Large Salem Sets Avoiding Nonlinear Patterns

UNIVERSITY OF WISCONSIN–MADISON

Jacob Denson
Advisor: Andreas Seeger

This material is based upon work supported by the National Science Foundation under Grant DMS-1764295.

Research Problem: Can Large Sets Avoid Patterns?

More specifically: If a set $X \subset \mathbb{R}^d$ has large *fractal dimension*, does it contain patterns? The main focus of this project is on the construction of counterexamples: for a given function f with domain $(\mathbb{R}^d)^n$, can we construct large sets X such that there are no distinct points $x_1, \dots, x_n \in X$ with $f(x_1, \dots, x_n) = 0$? We often study functions f which vanish on the diagonal $\Delta = \{(x, \dots, x) : x \in \mathbb{R}^d\}$, which makes it difficult to avoid zeroes if X is 'thick', i.e. has large fractal dimension.

A set avoiding 3-term APs

A subset of the parabola avoiding isosceles triangles

Example choices of f:

- If $f(x_1, x_2, x_3) = (x_1 - x_2) - (x_2 - x_3)$, then sets avoiding zeroes of f do not contain three term arithmetic progressions.
- If $f(x_1, x_2, x_3) = |x_1 - x_2|^2 - |x_2 - x_3|^2$, then sets in $\mathbb{R}^d$ avoiding zeroes of f do not contain the vertices of any isosceles triangle.

Mainly, this project constructs large *Salem sets* avoiding zeroes of *nonlinear* functions. Here are some results taken from (D., 2021):

Theorem 1. *Suppose $f : (\mathbb{R}^d)^n \to \mathbb{R}^i$ is a submersion. Then we can construct a Salem set $X \subset \mathbb{R}^d$ avoiding solutions to f with $\dim(X) = i/(n - 1/2)$.*

Theorem 2. *Let $g : (\mathbb{R}^d)^{n-1} \to \mathbb{R}^d$ be smooth, such that $D_{x^k} g = (\partial g_i / \partial x^k_j)$ is an invertible matrix for each $1 \leq k \leq n-1$. If*

$$f(x^1, \dots, x^n) = x^n - g(x^1, \dots, x^{n-1}),$$

then we can construct a Salem set $X \subset \mathbb{R}^d$ avoiding solutions to f with $\dim(X) = d/(n - 3/4)$, larger than that guaranteed by Theorem 1.

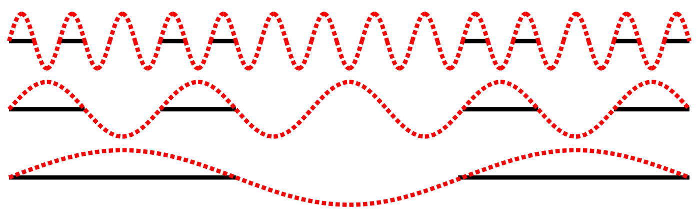
For instance, we can use these results to construct, for any smooth $\gamma : [0,1] \to \mathbb{R}^d$, a Salem set $X \subset [0,1]$ with dimension $4/9$ such that $\gamma(X)$ avoids vertices of isosceles triangles.

Salem Sets: Structure vs. Randomness

There are several fractal dimensions, and they differ subtly in the properties they measure:

- The *Hausdorff dimension* $\dim_{\mathbb{H}}(X)$ of a set $X \subset \mathbb{R}^d$ measures the ability to distribute mass onto X in a way that does not concentrate too strongly around individual points.
- The *Fourier dimension* $\dim_{\mathbb{F}}(X)$ of a set $X \subset \mathbb{R}^d$ measures the ability to distribute mass avoiding concentration 'at a particular frequency', as measured quantitatively through the Fourier transform.

One always has $\dim_{\mathbb{F}}(X) \leq \dim_{\mathbb{H}}(X)$ for any set $X \subset \mathbb{R}^d$, but the reverse is often *not true* if the set is clustered 'near particular frequencies', like if X is a flat surface (clustered near frequencies travelling tangent to the hyperplane), or a Cantor set (clustered near frequencies of the form 3^n), both sets with Fourier dimension zero. On the other hand, a curved hypersurface in $\mathbb{R}^d$ has Fourier dimension equal to $d - 1$, also it's Hausdorff dimension.



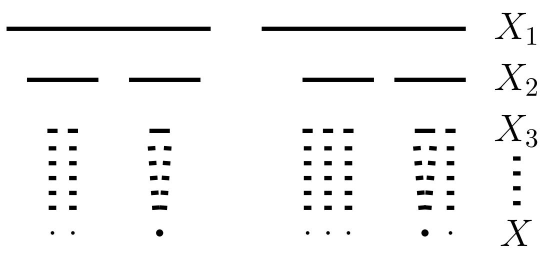
The Cantor Set correlates near frequencies of the form 3^n

We say a set X is *Salem* if $\dim_{\mathbb{F}}(X) = \dim_{\mathbb{H}}(X)$. *Random sets* are often almost surely Salem, since pure randomness prevents clustering at frequencies with high probability. But it is *suprisingly difficult* to control the Fourier dimension of sets when one introduces *structure* to sets, which may introduce subtle clustering near particular frequencies. In particular, the relation between *nonlinear structure* and Fourier dimension is especially difficult to understand. For instance, even determining the Fourier dimension of the set $\{x + x^2 : x \in C\}$, where C is the Cantor set, remains an open problem.

There are many constructions of sets with large *Hausdorff dimension* avoiding zeroes of nonlinear functions f (e.g. Máthé, 2017 or Fraser and Pramanik, 2018), but most constructions of large Salem sets avoiding functions f focus on linear functions f (e.g. Shmerkin, 2015 or Liang and Pramanik, 2020). Here we describe techniques to deal with the introduction of nonlinear structure to a random set via *probabilistic concentration inequalities*, and *oscillatory integrals*.

Constructing Salem Sets

We construct sets avoiding zeroes via a Cantor-type construction, i.e. iteratively defining sets $\{X_k\}$ by dissecting cubes (intervals if $d = 1$) at each stage into smaller cubes, and keeping a union of smaller cubes chosen carefully so they have *good Fourier analytic properties*, and *avoid a discretized version of the pattern*.



If, at each stage of the construction, we choose a large $N > 0$, subdivide each cube into smaller sidelength $1/N^{1/s}$ cubes, and take N of these cubes from each of the original intervals, then iteration for a fixed s should yield a set with Hausdorff dimension s. Since the subdivided set 'lives at a scale $1/N^{1/s}$', the uncertainty principle tells us to care about frequencies $|\xi| \lesssim N^{1/s}$. And indeed, obtaining a Salem set reduces to verifying the following exponential sum square root cancellation bound can be obtained:

Lemma. *For arbitrarily large $N > 0$, there exists an N element subset S of $[0,1]^d$ such that for any $\xi \in \mathbb{Z}^d$ with $|\xi| \lesssim N^{1/s}$*

$$\left| \frac{1}{N} \sum_{x \in S} e^{2\pi i \xi \cdot x} \right| \lessapprox N^{-1/2},$$

and for distinct $y_1, \dots, y_n \in S$, $|f(y_1, \dots, y_n)| \gtrsim N^{-1/s}$ (S contains no 'near zeroes').

Let us illustrate how the problem becomes harder as we *increase* s, i.e. we try and construct larger Salem sets. To do this, pick $10N$ points $\{x_1, \dots, x_{10N}\}$ uniformly at random from $[0,1]^d$. There are roughly $O(N^n)$ tuples $(y_1, \dots, y_n)$, where each y_i is taken from the points x_i. Each tuple has probability $O(N^{-i/s})$ of forming a near zero of f, since the zero set of f is a $dn - i$ dimensional hypersurface in $(\mathbb{R}^n)^d$. Thus we expect there to be roughly $O(N^{n-i/s})$ tuples formed from the points $\{x_i\}$ which give near zeroes.

- If $s \leq i/n$, we expect no tuples will give near zeroes, so setting $S = \{x_1, \dots, x_N\}$ will satisfy the contraints of the Lemma with high probability. Easy!
- If $s > i/n$, we expect there to be tuples giving near zeroes. So we let S be the set of points from the set $\{x_i\}$ which remain after *pruning*, i.e. after removing any point x_i which equals y_n for some tuple $(y_1, \dots, y_n)$ forming a near zero of f. If $s \leq i/(n-1)$, then we will prune at most $O(N^{n-i/s}) \ll 10N$ points, which means we can still guarantee S contains N points. For $s > i/(n-1)$, we cannot guarantee S contains any points, so $i/(n-1)$ is the limiting dimension we can expect.

For $s \leq i/n$, the selection process above is completely random, and so the square root cancellation property is almost automatic. But the pruning we must perform for $s > i/n$ is *structured*, i.e. it removes points clustered near zeroes of f, which may cause subtle problems with the Fourier dimension / square root cancellation.

Dealing With Pruning

Random collections of points satisfy square root cancellation – it is the pruning which makes the required Lemma difficult to prove. In other words, it suffices to prove the following 'pruning inequality'

$$\left| \frac{1}{N} \sum_{x_k \text{ pruned}} e^{2\pi i \xi \cdot x_k} \right| \lessapprox N^{-1/2}.$$

For $s \leq i/(n - 1/2)$, we can guarantee $O(N^{n-i/s}) = O(N^{1/2})$ points have been pruned, so the pruning inequality follows trivially from the triangle inequality. For $s > i/(n - 1/2)$ we work harder. Let us now make the assumption that $f(x) = x^n - g(x^1, \dots, x^{n-1})$ as in Theorem 2, and that $i = n$.

The left hand side of the pruning bound can be viewed as a very nonlinear function $F_\xi(x) = F_\xi(x_1, \dots, x_{10N})$ of the initial uniformly random points chosen. The theory of *probabilistic concentration inequalities* gives various tools guaranteeing that $|F_\xi(x) - \mathbb{E}[F_\xi(x)]|$ is bounded with high probability provided the maximum 'influence' of each variable x_i on F is not too large. Since we remove the points corresponding to the last coordinate of $(y_1, \dots, y_n)$, these points have 'a little too much influence' relative to the other points, but this can be dealt with because these variables are 'linear' in f (because of the extra structure assumed in Theorem 2), so we obtain $|F_\xi(x) - \mathbb{E}[F_\xi(x)]| \lessapprox N^{-1/2}$ with high probability for $s \leq d/(n-1)$.

Finally, we can use some inclusion-exclusion bounds to reduce the study of $\mathbb{E}[F_\xi(x)]$ to an oscillatory integral and apply non-stationary phase. But the inclusion-exclusion bounds obtained only work for $s \leq d/(n - 3/4)$ – one must understand the 'exclusion' in more detail past this range, which is why Theorem 2 only obtains a Salem set of dimension $d/(n - 3/4)$ rather than dimension $d/(n-1)$.

What's Next

Here are some problems to improve the results described in this poster:

- Can one improve the inclusion-exclusion bounds in the analysis of pruned sets to improve the dimension $d/(n - 3/4)$ in Theorem 2 to $d/(n-1)$, the best possible bound we can expect purely via pruning random points.
- Is there a nontrivial concentration argument for general f as in Theorem 1?
- Can we consider 'fractal domain' avoidance problems: Given a Salem set S and a nice function $f : S^n \to \mathbb{R}$, it is possible to construct a large Salem subset $X \subset S$ avoiding zeroes of f? In the simplest nontrivial example, S could be a curved hypersurface.
- Can we use modern 'square root cancellation methods', e.g. decoupling, to construct more structured Salem sets?