Squarefree Sets

Jacob Denson

May 2, 2018

Contents

Chapter 1

Background

1.1 Rusza: Difference Sets Without Squares

In this section, we describe the work of Ruzsa on the discrete squarefree difference problem, which provides inspiration for our speculated results for the squarefree subset problem in the continuous setting. 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 $1, 2, 3, 4, \ldots, N$. 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 $x_* + m^2 \leqslant N$. This implies $m \leqslant \sqrt{N - x_*} \leqslant \sqrt{N - 1}$, hence we cross out at most $\sqrt{N-1} + 1$ integers whenever with add a new element x_* to X. When the algorithm terminates, all integers must be crossed out, and if the algorithm runs n iterations, a union bound gives that we cross out at most $n[\sqrt{N-1} + 1]$ integers, hence $n[\sqrt{N-1} + 1] \geqslant N$. It follows that the set X we end up with satisfies

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

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

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

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 squares[1], this is incredibly pessimistic. Ruzsa's paper shows we should be more optimistic, taking advantage of the structure of perfect suqares 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*

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

where

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

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

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

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

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

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

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

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

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

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

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

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

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 dth powers and*

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

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 dth 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*

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

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

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

Hence

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

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 dth power modulo p with $|R| \geqslant d$.*

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

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

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

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

as $p \to \infty$, which is greater than $k-1$ for large enough p (this shows the theorem is essentially trivial for large 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 $b_1, \ldots, b_k \in \mathbf{Z}_p$ such that $b_i - b_j \notin Q$ for any i, j and

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

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

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

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

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

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 $1 + j(q-1) \geqslant p$, and thus $j \geqslant k$. □

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

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

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

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

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

or even whether it exists at all. The consequence of this would essentially solve the squarefree integers problem, since it 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 $X_0 \supset X_1 \supset \ldots$, 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, $X_0 = [0,1]$ is the unit interval. To aid in this process, Keletti 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 separated by a value 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 from the startpoint.

In order for this process to be well defined, we must have $l_0 = N_0 = 1$, and

$$\Delta_{n+1} + (N_{n+1} - 2)\varepsilon_{n+1} + N_{n+1} l_{n+1} \leqslant l_n$$

Of course, we must also have $l_n \to 0$, for otherwise we have no fractal structure in our sets. We then obtain a set X which has Hausdorff dimension one. For completeness, and to introduce readers to the methods of calculating Hausdorff dimension, we will prove this explicitly.

Lemma 2. *The set $X = \lim X_n$ has Hausdorff dimension one.*

Proof. Recall Frostman's lemma, which says that $\dim_{\mathbf{H}}(X) \geqslant s$ if and only if there is a finite positive Borel measure μ supported on X with $\mu(B_r(x)) \lesssim r^s$, for a universal constant depending only on μ. We can construct a probability measure on our set X with the required properties fairly simply, using what is called the *mass distribution principle*. For each n, we construct a probability measure μ_n supported on X_n by letting μ_0 by the uniform measure over $X_0 = [0,1]$, and then, recursively, let μ_{n+1} be defined from μ_n by uniformly distributing the mass of each interval J in X_n uniformly over the intervals in X_{n+1} formed from the parts of J. The probability measures μ_n weakly converge to a probability measure μ supported on X, because the distribution functions of the μ_n converge uniformly to μ. If I is an interval in X_n, and J is an interval in X_{n-1} with $J \subset I$, then we know that

$$\mu(J) = \mu_n(J) = \frac{\mu_{n-1}(I)}{N_n} = \frac{\mu(I)}{N_n}$$

and so by induction, we find

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

If J is any interval of length l_n, then $\mu(J)$ can intersect at most two intervals of length l_n in X_n, and so we obtain the general bound

$$\mu(J) \leqslant 2 \prod_{m=1}^{n} \frac{1}{N_m}$$

Now fix $\varepsilon > 0$, and let I be any interval with $l_{n+1} \leqslant |I| \leqslant l_n$. Then I can be covered by as few as $|I| l_{n+1}^{-1}$ intervals of length l_{n+1}, and so

$$\mu(I) \leqslant 2|I| l_{n+1}^{-1} \prod_{m=1}^{n} \frac{1}{N_m} = \left(2 l_n^{\varepsilon} l_{n+1}^{-1} \prod_{m=1}^{n} \right) |I|^{1-\varepsilon}$$

We see that it suffices to pick the constants l_n and N_n so there is a constant C_ε depending only on ε with

$$2 \frac{l_n^{\varepsilon}}{l_{n+1}} \prod_{m=1}^{n} \frac{1}{N_m} \leqslant C_\varepsilon$$

which effectively says that l_{n+1} is proportional to the total number of intervals at step X_n, if we weaken the inequality by the l_n^{ε} term. This is satisfied by the choice of constants we give at the end of the section (skip to the end and come back if you wish to check). Frostman's lemma then implies $\dim_{\mathbf{H}}(X) \geqslant 1 - \varepsilon$, and we can then let $\varepsilon \to 0$ to conclude $\dim_{\mathbf{H}}(X) = 1$. □

We claim that, with another 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 3. *A set X avoids translates if and only if there do not exists values $x_1 < x_2 \leqslant x_3 < x_4$ 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 $a \leqslant b - t$, then the equation

$$a - (a - t) = t = b - (b - t)$$

satisfies the constraints, since then $a - t < a \leqslant b - t < b$ are all elements of X. We also have

$$(b - t) - (a - t) = b - a$$

which satisfies the constraints if $a - t < b - t \leqslant a < b$. This covers all possible cases. Conversely, if there are $x_1 < x_2 \leqslant x_3 < x_4$ 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 $x_4 = x_3 + (x_4 - x_3)$. □

To determine the further properties of the constants l_n, N_n, ε_n, and Δ_n which prevent translation avoidance, we suppose there are $x_1 < x_2 \leqslant x_3 < x_4 \in X$ with $x_2 - x_1 = x_4 - x_3$. Since $l_n \to 0$, 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°, x_2°, x_3°, and x_4° to the length l_N intervals containing x_1, x_2, x_3, and x_4. Thus $0 \leqslant x_n - x_n^\circ \leqslant l_N$. If we choose our constants such that the startpoints $x_1^\circ, \ldots, x_4^\circ$ all lie at integer multiples of $l_N + \varepsilon_N$, and $4l_N \leqslant l_N + \varepsilon_N$, i.e. $3l_N \leqslant \varepsilon_N$, 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 M_N, and x_1° to lie at an odd multiple of M_N (possible by the shifting in the procedure), then it is impossible for this equation to hold, so by contradiction, X must be translation invariant.

Remark. *Can we adopt Ruzsa's squarefree discrete strategy combined with Keleti's approach to find a high dimensional continuous squarefree set? NOTE: We tried this but could only get 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.*

To summarize, in order for X to be translation invariant, we require $3l_n \leqslant \varepsilon_n$, for the nonshifted startpoints of X_n to lie at even multiples of M_n, and for the shifted startpoints to lie at odd multiples. This will be satisfied if

$$M_n \mid M_{n-1}, \varepsilon_n, \Delta_n \quad 2M_n \mid M_{n-1}, \varepsilon_n \quad 2M_n \nmid \Delta_n$$

In Keleti's paper, he chooses

$$N_n = n \quad l_n = 1/6^{n-1}n! \quad M_n = 3l_n \quad \Delta_n = 3l_n \quad \varepsilon_n = 5l_n$$

As an exercise, go back and check these constants give the required bounds for the Hausdorff dimension. We have avoided giving these constants until the end to emphasize *why* Keleti needed to choose these particular constants. Once they are chosen, we end up with a Hausdorff dimension one set. In Keleti's paper, he also remarks that by replacing the 6 in l_n with a slowly increasing set of even numbers, one can obtain a Hausdorff dimension one set which is linearly independant over the rational numbers.

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 $f : \mathbf{R}^d \to \mathbf{R}$ 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 $A \leqslant |\partial_d f(x)|$ and $|\nabla f(x)| \leqslant B$ for $x \in T_1 \times \cdots \times T_d$. There there exists a rational constant c and arbitrarily large integers $N \in M\mathbf{Z}$ for which there exist subsets $S_n \subset T_n$ such that*

(i) $f(x) \neq 0$ for $x \in S_1 \times \cdots \times S_d$.

(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 (but this portion need not be a complete interval, like in the last property).

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 $|\mathbf{A}| \leqslant N^{d-1}$, 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 $|\mathbf{B}| \leqslant MN^{d-1}$. Now we filter out a subcollection of intervals I if $|\mathbf{B} \cap I| \leqslant M^3 N^{d-2}$. Then we throw away at most $MN^{d-1}/M^3N^{d-2} = N/M^2$ 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 into intervals of length $C_0 c/N^{d-1}$. If we use the constant C_0 obtained from the lemma below, and we now discard any interval that intersects $\mathbf{B}$, or is adjacent to an interval intersecting $\mathbf{B}$. Since $\mathbf{B} \cap I$ has at most M^3N^{d-2} points, and I is decomposed into N^{d-2}/C_0c intervals, what remains is $N^{d-2}/C_0c - 3M^3N^{d-2} = (1/C_0c - 3M^3)N^{d-2}$ intervals, which in total has length

$$(C_0c/N^{d-1})(1/C_0c - 3M^3)N^{d-2} = (1 - 3M^3C_0c)N$$

Choosing c such that $3M^3C_0c < 1 - c$ completes the proof. □

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

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 $g : J' \to \mathbf{R}$ such that $f(x, y) = 0$ for $(x, y) \in J$ if and only if $y = g(x)$. We know that $|\nabla g| \leqslant C/\sqrt{d}$, 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 $x \in T_1, \ldots, T_{d-1}$ with $g(x) \in T_d$,

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

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

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

and now we know

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

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

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 $T_1, \ldots, T_d$ 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 points.

What remains is to bound the Hausdorff dimension of X by constructing a probability measure supported on X with suitable decay. Since every interval

is covered by a fixed number of intervals in some dyadic decomposition, we can focus on the intervals in our decomposion. 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.

Lemma 5. *Let I be an interval of length $1/M$ in a certain iteration of the construction, and J an interval of length $1/N$ which contains an interval in the next iteration. Then $\mu(I)/|I| \leqslant \mu(J)/|J| \leqslant 2\mu(I)/|I|$*

Proof. We subdivide I into $N|I|$ intervals of length $1/N$, and we know at least a fraction $1-1/M$ contain a union of smaller intervals with measure at least c/N contained in the next iteration. Enumerate these unions as $K_1, \dots K_L$, where $L \geqslant N|I|(1-1/M)$. Then

$$\mu(K_n) = \frac{|K_n|}{\sum_{n=1}^{L}|K_n|}\mu(I) \leqslant \frac{|K_n|}{Lc/N}\mu(I) \leqslant \frac{|K_n|}{c(1-1/M)}\frac{\mu(I)}{|I|}$$

and if $K_n \subset J$, then $\mu(J) = \mu(K_n)$ and $|K_n| \geqslant c|J|$, from which we obtain that

$$\frac{\mu(J)}{|J|} \leqslant \frac{1}{(1-1/M)}\frac{\mu(I)}{|I|} \leqslant 2\frac{\mu(I)}{|I|}$$

where the last point followed because we know $M \leqslant 1/2$, since we subdivide our intervals right at the beginning of the algorithm at least in two.

The interval I decomposes into $|I|/|J|$ intervals of length $|J|$, and we know at least $1-|I|$ of them contain a portion $c|J|$ of the length $|J|$. Thus

$$\mu(I) \geqslant (1-|I|)\frac{|I|}{|J|}\mu(J)$$

and so, since $|I| \leqslant 1/2$ after the first division

$$\frac{\mu(J)}{|J|} \leqslant \frac{1}{1-|I|}\frac{\mu(I)}{|I|} \leqslant \frac{2\mu(I)}{|I|}$$

The obvious bound of $|I|/|J|$ intervals implies

$$\mu(I) \leqslant (|I|/|J|)\mu(J)$$

which implies the lower bound

$$\frac{\mu(I)}{|I|} \leqslant \frac{\mu(J)}{|J|}$$

so the ratios are comparable to one another. □

Applying this result iteratively, if we let l_n denote the lengths of the intervals on the n'th iteration, and we let N_n denote the value we use to divide division. Then if we have a nested family of intervals $J_1, \ldots, J_N$ formed by the $1/N_n$ partitions of $I_1, \ldots, I_N$, then

$$\frac{\mu(J_N)}{|J_N|} \leqslant 2\frac{\mu(I_{N-1})}{|I_{N-1}|} = \frac{2|J_{N-1}|}{|I_{N-1}|}\frac{\mu(J_{N-1})}{|J_{N-1}|} = \frac{2}{l_{N-1}N_{N-1}}\frac{\mu(J_{N-1})}{|J_{N-1}|}$$

Applying this result iteratively, we find that if we have a family of nested family of intervals $I_1, \ldots, I_N$, then we can apply the lemma above iteratively to find

$$\frac{\mu(J_N)}{|J_N|} \leqslant \frac{2}{l_{N-1}N_{N-1}}\frac{\mu(J_{n-1})}{|J_{n-1}|} \leqslant \cdots \leqslant 2^N \prod_{n=1}^{N-1} \frac{1}{l_n N_n}$$

Now

Chapter 2

Ideas for New Work

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 anology to the squarefree subset problem on the interval $[0,1]$, because there is no canonical subset of $[0,1]$ which can be identified as 'perfect squares', unlike in $\mathbf{Z}$. If we only restrict ourselves to perfect squares of a countable set, like perfect squares of rational numbers, a result of Keleti gives us a set of full Hausdorff dimension avoiding this set. Thus, instead, we say a set $X \subset [0,1]$ is (continuously) **squarefree** if there are no nontrivial solutions to the equation $x - y = (u - v)^2$, in the sense that there are no $x, y, u, v \in X$ satisfying the equation for $x \neq y, u \neq v$. In this section we consider some blue sky ideas that might give us what we need.

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

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

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

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

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

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

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

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

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

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

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 $\{0, \ldots, m-1\}$ 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

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

(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

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

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

So it suffices to choose ε such that

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

This is equivalent to picking

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

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 $1/m^n$),

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

  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 $\alpha^2(x-z)^2 \geqslant \alpha^2$, 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 $\varepsilon + 2\alpha(m-1)\varepsilon \leqslant \alpha^2$, which is equivalent to saying

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

- 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 $|\alpha(x-y)| \geqslant \alpha$, 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 $\varepsilon^2 + \varepsilon \leqslant \alpha$, or

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

Provided ε is chosen as above, all solutions in E must occur in a common interval. Thus, if we now replace the intervals with a recursive fattened scaling of X, all solutions must occur in smaller and smaller intervals. If we choose the size of these scalings to go to zero, these solutions are required to lie in a common interval of length zero, and thus the three values must be equal to one

another. Rigorously, we set $\varepsilon \approx 1/m^2$, and $\alpha \approx 1/m$, we can define a recursive construction by setting

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Thus

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

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

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

If $N^\varepsilon \geqslant M^2/\delta_0$, then we can force $d(\alpha, \mathbf{Z}/NM^2) \leqslant \delta_0/NM^2$ 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 $r \leqslant \sqrt{\delta_0 N/2}$ and $2qr = Q\sqrt{N} + R$ with $R \leqslant (\delta_0/2)\sqrt{N}$. 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 $r \leqslant \sqrt{\delta_0 N/2}$, $2qr = Q\sqrt{N} + R$, and $R \leqslant (\delta_0/2)\sqrt{N}$. 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 $r \leqslant \varepsilon\sqrt{N}$, and $q \leqslant \sqrt{N}$. The following result reduces our problem to understanding the distribution of the smooth integers.

Lemma 6. *Fix constants A, B, and let $n \leqslant AN$ be an integer. If all prime factors of n are $\leqslant BN^{1-\delta}$, then n can be decomposed as qr with $r \leqslant \varepsilon\sqrt{N}$ and $q \leqslant \sqrt{N}$.*

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

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

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

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

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 $AB/\varepsilon \leqslant 1$. □

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 $n \leqslant N$ 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 $\sqrt{N}$,

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

Then

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

So

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

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,

2.7 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 $T : \mathbf{R}^n \to \mathbf{R}^m$ is a rank k linear transformation with integer

coefficients with respect to standard coordinates. Fix a constant A bounding the operator norm of T, in the sense that $\|Tx\|_\infty \leqslant A\|x\|_\infty$ for all $x \in \mathbf{R}^n$. If we fix sets $T_0, T_1, \ldots, T_n, \subset [0,1]$, which are unions of intervals of length $1/M$, with startpoints lying on integer multiples of $1/M$. Split each interval of T_a into length $1/N$ intervals, and then set

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

Since the startpoints of the intervals are integer multiples of $1/N$, $T(\mathbf{A})$ is contained with a rank k sublattice of $(\mathbf{Z}/N)^m$. The operator norm also guarantees $T(\mathbf{A})$ is contained within $[-A, A]^m$. Because of the lattice structure of the image, $\|x-y\| \geqslant 1/N$ for each distinct pair $x, y \in T(\mathbf{A})$. Furthermore, since $\mathrm{Im}(T)$ is k dimensional, we can cover $\mathrm{Im}(T) \cap [-A, A]^m$ by $(4AN)^k = O_{A,k}(N^k)$ cubes of size $1/2N$, and so there are at most $O_{A,k}(N^k)$ points in $T(\mathbf{A})$. 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

$$|\mathbf{B}| = |f(T(\mathbf{A}))| \leqslant |T(\mathbf{A})| = (4A)^k N^k = O_{A,k}(N^k)$$

We can now split each length $1/M$ interval in T_0 into $O_{A,k}(N^k)$ intervals of length $\Omega_{A,k}(1/MN^k) = \Omega_{A,k,M}(1/N^k)$, and because of the bounds on $|\mathbf{B}|$, at least one of these intervals must avoid elements of $\mathbf{B}$. In fact, by making the length of these intervals a constant amount smaller, we can assume that we can pick an interval I with $\mathrm{dist}(I, \mathbf{B}) \geqslant \Omega_{A,k,M}(1/N^k)$. The union of these intervals forms S_0. If f is C^1, and $\|\nabla f\|_\infty \leqslant B$, then

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

and so we may choose $S_n \subset T_n$ by thickening each startpoint $x \in T_n$ to a length $O_{A,k,M}(1/N^k)$ interval with no solutions $y = f(T(x))$, with $x_n \in S_n$, $y \in S_0$. Following the Hausdorff dimension computations in Pramanik and Fraser's paper, this gives a Hausdorff dimension $1/k$ set.

Remark. *If T is a rank k linear transformation with rational coefficients, then there is some integer λ such that λT has integer coefficients, and then the equation $y = f(Tx)$ is the same as the equation $y = f'((\lambda T)(x))$, where $f'(x) = f(x)/\lambda$. Since f' satisfies the same required regularity conditions that f does in the proof above, we conclude that we still get the dimension $1/k$ bound if T has rational rather than integral coefficients.*

Remark. *In Malabika's argument, she used the following approach:*

"We now fix an integer C depending only on A and k, and M, and split S into length $1/N$ subintervals I such that $|\mathbf{B} \cap I| \leqslant N^{k-1}/C$. In this step, we throw away at most $C(4A)^k N = O_{A,k,M}(N)$ intervals. Since $|T_a| \geqslant 1/M$, T_a contains at least N/M intervals, and so if we choose C small enough that $C(4A)^k \leqslant 1/M$, then there is at least one interval in T_a that we keep. If we

now split the remaining intervals into $2N^{k-1}/C$ intervals, which will have length $C/2N^k$,... "

But I don't think this is necessary in the current proof. Is this method important when we look at higher dimensions, or is this a universal simplification?

Remark. *Maybe I should apply my results to augment Mathé's result rather than Pramanik/Fraser, since their results work better in high dimensional space, where we can better employ the low degree polynomials.*

Using the same approach, if we now try to avoid solutions to $y = f(Tx)$, where y is now a vector in $\mathbf{R}^m$, and T has rank k. Consider unions of $1/M$ cubes $T_0, \ldots, T_n$. If we fix startpoints of each x_k forming a full rank lattice spaced by $1/N$, and consider the space $\mathbf{A}$ of products, then there are $O_{A,k}(N^k)$ points in $T(\mathbf{A})$. If we let $\mathbf{B}$ denote the bad points as before, then there can be at most $O_{A,k}(N^k)$ of those. If we now split each $1/M$ cube in T_0 into $O_{A,k}(N^k)$ cubes with sidelengths $\Omega_{A,k,M}(1/N^{k/m})$, then at least one of these cubes must avoid elements of $\mathbf{B}$, and by taking a slightly smaller interval (by a constant again), we may assume that the interval satisfies $\text{dist}(I, \mathbf{B}) \geqslant \Omega_{A,k}(1/MN^{k/m})$. If f is C^1, and $\|Df\|_\infty \leqslant B$, then

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

so we can thicken the startpoints to cubes with sidelengths $O_{A,k,M}(1/N^{k/m}\sqrt{m}) = O_{A,k,M,m}(1/N^{k/m})$. This should give a dimension m/k set.

2.8 Applications of Low Rank Coordinate Changes

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

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*