the more efficiently we will be able to dissect our intervals, the higher the Hausdorff dimension we should obtain. This discrete problem can be encompassed in what we call a *building block lemma*.

Note that since the interval dissection method is a discrete series of problems, it is no more difficult to avoid the zero sets of a countable collection of functions than to avoid a single function. We just make sure that for each function we have a covering of the domain by intervals of arbitrary scales that are eventually dissected to avoid the particular function.

Example. *Keleti's original problem can be seen as interval dissection. His building block lemma takes a union of intervals, and dissects them into a union of disjoint intervals containing a fraction $\Omega(1)$ of the mass of each interval at each stage. Because up to a constant, no mass is lost at each stage, Keleti obtains a Hausdorff dimension one set in the limit.*

Example. *Mathé essentially disguised an interval dissection method behind a strange topological argument to avoid configurations on a degree m polynomial. At each stage, he considers a union of intervals, divides them into length $1/N$ intervals, and takes a length $\Omega(1/N^m)$ portion of each such interval. He thus obtains a Hausdorff dimension $1/m$ solution.*

We consider a fractal avoidance problem, where we have to avoid the locus of an α dimensional set in \mathbf{R}^d . Our interval dissection method will take a union of intervals, divide them into length $1/N$ intervals, and essentially take a length $\Omega(1/N^{(d-1)/(d-\alpha)})$ of each such interval. We therefore obtain a solution to the problem with Hausdorff dimension $(d - \alpha)/(d - 1)$.

0.2 Fractal Avoidance and Random Selection

By rephrasing a configuration avoidance problem in a different light, we can view the problem from a different perspective. Once the problem is specified, the precise function f is irrelevant. We really only want to look at the zero set $Y = f^{-1}(0)$. If f is a smooth function with non-vanishing gradient, then Y will be a hypersurface, and one can apply various differential techniques to avoid the zeroes of Y . This is done, for instance, in Fraser and Pramanik's paper. But we now want to consider the case where f is non smooth, and Y can be as irregular as we wish. Thus we forget the function completely, and study what we call a *fractal avoidance problem*; we wish to find $X \subset \mathbf{R}$

For two sets $A, B \subset \mathbf{R}^n$, $A \overset{\circ}{-} B = (A \cap B) - \Delta$, where

$$\Delta = \{x : \text{there exists } i, j \text{ such that } x_i = x_j\}$$

is a union of hyperplanes. We consider a set $Y \subset \mathbf{R}^n$, and we wish to find $X \subset \mathbf{R}$ such that $X^n \overset{\circ}{-} Y = \emptyset$, i.e. for any *distinct* $x_1, \dots, x_n \in X$, $x \notin Y$ (we

say that X^d *primarily* avoids Y). Since Y is no longer the zero set of a smooth function, it can be arbitrarily fractal in nature. The *fractal avoidance problem* is to find X with a high Hausdorff dimension, characterized by the dimension of Y . Since Y now has essentially no geometric structure, we cannot rely on the previous techniques of Fraser/Pramanik and Mathé, and we shall find that a random construction finds a tight result in terms of the dimension of Y (Our current result finds X if Y is the countable union of Minkowski dimension α , but we are close to obtaining the same dimension of X if Y has Hausdorff dimension α). From the description of the interval dissection technique, it suffices to come up with an algorithm which, given a disjoint unions of intervals, dissects these intervals into smaller intervals avoiding the elements of Y . We now describe the gist of the technique we use for fractal dissection. The following lemma fullfills the requirement of a building block lemma for the construction.

Lemma 1. *Suppose $\dim_M(Y) < \alpha$, and let I_1, \dots, I_d be sets which are unions of length $1/M$ intervals. If $\beta < (d-1)/(d-\alpha)$, then for sufficiently large N , there exists subsets J_1, \dots, J_d of I_1, \dots, I_d such that if we split each I_n into length $1/N$ intervals, then J_n contains a length $1/N^\beta$ section of a fraction $1 - o(1)$ of all length $1/N$ intervals in I_n , and $J_1 \times \dots \times J_d$ avoids elements of Y .*

Proof. Since Y has Minkowski dimension bounded above by α , then for certain choice of sufficiently large N , if we divide $[0, 1]^d$ uniformly into sidelength $1/N^\beta$ cubes, Y intersects $o(N^{\alpha\beta})$ of them. Denote the set of such intervals by B , so $|B| = o(N^{\alpha\beta})$. Now split each length $1/N$ interval in each of the I_n into length $1/N^\beta$ intervals, and uniformly randomly select a single length $1/N^\beta$ interval from each of them, the union of which forming a random set S_n . We let $S = S_1 \times \dots \times S_d$. For any sidelength $1/N^\beta$ cube $C = C_1 \times \dots \times C_d$ in $I_1 \times \dots \times I_d$,

$$\mathbf{P}(C \subset S) = \prod_{n=1}^d \mathbf{P}(C_n \subset S_n) = \prod_{n=1}^d N^{1-\beta} = N^{d(1-\beta)}$$

And so if B' denotes the set of all cubes in B which are contained in S , then

$$\mathbf{E}|B'| = \sum_{C \in B} \mathbf{P}(C \subset S_n) = \sum_{C \in B} N^{d(1-\beta)} = o\left(N^{\alpha\beta + d(1-\beta)}\right)$$

Because $\alpha\beta + d(1-\beta) < 1$, $\mathbf{E}|B'| = o(N)$. In particular, we can now choose a particular instance of S for which $|B'| = o(N)$, so S is now no longer a random set. If we form the sets J_n by removing all sidelength $1/N^\beta$ intervals in S_n which form the sides of cubes intersecting Y , the sets J_1, \dots, J_n satisfy the conditions of the lemma. \square

Slight technical modifications to the queuing method using this lemma produce a dimension $(d-\alpha)/(d-1)$ set. This extends the result of Pramanik Fraser, where they construct dimension $1/(d-1)$ configurations avoiding zero sets of smooth functions with non-vanishing gradient, because the locus of such a function is an $\alpha = d-1$ dimensional hyperplane.

0.3 Low Rank Functions and Shifting?

0.4 Future Work?

After coming up with the building block lemma used in the fractal avoidance problem, we realized that the discrete lemma is really analogous to the construction of independent sets in hypergraphs, ala Turán's theorem. We are currently exploring how other techniques on construction independent sets in certain families of hypergraphs might be used to improve the dimension of X under certain assumptions on Y .

Chapter 1

Background

1.1 Ruzsa: Difference Sets Without Squares

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

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

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

we end up with satisfies

$$|X| \geq \frac{N}{\sqrt{N-1}+1} = \Omega(\sqrt{N})$$

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

In 1978, Sárközy proved an upper bound on the size of squarefree subsets of the integers, showing $D(N) = O(N(\log N)^{-1/3+\varepsilon})$ for every $\varepsilon > 0$. In particular, this proves a conjecture of Lovász that every infinite squarefree subset has density zero, because if X is any infinite squarefree subset, then

$$\frac{|X \cap [1, N]|}{N} \leq \frac{D(N)}{N} = O(\log(N)^{-1/3+\varepsilon}) = o(1)$$

Sárközy even conjectured that $D(N) = O(N^{1/2+\varepsilon})$ for all $\varepsilon > 0$. Thus the sieve technique is essentially optimal, an incredibly pessimistic point of view, since the Sieve method doesn't depend on any properties of the set of perfect squares. Ruzsa's results shows we should be more optimistic, taking advantage of the digit expansion of numbers to obtain infinite squarefree subsets X with $|X \cap [1, N]| = \Omega(N^{0.73})$. The method reduces the problem to a finitary problem of maximizing squarefree subsets modulo a squarefree integer m .

Theorem 1. *If m is a squarefree integer, then*

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

where

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

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

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

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

Remark. *Let us look at the analysis of the sieve method backwards. Rather than fixing N and trying to find optimal solutions of $[1, N]$, let's fix a particular strategy (to start with, the sieve strategy), and think of varying N and seeing how the size of the solution given by the strategy on $[1, N]$ increases over time. In our analysis, the size of a solution is directly related to the number of iterations the strategy can produce before it runs out of integers to add to a solution set. Because we apply a union bound in our analysis, the cost of each particular new iteration is the same as the cost of the other iterations. If the cost of each*

iteration was independant of N , we could increase the solution size by increasing N by a fixed constant, leading to family of solutions which increases on the order of N . However, as we increase N , the cost of each iteration increases on the order of \sqrt{N} , leading to us only being able to perform $N/\sqrt{N} = \sqrt{N}$ iterations for a fixed N . Rusza's method applies the properties of the perfect squares to perform a similar method of expansion. At an exponential cost, Rusza's method increases the solution size exponentially. The advantage of exponentials is that, since Rusza's is based on a particular parameter, a squarefree integer m , we can vary m to improve the iteration numbers more naturally.

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

Theorem 2. *Let $R \subset [1, m]$ be a subset of integers such that no difference is a power of d modulo m , where m is a squarefree integer. Construct the set*

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

Then A is squarefree.

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

For any n , we find

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

which therefore has cardinality

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

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

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

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

Theorem 3. *For every d and squarefree integer m , we can construct a set X whose differences contain no d th powers and*

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

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

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

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

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

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

Theorem 4. *If p is the least prime congruent to one modulo $2d$, then*

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

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

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

Hence

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

The claim is then completed by the following lemma. □

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

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

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

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

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

as $p \rightarrow \infty$, which is greater than $k-1$ for large enough p (this shows the theorem is essentially trivial for large enough primes, because we don't need to use any particularly interesting properties of the squares to prove the theorem). However, for smaller primes a more robust analysis is required. We shall construct a sequence $b_1, \dots, b_k \in \mathbf{Z}_p$ such that $b_i - b_j \notin Q$ for any i, j and

$$|B_j + Q| \leq 1 + j(q-1)$$

Given b_1, \dots, b_j , let b_{j+1} be any element of

$$(B_j + Q + Q) - (B_j + Q)$$

Since $b_{j+1} \notin B_j + Q$, $b_{j+1} - b_i \notin Q$ for any i . Since $b_{j+1} \in B_j + Q + Q$, the sets $B_j + Q$ and $b_{j+1} + Q$ are not disjoint (note $Q = -Q$ because $p \equiv 1 \pmod{2k}$), and so

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

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

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

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

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

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

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

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

1.2 Cantor Sets and Discrete Configurations

TODO: PROVIDE A PROOF THAT THE CANTOR SET AVOIDS ALL NUMBERS WITH TRINARY EXPANSION INCLUDING A TWO OR A THREE, AS INTUITION FOR HOW THE QUEUING PROCESS AVOIDS CONFIGURATIONS IN THE GENERAL CASE.

1.3 Keleti's Translate Avoiding Set

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

Lemma 3. *A set X avoids translates if and only if there do not exist 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 sets $X_0 \supset X_1 \supset \dots \rightarrow X$, with each X_N a union of disjoint intervals in $\mathcal{B}(L_N)$, for a decreasing sequence of lengths L_N converging to zero, to be chosen later. The set X_{N+1} is constructed from X_N

We initialize $X_0 = [0, 1]$, and $L_0 = 1$. The sets X_N are then defined recursively from the previous sets.

To aid the translate avoiding properties of the set, Keleti utilizes a queue of intervals, initially just containing $[0, 1]$. To construct X_1, X_2, \dots , Keleti iteratively performs the following procedure: After each iteration of the algorithm, we obtain a new set X_{N+1} , and so infinitely repeating the algorithm gives a sequence of sets X_1, X_2, \dots converging to a set X .

We claim that with the appropriate choice of parameters, X is a set avoiding translates. If X is not translate avoiding, there is $x_1 < x_2 \leq x_3 < x_4$ with

¹We always assume L_n/L_{n+1} is an integer so that intervals in $\mathcal{B}(L_n)$ are either almost disjoint from intervals in $\mathcal{B}(L_{n+1})$ or contained completely within such an interval

Algorithm 1 Construction of the Sets X_N

Set $N = 0$

Repeat

Take off an interval I from the front of the queue

For all $J \in \mathcal{B}(L_n)$ completely contained in X_N

Order the intervals in $\mathcal{B}(L_{n+1})$ contained in J as J_0, J_1, \dots, J_M

If $J \subset I$, add all odd indexed intervals J_1, J_3, \dots to X_{N+1}

Else Add all even indexed intervals J_0, J_2, \dots

Increase N by 1

$x_2 - x_1 = x_4 - x_3$. Since $L_n \rightarrow 0$, there is N such that x_1 is contained in an interval $I \in \mathcal{B}(L_N)$ that x_2, x_3, x_4 are not contained in. At stage N of the algorithm, the interval I is added to the end of the queue, and at a much later stage M , the interval I is retrieved.

If we follow the procedure to a suitable depth N , when the interval J is taken off the front of the queue Q , the intervals containing x_1 are shifted, whereas the intervals for x_2, x_3 , and x_4 are not shifted. To understand what this means, we find the startpoints $x_1^\circ, x_2^\circ, x_3^\circ$, and x_4° to the length L_N intervals containing x_1, x_2, x_3 , and x_4 . Thus $0 \leq x_n - x_n^\circ \leq L_N$. If we choose our parameters such that the startpoints $x_1^\circ, \dots, x_4^\circ$ all lie at integer multiples of ε_N , and $4L_N < \varepsilon_N$, then the equation $x_2 - x_1 = x_4 - x_3$ forces the discrete equation $x_2^\circ - x_1^\circ = x_4^\circ - x_3^\circ$. Now if we also force $x_2^\circ, x_3^\circ, x_4^\circ$ to lie at even multiples of Δ_N , and x_1° to lie at an odd multiple of Δ_N , which is possible by the shifting in the construction, then it is impossible for this equation to hold, so by contradiction, X must be translation invariant.

If we want to determine what makes the startpoints of the intervals in X_n lie at the appropriate multiples of Δ_n , we may assume by induction that the startpoints of X_n lie at integer multiples of Δ_n , and we now try and make the startpoints of X_{n+1} lie at integer multiples of Δ_{n+1} . If $I \subset J$ are two intervals, with $J \subset X_n$ and $I \subset X_{n+1}$, then J starts at some point $A\Delta_n$, and either I is unshifted, and starts at a point $A\Delta_n + B\varepsilon_{n+1}$, or it shifted, and is of the form $A\Delta_n + B\varepsilon_{n+1} + \Delta_{n+1}$. It is clear that the required constraints hold if $A\Delta_n + B\varepsilon_{n+1}$ is an even multiple of Δ_{n+1} for all A and B , and so we obtain the required results if $2\Delta_{n+1} \mid \Delta_n, \varepsilon_{n+1}$.

Thus to summarize, X is well defined and translation invariant provided

$$\Delta_{n+1} + N_{n+1}\varepsilon_{n+1} \leq L_n \quad 4L_n < \varepsilon_n \quad 2\Delta_{n+1} \mid \Delta_n, \varepsilon_{n+1}$$

Playing around with these inequalities shows the most elegant choices of parameters for a particular sequence N_n is

$$\Delta_n = L_n \quad \varepsilon_n = 6L_n \quad L_n = \frac{L_{n-1}}{8N_n}$$

At the N 'th scale of the algorithm, we have $\prod N_n$ intervals, each of length $(8^n \prod N_n)^{-1}$. The fact that $\prod N_n / (8^n \prod N_n) = o(1)$ implies that our set will

have Lebesgue measure zero. However, if the product $\prod N_n$ grows suitably fast, then the 8^n will be overwhelmed, so that for each ε , $(8^n \prod N_n)^{1-\varepsilon} \lesssim_\varepsilon \prod N_n$, i.e. $8^n \lesssim_\varepsilon (\prod N_n)^\varepsilon$ for each ε . This hints at the fact that $H^{1-\varepsilon}(X) = \infty$ for all $\varepsilon > 0$, so X has Hausdorff dimension one. This means setting $N_n = M$ for some constant M is *not* sufficient to obtain a Hausdorff dimension one set. Choosing a fast growing sequence, such as $N_n = n$ or $N_n = 2^{n-1}$, suffices to obtain a Hausdorff dimension one set.

Lemma 4. *The set $X = \lim X_n$ has Hausdorff dimension one.*

Proof. We use the mass distribution principle, as used in our note on calculating Hausdorff dimensions. If $X = \lim X_N$ is our set, then each X_N consists of $\prod_{n \leq N} N_n$ intervals, and so any interval I at this scale has

$$\mu(I) = \prod_{m=1}^n \frac{1}{N_m}$$

We then calculate

$$\prod_{m=1}^n \frac{1}{N_m} = \frac{8^{(n-1)(1-\varepsilon)}}{(\prod_{m=1}^n N_m)^\varepsilon} L_n^{1-\varepsilon}$$

A bound $\mu(I) \lesssim_\varepsilon L_n^{1-\varepsilon}$ is thereby obtained if for every $\varepsilon > 0$,

$$8^{(n-1)(1-\varepsilon)} \lesssim_\varepsilon \left(\prod_{n=1}^N N_n \right)^\varepsilon$$

This is true, for instance, if $N_n = n$. Since $L_n/L_{n+1} = 8N_{n+1}$, we can apply the note in the appendix to conclude $\mu(I) \lesssim_\varepsilon |I|^{1-\varepsilon}$ for all intervals I assuming that $N_{N+1} \lesssim_\varepsilon \prod_{n \leq N} N_n^\varepsilon$ for every $\varepsilon > 0$. This is true, in particular, if $N_n = n$, which applies because $n \lesssim_\varepsilon (n!)^\varepsilon$, and if $N_n = 2^{n-1}$, in which case we can use the fact that $2^n \lesssim_\varepsilon 2^{\varepsilon n^2}$. □

Remark. In Keleti's paper, he also remarks that by replacing the 8 in the choice of L_n with a slowly increasing set of even numbers, one can obtain a Hausdorff dimension one set which is linearly independent over the rational numbers. To see why this works, the condition of linear independence would fail if $a_1 x_1 + \dots + a_M x_M = 0$, where $x_1 < x_2 < \dots < x_M$, and a_1, \dots, a_M are integers with no common factor. Just as before, we find x_n° with $0 \leq x_n - x_n^\circ \leq L_N$. Provided that $(a_1 + \dots + a_M)L_N < \varepsilon_N$, and the x_n° lie at integer multiples of ε_N , we conclude that $a_1 x_1^\circ + \dots + a_M x_M^\circ = 0$. If K is an integer not dividing a_1 , then for suitably large N we assume that each $x_2^\circ, \dots, x_M^\circ$ lies at multiples of $K\Delta_N$. Shifting x_1° by a single multiple of Δ_N then breaks the equation from ever occurring in the first place. In order to guarantee this, we must first set $\varepsilon_n = A_n!L_n$ where A_n is an increasing sequence with $A_n \rightarrow \infty$.

We also guarantee that $x_2^\circ, \dots, x_M^\circ$ lies at multiples of $A_n! \Delta_n$. This can be guaranteed by induction if $A_{n+1}! \Delta_{n+1} \mid \Delta_n, \varepsilon_{n+1}$. Thus the parameters

$$\Delta_n = L_n \quad \varepsilon_n = A_n! L_n \quad L_{n+1} = \frac{L_n}{2N_{n+1} A_{n+1}!}$$

give a linearly independent set. Assuming the A_n grow incredibly slowly relative to the N_n , i.e.

$$N_n = n \quad A_n = \log \log n + O(1) \quad N_n = 2^n \quad A_n = \log n + O(1) \quad N_n = 2^{n^2} \quad A_n = n$$

then we obtain a set with Hausdorff dimension one.

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

- (i) $f(x) \neq 0$ for $x \in S_0 \times \dots \times S_d$.
- (ii) For $n \neq 0$, if we divide each interval T_n into length $1/N$ intervals, then S_n contains an interval of length C/N^d of each of these intervals.
- (iii) If T_0 is split into length $1/N$ intervals, then for a fraction $1 - 1/M$ of such intervals, S_0 is a union of length C/N^d intervals with total length C/N .

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

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

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

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

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

Remark. Suppose our equation is $f(x_0, Tx_1) = 0$, where T is a rational matrix of rank k , and x_0 is a scalar. Then we can actually obtain that $|\mathbf{B}| \lesssim |\mathbf{A}| \lesssim N^k$, since the points of \mathbf{A} are contained in a rank k lattice. Then the same lemma implies that we can actually build intervals of size $1/N^k$, leading to a dimension $1/k$ set rather than a dimension $1/(d-1)$ set (these dimensions agree if T has full rank). If our equation is $f(x_0, g_1(x_1), \dots, g_k(x_1)) = 0$, where each g_n is a rational polynomial with rank at most m , the same techniques should show that we get a dimension $1/mk$ set.

Is this true for equations of the form $f(g_1(x), \dots, g_k(x))$ as well? If we define \mathbf{B} as the set of bad points, and if $f(g_1(x_1, x_2), \dots, g_k(x_1, x_2)) = 0$. This can only be true if the zero set of f has some special structure with respect to the polynomials g_k .

Remark. The length $1/N$ portion of each interval guaranteed by (iii) is unnecessary 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 5. 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 $x \in S_1 \times \dots \times S_{d-1}$, if $f(y, x) = 0$, then $d(y, \mathbf{B}) \leq C_0 C_1 / N^{d-1}$.

Proof. Since $T_0 \times \dots \times T_d$ 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 $g : U \rightarrow I$ 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 receiving 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 receives 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

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

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

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

Thus, taking $n \rightarrow \infty$, 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 $1/(d-1)$.

1.5 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 $f : \mathbf{R}^d \rightarrow \mathbf{R}$ satisfying with $(\partial_n f)(0) = \alpha_n$,

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

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 6. 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 $|x_n - a_n| \leq \varepsilon\lambda$,

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

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

1.6 Equidistribution Results

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

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

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

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

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

$$\frac{1}{N} \sum_{n \leq N} f(n\alpha) \rightarrow \int_G f(x) dx$$

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

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

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

Theorem 6. A sequence x_1, x_2, \dots is equidistributed in \mathbf{T}^d if and only if for every nonzero $\xi \in \mathbf{Z}^n$,

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

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

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

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

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

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

Theorem 7 (Ratner). If $x_n = n\alpha + \beta$, then we can write $\alpha = \alpha_0 + \alpha_1$, where the sequence $n\alpha_0 + \beta$ is periodic in \mathbf{T}^d , and $n\alpha_1 + \beta$ is equidistributed over a subtorus of \mathbf{T}^d . In particular, if $\beta = 0$, then Ratner's theorem says that x_n has an evenly spaced subsequence equidistributed over a subtorus of the space containing the origin.

Proof. We induct on the dimension. For $d = 1$, the theorem is obvious, since either α is rational, and therefore $n\alpha + \beta$ is periodic, or α is irrational, and $n\alpha + \beta$ is equidistributed on \mathbf{T} . So now let us consider a sequence on the torus \mathbf{T}^{d+1} . If α is irrational, then x_n is equidistributed, and the theorem is obvious. Otherwise, there exists $\xi \in \mathbf{Z}^n$ such that $\xi \cdot \alpha \in \mathbf{Z}$. We may write $\alpha = \alpha_0 + \alpha_1$, where $\alpha_0 \in \mathbf{Q}^d$, and $\xi \cdot \alpha_1 = 0$. The sequence $n\alpha_0 + \beta$ is periodic, whereas $n\alpha_1$ takes values in the subtorus T of points x with $\xi \cdot x = 0$. Since $\xi \neq 0$, T is a d dimensional compact subgroup of \mathbf{T}^{d+1} isomorphic to \mathbf{T}^d . To see this, we assume for simplicity that $\xi_d \neq 0$. Then ξ^\perp is generated by the basis of d vectors $v_n = \xi_d e_n - \xi_n e_d$, for $1 \leq n \leq d$. Thus we get a homomorphism between \mathbf{R}^d and T given by the map $x \mapsto \sum x_n v_n$. If $\sum x_n v_n \in \mathbf{Z}$, so that $\sum x_n v_n = 0$ on T , then this means that $x_n \in \mathbf{Z}/\xi_d$ for each n , implying that

the kernel of this homomorphism is discrete. Lattice theory implies that we can write the kernel as $\bigoplus \mathbf{Z}\langle w_n \rangle$, for d generating vectors w_1, \dots, w_d . But then the map $f(x) = \sum x_n w_n$ gives an isomorphism between \mathbf{T}^d and T . If we set $\beta = f^{-1}(\alpha_0)$, then $f^{-1}(n\alpha_0) = n\beta$, and so by induction, we can write $\beta = \beta_0 + \beta_1$, where $n\beta_0$ is periodic, and $n\beta_1$ is equidistributed over a subtorus of \mathbf{T}^d . But then $nf(\beta_0) = f(n\beta_0)$ is periodic on T , and $nf(\beta_1) = f(n\beta_1)$ is equidistributed over a subtorus of T . Since the sum of two periodic sequences is periodic, $n\alpha_0 + nf(\beta_0)$ is periodic, and $nf(\beta_1)$ is equidistributed over a subtorus. We have $\alpha_0 + f(\beta_0) + f(\beta_1) = \alpha_0 + f(\beta) = \alpha_0 + \alpha_1$, completing the proof. \square

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

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

Theorem 8. *If x_n is equidistributed over a torus \mathbf{T}^d , and*

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

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

Proof. For notational simplicity, we let

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

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

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

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

1.7 Results about Hypergraphs

Lemma 7 (Turán). *For any k uniform hypergraph $H = (V, E)$ with $|E| \leq |V|^\alpha$, V contains an independent set of size $\Omega(|V|^{(k-\alpha)/(k-1)})$.*

Proof. We create an independent set I by the following procedure. First, select a subset S of vertices, including each independently with probability p . Delete a single vertex from each edge in each hypergraph entirely contained in S ,

obtaining an independant set I . We find that each edge in V is entirely included in S with probability p^k , and S has expected size $p|V|$, so $\mathbf{E}|I| = p|V| - p^k|E|$. If $|E| = |V|^\alpha$ for $\alpha \geq 1$, then setting $p = (1/2)|V|^{(1-\alpha)/(k-1)}$ induces a set I with size

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

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

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

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

therefore

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

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

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

which completes the proof. \square

1.8 Hyperdyadic Covers

Recall the definition of the Hausdorff measure $H^\alpha(E) = \lim_{\delta \rightarrow 0} H_\delta^\alpha(E)$, where

$$H_\delta^\alpha(E) = \inf \left\{ \sum r_n^\alpha : E \subset \bigcup B(x_n, r_n), r_n \leq \delta \right\}$$

The Hausdorff measure is well behaved with respect to the restricted class of dyadic covers, those covers with balls taking radii in the form $1/N, 1/N^2, \dots$. If $\delta \leq 1/N$, then for any $r_n \leq \delta$, we can find $1/N^{m_n}$ with $1/N^{m_n+1} \leq r_n \leq 1/N^{m_n}$, so

$$\sum 1/N^{\alpha m_n} \leq \sum r_n^\alpha \leq \sum 1/N^{\alpha(m_n+1)} = 1/N^\alpha \sum 1/N^{\alpha m_n}$$

Thus if we define the dyadic Hausdorff measure

$$H_\delta^\alpha(E) = \inf \left\{ \sum 1/N^{\alpha m_n} : E \subset \bigcup B(x_n, 1/N^{m_n}), 1/N^{m_n} \leq \delta \right\}$$

whose limit as $\delta \rightarrow 0$ is $H_N^\alpha(E)$, then $H_\delta^\alpha(E) \leq H_{\delta,N}^\alpha(E) \leq 1/N^\alpha H_\delta^\alpha(E)$, and so $H^\alpha(E) \leq H_N^\alpha(E) \leq H^\alpha(E)/N^\alpha$. This means that if $\dim_{\mathbf{H}}(E) > \alpha$, then we can cover E by a collection of dyadic sets, consisting of A_1 size $1/N$ balls, A_2

size $1/N^2$ balls, A_3 size $1/N^3$ balls, and so on and so forth, with $\sum A_n/N^{\alpha n}$ as small as we desire. Provided that we can work at dyadic scales, this bound shows that our results will connect strongly to the Hausdorff dimension.

Katz and Tao have also found a way to work at scales at exponential rather than dyadic differences. We fix a positive ε , and consider **hyperdyadic scales**, which are of the form $2^{-(1+\varepsilon)^k}$. For any n , and $\delta > 0$, the number of hyperdyadic scales between δ and δ^n is $O(\log n / \log(1/\varepsilon))$, most importantly, independent of δ . which makes things very easy to work with, rather than $O(n \log(1/\delta) / \log(1/\varepsilon))$, as in the dyadic scales. Unfortunately, a naive covering lemma does not give us good results about the Hausdorff measure with respect to hyperdyadic covers. Here choosing $2^{-(1+\varepsilon)^{k+1}} \leq r_n \leq 2^{-(1+\varepsilon)^k}$, and the multiplicative gap between the two values is not even independent of k . However, Katz and Tao found a way to work with such bounds. We assume δ and ε are some fixed parameters. If $A(\varepsilon, \delta)$ and $B(\varepsilon, \delta)$ are two quantities depending on ε and δ , we write $A \leq B$ mean $A \lesssim_{\varepsilon} \delta^{-C\varepsilon} B$ for some C , and for every ε . We let $A \approx B$ mean $A \leq B$ and $B \leq A$ hold simultaneously. We say a union of balls is δ discretized if it is the union of balls with radius $\approx \delta$. Thus there exists C_{ε} and C such that for each ball B_r of radius r , $|r - \delta| \leq C_{\varepsilon} \delta^{1-C\varepsilon}$. Thus

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

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

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

Theorem 9. *Let E be a compact subset of \mathbf{R}^n . If $0 < \alpha < n$, and $\dim(E) \leq \alpha$, then for each hyperdyadic number δ , we can associate a δ discretized set X_{δ} with $|X_{\delta} \cap B(x, r)| \leq \delta^n (r/\delta)^{\alpha}$ for all $\delta \leq r \leq 1$ and $x \in \mathbf{R}^n$, and every element of E is contained in infinitely many of the X_{δ} .*

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

$$\sum_n r_{\delta n}^{\alpha + C\varepsilon} \lesssim 1 \quad (1.1)$$

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

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

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

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

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

For a pair of hyperdyadic numbers δ and γ we set

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

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

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

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

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

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

Chapter 2

Ideas on Squarefree Subsets

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

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

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^∞ 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^∞ 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^∞ .

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 Ruzsa's method utilize the fact that the problem is a single optimization? Can we adapt Ruzsa's method work to give better results about finding subsets X of the integers such that $X - X$ is disjoint from $(X - X)^2$?

2.2 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 its dimension?

2.3 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.

2.4 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

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

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

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

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

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

Write $m = q\sqrt{N} + r$ (remember that we chose N so its square root is an integer), with $0 \leq r < \sqrt{N}$. 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.

2.5 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 8. *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 $p_1 \leq p_2 \leq \dots \leq p_K$. 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".

2.6 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

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

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 3

Low Rank Coordinate Changes

3.1 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 within 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_0N^{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 $1/N$ in-

intervals, we might be able to make points cluster more than in the original algorithm. Maybe the probabilistic 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)

3.2 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.*

3.3 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 $\pi : \mathbf{R}^k \rightarrow T$ 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 length $\Omega(1/N^k)$ 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.

3.4 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*

$$a_1x_1 + \dots + a_nx_n = 0$$

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 $n \rightarrow \infty$ to obtain a linearly independant set over the rational numbers. But since the Hausdorff dimension of our sets now only depends on the rank of T , rather than the total number of variables in T , we can let $n \rightarrow \infty$ to obtain full sets linearly independant over the rationals. More generally, for any Lipschitz

continuous function $f : \mathbf{R} \rightarrow \mathbf{R}$, we can find a full Hausdorff dimensional set such that there are no solutions

$$f(a_1x_1 + \cdots + a_nx_n, y)$$

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 isocles triangles. That is, if the curve is parameterized by $\gamma : [0, 1] \rightarrow \mathbf{R}^n$, 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 isocles triangle. This is, in a sense, a non-linear generalization of sets avoiding arithmetic progressions, since if Γ is a line, an isocles 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 isocles 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 progressions, 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 $a_1 E + \dots + a_N E$ 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 $f(a_1 E + \dots + a_N E)$ 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-vanishing derivative which vanishes over F , or a family of smooth functions with non-vanishing derivative around F .

3.5 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 $\mathbf{R}^n \rightarrow \mathbf{R}^m$ 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 $h_\alpha : U^k \rightarrow \mathbf{R}^{n-k}$ 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

$$\nabla f$$

3.6 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.

3.7 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 $T_0 \times \dots \times T_d$, 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 $x \in S_0 \times \dots \times S_d$.
- 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 $T_0 \times \dots \times T_d$, 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 $d(f(\mathbf{A} + \delta_0 e_0), \mathbf{Z}/N^m) \geq 1/K N^m$. 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?

3.8 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 $u, v \in \mathbf{Z} \cap NJ$, $(u - v)^2 \notin \mathbf{S} - \mathbf{S}$. 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 4

Extension to Arbitrary Cartesian Avoidance Problems

Using a pidgeonhole construction, we can actually extend Pramanik and Fraser's result to include avoiding the nullsets of arbitrary, nonsmooth functions. More generally, we may restructure the problem as taking a set $Y \subset [0, 1]^n$, and finding a set $X \subset [0, 1]$ such that X^n is disjoint from Y , at least when all variables are distinct. We will find that the lower the Minkowski dimension of the set Y , the more easily we will be able to avoid the nullset. We assume $\dim_M Y \geq 1$, for otherwise the problem is trivial.

Lemma 9. *Suppose $\dim_M(Y) < \alpha$, and let I_1, \dots, I_n be sets which are unions of length $1/M$ intervals. If $\beta < (n-1)/(n-\alpha)$, then for sufficiently large N , there exists subsets J_1, \dots, J_n of I_1, \dots, I_n such that if we split each I_k into length $1/N$ intervals, then J_k contains a length $1/N^\beta$ section of a fraction $1 - o(1)$ of all length $1/N$ intervals in I_k , and $J_1 \times \dots \times J_n$ avoids elements of Y .*

Proof. Since Y has Minkowski dimension bounded above by α , then for certain choice of sufficiently large N , if we divide $[0, 1]^n$ uniformly into sidelength $1/N^\beta$ cubes, Y intersects $o(N^{\alpha\beta})$ of them. Denote the set of such intervals by B , so $|B| = o(N^{\alpha\beta})$. Now split each length $1/N$ interval in each of the I_k into length $1/N^\beta$ intervals, and uniformly randomly select a single length $1/N^\beta$ interval from each of them, the union of which forming a random set S_k . We let $S = S_1 \times \dots \times S_n$. For any sidelength $1/N^\beta$ cube $C = C_1 \times \dots \times C_n$ in $I_1 \times \dots \times I_n$,

$$\mathbf{P}(C \subset S) = \prod_{k=1}^n \mathbf{P}(C_k \subset S_k) = \prod_{k=1}^n N^{1-\beta} = N^{n(1-\beta)}$$

And so if B' denotes the set of all cubes in B which are contained in S , then

$$\mathbf{E}|B'| = \sum_{C \in B} \mathbf{P}(C \subset S) = \sum_{C \in B} N^{n(1-\beta)} = o\left(N^{\alpha\beta+n(1-\beta)}\right)$$

Because $\alpha\beta + n(1-\beta) < 1$, $\mathbf{E}|B'| = o(N)$. In particular, we can now choose a particular nonrandom instance of S for which $|B'| = o(N)$, so S is now no longer a random set. If we form the sets J_k by removing all sidelength $1/N^\beta$ intervals in S_k which form the sides of cubes intersecting Y , the sets J_1, \dots, J_n satisfy the conditions of the lemma. \square

Remark. Recall Turán's theorem for hypergraphs, which says if a n uniform hypergraph on a vertex set V contains $|V|^\alpha$ edges, then there exists an independent subset of V of size $\Omega(|V|^{(n-\alpha)/(d-1)})$. If we take the length $1/N^\beta$ cubes in the lemma above, and form a hypergraph containing the edge I_1, \dots, I_d if and only if $I_1 \times \dots \times I_d$ intersects Y , then our lemma then becomes a variant of Turán's theorem, where because the edges between vertices of our hypergraph are 'localized' into $1/N$ quadrants, we can find an independent set which is locally large essentially everywhere.

Remark. The soft bounds in the inequality depend on the ability to cover Y efficiently for small N . If $\kappa = 1 - \alpha\beta - n(1-\beta)$, we need $N \geq 2/\log(\kappa)$ to maintain half of all intervals, which can be increased as slowly as possibly even for dyadic changes in scales. If we are only considering a single Y , we increase scales by a polynomial factor x^β each time.

If we consider a queueing process analogous to the queueing processes considered in the lemmas above, then we can apply the lemma iteratively. If we apply different parameters α_n and β_n with α_n converging to $\dim_M(Y)$ rapidly, and β_n converging rapidly to $(n-1)/(n-\alpha)$, then we end up with a set X of Hausdorff dimension $(n-\alpha)/(n-1)$ with X^n avoiding elements of Y when variables are distinct. Of course, since the queueing process is sequential, we can actually stack multiple avoidance problems together, to avoid sets Y which are the countable union of sets of Minkowski dimension α , i.e. sets Y with packing dimension $\dim_P(Y) < \alpha$.

Remark. The theorem naturally generalizes to a multivariate setting of the problem, where X lies in $[0,1]^d$ and we must find such a set X with $X^n \cap Y$ essentially avoiding one another. If $\dim_M(Y) < \alpha$, we can cover Y by $o(N^{\alpha\beta})$ sidelength $1/N^\beta$ cubes. If we consider I_1, \dots, I_n as in the previous problem, and select a sidelength $1/N^\beta$ cube from a length $1/N$ cube in I_k . Then the probability of selecting any cartesian product of cubes is $N^{dn(1-\beta)}$. Thus the expected number of intersections of these cubes with Y is $O(N^{\alpha\beta+dn(1-\beta)})$. Provided that $\alpha\beta + dn(1-\beta) < d$, the vast majority of intervals don't conflict, so we can set $\beta > d(n-1)/(dn-\alpha)$. This leads to a Hausdorff dimension $(nd-\alpha)/(n-1)$ set X . If $dn-\alpha > d(n-1)$, $\alpha < d$, which makes the problem trivial.

Remark. We can avoid the necessity that the variables are distinct in the avoidance problem above by repeating the lemma with slightly different sets Y . For

instance, in order to allow for two of the variables in X to be equal, while still avoiding elements of Y , we can consider the problem of ensuring X^{d-1} avoids a set of the form

$$\tilde{Y}_{ij} = \{y : (y_1, \dots, y_{i-1}, y_j, y_i, \dots, y_{d-1}) \in Y\}$$

Thus $\tilde{Y}_{ij} = f_{ij}^{-1}(Y)$, where $f_{ij}(y) = (y_1, \dots, y_{i-1}, y_j, y_i, \dots, y_{d-1})$. Thus we have $\dim_M(\tilde{Y}_{ij}) < \dim_M(Y)$, and so we can avoid \tilde{Y}_{ij} with dimension $(d-1-\alpha)/(d-2)$. In particular, if $\alpha < 1$, we can find a set X of Hausdorff dimension one such that X^d avoids elements of Y , except when all the variables are equal to one another. But in this case the dimension is trivial, since the complement of the projection on each coordinate is essentially the whole set.

Remark. Our random construction is essentially a disguised use of pidgeonholing. Define $\mathbf{A}_1, \dots, \mathbf{A}_d$ to be the space of all choices of all sets obtained from taking a length $1/N^\beta$ interval from each length $1/N$ interval in the respective sets I_1, \dots, I_n . Then each set \mathbf{A}_n has cardinality $(N^{\beta-1})^{N|I_n|}$. Define a positive integer-valued function f on $\mathbf{A}_1 \times \dots \times \mathbf{A}_d$ by letting $f(J_1, \dots, J_d)$ be the number of length $1/N^\beta$ boxes in $J \cap Y$. We will now calculate

$$\sum_{J_n \in \mathbf{A}_n} f(J_1, \dots, J_d)$$

For each length $1/N^\beta$ interval $J_0 \subset I_n$, the number of J_n with $J_0 \subset J_n$ is $(N^{\beta-1})^{N|I_n|-1}$, since we obtain all such J_n by fixing the choice of J_0 in the length $1/N$ interval it is contained within in I_n , and then considering arbitrary choices in the other length $1/N$ intervals. Thus if Y intersects K length $1/N^\beta$ cubes in $I_1 \times \dots \times I_d$, each cube is contained in precisely $\prod (N^{\beta-1})^{N|I_n|-1}$ tuples (J_1, \dots, J_d) , and so

$$\sum_{J_n \in \mathbf{A}_n} f(J_1, \dots, J_d) = K \prod (N^{\beta-1})^{N|I_n|-1}$$

Thus the average value of f on $\mathbf{A}^1 \times \dots \times \mathbf{A}_d$ is

$$\frac{K \prod (N^{\beta-1})^{N|I_n|-1}}{\prod (N^{\beta-1})^{N|I_n|}} = \frac{K}{N^{d(\beta-1)}}$$

In particular, we can find some J_1, \dots, J_d such that $f(J_1, \dots, J_d) \leq K/N^{d(\beta-1)}$. If $K \lesssim N^{\alpha\beta}$,

$$K/N^{d(\beta-1)} \lesssim N^{\alpha\beta-d(\beta-1)} = o(N)$$

Discarding all intervals $I \subset J_n$ which form the sidelength of a cube in J intersecting Y , we only lose a fraction $1-o(1)$ of the length $1/N^\beta$ intervals originally contained in J_n , and the resulting set satisfies the required properties of the map.

We now consider some applications of the theorem above to particular scenarios of configuration avoidance problems. TODO: CALCULATE THE DIMENSIONS BELOW MORE PRECISELY.

Example. Let Y have dimension α . If $f_1(x, y) = x + y$, $f_2(x, y) = x - y$, then $f_1^{-1}(Y) \cup f_2^{-1}(Y)$ has dimension $d + \alpha$, so we can find a set X of dimension $d - \alpha$ with $X + X$ and $X - X$ avoiding the elements of Y . We hope to improve this to where X is a vector space over \mathbf{Q} .

Example. If $Y \subset [0, 1]$ has dimension α , given the function $f(x, y) = |x - y|$, $f^{-1}(Y)$ has dimension bounded by $\alpha + 2d - 1$, so we can find a set X with dimension $1 - \alpha$ avoiding elements of Y .

Example. If we consider $Y \subset S^{d-1}$ a countable union of sets of Minkowski dimension α , we can find $X \subset \mathbf{R}^d$ such that $f(x, y) = (x - y)/|x - y|$ avoids elements of Y . We find $f^{-1}(Y)$ is $\alpha + d + 1$ dimensional, so we can find X with dimension $d + 1 - \alpha$ avoiding the directions Y .

Example. Given a set $Y \subset \mathbf{R}$, consider the problem of finding a set X in \mathbf{R}^d such that the angles formed by elements of X avoid elements of Y . Then the function is

$$f(x, y, z) = \frac{(x - y) \cdot (x - z)}{|x - y||x - z|}$$

and we want to avoid elements of $f^{-1}(\cos(Y))$. If Y is α dimensional, then $f^{-1}(\cos(Y))$ is $\alpha + 2d + (d - 1)$ dimensional, and so we can find a set X of Hausdorff dimension $1 - \alpha$ avoiding the angles in Y .

4.1 From Minkowski to Hausdorff dimension

We now try to extend the algorithm to the case where Y has Hausdorff dimension of Y , we need to avoid sets using the following construction: For any length $1/N$, we can cover Y by a family of cubes, each of length $1/N^n$ for some k , with union Y_0 . If this cover is by A_n length $1/N^n$ cubes, and $H_\alpha(Y) < \infty$, then for sufficiently large N we can choose the cover with

$$\sum \frac{A_n}{N^{\alpha n}} = H_\alpha(Y) + o(1)$$

Suppose that we can choose J_1, \dots, J_d controlling the area of the intersection $|J \cap Y_0|$. Then, as we consider the queueing process, the resulting set may not have nonempty intersection with Y , but the control on the intersection at the discrete scale may imply that for the end set X , the Hausdorff dimension of $X^d \cap Y$ is bounded by β . Provided that the Hausdorff dimension of X is greater than β , we can just remove the projections of Y onto each coordinate from X , without decreasing the Hausdorff dimension for X , leading to a set with X^d whose distances avoid elements of Y . Now consider the same random choice J_1, \dots, J_d as in the last argument. Then for any cube with sidelength $1/N^n$, with β dividing n , the probability that the cube is contained in J is the same as the probability that the unique cube with sidelength $1/N^\beta$ containing this cube is contained in J , which is $N^{d(1-\beta)}$. Thus $J \cap Y$ is covered by a random

subset of the family of cubes Y_0 , with each cube appearing in the cover with probability $N^{d(1-\beta)}$. Thus

$$\mathbf{E}[H_\alpha^{1/N}(J \cap Y)] \leq \sum \frac{A_n N^{d(1-\beta)}}{N^{\alpha n}} \leq N^{d(1-\beta)} H_\alpha(Y)$$

Applying Jensen's inequality, we find

$$\mathbf{E}[H_\alpha(J \cap Y)] \leq \mathbf{E}[H_\alpha^{1/N}(J \cap Y)] \leq N^{d(1-\beta)} H_\alpha(Y)$$

In particular, we can choose J such that $H_\alpha(J \cap Y) \leq N^{d(1-\beta)} H_\alpha(Y)$. Can we use the decay condition here to help us?

Josh suggested an alternate idea. Given Y , we consider a quantity ε , take the dyadic numbers $\delta_n = 2^{-(1+\varepsilon)^n}$, and then consider a hyperdyadic cover of A_n sidelength δ_n cubes. The condition that Y has Hausdorff dimension α guarantees that we have $A_n \lesssim_\varepsilon (1/\delta)^\alpha$, for some constant C independent of ε . The building block lemma then takes the following form. We are given a union of intervals X_n , each of length δ_n , and we must dissect this set into a union X_{n+1} of intervals of length δ_{n+1} , such that the cartesian product of any two disjoint set of intervals in X_{n+1} avoids the A_{n+1} cubes of length δ_{n+1} in the cover of Y . Suppose we take β length δ_{n+1} intervals from each length δ_n interval in X_n at random to form X_{n+1} . What does the distribution of $X_{n+1}^d - \Delta_{n+1}$ look like, where Δ_{n+1} is the union of cartesian products of length δ_{n+1} cubes containing a duplicate of the same interval. Any interval in $[0, 1]^d - \Delta_{n+1}$ is contained in X_{n+1}^d essentially with probability $(\beta \delta_{n+1}/\delta_n)^d$ (duplicates will be chosen with negligible probability), and so $X_{n+1}^d - \Delta_{n+1}$ contains, on average

$$A_{n+1}(\beta \delta_n/\delta_{n+1})^d \lesssim_\varepsilon \beta^d (1/\delta_{n+1})^{\alpha+C\varepsilon} (\delta_{n+1}/\delta_n)^d = \beta^d \delta_{n+1}^{d+\alpha+C\varepsilon} / \delta_n^d$$

elements of the cover of Y at the scale δ_{n+1}

? The probability that $X_{n+1}^d - \Delta$

Chapter 5

Appendix: Dimensions of Cantor Like Sets

This section provides tools to calculate the Hausdorff dimension of Cantor like sets. For simplicity, we work in one dimension. We analyze the dimension of sets X constructed as the limit of a nested family of sets $X_0 \supset X_1 \supset \dots$, each the union of intervals of a fixed length. For such a construction, we can naturally associate a sequence of Borel probability measures μ_N supported on X_N , weakly converging to a measure μ supported on X . Our main tools will establish estimates about μ through estimates on μ_N , which then allows us to apply Frostman's lemma.

To be precise, we consider a decreasing sequence of lengths L_N which for simplicity we assume is such that $L_N/L_{N+1} \in \mathbf{N}$. Then we require that X_N is a union of intervals in $\mathcal{B}(L_N)$, and each interval in $\mathcal{B}(L_N)$ contained in X_N contains at least one interval in $\mathcal{B}(L_{N+1})$ contained in X_{N+1} . We let μ_0 be the uniform probability measure on X_0 . Then μ_{N+1} is constructed from μ_N by taking each interval $I \in \mathcal{B}(L_N)$ and distributing the mass $\mu_N(I)$ equally over each interval $J \in \mathcal{B}(L_{N+1})$ contained in X_{N+1} . It is easy to see from the distribution functions of the μ_N , which converge pointwise, that the measures μ_N must converge weakly to some measure μ supported on X . Most importantly, it satisfies $\mu(I) = \mu_N(I)$ for each $I \in \mathcal{B}(L_N)$. This enables us to obtain estimates on μ from estimates on the μ_N , which follow essentially from combinatorial arguments.

Remark. *I am still not sure how ‘natural’ μ is to the problem. If $\mu(I) \lesssim |I|^\alpha$ is not satisfied for all I , does it follow that $\dim_{\mathbf{H}}(X) \leq \alpha$? It seems this probability measure is strongly related to the structure of the Cantor set so this isn't entirely implausible.*

5.1 Basic Covering

It is mostly easy to establish estimates of the form $\mu(I) \lesssim |I|^\alpha$ for $I \in \mathcal{B}(L_N)$, independent of N ; the measures μ_N are discrete, and therefore bounds on them can be obtained from combinatorial arguments. To obtain general bounds for all intervals I , we must rely on covering arguments. For any interval I with $L_{N+1} \leq |I| \leq L_N$, there are two obvious options to choose from. We can cover I by $O(1)$ intervals in $\mathcal{B}(L_N)$, obtaining $\mu(I) \lesssim L_N^\alpha = (L_N |I|^{-1})^\alpha |I|^\alpha$. Alternatively, we can cover I by $O(|I| L_{N+1}^{-1})$ intervals in $\mathcal{B}(L_{N+1})$, so we find $\mu(I) \lesssim (|I| L_{N+1}^{-1}) L_{N+1}^\alpha = (|I| L_{N+1}^{-1})^{1-\alpha} |I|^\alpha$. By considering both covering techniques simultaneously, optimizing for any particular interval length $|I|$, we obtain the tighter bound

$$\begin{aligned} \mu(I) &\lesssim \min \left((L_N |I|^{-1})^\alpha, (|I| L_{N+1}^{-1})^{1-\alpha} \right) |I|^\alpha \\ &\lesssim (L_N / L_{N+1})^{\alpha(1-\alpha)} |I|^\alpha \end{aligned}$$

Thus the general estimate $\mu(I) \lesssim |I|^\alpha$ is obtained provided that $L_N / L_{N+1} = O(1)$, which means L_N can grow at most exponentially. Unfortunately, this rarely occurs even in the most basic constructions.

5.2 An Epsilon of Room

We can obtain slightly more useful results by giving ourselves an epsilon of room. Factoring in an extra $|I|^\varepsilon$ into the minimization before optimizing our choice of covering, we find

$$\begin{aligned} \mu(I) &\lesssim \min \left((L_N |I|^{-1})^\alpha |I|^\varepsilon, (|I| L_{N+1}^{-1})^{1-\alpha} |I|^\varepsilon \right) |I|^{\alpha-\varepsilon} \\ &= \min \left((L_N |I|^{\varepsilon/\alpha-1})^\alpha, (|I|^{1+\varepsilon/(1-\alpha)} L_{N+1}^{-1})^{1-\alpha} \right) |I|^{\alpha-\varepsilon} \\ &\leq (L_N / L_{N+1})^{\alpha(1-\alpha)} (L_N^\alpha L_{N+1}^{1-\alpha})^\varepsilon |I|^{\alpha-\varepsilon} \end{aligned}$$

Thus if $L_N / L_{N+1} = O_\varepsilon(L_N^{-\varepsilon})$ for every $\varepsilon > 0$, then we obtain $\mu(I) \lesssim_\varepsilon |I|^{\alpha-\varepsilon}$ for each $\varepsilon > 0$, which is good enough to conclude $\dim_{\mathbf{H}}(X) \geq \alpha$. We should expect this bound to be much more versatile than the previous bound. The ratios L_N / L_{N+1} are obtained at a single scale of the iteration, whereas the lengths L_N are obtained from compounding lengths over many, many scales, and as such should have enough kick to overwhelm the ratio even when ε is arbitrarily small. In particular, this method applies if $L_{N+1} = \Omega(L_N^\alpha)$ for some universal parameter $\alpha \geq 1$. So a polynomial rate of decay is now allowed.

Example. If $L_N = e^{-r_N}$, then $L_N / L_{N+1} = e^{r_{N+1} - r_N}$, and $L_{N+1}^{-\varepsilon} = e^{\varepsilon r_{N+1}}$. Thus it suffices to show $r_{N+1} - r_N - \varepsilon r_{N+1} \leq 0$ for suitably large N depending on ε . If r_N is differentiable in N , and accelerating as N increases, the mean value theorem implies that the bound is implied by the fact that $r'_{N+1} - \varepsilon r_N \leq 0$. Thus, in particular, we may apply the bound if $L_N = \exp(-x^M)$, for any fixed $M > 0$.

Example. *Kaleti constructs a full dimensional fractal set avoiding translates of itself using the Cantor set method. He has $L_N = 1/8^{N-1}N!$, in which case $L_N/L_{N+1} = 8N$, which is eventually bounded by $L_{N+1}^{-\varepsilon} = (8^N(N+1)!)^\varepsilon$ for any $\varepsilon > 0$. The discrete bounds for $I \in \mathcal{B}(L_N)$ are easily established in this case, enabling us to conclude Kaleti's set is full dimensional.*

It is, of course, easy to construct superexponential examples of L_N which increase fast enough that we cannot use this technique. For instance, if $L_N = e^{-N!}$, or $L_N = e^{-e^N}$, then these results cannot be implied. In the next section, we provide a method which works when the lengths L_N have an *arbitrary* rate of decay, given additional *uniform* selection assumptions on the Cantor set construction.

5.3 Bounds on Uniform Constructions

Our final method for interpolating requires extra knowledge of the dissection process, but enables us to choose the L_N arbitrarily rapidly, so the mass of μ is distributed at lacunary lengths. The idea behind this is that there is an additional sequence of lengths R_N with $L_N \leq R_N \leq L_{N-1}$. The difference between R_N and L_{N-1} is arbitrary, but the difference between L_N and R_N is polynomial, which enables us to use the covering methods of the previous section. In addition, we rely on a 'uniform mass bound' between R_N and L_N to cover the remaining classes of intervals. Because we can take R_N arbitrarily large relative to L_N , this renders any constants that occur in the construction to become immediately negligible. This means we can choose L_N slightly more freely with relation to R_N .

One important feature of this method is that the Cantor set construction depends on parameters which can be freely changed to obtain the required bounds. We consider sequences $\beta = \{R_1 \geq R_2 \geq \dots\}$ of parameters determining parameters L_1, L_2, \dots , with L_N depending only on R_1, \dots, R_N , and a Cantor set X^β with a corresponding probability measure μ_β supported on X^β . Our goal is to find a particular sequence β such that X^β has Hausdorff dimension exceeding α . For two quantities A and B , we will let $A \lesssim_N B$ stand for an inequality with a hidden constant depending only on parameters with index smaller than N , i.e. $A \leq C(R_1, \dots, R_N)B$ for some constant $C(R_1, \dots, R_N)$.

Theorem 12. *If $\mu_\beta(I) \lesssim_{N-1} L_N^\alpha$ for $I \in \mathcal{B}(L_N)$, we have a controlled change of scales $R_N^{1+\varepsilon} \lesssim_{N-1,\varepsilon} L_N$ for all $\varepsilon > 0$, and we have a uniform distribution bound $\mu_\beta(J) \lesssim_{N-2} (R_N/L_{N-1})\mu_\beta(I)$ for $J \subset I$ with $J \in \mathcal{B}(R_N)$ and $I \in \mathcal{B}(L_{N-1})$, then for suitably fast decaying sequences β , $\dim_{\mathbf{H}}(X^\beta) \geq \alpha$.*

Proof. If I is an interval with $R_{N+1} \leq |I| \leq L_N$, we can cover I by elements of $\mathcal{B}(R_{N+1})$, then applying the uniform distribution bound to obtain that for any $\varepsilon > 0$,

$$\mu_\beta(I) \lesssim_{N-1} (|I|/R_{N+1})(R_{N+1}/L_N)L_N^\alpha = |I|/L_N^{1-\alpha} \leq |I|^{\alpha-\varepsilon} L_N^\varepsilon$$

Since the hidden constants only depend on R_1, \dots, R_{N-1} , and therefore not on R_N , we can pick R_N so that L_N is suitably large to give a parameter independent

bound. To be rigorous, for each N , we pick R_N suitably large, depending on R_1, \dots, R_{N-1} such that the bound $\mu_\beta(I) \leq |I|^{\alpha-\varepsilon}$ holds for all $\varepsilon \geq 1/N$ when $R_{N+1} \leq |I| \leq L_N$. By doing this, we obtain that $\mu_\beta(I) \lesssim_\varepsilon |I|^{\alpha-\varepsilon}$ for all intervals I where there is N such that $R_{N+1} \leq |I| \leq L_N$.

On the other hand, the uniform distribution bound implies that $\mu_\beta(I) \lesssim_{N-1} |I|^\alpha$ for $I \in \mathcal{B}(R_N)$. If there is N such that $L_N \leq |I| \leq R_N$, this means we apply the epsilon of room trick in the last section to find

$$\mu_\beta(I) \lesssim_{N-1} (R_N/L_N)^{\alpha(1-\alpha)} R_N^\varepsilon |I|^{\alpha-\varepsilon}$$

Thus if $R_N/L_N \lesssim_{N-1,\varepsilon} R_N^{-\varepsilon}$ for all ε , we can take R_N suitably large to annihilate the constant in the bound above. This means if β is suitably fast decreasing, X^β is an α dimensional set. \square

Example. In Fraser and Pramanik's paper, for $\alpha = 1/(d-1)$, they choose $R_N = A_N L_{N+1}^\alpha$ for some constant A_N , and choose L_N suitably rapidly decaying to establish the inequalities $\mu_\beta(I) \lesssim |I|^\alpha$ for $|I| = L_N$. The decomposition algorithm naturally leads to the uniform distribution inequality $\mu_\beta(J) \lesssim (R_N/L_N) \mu_\beta(I)$, which shows that we can choose L_N giving an α dimensional set.

Bibliography

- [1] I. Z. Ruzsa *Difference Sets Without Squares*
- [2] Tamás Keleti *A 1-Dimensional Subset of the Reals that Intersects Each of its Translates in at Most a Single Point*
- [3] Robert Fraser, Malabika Pramanik *Large Sets Avoiding Patterns*
- [4] Karl Dickman *On the Frequency of Numbers Containing Prime Factors of a Certain Relative Magnitude*
- [5] B. Sudakov, E. Szemerédi, V.H. Vu *On a Question of Erdős and Moser*