

Squarefree Sets

Jacob Denson

July 25, 2018

Contents

1	Background	2
1.1	Rusza: Difference Sets Without Squares	2
1.2	Keleti: Translate Avoiding Sets	7
1.3	Fraser/Pramanik: Extending Keleti Translation to Smooth Configurations	10
1.4	A Set Avoiding All Functions With A Common Derivative	13
2	Ideas on Squarefree Subsets	15
2.1	Ideas For New Work	15
2.2	Squarefree Sets Using Modulus Techniques	16
2.3	Idea; Delaying Swaps	19
2.4	Squarefree Subsets Using Interval Dissection Methods	19
2.5	Finding Many Startpoints of Small Modulus	21
2.6	A Better Approach	22
3	Low Rank Coordinate Changes	23
3.1	Boosting the Dimension of Pattern Avoiding Sets by Low Rank Coordinate Changes	23
3.2	Extension to Well Approximable Numbers	25
3.3	Equidistribution on the Torus To Extend the Result to Irrational Matrices	26
3.4	Equidistribution and Real Valued Matrices	26
3.5	Applications of Low Rank Coordinate Changes	27
3.6	Idea: Generalizing This Problem to low rank smooth functions .	27
3.7	Idea: Algebraic Number Fields	28
4	Appendix: Calculating Dimensions of Cantor Like Sets	29

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. Recall that if X and Y are subsets of integers, we let

$$X \pm Y = \{x \pm y : x \in X, y \in Y, x \pm y > 0\}$$

The *differences* of a set X are elements of $X - X$. The *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.

It is easily to greedily construct fairly large subsets of the integers by 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 $|X \cap [1, N]| = \Omega(\sqrt{N})$.

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 $\varepsilon > 0$. This effectively implies the greedy sieve technique of selecting squarefree subsets of the integers is asymptotically optimal. Since the Sieve method doesn't depend on any properties of the set of perfect squares¹, this is incredibly pessimistic. Ruzsa's paper shows we should be more optimistic, taking advantage of the structure of perfect squares 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*

¹In general, if $S = \{x_1, x_2 < \dots\}$ is a sequence of positive integers, the sieve strategy on $[1, N]$ produces a set containing no ' S differences' with at least $N(1 + K(N))^{-1}$ elements, where $K(N)$ is the greatest integer with $x_{K(N)} \leq N - 1$

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 make the exponentials match up how we like to obtain a better polynomial lower bound.

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 generally, Rusza's method works on the problem of constructing subsets of the integers whose differences are d 'th powers-free. 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

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

we claim that A is squarefree. 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 $s = ds_0$, then

$$(N/m^{s_0})^d = (r_s - r'_s) + Mm$$

and this contradicts the fact that $r_s - r'_s$ cannot be a d 'th power modulo m . On the other hand, we know m^s divides N^d , but m^{s+1} does not. This is impossible if s is not divisible by d , because primes in N^d occur in multiples of d , and m is squarefree. For any n , we find

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

which therefore has cardinality

$$\begin{aligned} |R|^{1+[n-1/d]} m^{n-1-[n-1/d]} &= m^n \left(\frac{|R|}{m} \right)^{1+[n-1/d]} \\ &\geq m^n \left(\frac{|R|}{m} \right)^{n+1/d} = \frac{(m^{n+1})^{1-1/d+\log_m |R|/d}}{m} \\ &= \frac{(m^{n+1})^{\gamma(m,d)}}{m} \end{aligned}$$

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

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

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

Theorem 2. *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} . 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 3. *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 1. *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 Keleti: Translate Avoiding Sets

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 $t \in \mathbf{R}$. Malabika has adapted this technique to construct high dimensional subsets avoiding nontrivial solutions to differentiable functions, and she thinks we can further exploit these ideas to obtain dimension one squarefree sets. The basic, but fundamental idea of Keleti is to introduce memory into Cantor set type constructions so the sets avoid progressions, and have dimension one. At each point in the process, he constructs a nested family of sets $X_0 \supset X_1 \supset \dots$, each set X_n a union of disjoint intervals of the same length l_n . It will be the set $X = \lim X_n$ which avoids translates. Initially, $X_1 = [0, 1]$ is the unit interval. To aid in this process, Keleti considers a queue Q with a history of all intervals created from the start to the end of the process, so Q initially just contains $[0, 1]$. We then perform the following procedure repeatedly:

- Take an interval I off the front of the queue Q .
- Let X_{n+1} be formed by taking N_{n+1} intervals of length l_{n+1} from each interval J in X_n , with the startpoints of each interval separated by ε_{n+1} . The intervals start directly at the beginning of the interval J , except in the case where $J \subset I$, where we shift the intervals to begin at a distance Δ_{n+1} from the startpoint.

In order for this process to be well defined, we must have $l_1 = N_1 = 1$, $\varepsilon_n \geq l_n$, and $\Delta_{n+1} + N_{n+1}\varepsilon_{n+1} \leq l_n$. Of course, we must also have $l_n \rightarrow 0$, for otherwise we have no fractal structure in our sets. We claim that, with the appropriate choice of parameters, X is a set avoiding translates. In this section, and in the sequel, we shall find it is most convenient to avoid certain configurations by expressing them as a particular equation, whose properties we can then exploit.

Lemma 2. *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 then $a - t < a \leq b - t < b$ 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

We now claim that, if the constants l_n, N_n, ε_n , and Δ_n possess suitable properties, then the set X is translation avoiding. We suppose the existence of

$x_2 - x_1 = x_4 - x_3$ as in the lemma, with $x_1 < x_2 \leq x_3 < x_4$. Since $l_n \rightarrow 0$, we know that eventually x_1 is contained in an interval J that x_2, x_3 , and x_4 are not contained in. 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 3. *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.

Remark. We attempted to obtain a squarefree subset of $[0, 1]$ by combining Ruzsa's squarefree discrete strategy with Keleti's decomposition approach to find a high dimensional continuous squarefree set. However, using these techniques we were only able to obtain a dimension $1/2$ set, which is only slightly better than a dimension $1/3$ set which exists from the general results given by Mathé's result, or Pramanik and Fraser's result, and is much less than the dimension 1 set that Malabika expects.

1.3 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 4. *Suppose that $f : \mathbf{R}^d \rightarrow \mathbf{R}$ is a C^1 function, and there are sets $T_1, \dots, T_d \subset [0, 1]$, which each T_n a union of almost disjoint closed intervals of length $1/M$ such that $A \leq |\partial_d f(x)|$ and $|\nabla f(x)| \leq B$ for $x \in T_1 \times \dots \times T_d$. There there exists a rational constant C and arbitrarily large integers $N \in \mathbf{M}\mathbf{Z}$ for which there exist subsets $S_n \subset T_n$ such that*

- (i) $f(x) \neq 0$ for $x \in S_1 \times \dots \times S_d$.
- (ii) For each $n \neq d$, If we split each interval T_n into $1/N$ intervals, then S_n contains an interval of length CN^{1-d} of each of these intervals.
- (iii) If we split T_d into $1/N$ intervals, then S_d contains a length C/N portion of at least a fraction $1 - 1/M$ of the total number of these intervals. This portion need not be a complete interval, like in the last property, but it is a union of length C/N^{d-1} intervals.

Proof. Choosing the sets S_n for $n \neq d$ is easy. We split the intervals T_n into length $1/M$ segments, and define S_n as the union of the first CN^{1-d} portion of them. This satisfies property (ii) of the theorem automatically. Now if $a \in \mathbf{R}^{d-1}$ is chosen, with each $a_k \in T_k$, then the total number of points $x \in T_d$ for which $f(a, x) = 0$ is M . This follows from the fact that T_d can contain at most M intervals of length $1/M$, and as we vary x , because the partial derivative $\partial_d f(a, x)$ is non-vanishing on the interval, the function is monotone on this interval. Define

$$\mathbf{A} = \{a : a_n \text{ is a startpoint of a length } 1/N \text{ segment in } T_n\}$$

Then $|\mathbf{A}| \leq N^{d-1}$, since T_n , contained in $[0, 1]$, can contain at most N almost disjoint intervals of length $1/N$. This means that if

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

is the set of 'bad points' in T_d , then $|\mathbf{B}| \leq MN^{d-1}$. Now we filter out a subcollection of intervals I if $|\mathbf{B} \cap I| \leq M^3 N^{d-2}$. Then we throw away at most $MN^{d-1}/M^3 N^{d-2} = N/M^2$ intervals. Thus we keep $N/M - N/M^2 = N/M(1 - 1/M)$ intervals, which is $1 - 1/M$ of the total number of intervals in the decomposition of T_d . Now we split each interval I in the subcollection into $(1/4C_0C)N^{d-2}$ subintervals of length $4C_0CN^{d-1}$, and remove any interval

which contains a bad point, or is adjacent to an interval containing a bad point, then what remains will consist of the interval T_d . It contains at least $M^3 N^{d-2}$ intervals from each interval I we didn't discard, so for any such interval I , $|T_d \cap I| \geq 1/4N$.

Now we split each interval I into intervals of length $C_0 C / N^{d-1}$. We now use the constant C_0 obtained from the lemma below, and discard any interval that intersects \mathbf{B} , or is adjacent to an interval intersecting \mathbf{B} . What remains is our set T_d . Since $\mathbf{B} \cap I$ has at most $M^3 N^{d-2}$ points, and I is decomposed into $N^{d-2}/C_0 C$ intervals, what remains is $N^{d-2}/C_0 C - 3M^3 N^{d-2} = (1/C_0 C - 3M^3)N^{d-2}$ intervals, which in total has length

$$(C_0 c / N^{d-1})(1/C_0 c - 3M^3)N^{d-2} = \frac{(1 - 3M^3 C_0 c)}{N}$$

Choosing C such that $3M^3 C_0 C < 1 - C$ completes the proof. \square

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 4. For any f and T_1, \dots, T_d , and C , there exists a constant C_0 depending on these quantities, such that for the choice of S_1, \dots, S_{d-1} , for any choice x_1, \dots, x_{n-1} in each of the sets S_n , if $f(x_1, \dots, x_{n-1}, x_n) = 0$, then $d(x_n, \mathbf{B}) \leq C_0 C / N^{d-1}$.

Proof. Let J be a d dimensional cube with sidelengths $1/M$ intersecting the zero set of f . The implicit function theorem can then be applied to conclude that, if we write $J = J' \times I$, where J_1 is a cube in \mathbf{R}^{d-1} and I is an interval, that there is a function $g : J' \rightarrow \mathbf{R}$ such that $f(x, y) = 0$ for $(x, y) \in J$ if and only if $y = g(x)$. We know that $|\nabla g| \leq C/\sqrt{d}$, because if $h(x) = f(x, g(x))$, then

$$0 = \partial_i h(x) = \partial_i f(x, g(x)) + \partial_d f(x, g(x)) \partial_i g(x)$$

Hence for $x \in T_1, \dots, T_{d-1}$ with $g(x) \in T_d$,

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

In particular, if $g(x) \in S_d$, then there is $a \in \mathbf{A}$ with

$$|x - a| \leq \sum |x_n - a_n| \leq \frac{C\sqrt{d}}{N^{d-1}}$$

and now we know

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

and so $g(a) \in \mathbf{B}$, and we can set $C_0 = BC\sqrt{d}/A$. \square

Remark. *In the paper, the authors state that the N^{1-d} bound for all but the last variable means we cannot get a Hausdorff dimension one set in the process. In certain cases where we can use properties of a particular function f , we can choose much larger sets and get a much better Hausdorff dimension bound. This is what we attempt to do with our strategies.*

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 J is a length $1/M_{n+1}$ interval, and $J \subset I$ a length $1/M_n$ interval, then we consider two cases:

- If J is subdivided in the non-specialized manner, then J contains N_n/M_n length C_n/N_n^{d-1} intervals, so $\mu(J) = (M_n/N_n)\mu(I)$.
- In the second case, we have at least $(1 - 1/M_n)(N_n/M_n)N_n^{d-2}$ length C_n/N_n^{d-1} intervals which we distribute the mass over, so

$$\mu(J) \leq \frac{M_n}{N_n^{d-1}(1 - 1/M_n)}\mu(I) \leq \frac{2M_n}{N_n^{d-1}}\mu(I)$$

In either case,

$$\mu(J) \leq (M_n/N_n)\mu(I) = \frac{M_n}{(C_n M_{n+1})^{1/d-1}}\mu(I)$$

so by induction, if I is a length $1/M_N$ interval considered in the process, then

$$\mu(I) \leq \prod_{n < N} \frac{M_n}{(C_n M_{n+1})^{\frac{1}{d-1}}} = \left(\prod_{n < N} \frac{M_{n+1}^{1-\frac{1}{d-1}}}{C_n^{\frac{1}{d-1}}} \right) \frac{1}{M_N} = \frac{A_N}{M_N}$$

If $J \subset I$ is any length $1/N_N$ interval considered in the algorithm, then either $\mu(J) = (M_N/N_N)\mu(I) \leq A_N/N_N$, as in the first case, or in the second case, $\mu(J) = (M_N/N_N(1 - 1/M_N))\mu(I)/M_N \leq 2A_N/N_N$. Thus $\mu(J) \leq 2A_N/N_N$. Now if $1/N_N \leq |I| \leq 1/M_N$, then I can be covered by $|I|N_N$ intervals of length $1/N_N$, and so

$$\mu(I) \leq 2|I|N_N \frac{A_N}{N_N} = 2A_N|I| = \frac{2A_{N-1}M_N^{1-\frac{1}{d-1}}}{C_{N-1}^{\frac{1}{d-1}}} |I| \lesssim_\varepsilon |I|M_N^{1-\frac{1}{d-1}-\varepsilon} \leq |I|^{\frac{1}{d-1}-\varepsilon}$$

Provided that we can choose M_N such that $A_N/C_N \lesssim_\varepsilon M_{N+1}^\varepsilon$ for all ε (this is why it is incredibly important that the values in the lemma are independent of N in the proof above). On the other hand, if $1/M_{N+1} \leq |I| \leq 1/N_N$, then I can be covered by a single length $1/N_N$ interval, hence

$$\mu(I) \leq \frac{2A_N}{N_N} = \frac{2A_N}{N_N} = \frac{2A_N}{(C_N M_{N+1})^{\frac{1}{d-1}}} \lesssim_\varepsilon \frac{1}{M_{N+1}^{\frac{1}{d-1}-\varepsilon}} \leq |I|^{\frac{1}{d-1}-\varepsilon}$$

Thus we obtain the theorem if $M_{N+1} = \exp(A_N/C_N)$, for instance.

1.4 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 5. *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 occurring 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 $\left| \sum \alpha_n a_n \right| > (\delta + \varepsilon C^*) \lambda$, then $\left| \sum \alpha_n x_n \right| \geq \delta \lambda$. If this does not occur \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 6. *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 high Hausdorff dimension.

Theorem 5. *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. We can skip the ‘double split’ of T_0 into length $1/N$ intervals, and then $1/N^k$ intervals, by just splitting these intervals immediately into length $\Omega(1/N^k)$ intervals, and discarding intervals containing or adjacent to intervals containing elements of \mathbf{B} . This guarantees that $\Omega(1/M)$ of each length $1/M$ interval is kept, but not that $\Omega(1/N)$ of each length $1/N$ interval is kept. Is this tighter quality of the decomposition really needed in the Hausdorff dimension argument?

Remark. If T is a rank k linear transformation with rational coefficients, then there is some number a such that aT has integer coefficients, and then the equation $y = f(Tx)$ is the same as the equation $y = f_0((aT)(x))$, where $f_0(x) = f(x)/a$. Since f_0 is also Lipschitz continuous, we conclude that we still get the dimension $1/k$ bound if T has rational rather than integral coefficients. More

generally, this trick shows the result applies unperturbed if all coefficients of T are integer multiples of some fixed real number. More generally, by varying the lengths of our length $1/N$ decomposition by a constant amount, we can further generalize this to the case where each column of T are integers multiples of some fixed real number.

Remark. To form \mathbf{A} , we take startpoints lying at equal spaced $1/N$ points. However, by instead taking startpoints at varying points in the length $1/N$ intervals, we might be able to make points cluster more than in the original algorithm. Maybe the 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 should give a set with full Minkowski dimension and Hausdorff dimension m/k avoiding solutions to $y = f(Tx)$.

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 on the Torus To Extend the Result to Irrational Matrices

Or maybe this shows that shifting wont help us if the points aren't rational?

3.4 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 , so $\Gamma = S((\mathbf{Z}/N)^n)$ forms a lattice. Since T has rank k , the image of U is contained within the image of S . We now wish to consider the distribution of the values of U within the quotient group V/Γ , where V is the image of T . Now given $x \in V/\Gamma$, $x, 2x, 3x, \dots$ is equidistributed in V/Γ if and only if x is **totally irrational**, in the sense that $\chi(x) \neq 0$ for all nonzero characters χ . LOOK UP VAN DER CORPUT TRICK.

3.5 Applications of Low Rank Coordinate Changes

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

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

The solution above gives a Hausdorff dimension $1/2$ set avoiding solutions to this equation. 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.

As a particular special case of our result, 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 independent 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 independent over the rationals. More generally, for any smooth function $f : \mathbf{R} \rightarrow \mathbf{R}$, we can find a full Hausdorff dimensional set such that there are no solutions

$$y = f(a_1x_1 + \dots + a_nx_n)$$

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

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 $m \ll n$.

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?*

3.6 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*

$$\nabla f$$

3.7 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.

Chapter 4

Appendix: Calculating Dimensions of Cantor Like Sets

This section considers an inconvenience in the calculation of the Hausdorff dimension of Cantor like sets, and an attempt to resolve it in the most likely of scenarios. That is, we consider the Hausdorff dimension of sets X constructed as the limit of a nested family of sets $X_0 \supset X_1 \supset \dots$, where $X_0 = [0, 1]$, and each X_{n+1} is a union of intervals of some length l_{n+1} obtained by dividing and removing some segments of the intervals X_n . To prove such sets are α dimensional, we rely on Frostman's lemma, which says that $H^\alpha(X) > 0$ if and only if there is a Borel probability measure μ supported on X with $\mu(I) \lesssim |I|^\alpha$ for each interval I . Since the construction is obtained as a limit of intervals, it is often to construct μ by the *mass distribution principle*. That is, we let μ denote the weak limit of the probability masses μ_n , where μ_0 is a uniform distribution over μ_0 , and μ_{n+1} is obtained from μ_n by distributing the mass $\mu_n(I)$ of each length l_n interval I contained in X_n over the portion of I that remains in X_{n+1} . The cumulative distribution functions of the μ_n uniformly converge, hence the μ_n converge weakly to some μ , which satisfy $\mu(I) = \mu_n(I)$ for each interval I as above. Because of this discreteness, it is most easy to establish a bound $\mu(I) \lesssim l_n^\alpha$ when $I \subset X_n$ is a length l_n interval. Since any interval I of length l_n is contained within at least two such intervals (or is contained in other length l_n intervals that μ assigns no mass to), we have the general bound $\mu(I) \lesssim l_n^\alpha$ for all intervals I of length l_n . Hausdorff dimension is a local property of a set¹, so it is natural to expect that we can obtain a general bound $\mu(I) \lesssim_\alpha |I|^\alpha$ given that one has established precisely the same estimate, but restricted to intervals I with $|I| = l_N$. This section concerns itself with ways that we can establish this general bound, and thus prove that $\dim_{\mathbf{H}}(X) \geq \alpha$.

¹If we define $\dim_{\mathbf{H}}(x) = \lim_{r \downarrow 0} \dim_{\mathbf{H}}(B_r(x) \cap X)$ then $\dim_{\mathbf{H}}(X) = \sup_{x \in X} \dim_{\mathbf{H}}(x)$.

Remark. *If this particular probability measure fails to satisfy the required bounds, does it follow that the set itself doesn't have the required Hausdorff dimension? Is there a counterexample to this statement? It seems this probability measure is strongly related to the structure of the Cantor set so it seems that if this measure fails it is doubtful any other measure will fail, at least locally.*

The main way of extending the results at particular scales to general results is by means of a covering argument. This is done in two ways for a given interval I with $l_{N+1} \leq |I| \leq l_N$.

- We can cover I by a single length l_N interval, so

$$\mu(I) \lesssim l_N^\alpha = \left(\frac{l_N}{|I|} \right)^\alpha |I|^\alpha \leq \left(\frac{l_N}{l_{N+1}} \right)^\alpha |I|^\alpha$$

- We can cover I by $|I|/l_{N+1}$ intervals of length l_{N+1} , so

$$\mu(I) \lesssim \frac{|I|}{l_{N+1}} l_{N+1}^\alpha = \left(\frac{|I|}{l_{N+1}} \right)^{1-\alpha} |I|^\alpha \leq \left(\frac{l_N}{l_{N+1}} \right)^{1-\alpha} |I|^\alpha$$

The first bound is better for $\alpha \leq 1/2$, the second for $\alpha \geq 1/2$, but only up to the exponent in the polynomial. By considering both covering techniques simultaneously, and then choosing the best such bound for any particular interval length $|I|$, we can actually obtain the tighter bound

$$\mu(I) \lesssim \left(\frac{l_N}{l_{N+1}} \right)^{(1-\alpha)\alpha} |I|^\alpha \leq \left(\frac{l_N}{l_{N+1}} \right)^{1/4} |I|^\alpha$$

Ultimately, however, we shall find that in general, without additional information about the measure μ , any of the three techniques suffice, because the polynomial power doesn't really help make a required bound more or less difficult to prove. For instance, here we find that using any of the three bounds we have obtained that our argument is only completed if l_N/l_{N+1} is upper bounded independently of N . In most cases this is very difficult to obtain, since it means that we can only restrict l_N to have exponential decay.

We can obtain slightly better results by giving ourselves an epsilon of room. Rewording the two bounds, we have

$$\mu(I) \leq \left(\frac{l_N}{|I|^{1-\varepsilon/\alpha}} \right)^\alpha |I|^{\alpha-\varepsilon} \quad \mu(I) \leq \left(\frac{|I|^{1+\varepsilon/(1-\alpha)}}{l_{N+1}} \right)^{1-\alpha} |I|^{\alpha-\varepsilon}$$

Taking the minimum of these two bounds and optimizing gives

$$\mu(I) \leq \left(\frac{l_N}{l_{N+1}} \right)^{\alpha(1-\alpha)} (l_N^\alpha l_{N+1}^{1-\alpha})^\varepsilon |I|^{\alpha-\varepsilon}$$

Thus we can easily obtain that $\dim_{\mathbf{H}}(X) \geq \alpha - \varepsilon$ for each ε , if $l_N/l_{N+1} \lesssim_\varepsilon l_{N+1}^\varepsilon$. This can be expected in a great many scenarios, since the lengths l_N/l_{N+1}

are obtained at a single scale, whereas the lengths l_{N+1} are obtained from compounding lengths over a great many scales. If we write $l_N = e^{-r_N}$, then $l_N/l_{N+1} = e^{r_{N+1}-r_N}$, so $l_{N+1}^{-\varepsilon} = e^{\varepsilon r_{N+1}}$, so it suffices to show $r_{N+1} - r_N - \varepsilon r_{N+1} \rightarrow -\infty$ for each ε . If r_N is differentiable and accelerating, this means precisely that $r'_N - \varepsilon r_N \rightarrow -\infty$, so in particular, if $l_N = e^{-x^M}$ for any M we can apply this principle.

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 $l_N^{-\varepsilon} = (8^{N-1}N!)^\varepsilon$ for any $\varepsilon > 0$.*

It is, of course, easy to construct examples of l_N for which we cannot use this technique. For instance, if $l_N = e^{-N!}$, or $l_N = e^{-e^N}$, the latter of which can be seen because e^N is it's own derivative. Our second method for interpolating requires extra knowledge of the dissection process, but enables us to choose the l_N decaying much more rapidly, when the filtration of the mass of μ only occurs at very sparse lengths. In particular, in Fraser and Pramanik's paper, the mass is essentially uniformly distributed over the length $1/N$ dissections, and then filtered to very particular length $1/N^k$ sections. Since the length $1/N$ and length $1/N^k$ differ in length only by a polynomial scale, the actual choice of N relative to M isn't really important, so we can let N become as large as we desire relative to the choice of M . In particular, we can choose the M_N to grow as fast as we want, which Fraser and Pramanik take full advantage of, choosing M_N even bigger than $2 \uparrow \uparrow N$, which enables us to overwhelm the arbitrary constants in the construction of the configuration avoiding sets they create.

We consider, for each $\beta = \{l_1 \geq r_1 \geq l_2 \geq r_2 \geq \dots \rightarrow 0\}$, a Cantor set X^β , and a corresponding probability measure μ_β supported on X^β . Our goal is to find a particular index set β such that X^β has Hausdorff dimension exceeding α . The reason for introducing this notation is so that we can choose bounds on the measures μ_β as functions of the l_1 and r_1 . In particular, we will let β_N denote the subset of parameters $\{l_1, r_1, l_2, r_2, \dots, l_N\}$. In this scenario, we assume the existence of a bound $\mu_\beta(J) \lesssim (r_N/l_N)\mu_\beta(I)$ for $J \subset I$ with $|J| = r_N$ and $|I| = l_N$. Essentially, this means that the probability mass of a length l_N interval considered in the algorithm is distributed essentially uniformly over the length r_N intervals it contains. The first advantage of this is to perform half of an interpolation; a bound $\mu_\beta(I) \lesssim |I|^\alpha$ for $|I| = l_N$ automatically implies that for $r_N \leq |I| \leq l_N$, we can cover I by length r_N intervals to obtain

$$\mu(I) \leq (|I|/r_N)(r_N/l_N)l_N^\alpha = \frac{|I|}{l_N^{1-\alpha}} \leq |I|^\alpha$$

Secondly, it means that the probability mass of a length l_N interval considered in the algorithm is distributed essentially uniformly over the length r_N intervals it contains, which, since we are able to adjust parameters, will enable us to take r_N arbitrarily large relative to the l_N , causing any constants that occur in the algorithm in previous steps to become immediately negligible. Given the existence of an inequality $\mu_\beta(I) \lesssim A_N(\beta_N)l_N$ for $|I| = l_N$ automatically

gives rise to a bound $\mu_\beta(I) \lesssim A_N(\beta|_N)r_N$ for $|I| = r_N$. If $r_N \leq B_N(\beta|_N)l_{N+1}^\alpha$, then for any $l_{N+1} \leq |I| \leq r_N$, we can cover I by a single length r_N interval to conclude that

$$\mu_\beta(I) \lesssim A_N(\beta|_N)r_N \leq A_N(\beta|_N)B_N(\beta|_N)l_{N+1}^{(1-\varepsilon)\alpha}r_N^\varepsilon \leq A_N(\beta|_N)B_N(\beta|_N)|I|^{(1-\varepsilon)\alpha}r_N^\varepsilon$$

Since all the values in this inequality are fixed, we can vary r_N to become incredibly small with respect to the previous constants to conclude that for certain rapidly decay β , $\mu_\beta(I) \lesssim_\varepsilon |I|^{\alpha-\varepsilon}$. This guarantees the existence of a sequence β with X^β α dimensional.

Example. In Fraser and Pramanik's paper, they choose $l_{N+1} = C_N r_N^{d-1}$, for some values C_N depending only on the values M_1, \dots, M_N and N_1, \dots, N_{N-1} . Thus $r_N/l_{N+1} = r_N^{2-d}/C_N$

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*