SOME OPEN PROBLEMS

STEFAN STEINERBERGER

Abstract. This list contains some open problems that I came across and that are not well known (no Riemann hypothesis...). Some are extremely difficult, others might be doable and some might be easy (one of the problems is that I do not know which is which). The presentation is pretty casual, the relevant papers/references usually have more details – if you have any questions, comments, remarks or additional references, please email me!

Contents

Last update: June 23, 2025. Partially supported by the NSF (DMS-1763179, DMS-2123224) and the Alfred P. Sloan Foundation.

Part 1. Combinatorics

1. The Motzkin-Schmidt problem

Let $x_1, \dots, x_n$ be n points in $[0,1]^2$. The goal is to find a line ℓ such that the $\varepsilon-$neighborhood of ℓ contains at least 3 points. How small can one choose ε (depending on n) to ensure that this is always possible? A simple pigeonholing argument shows that $\varepsilon \le 3/n$ always works. As far as I know, this trivial bound has never been improved. The question, due to T. Motzkin and W. M. Schmidt (independently), is whether $\varepsilon = o(1/n)$ is possible. J. Beck has established this result assuming that the points are somewhat evenly distributed in the square (J. Beck, Almost collinear triples among N points on the plane). I would also be interested in what happens when the points lie on $\mathbb{S}^2$ and one wants to capture at least 3 using neighborhoods of great circles.

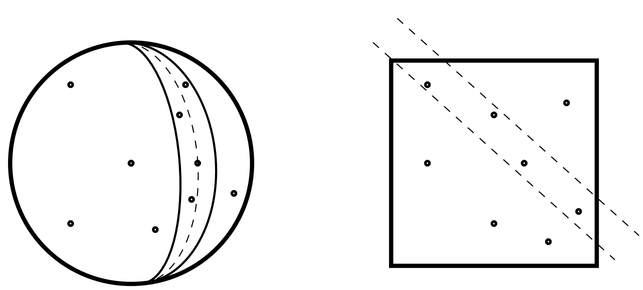


Figure 1. Finding strips that contains three points.

There is a harder version of the question (in the sense that if one could show that no such constant exists, then this implies Motzkin-Schmidt).

> Jean was very generous with problems. [...] He told me there the following problem, which to the best of my knowledge, is still open. Is it possible to find n points in the unit square such that the

> $1/n-$neighborhood of any line contains no more than C of them for some absolute constant C? The motivation for this problem comes from a possible construction of spherical harmonics as a combination of Gaussian beams, which would have $L^\infty-$norm bounded by a constant independently of the degree. (Varjú, Remembering Jean Bourgain (1954-2018), AMS Notices June 2021, p. 957)

Using suitable rectangles centered at the points of dimensions $\delta_n \times 2$ and then rotating them and controlling the L^2-norm in the usual manner from above and below one sees that there exists a strip of width

$$\delta_n \leq n \left(\sum_{\substack{i,j=1 \\ i \neq j}}^{n} \frac{1}{\|x_i - x_j\|} \right)^{-1}$$

containing at least 3 points. We note that this bound is always $\leq c/n$ but not necessarily better than that if the points are nicely distributed in the square.

2. Approximate Hadamard matrices

A Hadamard matrix is a matrix $A \in \{-1,1\}^{n \times n}$ such that for all $x \in \mathbb{R}^n$

$$\|Ax\| = \sqrt{n} \cdot \|x\|.$$

This is maybe not the usual definition but it is one that nicely generalizes: Dong-Rudelson (Approximately Hadamard matrices..., IMRN 2024) prove that there exists a universal constant C such that for every n there exists $A \in \{-1,1\}^{n \times n}$ such that for all $x \in \mathbb{R}^n$

$$c_1\|x\| \leq \|Ax\| \leq c_2 \cdot \|x\|$$

and $c_2/c_1 \leq C$. This means that A is not a (scaled) isometry but it is pretty close. I proved ('A note on approximate Hadamard matrices', Designs, Codes and Cryptography) that one can even choose A to be circulant. So in a certain sense, there should be *many* approximate Hadamard matrices.

(1) Can one get closer to an isometry as n increases? For example, is it true that for every n there exists $A \in \{-1,1\}^{n \times n}$ such that for all $x \in \mathbb{R}^n$

$$c_1\|x\| \leq \|Ax\| \leq c_2 \cdot \|x\|$$

and

$$\frac{c_2}{c_1} \leq 1 + \frac{1}{\sqrt{n}}?$$

(2) There should be nice, simple and completely explicit and completely elementary constructions of approximate Hadamard matrices. Dong-Rudelson use Vinogradov's Theorem about sums of three primes, my argument uses the existence of Littlewood polynomials, neither of those is really simple.

3. Great Circles on $\mathbb{S}^2$

Let $C_1, \dots, C_n$ denote the $1/n-$neighborhood of n great circles on $\mathbb{S}^2$. Here's a natural question: how much do they have to overlap? I proved (Discrete &

Computational Geometry, 2018) that

$$\sum_{\substack{i,j=1\\ i\neq j}}^{n} |C_i \cap C_j|^s \gtrsim_s \begin{cases} n^{2-2s} & \text{if } 0 \leq s < 2 \\ n^{-2}\log n & \text{if } s = 2 \\ n^{1-3s/2} & \text{if } s > 2. \end{cases}$$

and these bounds are sharp.

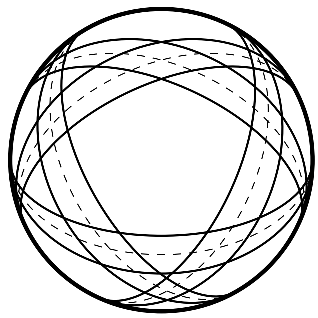


FIGURE 2. Great circles, their $1/n-$neighborhoods and intersection pattern.

The case $s = 1$ is interesting: it is essentially equivalent to the L^2-norm of the sum of characteristic functions: using χ_{C_i} to denote the characteristic function of C_i on $\mathbb{S}^2$, we see that

$$1 + \sum_{\substack{i,j=1\\ i\neq j}}^{n} |C_i \cap C_j| = \int_{\mathbb{S}^2} \sum_{i=1}^{n} \chi_{C_i}^2 + \sum_{\substack{i,j=1\\ i\neq j}}^{n} \chi_{C_i}\chi_{C_j}\, dx = \int_{\mathbb{S}^2} \left(\sum_{i=1}^{n} \chi_{C_i} \right)^2 dx.$$

This means there are arrangements of great circles where

$$\left\| \sum_{i=1}^{n} \chi_{C_i} \right\|_{L^1(\mathbb{S}^2)} \sim 1 \sim \left\| \sum_{i=1}^{n} \chi_{C_i} \right\|_{L^2(\mathbb{S}^2)}.$$

It would be very interesting to understand the behavior of the L^p-norm for some $p > 2$ and this seems to be a very difficult problem.

Open Question. What is the best lower bound on

$$\left\| \sum_{i=1}^{n} \chi_{C_i} \right\|_{L^p(\mathbb{S}^2)} \qquad \text{as } n \text{ increases?}$$

These results are partially inspired by trying to understand how curvature impacts the Kakeya phenomenon. By the Kakeya phenomenon in $d = 2$ dimensions, we mean the following relatively elementary proposition.

Proposition (Folklore). *Let $\ell_1, \ldots, \ell_n$ be any set of n lines in $\mathbb{R}^2$ such that any two lines intersect in some point and denote their $1/n-$neighborhoods by $T_1, \ldots, T_n$.*

Let $s \geq 0$, then there exists $c_s > 0$ such that

$$\sum_{\substack{i,j=1\\ i\neq j}}^{n} |T_i \cap T_j|^s \geq c_s \begin{cases} n^{2-2s} & \textit{if } 0 \leq s < 1 \\ \log n & \textit{if } s = 1 \\ n^{1-s} & \textit{if } s > 1. \end{cases}.$$

We see, essentially, that there has to be some unavoidable overlap, this is shown by the $\log n$ for $s = 1$. The results cited above show that on the sphere the log is necessary for $s = 2$. What happens in negative curvature? Poincaré disk?

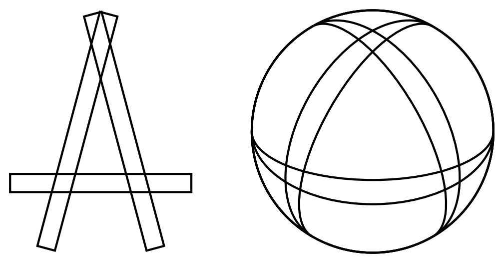


FIGURE 3. The difference between $\mathbb{R}^2$ and $\mathbb{S}^2$ illustrated: the curvature of $\mathbb{S}^2$ increases transversality, which decreases the area of intersection.

If we consider the $\delta-$neighborhoods $C_{1,\delta}, C_{2,\delta}, \ldots, C_{n,\delta}$ of n fixed great circles where no two great circles coincide and if$p_1, \ldots, p_n$ denote one of their 'poles', then

$$\lim_{\delta \to 0} \frac{1}{\delta^2} \sum_{\substack{i,j=1\\ i\neq j}}^{n} |C_{i,\delta} \cap C_{j,\delta}|^s = \sum_{\substack{i,j=1\\ i\neq j}}^{n} \frac{1}{(1 - \langle p_i, p_j \rangle^2)^{s/2}}.$$

This seems like an interesting minimization problem in its own right.

Update (Nov 2020)**.** This notion of Riesz energy

$$\sum_{\substack{i,j=1\\ i\neq j}}^{n} \frac{1}{(1 - \langle p_i, p_j \rangle^2)^{s/2}}$$

has been investigated by Chen, Hardin, Saff ('On the search for tight frames...', arXiv 2020) who show that minimizing configurations are well seperated.

4. Strange Patterns in Ulam's Sequence

In the 1960s, Stanislaw Ulam defined the following integer sequence: start with $1, 2$ and then add the smallest integer that can be uniquely written as the sum of two distinct earlier terms. It is not quite clear why he defined this sequence. It starts

$$1, 2, 3, 4, 6, 8, 11, 13, \ldots$$

I found (Experimental Mathematics, 2017) that this sequence seems to obey a strange quasi-periodic law: indeed, the first 10^7 terms of the sequence satisfy

$$\cos(2.5714474995 a_n) < 0 \qquad \text{except for } a_n \in \{2, 3, 47, 69\}.$$

I don't know what that number $2.571\ldots$ is or why this should be true. This type of computation has since been extended to higher values, it seems to hold true up to at least 10^{12} terms. Moreover, the sequence $2.571447\ldots a_n \mod 2\pi$ seems to have a limiting distribution that is compactly supported and has a strange shape (see the Figure). What is going on here?

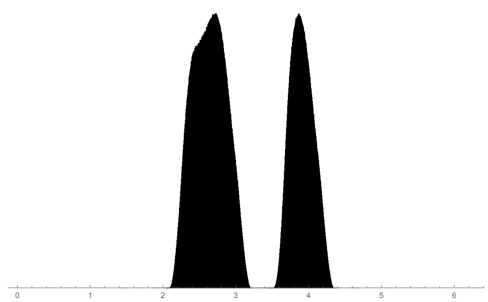


FIGURE 4. The empirical density of $2.571\ldots a_n \mod 2\pi$ seems to be compactly supported.

The same seems to be true if one starts with initial values different from $1, 2$. For some choices, the arising sequence seems to be a union of arithmetic progressions: whenever that is not the case, it seems to be 'chaotic' in the same sense: there exists a constant α (depending on the initival values) such that $\alpha a_n \mod 2\pi$ has a strange limiting distribution. There are now several papers showing that these strange type of phenomenon seems to persist even in other settings. I would like to understand what this sequence does – even the most basic things are not known: the sequence is known to be infinite since $a_n + a_{n-1}$ can be uniquely written as the sum of two earlier terms and thus

$$a_{n+1} \leq a_n + a_{n-1}$$

This also shows that the sequence grows at most exponentially and that is the best bound I am aware of. Empirically, the sequence has density $\sim 7\%$ and it seems like $a_n \leq 14n$ for all n sufficiently large.

Update (May 2022)**.** Rodrigo Angelo ('A hidden signal in Hofstadter's H sequence') discovered another example of a sequence with this property and gives rigorous proofs for this example.

Update (Jan 2025)**.** François Clément and I showed (arXiv:2501.16285) that

$$\min_{1 \leq i \leq n} \frac{a_{i+1}}{a_i} \leq 1 + c\frac{\log n}{n},$$

this means that some small gaps have to exist. We also improved the growth rate estimate $a_n \leq 1.45\ldots^n$ which was surprisingly tricky. It's clear that new ideas are needed.

5. The Number of Ulam Words

We consider strings of binary digits. We say that 0 and 1 are both Ulam. More generally, a string of binary digits is Ulam if it can be written uniquely as the concatenation of two smaller *distinct* Ulam strings. This means that 01 and 10 are Ulam. The Ulam binary strings of length 3 sare

$$001, 011, 100, 110$$

and 101 is missing because $101 = 10 \oplus 1 = 1 \oplus 01$. Adutwum, Clark, Emerson, Sheydvasser, Sheydvasser, Tougouma (arXiv:2410.01217) constructed all Ulam words up to length 30 and conjecture

$$\frac{\#\left\{\text{Ulam words of length } n\right\}}{2^n} \sim \frac{0.526}{n^{3/10}}.$$

The exponent 3/10 is most unusual but the numerics are somewhat convincing (as much as numerics up to $n \leq 30$ can be). What can be said about the number of Ulam words?

6. Topological Structures in Irrational Rotations on the Torus

Let $x_n = n\alpha \mod 1$ where $\alpha \in \mathbb{R}\setminus\mathbb{Q}$. Let us consider the first n elements $x_1, \ldots, x_n$ and then construct the following graph G on n vertices $\{1, 2, \ldots, n\}$: first, we connect $(1, 2)$, then $(2, 3)$, then $(3, 4)$ and so on until $(n, 1)$. This results in a cycle on n elements. Then we find the permutation $\pi : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ for which

$$x_{\pi(1)} < x_{\pi(2)} < \cdots < x_{\pi(n)}.$$

We then connect $\pi(1)$ to $\pi(2)$ and $\pi(2)$ to $\pi(3)$ and so on until $\pi(n)$ is being connected to $\pi(1)$. I used (arXiv, August 2020) these types of graphs to construct a test for whether $x_1, \ldots, x_n$ are i.i.d. samples of a random variable: in that case, the arising Graphs are expanders and close to Ramanujan. However, if one builds this graph from the sequence of irrational rotations on the torus (certainly **not** i.i.d.), very interesting graphs arise. It seems like they correspond to nice underlying limiting objects? What are those?

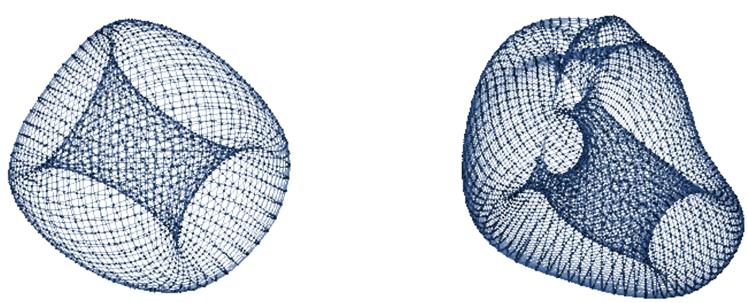


Figure 5. The Graph arising from $x_n = \phi n \mod 1$, where $\phi = (1+\sqrt{5})/2$ (left) and the van der Corput sequence in base 2 (right).

Simultaneously, if one takes the standard van der Corput sequence in base 2, one seems to end up with a really nice manifold with some strange holes where things are glued together in a fun way.

Update (Dec 2020)**.** We found ('Finding Structure in Sequences of Real Numbers via Graph Theory: a Problem List', arXiv, Dec 2020) that there are many sequences that lead to interesting Graphs when applying this construction. For most of them it is not at all clear why this happens.

7. Graphical Designs

This problem is somewhere between PDEs and Combinatorics. Let $G = (V, E)$ be a finite, undirected, simple graph. Then we can define functions on the Graph as mappings $f : V \to \mathbb{R}$. We can also define a Laplacian on the Graph, this is simply a map that sends functions to other functions. One possible choice is

$$(Lf)(v) = f(v) - \frac{1}{\deg(v)} \sum_{w \sim_E v} f(w).$$

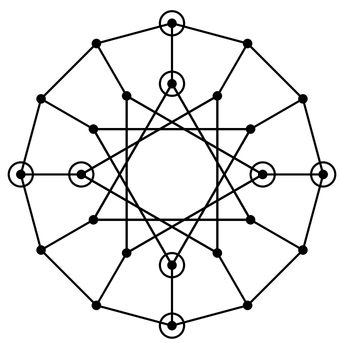


Figure 6. Generalized Petersen Graph (12,4) on 24 vertices: a subset of 8 vertices integrates the first 22 eigenfunctions exactly.

There are a couple of different definitions of the Laplacian and I don't know what's the best choice for this problem. However, these different definitions of a Laplacian all agree on $d-$regular graphs and the phenomenon at hand is already interesting for $d-$regular graphs. Once one has a Laplacian matrix, one has eigenvectors and eigenvalues. We will interpret these eigenvectors again as functions on the graph. What one observes is that for many interesting graphs, there are subsets of the vertices $W \subset V$ such that for many different eigenvectors ϕ_k of the Graph Laplacian

$$\sum_{v \in W} \phi_k(v) = 0.$$

What is interesting is that such sets seem to inherit a lot of rich structure. I proved (Journal of Graph Theory) that if there are large subsets W such that the equation holds for many different eigenvectors k, then the Graph has to be 'non-Euclidean' in the sense the volume of balls grows quite quickly (exponentially

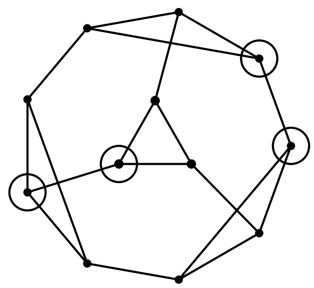


Figure 7. The Truncated Tetrahedral Graph on 12 vertices: a subset of 4 vertices integrates the first 11 eigenfunctions exactly.

depending on the precise parameters). This shows that we expect graphs with nice structure like this to be more like an expander than, say, a path graph. Konstantin Golubev showed (Lin. Alg. Appl.) that this framework naturally encodes some classical results from classical combinatorics: both the Erdős-Ko-Rado theorem and the Deza-Frankl theorem can be stated as being special types of these 'Graphical Designs'. There are many other connections: (1) for certain types of graphs, this seems to be related to results from coding theory and (2) such points would also be very good points when one tries to sample an unknown function in just a few vertices; this is because the definition can be (3) regarded as a Graph equivalent of the classical notion of 'spherical design' on $\mathbb{S}^{d-1}$. In fact, seeing as designs can be regarded as an algebraic definition that gives rise to Platonic bodies in $\mathbb{R}^3$, I like to think of these 'graphical designs' as the analogue of 'Platonic bodies in a graph'. There are a great many questions:

(1) when do such sets exists?
(2) are there nice extremal examples?
(3) how does one find them quickly?
(4) how can one prove non-existence?

The theory of spherical designs on $\mathbb{S}^{d-1}$ is quite rich and full of intricate problems so I would assumes the Graph analogue to be at least as difficult and probably more difficult. But there are many more graphs than there are spheres (only one per dimension: $\mathbb{S}^d$), so there should be many more interesting examples that may themselves be tied to interesting algebraic-combinatorial structures.

Update (Jan 2025)**.** There exist a fair number of results in this direction at this point (cf. work of Catherine Babecki, Rekha Thomas, Konstantin Golubev and others).

8. How big is the boundary of a graph?

Let $G = (V, E)$ be a graph. In 'The Boundary of a Graph and its Isoperimetric Inequality' (Jan 2022) we introduced the following notion of a 'boundary' of a graph: we say that $u \in V$ is part of the boundary ∂G if there exists another vertex $v \in V$ such that

$$\frac{1}{\deg(u)} \sum_{(u,w) \in E} d(w, v) < d(u, v).$$

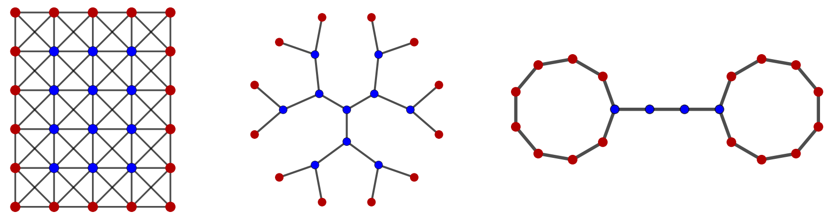


FIGURE 8. Graphs, their boundary ∂G (red) and $V \setminus \partial G$ (blue).

The same paper an isoperimetric inequality: each vertex $v \in V$ will detect a 'large' number of vertices as boundary vertices.

Theorem. *If G is a connected graph with maximal degree Δ, then for all $v \in V$*

$$\left| \left\{ u \in V \Big| \quad \frac{1}{\deg(u)} \sum_{(u,w) \in E} d(w,v) < d(u,v) \right\} \right| \geq \frac{1}{2\Delta} \frac{|V|}{\operatorname{diam}(G)}.$$

This inequality is presumably close to optimal and it implies

$$|\partial G| \geq \frac{1}{2\Delta} \frac{|V|}{\operatorname{diam}(G)}.$$

I would be interested in understanding whether one can other results of this flavor (perhaps invoking other graph parameters). I would also be interested in the isoperimetric problem: what are the graphs with minimal boundary? Put differently, what are the 'balls' in the graph universe?

Update (Sep 2022). Chiem, Dudarov, Lee, Lee & Liu ('A characterization of graphs with at most four boundary vertices') have characterized all graphs with at most 4 boundary vertices.

9. The Constant in the Komlos Conjecture

Let $A \in \mathbb{R}^{n \times n}$ have the property that all columns have ℓ^2-norm at most 1.

> **Komlos Conjecture.** There exists a universal constant $K > 0$ such that for all such matrices A
>
> $$\min_{x \in \{-1,1\}^n} \|Ax\|_{\ell^\infty} \leq K.$$

The best known result is $K = \mathcal{O}(\sqrt{\log n})$ (Banaszczyk) and the conjecture is that there exists a universal $K = \mathcal{O}(1)$ independent of the dimension. What makes the conjecture even more charming is the constant K might actually be quite small.

> **Question.** Is it possible to get good lower bounds on K?

Remarkably little seems to be known about this. It is not that easy to construct a matrix showing that $K > 1.5$ and apparently for a while it was considered a unreasonable guess that $K = 2$. The best known result that I know of is due to Kunisky ('The discrepancy of unsatisfiable matrices and a lower bound for the Komlos conjecture constant') showing that

$$K \geq 1 + \sqrt{2} = 2.4142\ldots$$

What makes the problem of constructing lower bounds hard is that given a $n \times n$ matrix, one needs to check 2^n vectors to verify that for all of them $\|Ax\|_{\ell^\infty}$ is large.

$$\frac{1}{\sqrt{3}}\begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \end{pmatrix}$$

FIGURE 9. A matrix showing $K \geq \sqrt{3}$.

What I found is that, at least with regards to numerical experimentation, **finite projective planes** seem to be interesting candidates (the matrix example showing $K \geq \sqrt{3}$ is the incidence matrix of the Fano plane). The problem then has completely combinatorial flavor. Given a finite projective plane X, what is the best constant c_X such that for every $2-$coloring $\chi : X \to \{-1, 1\}$, there always exists a line ℓ in the projective such that

$$\left| \sum_{x \in \ell} \chi(x) \right| \geq c_X \qquad ?$$

It is well-known (follows from the spectral bound, for example), that

$$c_X \geq (1 - o(1)) \cdot \sqrt{\# X}.$$

It is known (Joel Spencer, Coloring the projective plane, 1988) that this is the correct order of magnitude and, for some universal $\alpha > 1$,

$$c_X \leq \alpha \sqrt{\# X}.$$

We are now interested in the value of α for the following reason.

Proposition. If X is a finite projective plane, then

$$K \geq \frac{c_X}{\sqrt{\# X}}.$$

Of course, unsurprisingly, the constant c_X is also not easy to compute. What one can do in practice is to try heuristic coloring schemes to obtain upper bounds and hope that they somehow capture the underlying behavior.

X	$PG(2,2)$	$PG(2,3)$	$PG(2,4)$	$PG(2,5)$	$PG(2,7)$	$PG(2,13)$	$PG(2,23)$
c_X	$= 3$	$= 2$	$= 3$	≤ 4	≤ 4	≤ 8	≤ 12

TABLE 1. Some upper bounds on c_X

If, for example, it were the case that for $X = PG(2, 23)$, we indeed have $c_X = 12$, then this would correspond to a lower $K \geq \sqrt{6} \sim 2.44\ldots$. I have seen people use SAT solvers for related problems but it's not clear to me whether there is any hope of doing something like this here. It might also be interesting to obtain an upper bound on how well one could hope to do with this kind of construction.

Question. Is it possible to make Spencer's bound
$$c_X \le \alpha\sqrt{\#X}$$
effective? Can one choose $\alpha = 10$ for example?

Update (Aug 2022)**.** Victor Reis reports the bounds in Table 2. They do seem to indicate that there are actually rather effective two-colorings (in the sense of the α being close to 2 or even smaller) which would suggest that finite projective plane will not lead to good lower bound for the Komlos conjecture.

X	$PG(2,5)$	$PG(2,13)$	$PG(2,23)$	$PG(2,31)$	$PG(2,41)$
c_X	$=4$	≤ 6	≤ 8	≤ 10	≤ 12

X	$PG(2,47)$	$PG(2,67)$	$PG(2,79)$	$PG(2,83)$
c_X	≤ 14	≤ 16	≤ 16	≤ 20

Table 2. Bounds on c_X by Victor Reis

10. The Inverse of the Star Discrepancy

A central problem in discrepancy theory is the challenge of distributing points $\{x_1, \dots, x_n\}$ in $[0,1]^d$ as evenly as possible. Naturally there are different measures of regularity, one such measure is
$$\text{star-discrepancy} = \max_{y \in [0,1]^d} \left| \frac{1}{n} \# \left\{ 1 \le i \le n : x_i \in [0,y] \right\} - \text{vol}([0,y]) \right|,$$
where $[0,y] = [0,y_1] \times \cdots \times [0,y_d]$ has volume $\text{vol}([0,y]) = y_1 \cdot y_2 \cdot \cdots \cdot y_d$. A fundamental question in the area is the following.

Problem. Suppose we are in $[0,1]^d$ and want the star-discrepancy to be smaller than $0 < \varepsilon < 1$, how many points do we need?

The cardinality of the smallest set of points in $[0,1]^d$ achieving a star-discrepancy smaller than ε is sometimes denoted by $N_\infty^*(d,\varepsilon)$. The best known bounds are
$$\frac{d}{\varepsilon} \lesssim N_\infty^*(d,\varepsilon) \lesssim \frac{d}{\varepsilon^2},$$
where the upper bound is a probabilistic argument by Heinrich, Novak, Wasilkowski & Wozniakowski (2001). The lower bound was established by Hinrichs (2004) using Vapnik-Chervonenkis classes and the Sauer–Shelah lemma.

The upper bound construction is relatively easy: take iid random points. This leads to a fascinating dichotomy

- either random points are essentially as regular as possible
- or there are more regular constructions we do not yet know about.

I could well imagine that random is best possible but am personally hoping for the existence of better sets (because such sets would probably be pretty interesting). I gave a new proof of the lower bound ('An elementary proof of a lower bound for the inverse of the star discrepancy') which is relatively simple and entirely elementary – can any of these ideas be used to construct 'good' sets of points?

11. Erdős Distinct Subset Sums Problem

This problem is fairly well known, indeed, in a 1989 Rostock Math Kolloquium survey of problems, Erdős calls it "perhaps my first serious conjecture which goes back to 1931 or 32". Let $a_1 < \cdots < a_n$ be a set of n positive number integers such that all subset sums are distinct: from the sum of the subset it is possible to uniquely identify the subset. The powers of 2, for example, have this property. Erdős conjectured (and offered \$ 500) that $a_n \geq c \cdot 2^n$. Currently, the best known bound is

$$a_n \geq (c - o(1)) \frac{2^n}{\sqrt{n}}$$

where different estimates for c have been given over the years

$c \geq 1/4$	Erdős & Moser
$\geq 2/3^{3/2}$	Alon & Spencer
$\geq 1/\sqrt{\pi}$	Elkies
$\geq 1/\sqrt{3}$	Bae, Guy
$\geq \sqrt{3/2\pi}$	Aliev
$\geq \sqrt{2/\pi}$	Dubroff, Fox & Xu.

I gave another proof for $\sqrt{2/\pi}$ using Fourier Analysis ('Some Remarks on the Erdős Distinct Subset Sums Problem') which suggests a couple of additional things about the structure of such sets. In particular, it seems to suggest that optimal sets may have several values rather close to the maximal value and then additional values at some of the points $(\max a_i)/2, (\max a_i)/3, (\max a_i)/4, \ldots$. Elkies (J Comb Theory A, 1986) already has the following gorgeous reformulation of the problem.

> **Erdős Distinct Subset Sum Problem, Fourier Version**. Suppose $a_1, \ldots, a_n$ are distinct integers such that
> $$\int_0^1 \prod_{i=1}^{n} \cos(2\pi a_i x)^2 dx = \frac{1}{2^n},$$
> does this imply $\max_i a_i \gtrsim 2^n$? We know $\max_i a_i \gtrsim 2^n/\sqrt{n}$.

There is another interesting problem that arises from relaxing the conditions: what if we do not require subset sums to be distinct but merely require them to attain many different values (some duplicates are allowed as long as there are few).

> **Problem.** Are there sets $\{a_1, \ldots, a_n\} \subset \mathbb{N}$ such that the subset sums attain $(1 - o(1)) \cdot 2^n$ distinct values and $a_n = o(2^n)$?

12. Finding Short Paths in Graphs with Spectral Graph Theory

This section describes a curious phenomenon that I do not understand (described in greater detail in 'A Spectral Approach to the Shortest Path Problem', arXiv April 2020). Given a (connected) graph $G = (V, E)$ and a vertex $u \in V$, we can look for the following function $\phi : V \to \mathbb{R}$

$$\phi = \arg \min_{\substack{f:V\to\mathbb{R} \\ f(u)=0, f\not\equiv 0}} \frac{\sum_{(w_1,w_2)\in E} (f(w_1) - f(w_2))^2}{\sum_{w\in V} f(w)^2}.$$

Basically, ϕ is a function that vanishes in the vertex u but changes as little as possible from one vertex to the next (subject to a normalization in ℓ^2). ϕ is actually easy to compute, it is an eigenvector of a square matrix that is explicit (essentially the Graph Laplacian after one has removed the row and column that belongs to u). We can assume w.l.o.g. that ϕ is positive everywhere except in u.

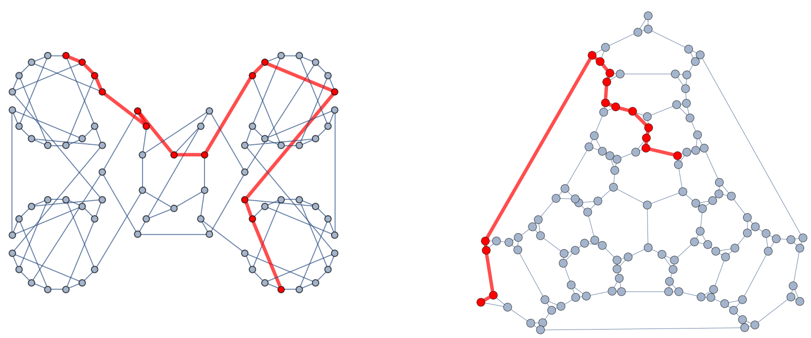


FIGURE 10. Paths taken by the Spectral Method.

If one then starts in a vertex $u \neq v \in V$ and is interested in a short path to the vertex u, one can do the following. Look among all neighbors of where you currently are for the one that has the smallest $\phi-$value. Go there and repeat. This will provably lead to a path from v to u.

> **Question.** Very often, this path will be fairly short (i.e. comparable in length to the shortest path). Why?

We emphasize that this not always the case; however, we found that in many cases these paths are quite good. What can be proven? Is it possible to find families of graphs for which these spectral paths always coincide with the shortest paths?

Part of the motivation is that this question can be interpreted as a discrete version of the Hot Spots conjecture (in particular, if the graph discretizes a convex domain in the usual grid-like fashion, then we expect ϕ to be monotonically increasing away from the vertex and to assume its maximum on the boundary).

Update (Nov 2020). Yancey & Yancey (arXiv:2011.08998) discuss some counterexamples and propose graph curvature as an interesting condition.

13. A Curve through the ℓ^1-norm of eigenvectors of Erdős-Renyi graphs

This is a strange phenomenon that Alex Cloninger and I discovered a while back. It is mentioned in our paper ('On The Dual Geometry of Laplacian Eigenfunctions', arXiv April 2018) but it's not widely known.

Let $G(n,p)$ be a standard Erdős-Renyi random graph and let $L = D - A$ be the associated Laplacian matrix. This matrix has eigenvalues $0 = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$. We only looked at the case where p is fixed and n is large and the Graph is usually (even highly) connected. In that case there is one trivial (constant) eigenfunction $\phi_1 = 1/\sqrt{n}$. We note that, since the eigenvectors are normalized in ℓ^2, we have

$$\|v_i\|_{\ell^1} \leq \sqrt{n} \cdot \|v_i\|_{\ell^2} = \sqrt{n}.$$

Moreover, $\|v_i\|_{\ell^1}$ is a measure for how localized an eigenvector is: the more it concentrates its mass on few vertices, the smaller the norm is. If the eigenvector is completely flat (i.e. constant), then the norm is maximal and given by $\sqrt{n}$.

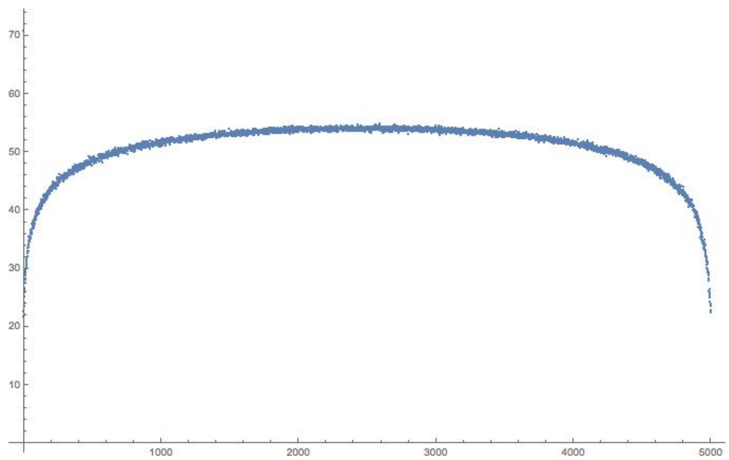


FIGURE 11. ℓ^1-norm of $v_1, \dots, v_n$ lies on a nice curve (the underlying graph is $G(n,p)$ with $n = 5000$ and $p = 0.4$).

Question. When we plot $\|v_i\|_{\ell^1}$ for $i = 1...n$, they seem to lie on a curve (see Fig. 11). Why would they do that?

This somehow means that eigenvectors at the edge of the spectrum are more localized: that is perhaps not too surprising. What is truly surprising is that the eigenvectors seem to **uniformly** lie very close to this curve. There seems to be a strong measure concentration phenomenon at work: the curve always looks the same for many different random realizations of $G(n,p)$ (for fixed n,p).

A second question is what happens in the middle. What we see there is that

$$\mathbb{E} \max_{1 \leq i \leq n} \frac{\|v_i\|_{\ell^1}}{\sqrt{n}} \sim 0.8.$$

The relevant i seems to be $i \sim n/2$. If $X \sim \mathcal{N}(0,1)$ is a standard Gaussian, then $\mathbb{E}|X| = \sqrt{2/\pi} \sim 0.7978\dots$. Coincidence?

Update (Dec 2020). I asked the question on mathoverflow. Ofer Zeitouni pointed out that works by Rudelson-Vershynin and Eldan et al. suggest the lower bound

$$\mathbb{E} \min_{1 \leq i \leq n} \frac{\|v_i\|_{\ell^1}}{\sqrt{n}} \gtrsim \frac{1}{(\log n)^c}.$$

14. Matching Oscillations in High-Frequency Eigenvectors.

This section discusses a phenomenon that is perhaps best introduced with an example: consider the Thomassen graph on 94 vertices, consider the Graph Laplacian $L = D - A$ with eigenvalues ordered as

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{94} = 0.$$

This graph is 3-regular, the three largest eigenvalues are distinct. The Figure shows the signs of ϕ_2, ϕ_3 (left and middle) and the sign of $\phi_2 \cdot \phi_3$.

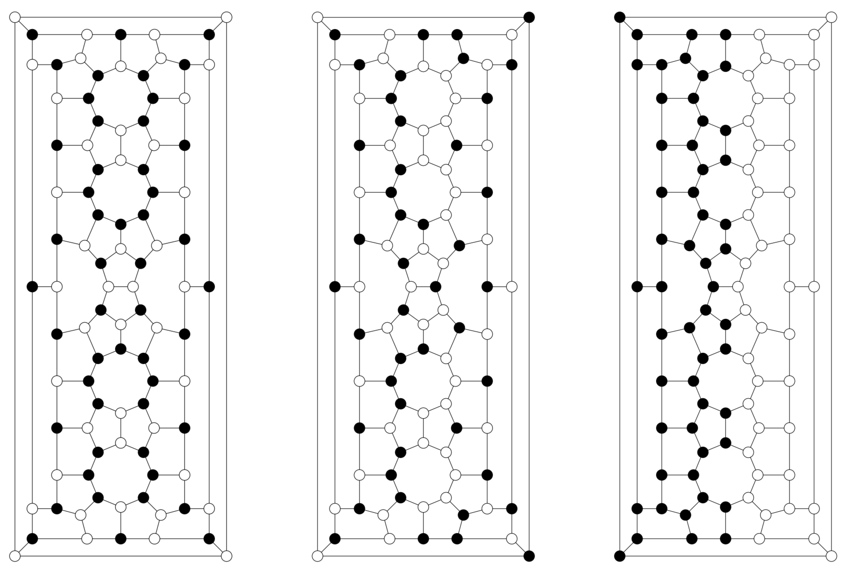


Figure 12. Sign of the 2^{nd} and the 3^{rd} eigenvector and of their product.

The second and the third eigenvector have sign changes across most of the edges: they oscillate essentially as quickly as the graph allows. In contrast, the (pointwise) product of these high-frequency eigenvectors appears to be much smoother and exhibits a sign pattern typical of low-frequency eigenvectors: positive and negative entries are clustered together and meet across a smooth interface.

I gave a theoretical explanation in ('The product of two high-frequency Graph Laplacian eigenfunctions is smooth', Discrete Mathematics) but it seems like it's a rather rich phenomenon and maybe I barely scratched the surface? It also seems as if this might actually be useful in applications...?

15. Balanced Measures are small?

Let $G = (V, E)$ be some finite, combinatorial graph. We can define an infinite sequence in the set of vertices by starting with some arbitrary vertex and then setting

$$\sum_{i=1}^{k} d(x_{k+1}, x_i) = \max_{v \in V} \sum_{i=1}^{k} d(v, x_i).$$

In words: the new vertex that is being added has the property of maximizing the sum of distances to the previous vertices. This procedure is also interesting in the continuous setting (see 'Sums of distances between points on the sphere')

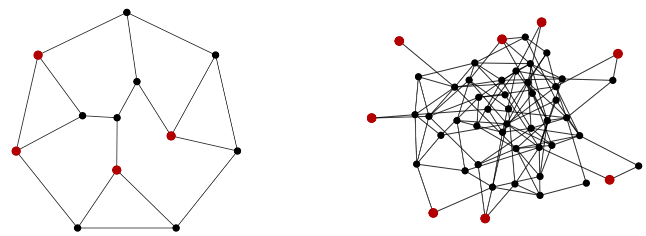


FIGURE 13. The Frucht graph (left) and an Erdős-Renyi random graph on $n = 50$ vertices (right). The rule ends up only selecting the red vertices (not necessarily with equal frequency).

One could now wonder about the long-term behavior of this construction and what one observes is that vertices are not taken with equal frequency, things are quite lopsided and the actual frequency with which a vertex is selected converges to a very special kind of probability distribution.

Definition 1. A probability measure μ on the vertices V is *balanced* if

$$\forall w \in V \qquad \mu(w) > 0 \implies \sum_{u \in V} d(w,u)\mu(u) = \max_{v \in V} \sum_{u \in V} d(v,u)\mu(u).$$

So basically the support of the measure is simultaneously the point to which the global transport of the entire measure would be the most expensive. We note that balanced measures are not necessarily unique.

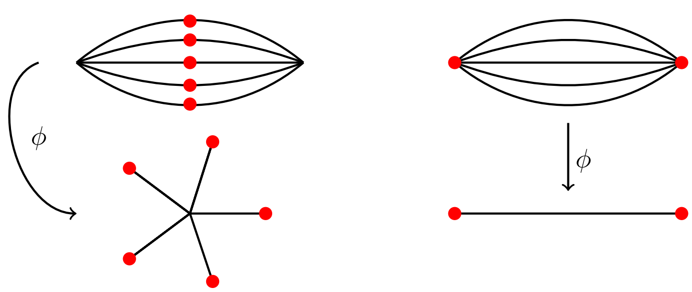


FIGURE 14. m path graphs of length $2\ell + 1$ glued together at the endpoints. Top: two different balanced measures leading to (bottom) two different embeddings emphasizing different aspects.

The thing that is now really interesting is the following.

> **Problem.** Balanced measures tend to be supported in a rather small number of vertices. Why?

There are certainly graphs where this is not the case: examples are given by the dodecahedral graph or the Desargue graphs for which the uniform measure is balanced. In practice, it seems very difficult for graphs to have the support of a balanced measure contain a large number of vertices and it would be interesting to have a more quantitative understanding of this.

Update (Jan 2025)**.** Baimetov, Bushling, Goh, Guo, Jacobs, Lee ('A decomposition theorem for balanced measures') provide some structural theory: a measure is balanced if and only if it is a convex combination of compatible balanced measures (see their paper for more details).

16. Different areas of triangles induced by points on $\mathbb{S}^1$

Suppose we are given n distinct points $x_1, x_2, \dots, x_n \in \mathbb{S}^1 = \left\{x \in \mathbb{R}^2 : \|x\| = 1\right\}$. Any three points x_i, x_j, x_k induce a triangle $\Delta(x_i, x_j, x_k)$ which has some area. The question is: how many different areas are we going to see?

Question. Is it true that

$$\#\left\{\text{area}\left[\Delta(x_i, x_j, x_k)\right] : 1 \leq i < j < k \leq n\right\} \gtrsim n^2 \qquad ?$$

One could even ask the stronger question whether any fixed point, being part of $\sim n^2$ triangles, is bound to see at least $\sim n^2$ different areas, i.e. whether

$$\#\left\{\text{area}\left[\Delta(x_1, x_j, x_k)\right] : 1 \leq j < k \leq n\right\} \gtrsim n^2 \qquad ?$$

There are a couple of different reasons why I like the problem (besides it being a totally elementary question). The first is that the extremizers are not obvious. It is completely clear that one would expect n equispaced points to generate the least number of different triangle areas and this is probably true. However, the set of points

$$X_n = \left\{e^{ik\alpha} : 1 \leq k \leq n\right\} \subset \mathbb{C} \cong \mathbb{R}^2$$

only differs up to a constant. It is easy to see that, since areas of a triangle are determined by the three sides and we constrain the points to lie on $\mathbb{S}^1$, that

$$\#\left\{\text{area}\left[\Delta(x_i, x_j, x_k)\right] : 1 \leq i < j < k \leq n\right\} \lesssim \left(\#\left\{\|x_i - x_j\| : 1 \leq i, j \leq n\right\}\right)^2$$

and irrational rotations on the torus only have $\sim n$ pairwise distances since

$$|e^{ik\alpha} - e^{im\alpha}| = |e^{i(k-m)\alpha} - 1|.$$

The second reason I like the problem is the formula

$$\text{area}\left[\Delta(e^{is}, e^{it}, e^{iu})\right] = 4 \sin\left(\frac{s-t}{2}\right)^2 \sin\left(\frac{u-s}{2}\right)^2 \sin\left(\frac{t-u}{2}\right)^2$$

This naturally calls for the angles to be in some form of generalized arithmetic progression but it's clear that there is a nonlinear twist to it. The third reason I like the problem is that it is known (Erdos-Purdy and others) that n points in the planar induce at least $\lfloor (n-1)/2 \rfloor$ different triangle areas: taking equispaced points on two parallel lines shows that this is sharp. This means the curvature of $\mathbb{S}^1$ has to somehow play a role.

Update (May 2023)**.** I asked the question on mathoverflow, no reply.

Part 2. **Analysis**

17. Sums of distances between points on the sphere

It is known that for any set $\{x_1, \dots, x_n\} \subset \mathbb{S}^2$, we have

$$\sum_{i,j=1}^{n} \|x_i - x_j\| \le \frac{4}{3}n^2 - c\sqrt{n}.$$

Here, $\| \cdot \|$ is the Euclidean distance between points in $\mathbb{R}^3$. The number $4/3$ is the average distance between two randomly chosen points on the sphere and $c > 0$. The inequality follows from the Stolarsky Invariance Principle (cf. works of Dmitriy Bilyk) and the lower bound on the L^2-spherical cap discrepancy due to Jozsef Beck.

No good explicit construction of points with this property is known (the only construction that is known to attain this bound are randomized constructions involving either jittered sampling or determinantal point processes). A strange mystery (discussed in 'Polarization and Greedy Energy on the Sphere') is the following: if we start with an arbitrary initial set of points $\{x_1, \dots, x_m\}$ and then define a greedy sequence by setting the next point so as to maximize the sum of the existing distances

$$x_{n+1} = \arg\max_{x \in \mathbb{S}^2} \sum_{i=1}^{n} \|x - x_i\|.$$

This seems to actually lead to a sequence of maximal growth. We know ('Polarization and Greedy Energy on the Sphere') that, for n sufficiently large (depending on the initial conditions),

$$\sum_{i,j=1}^{n} \|x_i - x_j\| \ge \frac{4}{3}n^2 - 100n$$

and it's completely clear from the proof that this far from optimal (in fact, it is optimal for a much wider class of sequences that is clearly not as well behaved). The procedure is very stable: one does not need to take the actual maximum, it is enough to pick a value that is sufficiently close. One can even sometimes replace x_{n+1} by some other arbitrary point (perhaps a really bad one) and, as long as one does not do that too often, this does not seem to cause any issues, it appears to be an incredibly robust procedure. It also seemingly works in higher dimensions. Why? We also note that there are at least two different problems here

(1) finding a good construction for fixed n
(2) finding a sequence $(x_k)_{k=1}^{\infty}$ such that $(x_k)_{k=1}^{n}$ is good uniformly in n.

It is clear that the second problem is harder than the first but it also appears, numerically, as if the greedy sequence leads to a sequence $(x_k)_{k=1}^{\infty}$ such that $(x_k)_{k=1}^{n}$ is optimal (up to constants) uniformly in n.

The problem is also interesting on graphs (see 'Balanced measures are small?')

18. A Directional Poincare Inequality: flows of vector fields

I showed (Arkiv Math. 2016) a curious refinement of the Poincaré inequality on the torus $\mathbb{T}^d$. A special case on the 2−dimensional Torus $\mathbb{T}^2$ reads as follows: for functions $f : \mathbb{T}^2 \to \mathbb{R}$ with mean value 0

$$\|f\|_{L^2(\mathbb{T}^2)}^2 \leq c \|\nabla f\|_{L^2(\mathbb{T}^2)} \left\| \frac{\partial f}{\partial x} + \sqrt{2} \frac{\partial f}{\partial y} \right\|_{L^2(\mathbb{T}^2)}$$

Moreover, the inequality fails when $\sqrt{2}$ is replaced by e and probably fails when $\sqrt{2}$ is replaced by π. This is a funny inequality because, as opposed to the classical Poincare inequality, this one does not have a difference in regularity: there are infinitely many orthogonal functions where the LHS is (up to a constant) comparable to the RHS. So in some sense it is an absolutely sharp form of Poincare where some portion of the derivative is exchanged against a directional derivative.

One would assume that this is generally possible. For example, let V be a vector field on $\mathbb{T}^2$. When do we have, for some fixed universal $\delta = \delta(V) > 0$ that for all $f \in C^\infty(\mathbb{T}^2)$ with mean value 0

$$\|\nabla f\|_{L^2(\mathbb{T}^2)}^{1-\delta} \| \left\langle \nabla f, V \right\rangle \|_{L^2(\mathbb{T}^2)}^{\delta} \geq c_\alpha \|f\|_{L^2(\mathbb{T}^2)}?$$

How does δ depend on V? One would expect that it depends on the mixing properties of V: the better it mixes, the larger δ can be. It should be connected to how quickly the flow of the vector field transports you from the vicinity of one point to the vicinity of another point. Is it true that δ can never be larger than $1/2$?

19. Auto-Convolution Inequalities and additive combinatorics

All these questions are relevant in additive combinatorics (see papers) but interesting in their own right. Let $f \in C^\infty(-1/4, 1/4)$ satisfy $f \geq 0$ and have total integral 1. Then the convolution $f * f$ is compactly supported on $(-1/2, 1/2)$ and, by Fubini, still has total integral 1. How large is $\|f * f\|_{L^\infty}$ going to be? By the Pigeonhole principle, we have $\|f * f\|_{L^\infty} \geq 1$. However, this is clearly lossy since it can only be sharp if $f * f$ was the characteristic function on $(-1/2, 1/2)$ which is the convolution of a function with itself (this can be seen by looking at the Fourier transform which assumes negative values).

Theorem (Alex Cloninger & S, Proc. Amer. Math. Soc.)**.** *Let $f : \mathbb{R} \to \mathbb{R}_{\geq 0}$ be supported in $[-1/4, 1/4]$. Then*

$$\sup_{x \in \mathbb{R}} \int_{\mathbb{R}} f(t) f(x-t) dt \geq 1.28 \left(\int_{-1/4}^{1/4} f(x) dx \right)^2.$$

It seems likely that the optimal constant is closer to ~ 1.5.

The question also has a dual formulation (which also has relevance in Additive Combinatorics): whenever we have an L^1-function f, then there exists a shift such that $f(x)$ and $f(x-t)$ have small inner product. We observe that, for $f \geq 0$,

$$\begin{aligned} \min_{0 \leq t \leq 1} \int_{\mathbb{R}} f(x) f(x-t) dx &\leq \int_0^1 \int_{\mathbb{R}} f(x) f(x-t) dx dt \\ &\leq \int_0^\infty \int_{\mathbb{R}} f(x) f(x-t) dx dt = \frac{\|f\|_{L^1}^2}{2}. \end{aligned}$$

So the statement itself is not complicated but the optimal constant seems to be quite complicated. We currently know that

Theorem (Rick Barnard & S., J. Number Theory)**.** *We have*

$$\min_{0 \le t \le 1} \int_{\mathbb{R}} f(x) f(x+t) dx \le 0.42 \|f\|_{L^1}^2$$

and 0.42 cannot be replaced by 0.37.

Noah Kravitz (arXiv, April 2020) proved that the question is equivalent to an old question about the cardinality of difference bases. We conclude with a related question of G. Martin & K. O'Bryant: if $f \in L^1(\mathbb{R})$ is nonnegative, can Hölder's inequality be improved? Is there a universal $\delta > 0$ such that

$$\frac{\|f * f\|_{L^\infty} \|f * f\|_{L^1}}{\|f * f\|_{L^2}^2} \ge 1 + \delta?$$

They produce an example ($f(x) = 1/\sqrt{2x}$ if $0 < x < 1/2$, $f(x) = 0$ otherwise) showing that δ cannot exceed 0.13.

Update (Aug 2022). Ramos & Madrid (Comm. Pure Appl Anal, 2021) proved that the optimal constant c in

$$\min_{0 \le t \le 1} \int_{\mathbb{R}} f(x) f(x+t) dx \le c \|f\|_{L^1}^2$$

is *strictly less* than what one obtains from the argument by Rick and myself.

Update (Nov 2022). Ethan White ('An almost-tight L^2 autoconvolution inequality') essentially solves an L^2-version of the convolution problem (in the sense of getting the sharp constant up to 4 digits).

Update (June 2025). There has been some improvement by the Google AlphaEvolve Team (arXiv:2506.13131) and Boyer-Li (arXiv:2506.16750). The numerical examples of extremizers suggest that this might be a very difficult problem indeed.

20. Maxwell's Conjecture on Point Charges

This is a strikingly simple question that can be traced to the work of James Clerk Maxwell. Let $x_1, x_2, x_3 \in \mathbb{R}^2$ and define $f : \mathbb{R}^2 \to \mathbb{R}$ via

$$f(x) = \sum_{i=1}^{3} \frac{1}{\|x - x_i\|}.$$

How often can ∇f vanish or, phrased differently, how many critical points can there be? It is easy to see that all the critical points have to be in the convex hull of the three points of x_1, x_2, x_3. Once one does some basic experimentation, one sees that there seem to be at most 4 critical points (if the triangle is very flat, it is possible that there are only 2). Gabrielov, Novikov, Shapiro (Proc. London Math, 2007) showed that there are at most 12. A more recent 2015 Physica D paper of Y.-L. Tsai (Maxwell's conjecture on three point charges with equal magnitudes) shows that there are at most 4 critical points – the proof is heavily computational and seems hard to generalize. Is there a 'simple' proof for $n = 3$? What about other potential functions? Now suppose there are 4 points x_1, x_2, x_3, x_4. In that case, it is not even known whether there is a *finite* number! There are even related one-dimensional problems that are wide open such as

Conjecture (Gabrielov, Novikov, Shapiro). Let $(x_1, y_1), \dots, (x_\ell, y_\ell) \in \mathbb{R}^2$. Then for any choice of $(\zeta_1, \dots, \zeta_\ell)$ and any $\alpha \geq 1/2$, the function $V_\alpha : \mathbb{R} \to \mathbb{R}$ given by

$$V_\alpha(x) = \sum_{i=1}^{\ell} \frac{\zeta_i}{(x - x_i)^2 + y_i^2)^\alpha}$$

has at most $2\ell - 1 - 1$ real critical points.

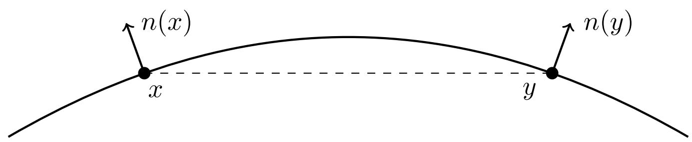
Update (July 2023) **.** Vladimir Zolotov ('Upper bounds for the number of isolated critical points via Thom-Milnor theorem') gives a number of strong bounds for this and related problems.

21. Quadratic Crofton and sets that barely see themselves

This is a problem at the interface of the calculus of variations and integral/random geometry. The problem is easily stated in any dimension but already the case of convex domains $\Omega \subset \mathbb{R}^2$. Let us first consider a one-dimensional rectifiable set $\mathcal{L} \subset \mathbb{R}^2$ with length $L > 0$. We define a notion of energy as

$$E(\mathcal{L}) = \int_{\mathcal{L}} \int_{\mathcal{L}} \frac{|\langle n(x), y - x \rangle \langle y - x, n(y) \rangle|}{\|x - y\|^3} \, d\sigma(x) d\sigma(y),$$

where x, y are elements in the set, $n(x)$ and $n(y)$ denote the normal vectors in x and y, respectively, and σ is the arclength measure.



One way of thinking about the functional is that it measures the behavior of the set when projected onto itself in the following sense: consider $x, y \in \mathcal{L}$ and let us take small neighborhoods around x and y (we may think of these as approximately being short line segments). We could then ask for the expected size of the projection of one such line segment onto the other under a 'random' projection. Equivalently, we can ask for the likelihood that a 'random' line intersects both line segments.

Questions. Given a convex set Ω, which $\mathcal{L} \subset \Omega$ minimize $E(\mathcal{L})$?

A partial answer ('Quadratic Crofton and sets that see themselves as little as possible') is given by the following Theorem that shows that, at least for a positive proportion of lengths L the answer is fairly simple.

Theorem. *Let $\Omega \subset \mathbb{R}^n$ be a bounded, convex domain with C^1-boundary. There exists a constant c_Ω such that if*

$$0 \leq L - m|\partial\Omega| \leq c_\Omega \qquad \text{for some} \quad m \in \mathbb{N},$$

then among all $(n-1)-$dimensional piecewise differentiable $\Sigma \subset \Omega$ with surface area $\mathcal{H}^{n-1}(\Sigma) = L$ the energy

$$E(\Sigma) = \int_{\Sigma} \int_{\Sigma} \frac{|\langle n(x), y - x \rangle \langle y - x, n(y) \rangle|}{\|x - y\|^{n+1}} \, d\sigma(x) d\sigma(y)$$

is minimized by m copies of the boundary and a segment of a hyperplane.

One can also rephrase the question (see the referenced paper). By Crofton's formula, the expected number of intersections of a 'random' line with a set is only determined by the co-dimension 1 volume of the set (i.e. the length in $\mathbb{R}^2$).

Questions. Which set minimizes the variance of intersections?

One motivating image could be the following. Suppose we cover Earth with sensors to detect cosmic radiation: how should we arrange the sensors so that a random ray is roughly captured by the same number of sensors? The real-life case of this scenario is covered by the above Theorem: since sensors are expensive, we are in the setting where $0 < L \ll 1$ and one should put the sensors in a single hyperplane.

22. Opaque Sets

This is very much related to the previous question (in fact, it inspired work that led to the problems in the previous section). The problem goes back to a 1916 paper of Mazurkiewicz and a 1959 paper of Bagemihl. Already the special case of the unit square is interesting.

> What is the length of the shortest one-dimensional set such that each line intersecting $[0,1]^2$ also intersects the set?

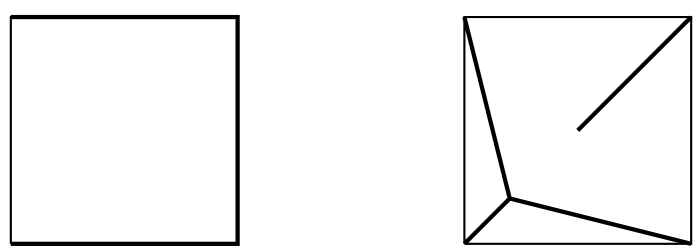


Figure 15. Left: three sides of the boundary give an opaque set with length 3. Right: the conjectured shortest opaque set for the unit square with length $\sqrt{2} + \sqrt{3/2} \sim 2.63$.

There is a general bound for convex $\Omega \subset \mathbb{R}^2$ going back to Jones (1962).

Theorem (Jones, 1962)**.** *If $\Omega \subset \mathbb{R}^2$ is convex, then any opaque set has length*

$$L \geq \frac{|\partial \Omega|}{2}.$$

It is known that this cannot be improved in general (take an extremely thin rectangle and use two short sides and one long sides). There are many proofs of this (often using variants of Crofton's Formula or, equivalently, averaging over projections). What is somewhat shocking is that, even for the unit square $[0,1]^2$, the best lower bounds aren't much better: the best bound is something like 2.00002[1]. It would be very nice to have some better bounds.

Update (Jan 2025)**.** I proved ('A Stability Version of the Jones Opaque Set Inequality') that if Jones' inequality is close to satisfied, then the corresponding

[1] Kawamura, Akitoshi; Moriyama, Sonoko; Otachi, Yota; Pach, Janos (2019), A lower bound on opaque sets, Computational Geometry, 80: 13-22

opaque set has to 'point in the same directions as the boundary $\partial\Omega$'. More concretely, if there was an opaque set for $[0,1]^2$ with length close to 2, then most of it would have to be nearly axis-parallel.

23. A Greedy Energy Sequence on the Unit Interval

This is a very curious phenomenon. Identify the one-dimensional Torus $\mathbb{T}$ with $\mathbb{T} \cong [0,1]$ and consider the function $f : \mathbb{T} \to \mathbb{R}$ given by

$$f(x) = x^2 - x + \frac{1}{6}.$$

(The full phenomenon seems to hold for much more general functions but this seems to be the easiest special case.) This function has a maximum in 0 and mean value 0. We can now consider sequences obtained in the following way

$$x_{n+1} = \arg\min_{x \in \mathbb{T}} \sum_{k=1}^{n} f(x - x_k).$$

What happens is that the arising sequence $(x_n)_{n=1}^{\infty}$ seems to be very regularly distributed in all the usual ways: for any subinterval $J \subset [0,1]$, we have

$$\#\left\{1 \leq i \leq N : x_i \in J\right\} \sim |J| \cdot N + \text{very small error term}.$$

There are many other ways of phrasing the phenomenon, for example it seems to be that

$$\sum_{k,\ell=1}^{N} f(x_k - x_\ell) \qquad \text{grows very slowly (logarithmically?) in } N.$$

We only know the much weaker bound

$$\sum_{k,\ell=1}^{N} f(x_k - x_\ell) \lesssim n.$$

Another observation is that

$$\left\| \sum_{k=1}^{n} f(x - x_k) \right\|_{L^\infty} \qquad \text{grows very slowly (logarithmically?) in } N.$$

The best known result is in a paper with Louis Brown (J. Complexity) that shows

$$\left\| \sum_{k=1}^{n} f(x - x_k) \right\|_{L^\infty} \lesssim n^{1/3} \qquad \text{for infinitely many } n.$$

We note that this is the sum of n functions of size ~ 1: for it to grow only logarithmically, a lot of cancellation has to take place. The function f has mean value 0, so cancellation implies that the x_k have to be somehow evenly spread. One could phrase many of these things in terms of exponential sum estimates which seem to be small, i.e.

$$\sum_{k=1}^{n} e^{2\pi i \ell x_k} \qquad \text{is relatively small.}$$

One explicit conjecture one could make is

$$\sum_{\ell=1}^{n} \frac{1}{\ell} \left| \sum_{k=1}^{n} e^{2\pi i \ell x_k} \right| \lesssim \log n$$

but I would be interested in anything that could be said. Extensions to higher dimensions or other domains would be very, very interesting. I first studied (Monatshefte Math. 2020) such sequences for the special case (long story, explained in the paper why)

$$f(x) = -\log|2\sin(\pi x)|.$$

Florian Pausinger then proved that when initialized on sets with exactly one element, then sequences of this type are always variations of the van der Corput sequence (Annali di Matematica Pura ed Applicata, 2020). I later realized that this function f can be much more elegantly phrased in the complex plane and that lead to nearly optimal results (arXiv, June 2020) for this particular function – but the proof is quite special and uses a number of tricks that are highly tailored to this particular function; the phenomenon seems to be much, much more robust. Louis Brown and I (J. Complexity, 2020) proved Wasserstein bounds that get really good in dimensions $d \geq 3$. But the one-dimensional problem seems to be hard and quite interesting.

24. The Kritzinger sequence

Ralph Kritzinger ('Uniformly distributed sequences generated by a greedy minimization of the L^2 discrepancy') defined the following sequence $(x_n)_{n=1}^{\infty}$. One starts with $x_1 = 1/2$ and then sets, in a greedy fashion,

$$x_{N+1} = \arg\min_{0 \leq x \leq 1} \qquad -2\sum_{n=1}^{N} \max\{x, x_n\} + (N+1)x^2 - x.$$

This seems maybe a bit arbitrary at first glance but arises naturally when trying to pick x_{N+1} in such a way that the L^2-distance between the empirical distribution and the uniform distribution is as small as possible (see the paper). What is particularly nice about this greedy sequence is that its consecutive elements are 'nice'

$$\frac{1}{2}, \frac{1}{4}, \frac{5}{6}, \frac{1}{8}, \frac{7}{10}, \frac{5}{12}, \frac{13}{14}, \dots$$

We observe that x_n can be written as $x_n = p/(2n)$ with p odd (additional cancellation may occur, so the denominator is always a divisor of $2n$). The sequence seems to be *very* regularly distributed in the sense that

$$\max_{0 \leq x \leq 1} |\#\{1 \leq i \leq N : x_i \leq x\} - Nx| \qquad \text{as a function of } N \text{ is very small.}$$

Kritzinger proves

$$\max_{0 \leq x \leq 1} |\#\{1 \leq i \leq N : x_i \leq x\} - Nx| \lesssim \sqrt{N}$$

but one could imagine the upper bound being as small as $\log N$. It doesn't seem to matter much whether $x_1 = 1/2$. In fact, even starting with an arbitrary initial set $\{x_1, \dots, x_m\} \subset [0, 1]$, one observes this high degree of regularity. Why?

Update (July 2022)**.** The Kritzinger sequence turns out to coincide with the sequence that one obtains when greedily minimizing the Wasserstein W_2 distance between the empirical measure and the Lebesgue measure on $[0, 1]$. Using some other ideas I was able to show (' On Combinatorial Properties of Greedy Wasserstein Minimization') that for infinitely many $N \in \mathbb{N}$

$$\max_{0 \leq x \leq 1} \left| \int_0^x \# \left\{ 1 \leq i \leq N : x_i \leq y \right\} - Ny \, dy \right| \lesssim N^{1/3}.$$

This in particular implies that the sequence is quite a bit more regular than iid random points (for which this quantity would be $\sim N^{1/2}$ with overwhelming likelihood).

25. A Special Property that some Lattices have?

This is a purely geometric problem that arose out of some calculus of variations considerations (see the paper). Consider the standard hexagonal lattice Λ in $\mathbb{R}^2$ and fix the density (say, the volume of each little triangle is 1). Let $r > 0$ be an arbitrary real number and consider

$$\Lambda_r = \left\{ x \in \Gamma : \|x\| = r \right\}.$$

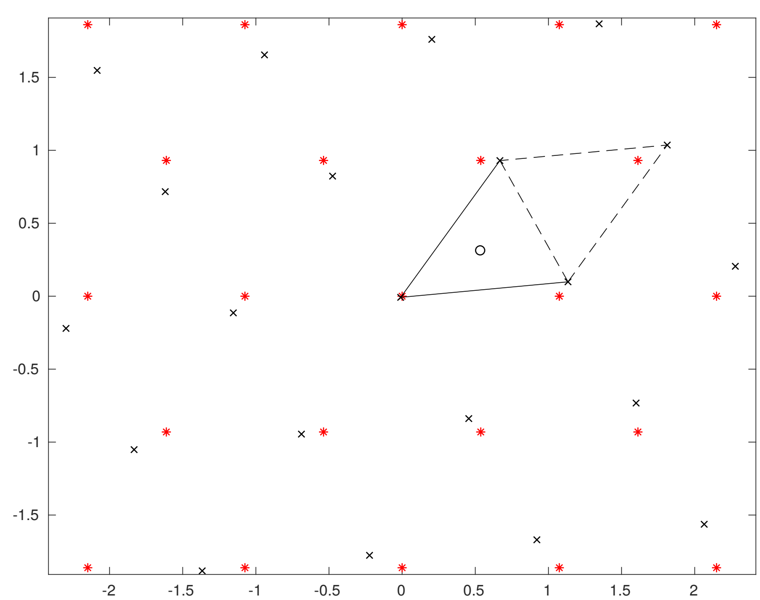


Figure 16. The Hexagonal Lattice and a slight perturbation.

We can now perturb the lattice a little bit: by this I mean that we perturb the basis vectors a tiny bit (but in such a way that the density, the volume of a fundamental cell, is preserved). This 'wiggling of the lattice' leads to a 'wiggling of the points' Λ_r (by this we mean exactly what it sounds like: each point in Λ_r has a basis representation $a_1 v_1 + a_2 v_2$ where v_1, v_2 are the basis vectors of the hexagonal lattice and we now consider $a_1 w_1 + a_2 w_2$ where w_1, w_2 are the slightly perturbed vectors). After wiggling the points in this way, some will move closer and some will move further away.

Theorem (Faulhuber & S, J. Stat. Phys). The sum of the distances increases under small perturbations.

I believe this to be quite a curious property: it shows, in a certain sense, 'points in the hexagonal lattice are, on average, closer to the origin than the points of any nearby lattice'. It seems a bit like optimal sphere packing but also like something else. I would believe that most of the lattices that are optimal w.r.t. sphere packing have this property but it's not clear to me whether there are others.

Question. Which other lattices have this property? Even in $\mathbb{R}^3$ this already seems tricky. What about D4 or E8? Leech?

Our proof for the hexagonal lattice is actually quite simple: the set Λ_r has a rotational symmetry by 120° so instead of studying Λ_r, it suffices to study a triple of points having this symmetry and then the computation becomes explicit. In principle this should also work for other lattices but one has to identify proper symmetries and then see whether one can do the computations.

Update (Dec 2022) Paige Helms ('Extremal Property of the Square Lattice', arXiv Dec 2022) has established a similar result for the square lattice $\mathbb{Z}^2$.

26. Roots of Classical Orthogonal Polynomials

Consider the differential equation $-(p(x)y')' + q(x)y' = \lambda y$, where $p(x)$ is a polynomial of degree at most 2 and $q(x)$ is a polynomial of degree at most 1. This setting includes the classical Jacobi polynomials, Hermite polynomials, Legendre polynomials, Chebychev polynomials and Laguerre polynomials.
In 1885, Stieltjes studied a special case, the Jacobi polynomials given by

$$(1-x^2)y''(x) - (\beta - \alpha - (\alpha+\beta+2)x)\, y'(x) = n(n+\alpha+\beta+1)y(x)$$

and proved that the solution, a polynomial of degree n, has the following nice interpretation: its roots are exactly the minimal energy configuration of

$$E = -\sum_{\substack{i,j=1 \\ i \neq j}}^{n} \log|x_i - x_j| - \sum_{i=1}^{n} \left(\frac{\alpha+1}{2} \log|x_i - 1| + \frac{\beta+1}{2} \log|x_i+1| \right).$$

Differentiating E, this results in an interesting relationship between the roots

$$\sum_{\substack{k=1 \\ k \neq i}}^{n} \frac{1}{x_k - x_i} = \frac{1}{2}\frac{\alpha+1}{x_i - 1} + \frac{1}{2}\frac{\beta+1}{x_i+1} \qquad \text{for all } 1 \leq i \leq n.$$

I managed to extend this result to all classical polynomials (Proc. AMS 2018).

Theorem. *Let $p(x), q(x)$ be polynomials of degree at most 2 and 1, respectively. Then the set $\{x_1, \dots, x_n\}$, assumed to be in the domain of definition, satisfies*

$$p(x_i) \sum_{\substack{k=1 \\ k \neq i}}^{n} \frac{2}{x_k - x_i} = q(x_i) - p'(x_i) \qquad \textit{for all} \quad 1 \leq i \leq n$$

if and only if

$$y(x) = \prod_{k=1}^{n} (x - x_k) \quad \textit{solves} \quad -(p(x)y')' + q(x)y' = \lambda y \quad \textit{for some } \lambda \in \mathbb{R}.$$

What's particularly interesting is that one can use this to define a system of ODEs for which the stationary state corresponds exactly to roots of classical orthogonal polynomials. More precisely, consider

$$\frac{d}{dt}x_i(t) = -p(x_i)\sum_{\substack{k=1\\k\neq i}}^{n}\frac{2}{x_k(t)-x_i(t)} + p'(x_i(t)) - q(x_i(t)) \qquad (\diamond)$$

We can then show that the underlying system of ODEs converges exponentially quickly to the true solution.

Theorem. *The system* $(\diamond)$ *converges for all initial values* $x_1(0) < \cdots < x_n(0)$ *to the zeros* $x_1 < \cdots < x_n$ *of the degree* n *polynomial solving the equation. Moreover,*

$$\max_{1\leq i\leq n}|x_i(t) - x_i| \leq ce^{-\sigma_n t},$$

where $c > 0$ *depends on everything and* $\sigma_n \geq \lambda_n - \lambda_{n-1}$.

This allows one to find roots of an orthogonal polynomial p_n by simply running an ODE. It is actually completely independent of p_{n-1} or p_{n+1}, there are no recurrence relations, no solution formulas, it's just an ODE.

> **Question.** Do analogous systems of ODEs exist for other types of orthogonal polynomials? Is it possible to get results in a similar spirit?

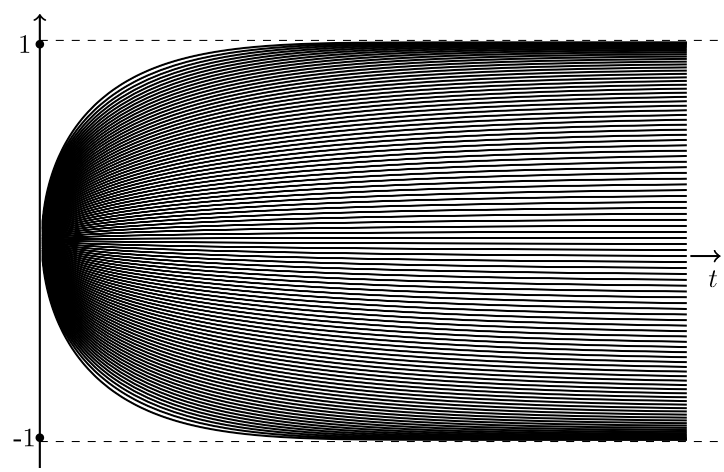


FIGURE 17. Evolution of the system of ODEs for $0 \leq t \leq 0.01$ approaches the zeros of the Legendre polynomial P_{100} in $(-1, 1)$.

27. An Estimate for Probability Distributions

This question seems quite elementary: it's really a question about real functions. Suppose we are given a probability distribution $f(x)dx$ on the *positive* real line $[0, \infty]$ and X, Y are independent random variables drawn from that distributions. We can try to analyze the event

$$\{X + Y \geq 2z\},$$

where z is some large parameter. There are two ways this event can happen: either one of the random variables is smaller than z (in which case the other one has to be bigger) or they are both bigger than z. A fascinating result of Feldheim & Feldheim (arXiv:1609.03004) says that

$$\limsup_{z\to\infty} \frac{\mathbb{P}(X+Y\geq 2z \quad \text{and} \quad \min(X,Y)\leq z)}{\mathbb{P}\left(X+Y\geq 2z \quad \text{and} \quad \min(X,Y)\geq z\right)} = \infty.$$

I would like to understand whether one can quantify how this result goes to infinity. Suppose we have a random variable that is not compactly supported (and maybe has a smooth density?)

- **Question 1.** Is there always a $z>0$ such that
$$\frac{\mathbb{P}(X+Y\geq 2z \quad \text{and} \quad \min(X,Y)\leq z)}{\mathbb{P}\left(X+Y\geq 2z \quad \text{and} \quad \min(X,Y)\geq z\right)} \geq \frac{(2\log 2)z}{\mathrm{med}(X)}?$$
- **Question 2.** Is there always a $z>0$ such that
$$\frac{\mathbb{P}(X+Y\geq 2z \quad \text{and} \quad \min(X,Y)\leq z)}{\mathbb{P}\left(X+Y\geq 2z \quad \text{and} \quad \min(X,Y)\geq z\right)} \geq \frac{2z}{\mathbb{E}X}?$$

The numbers are coming from assuming that exponential distributions are the worst case (they might not be). In case the constants are wrong: is the growth of the RHS linear in z? If that is wrong: what is it? Note that all these probabilites can be written explicitly as integrals over $f(x)f(y)dxdy$ over certain regions.

I was originally interested in whether the assumption of the random variable not having a compactly supported distribution is necessary. It turns out that it is: I proved (Stat. Prob. Lett.)

Theorem. *If X,Y are i.i.d. random variables drawn from an absolutely continuous probability distributions with density $f(x)dx$ on $\mathbb{R}_{\geq 0}$, then*

$$\sup_{z>0} \mathbb{P}\left(X\leq z \text{ and } X+Y\geq 2z\right) \geq \frac{1}{24+8\log_2\left(\mathrm{med}(X)\|f\|_{L^\infty}\right)},$$

where $\mathrm{med}\, X$ *denotes the median of the probability distribution. This estimate is sharp up to constants and the supremum can be restricted to* $0\leq z\leq \mathrm{med}(X)$.

It would be interesting to know whether it is possible to determine the sharp constants and the extremal distribution.

28. Hermite-Hadamard Inequalities

The Hermite-Hadamard inequality states that if $f:[a,b]\to\mathbb{R}$ is convex, then

$$\frac{1}{b-a}\int_a^b f(x)dx \leq \frac{f(a)+f(b)}{2}.$$

It is not difficult: a convex function stays below a line. However, once one goes to higher dimensions, things become extremely difficult. I proved (J. Geom. Anal, 2018) that if $\Omega\subset\mathbb{R}^n$ is convex and $f:\Omega\to\mathbb{R}$ is convex and positive on $\partial\Omega$, then

$$\frac{1}{|\Omega|}\int_\Omega f dx \leq \frac{c_n}{|\partial\Omega|}\int_{\partial\Omega} f d\sigma,$$

where c_n is a universal constant. It was then shown by Beck, Brandolini, Burdzy, Henrot, Langford, Larson, Smits and S (J. Geom. Anal.) that this inequality does indeed hold for subharmonic functions as well and that $c_n \leq 2n^{3/2}$. The sharp constant was obtained by Simon Larson (2020) who proved that $c_n = n$.

I proved in the original paper (J. Geom. Anal, 2018) that if $f : \Omega \to \mathbb{R}$ is merely subharmonic, i.e. $\Delta f \geq 0$, then we still have

$$\int_\Omega f dx \leq c_n |\Omega|^{1/n} \int_{\partial\Omega} f d\sigma.$$

Jianfeng Lu and I then proved (Proc. AMS 2020) that one can take $c_n = 1$. Jeremy Hoskins and I proved that (arxiv, Dec 2019) $c_2 < 1/\sqrt{2\pi} \sim 0.39\ldots$ and have obtained a candidate domain that leads to a constant of ~ 0.358. We believe that this is probably close to the best possible domain, it is shown in the Figure. It's currently not even known whether an extremal shape exists. Does it exist? And does the curvature of its boundary vanish at exactly one point?

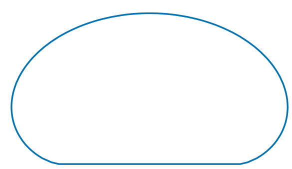


FIGURE 18. A candidate for the extremal shape in $n = 2$ dimensions.

Question. What can be said about the extremal domain? Does its curvature vanish in exactly one point?

There is also another interesting phenomenon: all these inequalities are proven for subharmonic functions. This is of course more general since every convex function is subharmonic but not vice versa. It is also clear from the characterization of these inequalities, that the extremal functions for the subharmonic Hermite-Hadamard inequalities are not going to be convex, they will merely be harmonic. So we would expect stronger statements in the case where the function f is convex.

Question. What are the optimal constants for the Hermite-Hadamard inequalities

$$\frac{1}{|\Omega|} \int_\Omega f dx \leq \frac{c_n}{|\partial\Omega|} \int_{\partial\Omega} f d\sigma$$

and

$$\int_\Omega f dx \leq c_n |\Omega|^{1/n} \int_{\partial\Omega} f d\sigma$$

when f is assumed to be convex (and $\Omega \subset \mathbb{R}^n$ is convex)?

We know, from the subharmonic case, that $c_n \leq n$ for the first inequality and $c_n \leq 1$ for the second inequality. But at this point even the growth/decay of these functions as a function of n is not clear when we restrict to convex functions. I mentioned in the original paper (J. Geom. Anal, 2018) that this problems seems to have some connection to an optimal transport problem where one transports the interior volume to the surface along lines in the most even way.

29. A Strange Inequality for the number of critical points

A while back I ran into a curious inequality – it sort of dropped out of other things I was doing ('Wasserstein Distance, Fourier Analysis …', 2018).

Theorem. *Let $f:\mathbb{T}\to\mathbb{R}$ be continuously differentiable with mean value* 0. *Then*

$$(\text{number of critical points of } f)\cdot \|f\|_{L^2(\mathbb{T})} \gtrsim \frac{\|f'\|^2_{L^1(\mathbb{T})}}{\|f'\|_{L^\infty(\mathbb{T})}}.$$

It says something interesting: if a function has large derivatives, then it is either big or it is has a lot of wiggles (= critical points). I always thought that this was a really curious kind of statement. I would like to understand this better. Note that there is a relatively easy inequality

$$(\text{number of critical points of } f)\cdot \|f\|_{L^\infty(\mathbb{T})} \gtrsim \|f'\|_{L^1(\mathbb{T})}.$$

Are there more such inequalities? Are they part of a family? I would be especially interested in higher-dimensional analogues.

30. An improved Isoperimetric Inequality?

The classical isoperimetric inequality in $\mathbb{R}^n$ says that a large set has a large boundary and, for $\Omega\subset\mathbb{R}^d$,

$$|\partial\Omega| \geq c\cdot|\Omega|^{\frac{d-1}{d}}$$

Let now $\Omega\subset\mathbb{R}^d$ be a bounded domain with smooth boundary and let $x\in\Omega$ be an arbitrary point in the domain. We define a subset $(\partial\Omega)_x\subseteq\partial\Omega$ via

$$(\partial\Omega)_x = \{y\in\partial\Omega : \text{the geodesic from } x \text{ to } y \text{ arrives non-tangentially}\}.$$

We note that the geodesic is defined as the shortest path $\gamma:[0,1]\to\Omega$ with $\gamma(0)=x$ and $\gamma(1)=y$. We say that it arrives non-tangentially if $\langle\gamma'(1),\nu\rangle\neq 0$, where ν is the normal vector of $\partial\Omega$ in y. Of course $(\partial\Omega)_x$ is a subset of the full boundary $\partial\Omega$. We were interested in whether this non-tangential boundary $(\partial\Omega)_x$ still obeys some form of isoperimetric principle.

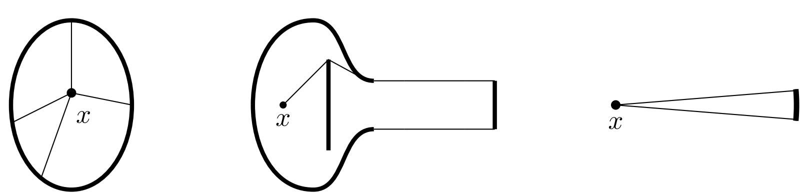


Figure 19. Various examples of $(\partial\Omega)_x$.

It is not terribly difficult to show (and was done in 'The Boundary of a Graph and its Isoperimetric Inequality', Jan 2022) that for *convex* domains $\Omega\subset\mathbb{R}^d$

$$\forall x\in\Omega \qquad |(\partial\Omega)_x| \geq (d-1)\frac{|\Omega|}{\operatorname{diam}(\Omega)}.$$

The constant $d-1$ cannot be optimal (but is optimal up to a factor of 2 in $d=2$). It seems natural to ask: what is the optimal constant c_d such that for *convex* $\Omega\subset\mathbb{R}^d$

$$\forall x\in\Omega \qquad |(\partial\Omega)_x| \geq c_d\frac{|\Omega|}{\operatorname{diam}(\Omega)}?$$

The other natural question is to ask what sort of conditions one needs on the domain Ω for this non-tangential isoperimetric principle to hold.

31. Geometric Probability

Let $\Omega \subset \mathbb{R}^d$ be a domain with finite volume. Suppose X, Y are two i.i.d. random variables that are uniformly distributed in Ω. It is clear by scaling that

$$\mathbb{E}\|X - Y\|_{\ell^2(\mathbb{R}^d)} \geq c_d |\Omega|^{1/d}.$$

An old result of Blaschke shows that the sharp constant c_d is given by the ball. This is perhaps not too surprising: two randomly chosen points from the ball are closer to each other than any other points (see also Bonnet, Gusakova, Thäle, Zaporozhets, arXiv 2020).

Problem. One would naturally expect an inequality of the flavor

$$\mathbb{V}\|X - Y\|_{\ell^2(\mathbb{R}^d)} \geq c_d |\Omega|^{2/d}.$$

Which domain minimizes the variance, which gives the smallest constant c_d? Is it even clear that the extremal domain is convex?

32. The Polarization Constant in $\mathbb{R}^n$

The following question is due to Benitez, Sarantopoulos & Tonge ('Lower bounds for norms of products of polynomials, Math. Proc. Cambridge Philos. Soc. 1998).

Problem. Let $\{v_1, \dots, v_n\} \subset \mathbb{R}^n$ be any set of n vectors all of which have length 1. Is

$$\max_{\|x\|=1} |\langle x, v_1 \rangle \cdot \langle x, v_2 \rangle \cdots \langle x, v_n \rangle| \geq n^{-n/2}?$$

The inequality, if true, would be sharp for orthonormal systems. Pinasco ('The $n-$th linear polarization constant...', arXiv Aug 2022) has proven the result for $n \leq 14$. It is also relatively easy to prove the result up to constants: by averaging over points on the sphere and linearity of expectation, we see that

$$\mathbb{E} \log \left(|\langle x, v_1 \rangle \cdot \langle x, v_2 \rangle \cdots \langle x, v_n \rangle| \right) = n \int_{\mathbb{S}^{n-1}} \log |\langle x, v_1 \rangle| \, dx.$$

This integral looks unpleasant but we know that, for sufficiently high-dimensions, the inner product $|\left\langle x, v \right\rangle|$ is distributed like a Gaussian with variance $1/n$. Thus we expect that

$$\int_{\mathbb{S}^{n-1}} \log |\langle x, v_1 \rangle| \, dx \sim \int_{\mathbb{R}} \frac{\sqrt{n} e^{-\frac{nx^2}{2}}}{\sqrt{2\pi}} \log |x| dx.$$

The integral can be evaluated

$$\int_{\mathbb{R}} \frac{\sqrt{n} e^{-\frac{nx^2}{2}}}{\sqrt{2\pi}} \log |x| dx = -\frac{\gamma + \log(2)}{2} - \frac{1}{2} \log n,$$

where $\gamma \sim 0.577\dots$ is the Euler-Mascheroni constant. Thus

$$\max_{\|x\|=1} |\langle x, v_1 \rangle \cdot \langle x, v_2 \rangle \cdots \langle x, v_n \rangle| \geq 0.52^n \cdot n^{-n/2}.$$

Note that several arguments leading to better constants are known (see Pinasco).

33. Bad Science Matrices

Suppose $A \in \mathbb{R}^{n \times n}$ has rows that are normalized in $\ell^2(\mathbb{R}^n)$. How large can

$$\beta(A) = \frac{1}{2^n} \sum_{x \in \{-1,1\}^n} \|Ax\|_{\ell^\infty} \qquad \text{be?}$$

This has a number of different interpretations: as an affine map that sends the unit cube $\{-1,1\}^n$ to points with, typically, at least one large coordinate. It can also be understood as a series of n statistical tests that, while individually fair, are likely to have at least one give a small $p-$value when tested on a sequence of fair flips of a coin. Various other geometric interpretations are conceivable. I proved ('Bad Science Matrices') that

$$\max_{A \in \mathbb{R}^{n \times n}} \frac{1}{2^n} \sum_{x \in \{-1,1\}^n} \|Ax\|_{\ell^\infty} = (1 + o(1)) \cdot \sqrt{2 \log n}.$$

The problem is that

(1) the $o(1)$ term is too big: the optimal rate is attained for random $\pm 1/\sqrt{n}$ matrices but the extremizers do not seem to behave like that at all, they seem to be very structured and very much not random
(2) needless to say, it would be nice to refine the $o(1)$ term but this might be extremely complicated; maybe it can be done when n is a power of 2 via some clever recursive construction
(3) it is not at all clear how to find candidate extremizers and, even if they are identified, it is not clear how to prove that they are extremal

The best bounds I currently know in dimensions $2 \le n \le 8$ are as follows.

n	2	3	4	5	6	7	8
$\beta_n \geq$	1.41	1.57	1.73	1.79	1.86	1.93	2

Figure 20. Lower bounds for $\beta_n = \max \beta(A)$ when $n \le 8$

To illustrate the difficulty, the best example for $n = 5$ I currently know is

$$A = \frac{1}{2\sqrt{3}} \begin{pmatrix} 2 & 2 & 0 & 0 & 2 \\ -2 & 2 & 0 & 2 & 0 \\ -2 & 0 & 0 & -2 & 2 \\ 0 & -\sqrt{3} & \sqrt{3} & \sqrt{3} & \sqrt{3} \\ 0 & \sqrt{3} & \sqrt{3} & -\sqrt{3} & -\sqrt{3} \end{pmatrix}$$

which shows that

$$\max_{A \in \mathbb{R}^{5 \times 5}} \beta(A) \geq \frac{2 + 3\sqrt{3}}{4} \sim 1.799 \ldots.$$

It appears magnificently structured but it's maybe less clear what the structure is.

Update (June 2025). Albors-Bhatti-Ganjoo-Guo-Kunisky-Mukherjee-Stepin-Zeng (arXiv:2408.00933) give several different explicit examples of matrices that reach around 18% of the extremal value.

34. Sums of Square Roots Close to an Integer

This problem is somewhere between Analysis, Fourier Analysis and Number Theory. It originates from, it seems, Computer Graphics (comparing distances in Euclidean spaces) and is a special case of a well-known problem in complexity theory.

If we have the sum of k square roots of integers $1 \leq a_i \leq n$, how close can this be to an integer? For example

$$\sqrt{3} + \sqrt{20} + \sqrt{23} \sim 11 + 0.0000182859.$$

Note that if all the integers are squares themselves, then this sum of square roots is an integer and we are not interested in that case, we are interested in 'near-misses'.

> **Problem.** Fix $k \in \mathbb{N}$ and use $\|x\|$ to denote the distance between $x \in \mathbb{R}$ and the nearest integer. How does
> $$\min\left\{\|\sqrt{a_1} + \cdots + \sqrt{a_k}\| : 1 \leq a_1, \dots, a_k \leq n, \sqrt{a_1} + \cdots + \sqrt{a_k} \notin \mathbb{N}\right\}$$
> scale as a function of n?

This case $k = 1$ is pretty easy: if n is not a square number, then

$$\|\sqrt{n}\| \gtrsim \frac{1}{n^{1/2}}.$$

The best case is clearly when n is merely 1 removed from a square number and then this follows simply from Taylor expansion. The case $k = 2$ was considered by Angluin-Eisenstat.

Theorem. *If $1 \leq a, b \leq n$ are integers and $\sqrt{a} + \sqrt{b} \notin \mathbb{N}$ then*

$$\|\sqrt{a} + \sqrt{b}\| \gtrsim \frac{1}{n^{3/2}}.$$

This can be seen by squaring $\sqrt{a} + \sqrt{b}$: if the number is extremely close to an integer, then so is its square. The square is $a + b + \sqrt{4ab}$ and we know from $k = 1$ how close $\sqrt{4ab}$ can be to an integer. This also gives us optimal cases: we want to pick a and b so that $4ab$ is very close to a square number.

$k = 3$ is the first open case. We know that if $1 \leq a, b, c \leq n$ are integers and $\sqrt{a} + \sqrt{b} + \sqrt{c} \notin \mathbb{N}$ then

$$\|\sqrt{a} + \sqrt{b} + \sqrt{c}\| \gtrsim \frac{1}{n^{7/2}}$$

with the following nice argument that we learned from Arturas Dubickas and Roger Heath-Brown (independently): multiplying over all 8 possible choices of signs

$$\prod \left(\pm\sqrt{a} \pm \sqrt{b} \pm \sqrt{c} \right) \in \mathbb{N}$$

and since all 8 products are of size $\sim \sqrt{n}$, none of these expressions can be closer than $n^{-7/2}$ to an integer without being one. Conversely, we have the following nice example due to Nick Marshall

$$\|\sqrt{(k-1)^2 + 2} + \sqrt{(k+1)^2 + 2} + \sqrt{(2k)^2 - 8}\| \sim \frac{4}{k^5}$$

which shows that

$$\frac{1}{n^{7/2}} \lesssim \|\sqrt{a} + \sqrt{b} + \sqrt{c}\| \lesssim \frac{1}{n^{5/2}}.$$

A natural guess would be n^{-3}: the numbers $\sqrt{a} \mod 1$ are pretty rigid but if we sum 3 of them, maybe there is enough randomness that they start behaving like iid random variables would?

I proved ('Sums of square roots that are close to an integer') that, when k gets large, there are examples with

$$0 <\lesssim \|\sqrt{a_1} + \cdots + \sqrt{a_k}\| \lesssim n^{-c \cdot k^{1/3}}.$$

Surely one would expect the true rate to be closer to n^{-k}.

Update (April 2024). Siddharth Iyer (Distribution of sums of square roots modulo 1) used an entirely different approach to show $\lesssim n^{-k/2}$.

Update (June 2024). Additional results were obtained by Arturas Dubickas (J. Complexity, 2024).

35. The Bambah-Chowla Theorem

Bambah-Chowla[2] proved that every interval

$$(n, n + \sqrt{8}\, n^{1/4} + 1)$$

contains an integer that can be written as the sum of two squares. It is quite clear that $n^{1/4}$ is way too large, the truth is probably $\log n$ (worst case $\log^c n$). Moreover, the Bamba-Chowla Theorem has a one line proof

$$n \leq \left\lfloor \sqrt{n} \right\rfloor^2 + \left\lceil \sqrt{n - \left\lfloor \sqrt{n} \right\rfloor^2} \right\rceil^2 \leq n + \sqrt{8} n^{1/4} + 1.$$

Nonetheless, it has never been significantly improved, the constant $\sqrt{8}$ is unchanged.[3]. Geometrically, this would mean that there exists a very wide annulus in $\mathbb{R}^2$ that is centered in the origin and avoids all lattice points.

36. A Beurling estimate for GFA?

Suppose we are given n disks of radius 1 in the Euclidean plane $\mathbb{R}^2$ which are arranged in such a way that the disks are all disjoint but touching in the sense that they form one connected component (see the picture below for two examples). We use $x_1, \ldots, x_n \in \mathbb{R}^2$ to denote the centers of these disks, let $\alpha > 0$ be some arbitrary real parameter and define the notion of energy $E : \mathbb{R}^2 \setminus \{x_1, \ldots x_n\} \to \mathbb{R}_{\geq 0}$

$$E(x) = \sum_{k=1}^{n} \frac{1}{\|x - x_k\|^{\alpha}}.$$

We now consider the following simple randomized procedure.

(1) Suppose $x_1, \ldots, x_n$ have their center of mass in the origin $(0, 0)$. (It's not important that it's exactly the origin, just fixing translation invariance).
(2) Pick $r \gg e^{e^n}$ to be an extremely large number.

[2] R. P. Bambah and S. Chowla, On numbers which can be expressed as a sum of two squares, Proc. Nat. Inst. Sci. India 13 (1947), 101–103.

[3] For improvement in the constant +1, see G. J. O. Jameson, More on the gaps between sums of two squares. Math. Gaz. 103 (2019).

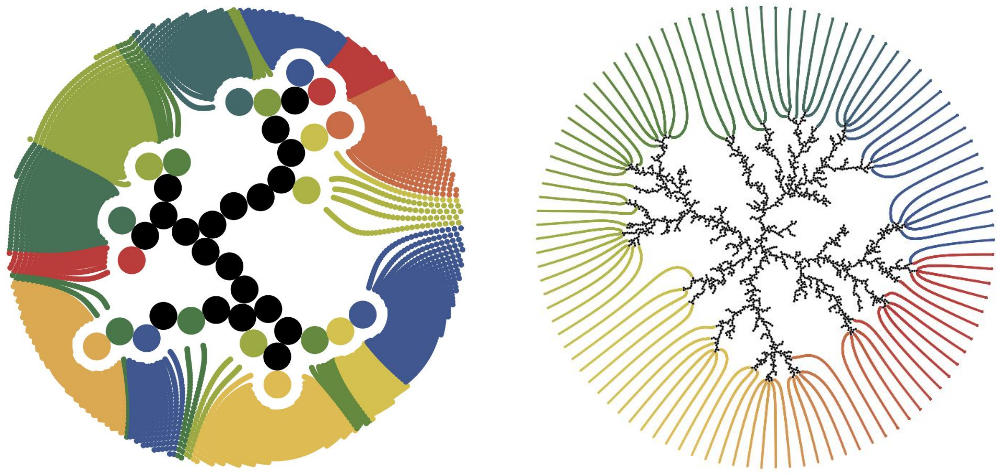


FIGURE 21. Left: a small number of disks colored by the color of the incoming gradient flows. Black disks are not hit by any gradient flows. More exposed disks are more likely to be hit. Right: several thousand disks, 250 incoming gradient flows (equispaced in angle).

(3) Pick, uniformly at random, a point x_0 on the circle $\left\{x \in \mathbb{R}^2 : \|x\| = r\right\}$.
(4) Consider the gradient *ascent* started in x_0 with respect to the energy E. This gives rise to a curve γ that starts in $\gamma(0) = x_0$ and satisfies

$$\dot{\gamma}(t) = (\nabla E)(\gamma(t)).$$

(5) We stop the curve once it is within distance 2 of one of the existing points $x_1, \dots, x_n \in \mathbb{R}^2$. Once the curve stops, we are at distance 2 of exactly one existing disk (with likelihood 1).

The main problem is now the following

> **Question.** The 'most popular' disk, the one most likely to be hit, how popular is it. What is the best possible upper bound on
>
> $$\max_{1 \le j \le n} \mathbb{P}\left(\text{gradient flow hits } x_j\right)$$
>
> in terms of n alone?

The question is a question of scaling in n, not about precise constants. It is kind of clear what one would probably expect: the worst case should be when the n disks are all arranged on a line and, in that case, the two endpoints are the most exposed and should be the ones most likely to be hit. How likely that is will then depend on α: the larger α, the more likely they are to be hit.

Motivation. Such a result could be understood as an analogue of Beurling's estimate for Brownian motion. I proved ('Random Growth via Gradient Flow

Aggregation') that when $0 < \alpha < 1$, then we have

$$\max_{1 \le j \le n} \mathbb{P}\left(\text{gradient flow hits } x_j\right) \le n^{\frac{\alpha-1}{2\alpha+2}}.$$

This estimate is nearly optimal when α is close to 0 but surely not optimal in general. It also gives no nontrivial results when $\alpha \ge 1$ but one would expect n^{-c_α} for some $c_\alpha > 0$ and all $0 < \alpha < \infty$. When $\alpha \to 0$, the optimal bound has to be close to $n^{-1/2}$ and when $\alpha \to \infty$, there is no nontrivial bound, one only gets $\lesssim 1$. Such Beurling estimates would, in turn, give rise to improved diameter estimate for the Gradient Flow Aggregation process, see the paper for more details.

Part 3. Fourier Analysis

37. An Exponential Sum for Sequences of Reals

This question is motivated by an inequality I proved for sequences exhibiting Poissonian Pair Correlation (J. Number Theory, 2020). The special role that $\sqrt{n} \mod 1$ plays in these types of gap statistics suggests that for $x_n = \sqrt{n}$, uniformly in N,

$$\sum_{k=1}^{N} \left| \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i k x_n} \right|^2 \lesssim 1.$$

Is this true? If so, then this would be best possible. Is it possible to describe other sequences having this property? Such sequences are candidates for having interesting gap statistics.

38. The Bourgain-Clozel-Kahane Root Uncertainty Principle

Let $f : \mathbb{R} \to \mathbb{R}$ be in $L^1(\mathbb{R})$ and even. We define

$$A(f) := \inf \{r > 0 : f(x) \geq 0 \text{ if } |x| > r\}$$
$$A(\widehat{f}) := \inf \{r > 0 : \widehat{f}(y) \geq 0 \text{ if } |y| > r\}.$$

We have

Theorem (Bourgain, Clozel & Kahane)**.** *Let $f : \mathbb{R} \to \mathbb{R}$ be a nonzero, integrable, even function such that $f(0) \leq 0$, $\widehat{f} \in L^1(\mathbb{R})$ and $\widehat{f}(0) \leq 0$. Then*

$$A(f)A(\widehat{f}) \geq 0.1687,$$

and 0.1687 *cannot be replaced by* 0.41*.*

Felipe Goncalves, Diogo Oliveira e Silva and I improved this lower bound to 0.2025 and showed that it cannot be replaced by 0.353. What's really quite interesting is that the extremal function has to have infinitely many double roots. It would be nice to understand how it behaves.

There are now several papers concerned with questions of this type. One that is especially worth emphasizing is a very nice result by Cohn - Goncalves ('An optimal uncertainty principle in twelve dimensions via modular forms'). They prove that the optimal constant is $\sqrt{2}$ in 12 dimensions.

39. An Uncertainty Principle

This question is motivated by a basic question: when averaging a function f by convolving with a function u (resulting in the 'averaged function' $u * f$), what function u should one consider? The question is intentionally vague and I would be interested in good axiomatic results ('the 'smoothest' average should satisfy properties $P_1, P_2, \dots$ and the only functions satisfying all these properties are ...').

One such axiomatic approach resulted in a really interesting uncertainty principle ('Scale Space...', arXiv, May 2020). It says that for $\alpha > 0$ and $\beta > n/2$, there exists $c_{\alpha,\beta,n} > 0$ such that for all $u \in L^1(\mathbb{R}^n)$

$$\| |\xi|^\beta \cdot \widehat{u} \|_{L^\infty(\mathbb{R}^n)}^\alpha \cdot \| |x|^\alpha \cdot u \|_{L^1(\mathbb{R}^n)}^\beta \geq c_{\alpha,\beta,n} \|u\|_{L^1(\mathbb{R}^n)}^{\alpha+\beta}.$$

These inequalities arise naturally when looking for the 'best' or 'smoothest' convolution kernel. I would be interested in what can be said about the extremizers.

> **Question.** What can be said about the extremizers of this functional? One interesting question would be whether the extremizer 'exploits' the $L^\infty-$bound fully and assumes it infinitely many times such that $|\widehat{u}(\xi)| \sim |\xi|^{-\beta}$. This would imply that the extremizer is not smooth.

When $n = 1$ and $\beta = 1$, then for many values of α, the characteristic function centered at the origin seems to play a special role. When $n = 1$ and $\beta = 2$, then $u(x) = 1 - |x|$ seems to play a special role (up to symmetries). It is not clear to me whether they are global extremizers but it seems conceivable.

Discrete versions of these statements on $\mathbb{Z}$ have been proven in joint work with Noah Kravitz (arXiv, July 2020). More precisely, we showed the following: suppose $u : \{-n, \dots, n\} \to \mathbb{R}$ is a symmetric function normalized to $\sum_{k=-n}^{n} u(k) = 1$. We show that every convolution operator is not-too-smooth, in the sense that

$$\sup_{f \in \ell^2(\mathbb{Z})} \frac{\|\nabla (f * u)\|_{\ell^2(\mathbb{Z})}}{\|f\|_{\ell^2}} \geq \frac{2}{2n+1},$$

and we show that equality holds if and only if u is constant on the interval $\{-n, \dots, n\}$. In the setting where smoothness is measured by the ℓ^2-norm of the discrete second derivative and we further restrict our attention to functions u with nonnegative Fourier transform, we establish the inequality

$$\sup_{f \in \ell^2(\mathbb{Z})} \frac{\|\Delta (f * u)\|_{\ell^2(\mathbb{Z})}}{\|f\|_{\ell^2(\mathbb{Z})}} \geq \frac{4}{(n+1)^2},$$

with equality if and only if u is the triangle function $u(k) = (n + 1 - |k|)/(n + 1)^2$. It would be interesting to have variants of this type of statements for other ways of measuring smoothness, other L^p-spaces.... – this seems to be quite interesting and quite unexplored!

I would also be quite interested in what can be said about the optimal function u when restricted to functions $u : [-\infty, 0] \to \mathbb{R}$. This would have practical applications: when smoothing some real numbers (say, the stock prize or the current temperature) we cannot look into the future. Thus the average has to be taken with respect to the past (see also S & Tsyvinski 'On Vickrey's Income Averaging').

Update (Feb 2024)**.** Sean Richardson ('A Sharp Fourier Inequality and the Epanechnikov Kernel') solved the problem for the second derivatives *without* the assumption that the kernel have a non-negative Fourier transform. The extremal kernel is explicit and very close in shape to a parabola $\max\left\{1 - |x|^2, 0\right\}$.

40. Littlewood's Cosine Root Problem

Let $A \subset \mathbb{N}$. How many roots does the function

$$f(x) = \sum_{k \in A} \cos(kx) \qquad \text{necessarily have on } [0, 2\pi]?$$

Littlewood originally conjectured that such a function should have $\sim |A|$ roots which is now known to be false (Borwein, Erdelyi, Ferguson, Lockhart, Annals). The best unconditional lower bound is due to Sahasrabudhe (Advances, 2016) which shows that

$$\text{number of roots} \gtrsim (\log \log \log |A|)^{1/2-}.$$

Erdelyi (2017) improved the $1/2-$ to $1-$. Surely it must be much bigger than that!

A warning example can be found in the paper of Sahasrabudhe: the trigonometric polynomial

$$2 \cos \theta + \sum_{r=2}^{2n} \sin \left(\frac{r\pi}{2} \right) \cos (rx)$$

has only 2 roots and all coefficients in $\{0, -1, 1, 2\}$. So it's not enough to work with the fact that the coefficients are small in the sense of having a small absolute value, it is actually important that they are in $\{0, 1\}$ which naturally restricts the number of approaches that one could try.

Update (July 2025)**.** Bedert (arXiv:2407.16075) removed a logarithm in the lower bound.

41. The 'Complexity' of the Hardy-Littlewood Maximal Function

Given a function $f : \mathbb{R} \to \mathbb{R}$, the Hardy-Littlewood maximal function is defined via

$$(\mathcal{M}f)(x) = \sup_{r>0} \frac{1}{2r} \int_{x-r}^{x+r} |f(z)| dz.$$

This is fairly classical object. The following object is less classical: define, for a given function $f : \mathbb{R} \to \mathbb{R}$ and a given $x \in \mathbb{R}$,

$$r_f(x) = \inf_{r>0} \left\{ \frac{1}{2r} \int_{x-r}^{x+r} f(z) dz = \sup_{s>0} \frac{1}{2s} \int_{x-s}^{x+s} f(z) dz \right\}.$$

So $r_f(x)$ is simply the shortest interval such that the average of f over that interval is the same as the largest possible average.

Vague Problem. r_f should assume many different values.

I proved (Studia Math, 2015) that if f is periodic and r_f assumes only two values $\{0, \gamma\}$ and r_{-f} also only assumes the same two values $\{0, \gamma\}$, then

$$f(x) = a + b \sin (cx + d)$$

and c is determined by γ. The proof requires transcendental number theory (the Lindemann-Weierstrass theorem), I always thought that was strange. Maybe we even have:

Conjecture. If $f \in L^{\infty}(\mathbb{R})$ and r_f assumes only finitely many values, then

$$f(x) = a + b \sin (cx + d).$$

Motivated by some heuristics (see paper), maybe we also have

Conjecture. Suppose $f : \mathbb{R} \to \mathbb{R}$ is C^1 and satisfies

$$f'(x+1) - f(x+1) = -f'(x-1) - f(x-1) \qquad \text{whenever } f(x) < 0.$$

Then

$$f(x) = a + b \sin(cx + d) \qquad \text{for some } a, b, c, d \in \mathbb{R}.$$

In general, it would be nice to have a better understanding of r_f and how it depends on f. Can $r_f(\mathbb{R})$ assume infinitely many values while not containing an interval?

42. A Compactness Problem

This is a phenomenon that I find really interesting: it should exist in many settings but I only know how to prove it on $\mathbb{T}^2$. Let $f \in C^\infty(\mathbb{T}^2)$ have mean value 0. Consider the problem of maximizing the average value of f over a closed geodesic (straight periodic lines). This means we are interested in

$$\sup_{\gamma \text{ closed geodesic}} \frac{1}{|\gamma|} \left| \int_\gamma f \, d\mathcal{H}^1 \right|,$$

where γ ranges over all closed geodesics $\gamma : \mathbb{S}^1 \to \mathbb{T}^2$ and $|\gamma|$ denotes their length.

The idea is that such an extremal geodesic somehow cannot be very long unless the function oscillates a lot. If the function is very nice and smooth, then that supremum should be attained by a relatively short geodesic.

Theorem (S, Bull. Aust. Math. Soc., 2019)**.** *Let $f : \mathbb{T}^2 \to \mathbb{R}$ be at least $s \geq 2$ times differentiable and have mean value 0. Then*

$$\sup_{\gamma \text{ closed geodesic}} \frac{1}{|\gamma|} \left| \int_\gamma f \, d\mathcal{H}^1 \right|,$$

is assumed for a closed geodesic $\gamma : \mathbb{S}^1 \to \mathbb{T}^2$ of length no more than

$$|\gamma|^s \lesssim_s \left(\max_{|\alpha|=s} \|\partial_\alpha f\|_{L^1(\mathbb{T}^2)} \right) \|\nabla f\|_{L^2} \|f\|_{L^2}^{-2}.$$

I always though this was a really interesting result. I would expect that it's not quite optimal (there should be a loss of derivatives on the right-hand side). I would also expect that there are analogous results on higher-dimensional tori $\mathbb{T}^d$. I would in fact expect that such results actually exist in a wide variety of settings: a natural starting point might be a setting where geodesics and Fourier Analysis work well together.

> **Question** What is the sharp form in $\mathbb{T}^2$? Is it possible to prove analogous results on $\mathbb{T}^d$ or in other settings? What is the correct formulation of this underlying phenomenon without geodesics?

It's not clear to me how to phrase this problem in a setting where geodesics don't make sense. What's a proper way to encode this principle in Euclidean space?

43. Some Line Integrals

This question is motivated by a result that Felipe Goncalves, Diogo Oliveira e Silva and I proved (Journal of Spectral Theory). In particular, any progress on this particular problem would lead to some refined statement about the $n-$point correlation of eigenfunctions of Schrödinger operators. The problem itself is completely elementary. Let $\mathbb{T}^d \cong [0,1]^d$ be the standard $d-$dimensional Torus and define the function

$$f_d(x) = \text{sign}\left(\prod_{k=1}^{d} \cos(2\pi x_k)\right),$$

where sign is simply the sign of the real number (with $\text{sign}(0) = 0$). This is simply a nice function that assumes values in $\{-1, 0, 1\}$ in a checkerbox pattern. Here's the question: let $a, b \in \mathbb{R}^d$ and let

$$\gamma(t) = at + b \qquad \text{mod } 1.$$

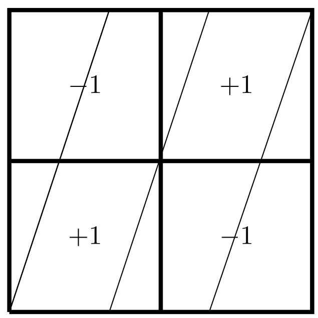


FIGURE 22. The sign of $\sin(x)\sin(y)$ for $(x,y) \in \mathbb{T}^2$ and a closed geodesic that spends significantly more time in the positive region than in the negative region. The flow $\gamma(t) = (t,t)$ would spend even more time in the positive region but that is not allowed: the coefficients have to be different.

We will also assume that all the entries of the vector a are distinct. These linear flows γ can be periodic or not periodic. We only care about the ones that are periodic, this means that $aL_\gamma \in \mathbb{Z}^d$ for some minimal $0 < L_\gamma \in \mathbb{R}$ in which case this linear flow has length $L_\gamma \|a\|$. What can be said about

$$\frac{1}{L_\gamma \|a\|} \int_{\text{one period}} f_d(\gamma(t)) dt.$$

Typically it will be close to 0. What is the largest value it can assume? For $d = 2$ we solve the problem explicitly and find some very short geodesic that is the unique maximizer. As $d \geq 3$, the techniques from our paper might still apply but it seems more challenging to get good values.

Update (Dec 2022). Dou, Goh, Liu, Legate, Pettigrew ('Cosine Sign Correlation', JFAA) proved the following result: if a_1, a_2, a_3 are three different positive integers and if $x \sim U[0, 2\pi]$ is a uniformly distributed random variables then the likelihood that $\cos(a_1 x), \cos(a_2 x), \cos(a_3 x)$ are all positive or all negative is $\geq 1/9$ with equality if and only if $\{a_1, a_2, a_3\} = \{m, 3m, 9m\}$ for some $m \in \mathbb{N}$. They conjecture that for 4 different numbers the extremal set $\{1, 3, 11, 33\}$ (and multiples thereof).

Update (June 2025). Regarding the Cosine Sign Correlation problem, the question how large the set in $[0, 2\pi]$ has to be where

$$\cos(a_1 x), \cos(a_2 x), \cos(a_3 x), \dots, \cos(a_n x) \qquad \text{all have the same sign,}$$

the argument from above Dou, Goh, Liu, Legate, Pettigrew ('Cosine Sign Correlation', JFAA) shows that it can be as small as $1/3^{n-1}$. Lars Becker came up with the following neat argument showing that it is always at least $1/4^n$. The measure of the set where all cosines are positive can be written as

$$\int_0^1 \prod_{k=1}^{n} 1_{[-1/4,1/4]}(a_k x).$$

Suppose that $\phi : \mathbb{T} \to \mathbb{R}$ satisfies $\widehat{\phi} \geq 0$ and $\phi \leq 1_{[-1/4,1/4]}$. Then, taking a Fourier transform,

$$\int_0^1 \prod_{k=1}^{n} 1_{[-1/4,1/4]}(a_k x) \geq \int_0^1 \prod_{k=1}^{n} \phi(a_k x) \geq \phi(0)^n.$$

The optimal choice of ϕ is given by

$$\phi = 4 \cdot 1_{[-1/8,1/8]} * 1_{[-1/8,1/8]}$$

and that gives a lower bound of $1/4^n$. So the answer is somewhere between $1/3^n$ and $1/4^n$.

44. A Cube in $\mathbb{R}^n$ hitting a lattice point

This problem is from a paper of Henk & Tsintsifas ('Lattice Point Coverings').

> **Problem.** Is there a universal constant c (independent of everything) such that each cube $Q \subset \mathbb{R}^n$ of side length c (possibly translated away from the origin and rotated) always intersects $\mathbb{Z}^d$?

It is clear that there exists such a constant c_n for each dimension and a result of Banaszczyk implies $c_n \lesssim \sqrt{\log n}$. But maybe there exists a uniform constant? (This can be understood as a relaxation of the Komlos conjecture).

There is a somewhat dual question: given a cube $Q \subset [0,1]^n$ whose center is in $0 \in \mathbb{R}^n$, can the cube be rotated in such a way so as to capture a lot more lattice points than predicted by its volume? For the sake of concreteness, we ask

> **Problem.** For which dimensions $n \in \mathbb{N}$ (if any) is it possible to rotate the cube centered at the origin with sidelength 1000 so that it contains 1001^n lattice points?

45. A kind of uncertainty principle?

Let $A \subset \mathbb{N}$ be a finite set of integers. What is the smallest $r > 0$ such that there exists a function $f : \mathbb{T} \to \mathbb{R}$ with

(1) $\operatorname{supp} f \subseteq (-r, r)$
(2) $f(x) \geq 0$ but is not identically 0 and $\int_{\mathbb{T}} f(x)dx > 0$
(3) for all $a \in A$ we have

$$\widehat{f}(a) = 0.$$

In words, what is the smallest possible support of a probability measure centered at 0 that vanishes on a prescribed set of integers? There is an easy lower bound

$$r \gtrsim \frac{1}{\min_{a \in A} a}.$$

One would be inclined to think if A is some lacunary set, say $A = \{2^n : 1 \leq n \leq 1000\}$, then probably this simple lower bound is sharp. But it is clear that if we have, say, $A = \{n, n+1, \dots, 2n\}$, then one would expect that the support has to be quite a bit bigger than $1/n$. Is there an estimate like

$$r \gtrsim \left(\sum_{a \in A} \frac{1}{a^2} \right)^{1/2}$$

or something along these lines?

Motivation. This question is motivated by the following particularly simple proof of a fascinating result of Kozma-Oravecz (see also 'Local sign changes of polynomials'): the function

$$g(x) = \sum_{a \in A} c_a \cos(ax) \qquad \text{has a root in each interval of length} \qquad \sum_{a \in A} \frac{1}{a}.$$

It is clear that if $A = \{b\}$, then there exists such a function f supported on an interval of length b^{-1}. Then the convolution is of the form

$$(f * g)(x) = \sum_{\substack{a \in A \\ a \neq b}} d_a \cos(ax)$$

and the result follows by induction. If one had a better understanding of the question above, one could try to remove more than 1 frequency at a time which might lead to stronger results.

It is clear that the question is equally interesting in higher dimension $\mathbb{T}^d$ or on the sphere $\mathbb{S}^d$ with trigonometric polynomials replaced by spherical harmonics.

Part 4. **Partial Differential Equations**

46. The Hot Spots Conjecture

Let $\Omega \subset \mathbb{R}^2$ be convex (or maybe only simply connected). Let u_2 be the smallest nontrivial eigenfunction of the Neumann Laplacian, i.e.

$$\begin{cases} -\Delta u_2 & = \mu_2 u_2 \qquad \text{in } \Omega \\ \frac{\partial}{\partial n} u_2 & = 0 \qquad\quad \text{on } \partial\Omega. \end{cases}$$

Are maximum and minimum assumed at the boundary? This famous conjecture of J. Rauch has inspired a lot of work. I proved (Comm. PDE, 2020), that if Ω is a convex domain of dimension $N \times 1$, then maximum and minimum are at most distance ~ 1 from a pair of points whose distance is the diameter of Ω. This is the optimal form of this statement (think of a rectangle), I always wondered whether the argument could possibly be sharpened to say more about Hot Spots.

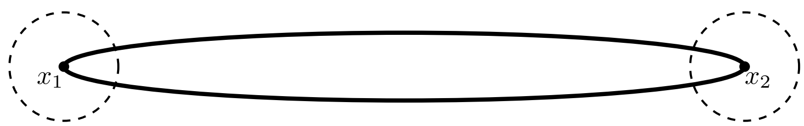


Figure 23. Maximum and minimum are attained close (at most a universal multiple of the inradius away) to the points achieving maximal distance (the 'tips' of the domain).

Update (Aug 2020). In a recent paper ('An upper bound on the hot spots constants'), it is shown that whenever the conjecture fails, it cannot fail too badly: if $\Omega \subset \mathbb{R}^d$ is a bounded, connected domain with smooth enough boundary, then

$$\|u\|_{L^\infty(\Omega)} \leq 60\|u\|_{L^\infty(\partial\Omega)}.$$

One naturally wonders about the optimal constant in this inequality. The proof shows that 60 can be replaced by 4 in sufficiently high dimensions. An example of Kleefeld shows that the constant is at least 1.001.

Update (May 2022). Mariano, Panzo & Wang ('Improved upper bounds for the Hot Spots constant of Lipschitz domains') have improved the constant in $\|u\|_{L^\infty(\Omega)} \leq c\|u\|_{L^\infty(\partial\Omega)}$ to $c \leq \sqrt{e} + \varepsilon$ in high dimensions.

47. Norms of Laplacian Eigenfunctions

Let (M, g) be some smooth compact $d-$dimensional Riemannian manifold (possibly without boundary if that makes life easier) and consider the sequence of Laplacian eigenfunctions

$$-\Delta \phi_k = \lambda_k \phi_k$$

which we assume to be normalized as $\|\phi_k\|_{L^2} = 1$. There has been a lot of work on trying to understand how large these eigenfunctions can be. A classical result is

$$\|\phi_k\|_{L^\infty} \lesssim \lambda_k^{\frac{d-1}{4}}$$

and this is sharp on the sphere. Another celebrated result is the local Weyl law: if the manifold is assumed to have volume 1, then

$$\sum_{k=1}^{n} \phi_k(x)^2 = n + \mathcal{O}(n^{\frac{d-1}{d}}).$$

Question. How large can

$$\sum_{k=1}^{n} \|\phi_k\|_{L^\infty} \qquad \text{be?}$$

Is there a bound along the lines of

$$\sum_{k=1}^{n} \|\phi_k\|_{L^\infty} \lesssim n(\log n)^c \qquad ?$$

The sum is, trivially, at least $\sim n$. If we consider a high-dimensional torus $\mathbb{T}^d$, $d \geq 5$, then the multiplicity of the eigenspaces should grow polynomially in the frequency and a random rotation of each eigenspace should lead to a basis for which the sum is $\sim n\sqrt{\log n}$ though it is not clear to me how easy it is to make this precise.

There is a naturally dual question. Considering that

$$1 = \|\phi_k\|_{L^2}^2 \leq \|\phi_k\|_{L^1} \|\phi_k\|_{L^\infty}$$

there is the equally interesting problem of bounding $\|\phi_k\|_{L^1}$ from below. This problem has received much less attention. It is not entirely clear to me to what extent it is really dual. It arose naturally in 'Quantum Entanglement and the Growth of Laplacian Eigenfunctions'.

Question. How small can

$$\sum_{k=1}^{n} \|\phi_k\|_{L^1} \qquad \text{be?}$$

Is there a bound along the lines of

$$\sum_{k=1}^{n} \|\phi_k\|_{L^1} \gtrsim \frac{n}{(\log n)^c} \qquad ?$$

Maybe $c = 0$? I would expect this to be true for some $c \leq 1$ and maybe even $c \leq 1/2$.

48. Laplacian Eigenfunctions with Dirichlet conditions and mean value 0

I learned about this charming problem from Raghavendra Venkatraman. Suppose we have some domain $\Omega \subset \mathbb{R}^n$ and consider

$$-\Delta u_k = \lambda_k u_k$$

with Dirichlet boundary conditions $u_k\big|_{\partial\Omega} = 0$. If the boundary conditions were Neumann, then the first eigenfunction would be constant and all the other ones would, by orthogonality, have mean value 0. The situation is quite different for Dirichlet: already the first eigenfunction is not constant and there is no reason why any of them should have mean value 0. One would expect that, for a generic domain, none of them have mean value 0. Conversely, on the $d-$dimensional ball, we see that only $\sim n^{1/d}$ of the first n eigenfunctions have a mean value different from 0. In general, having many of these eigenfunctions have mean value 0 should be a sign of great symmetry in the domain.

> **Problem.** Is it true that among the first n Dirichlet eigenfunctions on a domain in $d-$dimensions at least $n^{1/d}$ have a mean value different from the ball? One might even conjecture that if there are $\leq Cn^{1/d}$, then it already has to be the ball independently of C (meaning that rate $\sim n^{1/d}$ already implies it is the ball).

The ball is the most symmetric and should be the worst. Raghav and I proved ('Dirichlet eigenfunctions with nonzero mean value ') that, up to logs, at least $n^{1/d}$ of the first n have a mean value different from 0. Clearly, the square root should not be there.

49. A PDE describing roots of polynomials under differentiation

Suppose we an initial probability measure $\mu = f(x)dx$. We can then consider n iid copies of this random variable and create the random polynomial

$$p_n(x) = \prod_{k=1}^{n} (x - x_i).$$

Note that in this model of randomness, the roots are random (other popular models for random polynomials choose random coefficients).

> **Question.** What is the distribution of roots of the $(t \cdot n)-$th derivative of the polynomial for $0 < t < 1$?

In 'A Nonlocal Transport Equation Describing Roots of Polynomials Under Differentiation' I gave a formal derivation of a PDE describing the evolution of the density

$$\frac{\partial u}{\partial t} + \frac{1}{\pi} \frac{\partial}{\partial x} \arctan\left(\frac{Hu}{u}\right) = 0,$$

where Hu is the Hilbert transform of u.

> **Question.** Does this PDE have a solution for smooth, compactly supported initial data on $\mathbb{R}$?

The best result in this direction is the *beautiful* work of Alexander Kiselev and Changhui Tan ('The Flow of Polynomial Roots Under Differentiation') who study the analogous problem on $\mathbb{S}^1$ (with polynomials replaced by trigonometric polynomials). However, nothing so far seems to be known on the real line.

The PDE seems to be of independent interest: using work of Shlyakhtenko-Tao, I gave another (again formal) that this process is also equivalent to fractional free convolution in Free Probability Theory (see 'Free Convolution Powers via Roots of Polynomials'). This claim has since been proven rigorously (see for example, Hoskins-Kabluchko or Arizmendi-Garza-Vargas-Perales). This naturally suggests that the PDE should have a solution, it should have infinitely many conservation laws etc. etc. etc.

50. The Khavinson-Shapiro conjecture

This conjecture is due to Khavinson & Shapiro, I first saw it in the charming article 'A Tale of Ellipsoids in Potential Theory' by Khavinson & Lundberg.

The starting point is the following basic fact.

Proposition. *Let $\Omega \subset \mathbb{R}^n$ be an ellipsoid and consider the Dirichlet problem*

$$\begin{cases} \Delta u &= 0 \\ u &= p \qquad \text{on } \partial\Omega, \end{cases}$$

where $p : \mathbb{R}^n \to \mathbb{R}$ is a polynomial. Then u is a polynomial.

So, basically, the harmonic extension of a polynomial to the inside of the domain remains a polynomial. The question is: does this property characterize ellipsoids?

It's an *extraordinarily* beautiful conjecture but it's not so clear how to approach it.

Part 5. Problems involving Optimal Transport

51. Optimal Transport with concave cost

Perhaps the easiest instance of this problem is the following: let $X = \{x_1, \ldots, x_n\}$ and $Y = \{y_1, \ldots, y_n\}$ be two sets of real numbers, what can be said about the optimal transport cost

$$W_p^p(X,Y) = \min_{\pi \in S_n} \sum_{i=1}^{n} |x_i - y_{\pi(i)}|^p,$$

where $\pi : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ ranges over all permutations? The answer is trivial when $p \geq 1$: order both sets in increasing order and send the $i-$th largest element from X to the $i-$th largest element from Y. The problem becomes highly nontrivial when the cost function is concave.

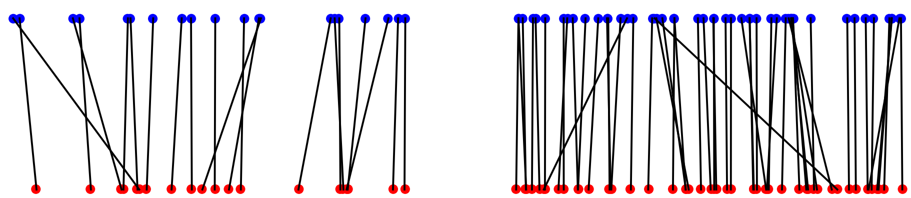


Figure 24. 20 (right: 50) red points on $\mathbb{R}$ being optimally matched to 20 (right: 50) blue points on $\mathbb{R}$ (shown displaced to illustrate the matching) and subject to cost $c(x,y) = |x-y|^{1/2}$.

> *For concave functions of the distance, the picture which emerges is rather different. Here the optimal maps will not be smooth, but display an intricate structure which – for us – was unexpected; it seems equally fascinating from the mathematical and the economic point of view. [...] To describe one effect in economic terms: the concavity of the cost function favors a long trip and a short trip over two trips of average length [...] it can be efficient for two trucks carrying the same commodity to pass each other traveling opposite directions on the highway: one truck must be a local supplier, the other on a longer haul.* (Gangbo & McCann)

What we (Andrea Ottolini and S, 'Greedy Matching in Optimal Transport with concave cost') found is that the obvious greedy algorithm (find the red and blue point at minimal distance among all pairs, connect the two, remove them both from their respective sets and repeat) does *amazingly* well when the cost function is very concave (say, $c(x,y) = d(x,y)^{0.01}$) with errors as small as 0.01%.

Another amazing matching is the Dyck matching introduced by Caracciolo-D'Achille-Erba-Sportiello (' The Dyck bound in the concave 1-dimensional random assignment model'). Their idea is to introduce $g : [0,1] \to \mathbb{Z}$

$$g(x) = \#\{1 \leq i \leq n : x_i \leq x\} - \#\{1 \leq i \leq n : y_i \leq x\}$$

The function is increasing whenever x crosses a new element of x while it decreases every time it crosses an element of y. The Dyck matching is then obtained by matching across level sets of the function g (see Fig. 25). The Dyck matching is independent of the cost function.

FIGURE 25. A collection of $n = 5$ red/blue points, the function $g(x)$ (left, rescaled for clarity) and the Dyck matching (right).

> **Question.** What we see empirically is that the greedy matching is very good for $c(x,y) = d(x,y)^p$ and p close to 0 while the Dyck matching is very good when $c(x,y) = d(x,y)^p$ when p is close to 1. Both matching are *okay* in the other regime for random points (with errors as small as 10%). Can this be made precise? One particularly fascinating regime is that of iid random points (where the Dyck matching is known to have optimal order but the constant is not known; for the greedy algorithm even the order is open).

Generally it seems that very little is known for optimal transport with concave cost. The associated transport plans seem to have all sorts of intriguing mixture of local and global behavior.

52. A Wasserstein transport problem

Consider the unit cube $[0,1]^d$. For which values of p, d is there a sequence $(x_n)_{n=1}^{\infty}$ such that, uniformly in N,

$$W_p\left(\frac{1}{N}\sum_{k=1}^{N}\delta_{x_k}, dx\right) \lesssim N^{-1/d},$$

where W_p is the $p-$Wasserstein distance? For each fixed dimension, the problem gets harder as p increases. Here is what I know:

(1) A result of Cole Graham ('Irregularity of distribution in Wasserstein distance', 2019) implies that for $d = 1$, no such value p exists since there is no such sequence even for $p = 1$.
(2) In $d \geq 2$, Louis Brown and I ('On the Wasserstein distance between classical sequences and the Lebesgue measure', 2020) constructed a sequence that has this rate for $p \leq 2$. The argument requires a nontrivial amount of Number Theory (the existence of certain badly approximable vectors) and it would be very desirable to have a more stable, robust, explicit, simple construction.
(3) Boissard & Le Gouic (On the mean speed of convergence of empirical and occupation measures in Wasserstein distance, 2014) have an argument showing that for $d > 2p$, points chosen uniformly at random satisfy the inequality. Can this be extended to a sequence?

Is it true that for, say, $d = 2$, this is impossible for p sufficiently large?

53. A Wasserstein Inequality in two dimensions

Let (M,g) be a smooth compact $d-$dimensional Riemannian manifold without boundary and let $G(x,y)$ to denote the Green function of the Laplacian, i.e. G has mean value 0 and

$$-\Delta_x \int_M G(x,y)f(y)dy = f(x).$$

I proved ('A Wasserstein Inequality and Minimal Green Energy on Compact Manifolds') that for any $\{x_1,\dots,x_n\} \subset M$

$$W_2\left(\frac{1}{n}\sum_{k=1}^{n}\delta_{x_k}, dx\right) \lesssim_M \frac{1}{n}\left|\sum_{\substack{k,\ell=1\\ k\neq \ell}}^{n} G(x_k,x_\ell)\right|^{1/2} + \begin{cases} \frac{\sqrt{\log n}}{\sqrt{n}} & \text{if } d=2\\ n^{-1/d} & \text{if } d\geq 3\end{cases}.$$

This inequality is sharp up to constants when $d \geq 3$.

Question. Is the $\sqrt{\log n}$ term for $d=2$ necessary?

I do not know and could well imagine that it is or is not necessary. It would be very interesting if it were not necessary, then the argument in (Brown & S, Positive-definite Functions, Exponential Sums and the Greedy Algorithm: a curious Phenomenon) would lead to an explicit greedy construction of an infinite sequence

$$W_2\left(\frac{1}{n}\sum_{k=1}^{n}\delta_{x_k}, dx\right) \lesssim n^{-1/2}$$

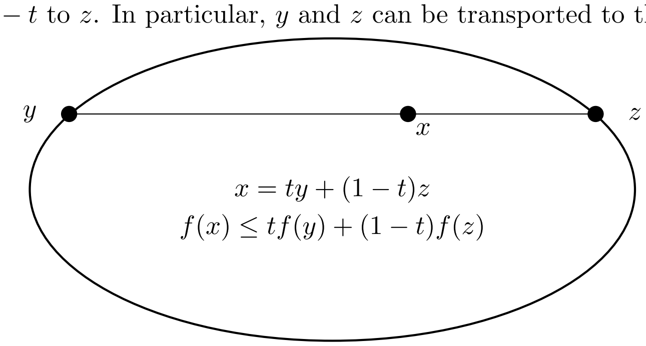
on general manifolds (which would partially answer the preceding question).

54. Transporting Mass from Ω to $\partial\Omega$

This is a curious problem that is naturally connected some rather interesting questions (see below). The setup is as follows: we are given a domain $\Omega \subset \mathbb{R}^n$. Let us assume for simplicity that the domain is convex and very nice. We start with the Hausdorff measure $\mathcal{H}^d$ on Ω. There is now a game that can be played: for any fixed $x \in \Omega$ and any two points $y,z \in \Omega$ such that

$$x = ty + (1-t)z$$

we are allowed to take the measure at x and transport a fraction of t to y and a fraction of $1-t$ to z. In particular, y and z can be transported to the boundary.



We play this game until all the measure is on the boundary, we call it μ. The total measure on the boundary then

$$\mu(\partial\Omega) = \mathcal{H}(\Omega).$$

We are interested in how 'evenly' it can be distributed: the question is therefore: how small can
$$\left\| \frac{d\mu}{d\mathcal{H}^{n-1}(\partial\Omega} \right\|_{L^\infty} \qquad \text{be?}$$
Here the derivative is understood as Radon-Nikodym. If the measure on the boundary was perfectly flat, then
$$\left\| \frac{d\mu}{d\mathcal{H}^{n-1}(\partial\Omega} \right\|_{L^\infty} = \frac{\mathcal{H}^d(\Omega)}{\mathcal{H}^{d-1}(\Omega)}.$$

This problem has a curious relationship with Hermite-Hadamard inequalities for convex functions: more precisely, for convex, nonnegative $f : \Omega \to \mathbb{R}$, we have
$$\int_\Omega f \; dx \leq \left\| \frac{d\nu}{d\sigma} \right\|_{L^\infty} \cdot \int_{\partial\Omega} f \; d\sigma.$$
In particular, if the centers of mass of Ω and $\partial\Omega$ are distinct, then the constant is strictly larger than $\mathcal{H}^d(\Omega)/\mathcal{H}^{d-1}(\Omega)$ and the best possible measure is *not* flat.

> **Question.** Which one is the flattest measure,
> $$\text{the smallest value of} \qquad \left\| \frac{d\nu}{d\sigma} \right\|_{L^\infty}$$
> that can be achieve by this type of transport?

This technique was used in (The Hermite-Hadamard inequality in higher dimension, J. Geom. Anal) to obtain some bounds; however, none of these arguments attempted to be optimal in any way (they are quite lose in terms of the constants).

55. Why is this Linear System usually solvable?

This problem is somewhere between Linear Algebra and Optimal Transport. It's about the existence of certain measures on discrete spaces so that the optimal transport has a certain invariance property, so we put it here. It can be described in a very easy way in the language of linear algebra.

This strange phenomenon was discovered in 'Curvature on graphs via equilibrium measures' (J. Graph Theory 2023). Let $G = (V, E)$ be some finite, connected graph and let D denotes its graph distance matrix,
$$D_{ij} = d(v_i, v_j).$$
This is a symmetric integer-valued matrix. We are interested in the linear system
$$Dx = \mathbf{1},$$
where $\mathbf{1} = (1, 1, 1 \ldots, 1)$ is the all 1's vector. For reasons that I cannot explain, it seems that this linear system tends to have a solution for **most** graphs. The smallest counterexamples are 2 graphs on 7 vertices and 14 graphs on 8 vertices. It is known that there are infinitely many examples of graphs for which the linear system does not have a solution: this is proven in this paper[4].

[4] William Dudarov, Noah Feinberg, Raymond Guo, Ansel Goh, Andrea Ottolini, Alicia Stepin, Raghavenda Tripathi, Joia Zhang, On the image of graph distance matrices, arXiv

> **Question.** Is there some sense in which graphs $G = (V, E)$ for which the linear system $Dx = \mathbf{1}$ does not have a solution is 'small'? Are such graphs exceptional?

It is also proven, in the paper cited above, that for Erdős-Renyi graphs $G(n, p)$ the matrix D is invertible with likelihood tending to 1 as $n \to \infty$.

Part 6. Linear Algebra

56. Eigenvector Phase Retrieval

Suppose $A \in \mathbb{R}^{n \times n}$ has eigenvalue $\lambda \in \mathbb{R}$, suppose that eigenvalue has multiplicity 1 and there is a unique eigenvector (up to sign) $Ax = \lambda x$. Knowing A and λ, I can find x by solving

$$(A - \lambda \cdot \mathrm{Id}_{n \times n})x = 0.$$

This can be done in $\mathcal{O}(n^3)$ time.

Suppose now someone, additionally, gives you the n numbers $(|x_i|)_{i=1}^n$, the absolute value of n of these numbers. Is it possible to quickly recover the missing signs $x_i = \varepsilon_i |x_i|$? Since we have strictly more information, the problem become easier and can be solved in $\mathcal{O}(n^3)$. But it somehow feels as if this additional information should help us (and potentially help us a lot). Hau-tieng Wu and I ('Recovering eigenvectors from the absolute value of their entries', on arXiv) propose an algorithm that works some of the time. The problem should become a lot easier when $|\lambda|$ is very large, i.e. when $|\lambda| \sim \|A\|$.

57. Matrix products

Let $A, B \in \mathbb{R}^{n \times n}$ be two symmetric, positive-definite matrices. Under which circumstances is it true that 'ordered products are always bigger than unordered products', i.e. for example

$$\|AABABA\| \leq \|AAAABB\|?$$

We know

(1) that inequalities of this type are always true when $n = 2$ (joint work with R. Alaifari and L. Pierce, Proc. AMS, 2020)
(2) that individual such inequalities can be true for all n
(3) that there are such inequalities that are false for all $n \geq 3$ (as shown by S. Drury, Electron. J. Linear Algebra, 2009)

I would assume that such inequalities are 'generally' true. There are many ways of making this precise: one way would be to say that for any fixed inequality, the measure of matrices (A, B) for which that fixed inequality fails becomes small as $n \to \infty$. Moreover, one would assume that, as the products gets longer, there should be less and less counterexamples.

58. The Kaczmarz algorithm

The Kaczmarz is an interesting algorithm for solving linear systems of equations $Ax = b$. It interprets such systems as the intersection of hyperplanes: using a_i to denote the $i-$th column of A. Then we are looking for a solution of

$$\langle a_i, x \rangle = b_i,$$

for all i. The Kaczmarz method is an iterative scheme: given an approximate solution x_k, let us pick an equation, say the $i-$th equation, and modify x_k the smallest possible amount necessary to make it correct: set $x_{k+1} = x_k + \delta a_i$, where δ is such that $\langle a_i, x_{k+1} \rangle = b_i$. Formally,

$$x_{k+1} = x_k - \frac{b_i - \langle a_i, x_k \rangle}{\|a_i\|^2} a_i.$$

Strohmer & Verhsynin proved that if the $i-$th equation is chosen with likelihood proportional to $\|a_i\|^2$, then this algorithm converges exponentially and

$$\mathbb{E}\, \|x_k - x\|_2^2 \leq \left(1 - \frac{1}{\|A\|_F^2 \|A^{-1}\|_2^2}\right)^k \|x_0 - x\|_2^2.$$

I proved (Randomized Kaczmarz..., arXiv, June 2020) that this algorithm has a particular connection to the singular vectors of the matrix A. More precisely,

Theorem. *Let v_ℓ be a (right) singular vector of A associated to the singular value σ_ℓ. Then*

$$\mathbb{E} \left\langle x_k - x, v_\ell \right\rangle = \left(1 - \frac{\sigma_\ell^2}{\|A\|_F^2}\right)^k \left\langle x_0 - x, v_\ell \right\rangle.$$

This suggests that $x_k - x$ will, for large value of k, be mainly a combination of small singular vectors (i.e. singular vectors associated to small singular values σ_ℓ). This has an interesting geometric combination that I would like to understand better: it basically means you bounce around the hyperplanes in a way that prefers certain angles. What I would like to understand better is more refined statistics of the vector

$$\left(\langle x_k - x, v_\ell \rangle\right)_{\ell=1}^n \qquad \text{as } k \text{ increases.}$$

The Theorem mentioned above analyzes the expectation of a fixed entry as k increases but surely there is no strong form of concentration. Presumably the variance is gigantic? What happens geometrically to the point? In the same paper, I also showed that

Theorem. *If $x_k \neq x$ and $\mathbb{P}(x_{k+1} = x) = 0$, then*

$$\mathbb{E} \left\langle \frac{x_k - x}{\|x_k - x\|}, \frac{x_{k+1} - x}{\|x_{k+1} - x\|} \right\rangle^2 = 1 - \frac{1}{\|A\|_F^2} \left\| A \frac{x_k - x}{\|x_k - x\|} \right\|^2.$$

This emphasizes the same principle: one bounces around randomly but at different speeds in different subspaces. This gives some insight into what is happening – in particular, can these ideas be somehow used to accelerate the convergence of the algorithm?

59. Approximate Solutions of linear systems

Let $A \in \mathbb{R}^{n \times n}$ be invertible, $x \in \mathbb{R}^n$ unknown and $b = Ax$ given. We are interested in *approximate* solutions: vectors $y \in \mathbb{R}^n$ such that $\|Ay - b\|$ is small. I proved ('Approximate Solutions of Linear Systems at a universal rate') that for $0 < \varepsilon < 1$ there is a composition of k orthogonal projections onto the n hyperplanes generated by the rows of A, where

$$k \leq 2 \log\left(\frac{1}{\varepsilon}\right) \frac{n}{\varepsilon^2}$$

which maps the origin to a vector $y \in \mathbb{R}^n$ satisfying $\|Ay - Ax\| \leq \varepsilon \cdot \|A\| \cdot \|x\|$. This upper bound on k is independent of the spectral properties of the matrix A.

The proof is probabilistic. This leads to a natural question.

Problem. Is the log factor necessary?

It would be very nice if it could be removed: note that, in some sense and after some rescaling, the quantity n/ε^2 is close to the effective numerical rank, the dimension of the space where the matrix can be large. Removing the log would mean that projections explore the space effectively. This might be too good to be true and may simply be a good indicator that the log cannot be removed.

60. Finding the Center of a Sphere from Many Points on the Sphere

Here is a particularly funny way of solving linear systems $Ax = b$ where $A \in \mathbb{R}^{n \times n}$ is assumed to be invertible (taken from 'Surrounding the Solution of a Linear System of equations from all sides', arXiv, Sep. 2020). Denote the rows of A by $a_1, \dots, a_n$. Then, for any $y \in \mathbb{R}^n$ and any $1 \leq i \leq n$,

$$y \qquad \text{and} \qquad y + 2 \cdot \frac{b_i - \langle y, a_i \rangle}{\|a_i\|^2} a_i$$

have the same distance to the solution x. This means we can very quickly generated points that all have the same distance from the solution by starting with a random guess for the solution and then iterating this procedure. Indeed, generating m points on a sphere around the solution x has computational cost $\mathcal{O}(n \cdot m)$, it is very cheap. In particular, it is very cheap to generate $c \cdot n$ points on the sphere like that, where c is a constant.

Problem. Given at least $n + 1$ points on a sphere in $\mathbb{R}^n$, how would one quickly determine an accurate approximation of its center? Does it help if one has $c \cdot n$ points?

The problem can, of course, be solved by setting up a linear system – the question is whether it can be done (computationally) cheaper if one is okay with only having an approximation of the center.

A very natural way to do is to simply average the points. This is not very good when the points are clustered in some region of space, though. I proved that if you pick the rows of A with likelihood proportional to $\|a_i\|^2$ and then average, then the arising sequence of points satisfies

$$\mathbb{E} \left\| x - \frac{1}{m} \sum_{k=1}^{m} x_k \right\| \leq \frac{1 + \|A\|_F \|A^{-1}\|}{\sqrt{m}} \cdot \|x - x_1\|.$$

This gives rise to an algorithm that is as fast as the Random Kaczmarz method. A better way of approximating the center would presumably give rise to a faster method!

61. Small Subsingular Values

Suppose $A \in \mathbb{R}^{m \times n}$ with $m > n$ is a tall rectangular matrix with many more rows than columns. We assume furthermore that the rows are all normalized in ℓ^2. We can now define, for any $0 < \alpha < 1$ the restricted singular value

$$\sigma_{\alpha,\min}(A) = \min_{\substack{S \subset \{1,2,\dots,m\} \\ |S| = \alpha m}} \inf_{x \neq 0} \frac{\|A_S x\|}{\|x\|},$$

where A_S is the restriction of A to rows indexed by S. It's clear that this quantity will grow as α grows and coincides with the classical smallest singular value of A when $\alpha = 1$. Haddock, Needell, Rebrova & Swartworth (Quantile Kaczmarz, SIMAX 2022) proved that for certain types of random matrices one has

$$\sigma_{\alpha,\min}(A) \gtrsim \alpha^{3/2} \sqrt{\frac{m}{n}} \qquad \text{with high likelihood.}$$

I'd be interested in understanding what the best kind of matrix for this problem would be, the one maximizing these quantities. Note that since the rows are all normalized in ℓ^2, we can think of the rows as points on the unit sphere.

Let us consider the case where $A \in \mathbb{R}^{m \times n}$ has each row sampled uniformly at random from the surface measure of $\mathbb{S}^{n-1}$ and suppose that the matrix is large, $m, n \gg 1$, and that the ratio m/n is large. Trying to find a subset $S \subset \{1, 2, \dots, m\}$ such that A_S has a small singular value might be difficult, however, we can flip the question: for a given $x \in \mathbb{S}^{n-1}$, how would we choose S to have

$$\|A_S x\|^2 = \sum_{i \in S} \langle x, a_i \rangle^2 \qquad \text{as small as possible?}$$

This is a much easier problem: compute $\langle x, a_i \rangle^2$ for $1 \leq i \leq m$ and then pick S to be the set of desired size corresponding to the smallest of these numbers. Using rotational invariance of Gaussian vectors, we can suppose that $x = (1, 0, \dots, 0)$. Then we expect, in high dimensions, that

$$\langle a_i, x \rangle \sim \frac{1}{\sqrt{n}} \gamma \qquad \text{where } \gamma \sim \mathcal{N}(0, 1).$$

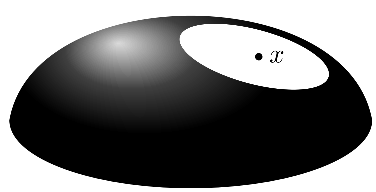


FIGURE 26. Removing a small spherical cap around the vector x.

This suggest a certain picture: large inner products are those where many rows a_i are nicely aligned with x and we know with which likelihood to expect them (these are just all the points in the two spherical caps centered at x and $-x$). This would then suggest that, in the limit as $m, n, m/n \to \infty$, we have an estimate along the lines of

$$\frac{\sigma^2_{\alpha,\min}(A)}{(m/n)} = \frac{1}{\sqrt{2\pi}} \int_{-\beta}^{\beta} e^{-x^2/2} x^2 \, dx$$

where the parameter α is implicitly defined via

$$\frac{1}{\sqrt{2\pi}}\int_{-\beta}^{\beta} e^{-x^2/2}dx = \alpha.$$

> **Question.** Is there a universal estimate, for $A \in \mathbb{R}^{m\times n}$ (maybe subject to $m, n \to \infty$ and maybe also $m/n \to \infty$) with all rows normalized to 1, along the lines of
>
> $$\sigma_{\alpha,\min}(A) \le c_\alpha \sqrt{\frac{m}{n}}$$
>
> where c_α is the number predicted by what one obtains when the rows are sampled uniformly at random on the sphere? Or is there an improvement by picking the rows to be a highly structured set of points?

Motivation. These questions arise naturally in the context of $q-$quartile Kaczmarz method (see Haddock, Needell, Rebrova & Swartworth (Quantile Kaczmarz, SIMAX 2022) and my paper on quantile Kaczmarz (Information and Inference)). However, I like them independently of that, it seems like a very nicely geometric question.

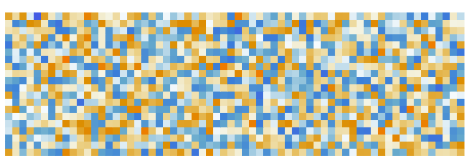
62. Solving Equations with More Variables than Equations.

This is a fun problem from joint work with Ofir Lindenbaum (arXiv March 2020, to appear in Signal Processing).

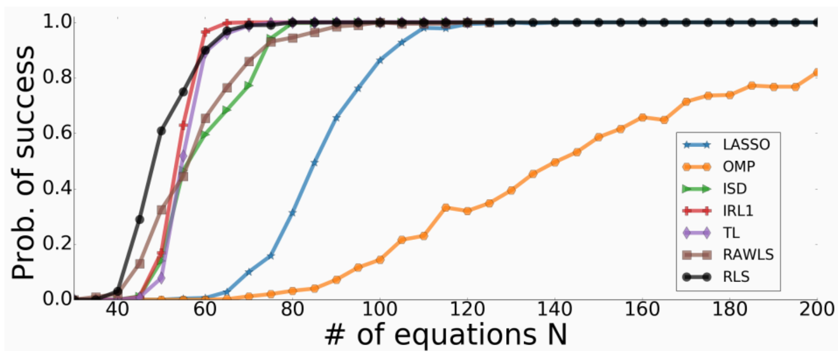
There is an underlying vector $x \in \mathbb{R}^d$ all of whose entries are either $-1, 0, 1$ and most of them are 0. In fact, we may assume that only a relatively small number is ± 1. We would like to understand how x looks like but we only have access to

$$y = Ax + \omega,$$

where $A \in \mathbb{R}^{n\times d}$ is a random matrix filled with independent $\mathcal{N}(0,1)$ Gaussians and $\omega \in \mathbb{R}^n$ is a random Gaussian vector.

$y =$  $x + w$

It is not terribly difficult to see that if n is very, very large, then it is fairly easy to reconstruct x. The question is: how small can you make n and still reconstruct x with high likelihood? What is remarkable is that this is doable even when n is smaller than the number of variables d. Ofir and I propose a fun algorithm: you take random subsets of the n rows, then do a least squares reconstruction and then average this over many random subsets. The method seems to differ from other methods and does work rather well even when n is quite small.



In fact, in some parameter regimes (small number of variables, little information), this method outperforms all the other methods. The method itself is quite simple and it seems like that one should be able to further improve it by playing with it.

Question. Are there natural variations on this idea?

Many of the other proposed method come with a wide variety of variations; our particular approach seems to have not been explored very much, so maybe there are some interesting variations that maybe work even better?

Update (Mar 2021)**.** Ofir Lindenbaum and I found a tweak of the method which we call RLS (Refined Least Squares for Support Recovery, arXiv, March 2021) which leads to state-of-the-art results in many regimes. It seems very likely that we have not yet fully exhausted the possibility of the method.

63. Kaczmarz Kac Walk

Consider the following procedure (see the paper 'Kaczmarz Kac Walk')

Kaczmarz Kac Walk. For $A \in \mathbb{R}^{m \times n}$ with rows normalized to length 1, pick two rows $i \neq j$ uniformly at random and if $A_i \neq \pm A_j$,

$$A_j \leftarrow \frac{A_j - \langle A_i, A_j \rangle A_i}{\sqrt{1 - \langle A_i, A_j \rangle^2}} \tag{1}$$

as well as $b_j \leftarrow (b_j - \langle A_i, A_j \rangle b_i)/\sqrt{1 - \langle A_i, A_j \rangle^2}$.

This process should make the rows 'more orthogonal' and the new linear system should be better conditioned and easier to solve. Moreover, each step is actually computationally very cheap: one has to compute only one inner product. The behavior of the procedure is interesting to say the least: it seems that if $m \leq n$, then the sequence of matrices that one gets in this way simply converges to some matrix with orthogonal rows (and an orthogonal matrix when $m = n$). When $m > n$, interesting things start to happen: when $m = n + 1$, it seems that it may converge to an orthonormal basis + 1 extra vector but the numerics is tricky. It is not even entirely clear what happens when $n = 2 \ll m$ and m is very large. In that case, it seems that one converges to all vectors being close to an orthonormal basis (and its negative elements, so $\pm e_1, \pm e_2$). I tried writing down a mean-field limit and got

$$\begin{aligned}\frac{\partial}{\partial t}u(t,x) &= -u(t,x)\\ &\quad + u(t, x+\pi/2)\int_{x-\pi/2}^{x+\pi/2} u(t,y)dy + u(t,x-\pi/2)\int_{x+\pi/2}^{x+3\pi/2} u(t,y)dy.\end{aligned}$$

for the evolution of the density (see 'Kaczmarz Kac Walk'). This is a curious evolution equation and maybe leads to singularity formation?

Update (Jan 2025)**.** Shah and Detherage (arXiv:2411.16101) prove, among other things, the convergence of n vectors to an orthogonal system.

Part 7. Miscellaneous/Recreational.

64. Geodesics on compact manifolds

This question is about whether the vector field $V(x,y) = (\sqrt{2}, 1)$ on the two-dimensional flat torus $\mathbb{T}^2$ has, in some sense, the best mixing properties. Let (M, g) be a smooth, compact two-dimensional Riemannian manifold without boundary: let $x \in M$ be a particular starting point and let $\gamma : [0, \infty] \to \mathbb{R}$ be a geodesic starting in x (in some arbitrary direction; parametrized according to arclength).

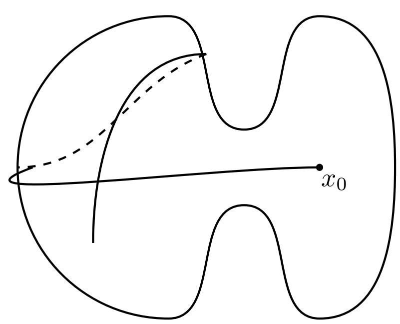


Figure 27. The best space-filling geodesic?

For any $\varepsilon > 0$, we can define L_ε as the smallest number such that

$$\{\gamma(t) : 0 \le t \le L_\varepsilon\} \qquad \text{is } \varepsilon-\text{dense on the manifold.}$$

Put differently, L_ε is how long we have to go along the geodesic so that it visits every point on the manifold up to distance at most ε. Here's the question: how long does L_ε have to be given ε? Since its $\varepsilon-$neighborhood is the entire manifold, we expect $L_\varepsilon \cdot \varepsilon \gtrsim \mathrm{vol}(M)$.

> **Problem.** Suppose (M,g) has the property that there exists a fixed geodesic such that
> $$L_\varepsilon \leq \frac{c}{\varepsilon}$$
> for one fixed universal c and all sufficiently small ε. What does this tell us about (M,g)?

One example would be $M = \mathbb{T}^2$ with the canonical metric and the geodesic moving in a direction whose ratio of x and $y-$coordinates is badly approximable. It seems reasonable to assume that one can glue many tori together but the queston is whether these are fundamentally the only type of examples. Does hyperbolicity help?

One would perhaps assume that in a very hyperbolic two-dimensional setting one has something like
$$L_\varepsilon \lesssim \frac{\log(1/\varepsilon)}{\varepsilon}$$
for fairly generic geodesics?.

Update (Dec 2022). Apparently this type of property is known as the existence of a 'superdense' geodesic, two very recent papers in this spirit are

- J. Beck and W. Chen, Generalization of a density theorem of Khinchin and diophantine approximation
- J. Southerland, Superdensity and bounded geodesics in moduli space

65. The Traveling Salesman Constant

Pick n points i.i.d. from $[0,1]^2$. The length of the shortest traveling salesman path is known to satisfy
$$\text{length of shortest path} \sim \beta\sqrt{n},$$
where β is a universal constant (this is the Beardwood-Halton-Hammersley theorem from 1948). They gave the estimates
$$\frac{5}{8} \leq \beta \leq 0.92116\ldots$$
The best known lower bound is due to Gaudio & Jaillet (Op. Rest. Lett., 2020) and is $\beta \geq 0.6277$. I proved (Adv. Appl. Prob.) that $\beta \leq \beta_{BHH} - 10^{-6}$ though, if numerical evaluation of integrals is permissible, the improvement is a bit bigger. Numerical experiments suggest that $\beta \sim 0.7$. It seems like such a fundamental question, it would be nice to understand this a bit better.

66. Number of Positions in Chess

This is a very old question going back to Shannon's estimate for the complexity of chess. C. Shannon roughly estimate the number of admissible positions in Chess to be
$$\sim \frac{64!}{32!(8!)^2(2!)^6} \sim 4.6 \cdot 10^{42}.$$
Shannon's way of counting is rough, it excludes some admissible positions and includes some impossible ones. The best known upper bound is $\leq 10^{46}$. I (Int. J. Game Theory, 2015) showed that if one excludes promotion (a pawn at the end of the board may be exchanged), one can bound the number from above by $\leq 2 \cdot 10^{42}$.

I believe the actual number is quite a bit smaller. None of these counting scheme's properly account for the pawns. A pawn in A2 can never move over and end up on H3. I think properly counting that should decrease the number a lot. The commonly established wisdom is that the truth is somewhere between 10^{40} and 10^{50} but I think it's actually less than that, maybe even less than 10^{38}. This is arguably not very important but I am slightly bothered by the fact that everybody seems to be so sure that it's $\geq 10^{40}$.

Update (Dec 2021). Gourion (arXiv:2112.09386) proposes a new upper bound of 4×10^{37} for number of states without promotion.

67. Ulam Sets

Motivated by the strange behavior of Ulam sequences, Noah Kravitz and I (Integers, 2018) looked into Ulam sets: for a set of elements in a vector space $\{x_1, \dots, x_n\}$, keep adding the shortest vector that can be uniquely written as the sum of two distinct earlier terms. We observed that even the simplest settings, $\mathbb{R}^2$, $\mathbb{R}^3$, $\mathbb{Z} \times \mathbb{Z}_5$,..., lead to very strange structures: some seemingly random, some extremely structured. What is happening here?

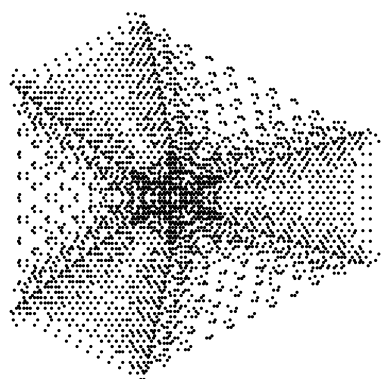


Figure 28. The set in $\mathbb{R}^3$ generated from $(1,0,0),(0,1,0),(0,0,1)$ (projected onto the plane that is orthogonal to $(1,1,1)$).

Update (Aug. 2020)**.** Bade, Cui, Labelle, Li (arXiv, August 2020) have looked at these types of sets in other settings as well. Lots and lots of structure!

Update (Dec. 2022)**.** Andrei Mandelshtam (On fractal patterns in Ulam words, arXiv:2211.14229) has found some *highly* intricate structure in Ulam words.

68. An amusing sequence of functions

Let us consider the sequence

$$f_n(x) = \sum_{k=1}^{n} \frac{|\sin(k\pi x)|}{k}.$$

This sequence arose out of some fairly unrelated questions (that were further pursued in a paper with X. Cheng and G. Mishne, J. Number Theory) but turned out to be quite curious.

Theorem (S, Mathematics Magazine 2018). *The function $f_n(x)$ has a strict local minimum in $x = p/q$ for all $n \geq q^2$.*

The asymptotically sharp scaling is given by $n \geq (1 + o(1))q^2/\pi$. It's not difficult to see that f_n grows like $\log n$ and thus f_∞ does not exist. But as n becomes large, there does seem to be some sort of universal function that emerges. Is it possible to make some more precise statements about f_n?

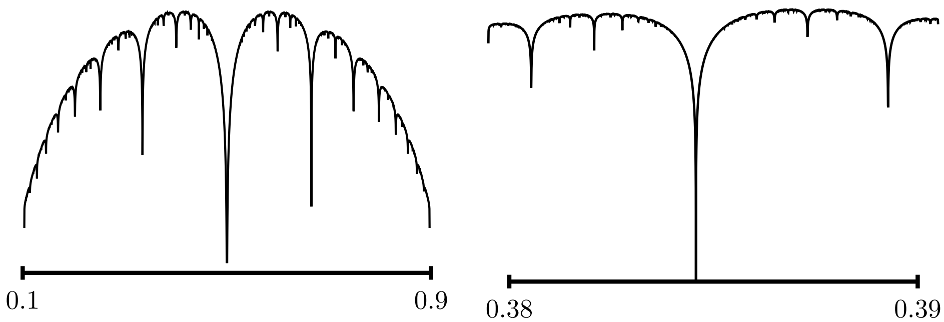


FIGURE 29. The function $f_{50.000}$ on $[0.1, 0.9]$ and zoomed in (right). The big cusp in the right picture is located at $x = 5/13$, the two smaller cusps are at $x = 8/21$ and $x = 7/18$.

More generally, if (M, g) is a compact manifold and

$$-\Delta\phi_k = \lambda_k\phi_k$$

is a sequence of L^2-normalized eigenfunctions, is it possible to say anything about the function $f_n : M \to \mathbb{R}$ given by

$$f_n(x) = \sum_{k=1}^{n} \frac{|\phi_k(x)|}{\sqrt{\lambda_k}}?$$

Part 8. Solved Problems

69. A Wasserstein Uncertainty Principle with Applications

This question arose out of understanding level sets of sums of Laplacian eigenfunctions (Calc. Var. PDE 2020) but is actually a topic that is of independent interest and has more to do with calculus of variations and geometric measure theory.

Let $\Omega = [0,1]^d$ (presumably this holds on much more general domains, manifolds, etc.) and let $f : [0,1]^d \to \mathbb{R}$ denote a function with mean value 0. Then

$$\mu = \max(f,0)dx \qquad \text{and} \qquad \nu = \max(-f,0)dx$$

are two measures with the same total mass (since f has mean value 0). How much does it cost to 'transport' μ to ν? If we assume that transporting a $\varepsilon-$unit of measure distance D costs $\varepsilon \cdot D$, then this naturally leads to the 'Earth-Mover' Wasserstein distance W_1. The size of $W_1(\mu,\nu)$ depends on the function, of course.

Here's a basic idea: if $W_1(\mu,\nu)$ is quite small, then the transport is cheap. But if the transport is cheap, then most of the positive part of f has to lie pretty close to most of the negative part of f. But that should somehow force the zero set $\{x : f(x) = 0\}$ to have large $(d-1)-$dimensional volume. In (Calc Var Elliptic Equations, 2020) I proved in $d = 2$ dimensions, i.e. for $f : [0,1]^2 \to \mathbb{R}$, that

$$W_1(f_+, f_-) \cdot \mathcal{H}^1 \left\{ x \in (0,1)^2 : f(x) = 0 \right\} \gtrsim \frac{\|f\|_{L^1}^2}{\|f\|_{L^\infty}}.$$

This result is sharp. Amir Sagiv and I generalized this to higher dimensions (SIAM J. Math. Anal). The currently sharpest form in higher dimensions is due to Carroll, Massaneda & Ortega-Cerda (Bull. London Math. Soc.) and reads

$$W_1(f_+, f_-) \cdot \mathcal{H}^{d-1} \left\{ x \in (0,1)^d : f(x) = 0 \right\} \gtrsim_d \left(\frac{\|f\|_{L^1}}{\|f\|_{L^\infty}} \right)^{2-\frac{1}{d}} \|f\|_{L^1}.$$

Here, it is not clear whether the power is optimal or not. Of course, for all these inequalities it would also be interested in having the same underlying thought expressed in other ways: certainly the idea behind these things can be expressed in many different ways.

Update (Nov. 2020)**.** A sharp form of this principle has been established in

> Fabio Cavalletti, Sara Farinelli, Indeterminacy estimates and the size of nodal sets in singular spaces, arXiv:2011.04409

70. A Sign Pattern for the Hypergeometric Function ${}_1F_2$

This is motivated by the immediately preceding section: some curious structure arises naturally when studying the local stability of the inequality.

Question. Let $\alpha > 0$. We define, for integers $k \geq 1$, the sequence

$$a_k = {}_1F_2\left(\frac{1+\alpha}{2}; \frac{3}{2}, \frac{3+\alpha}{2}; -\frac{\pi^2}{16}(2k-1)^2 \right).$$

For which α is it true that $a_k \geq 0$ for odd values of k and $a_k \leq 0$ for even values of k?

If α is an integer, the hypergeometric function simplifies tremendously and it is not hard to check that the desired property is satisfied for $\alpha \in \{2, 3, 4, 5, 6\}$. It should be true for all integers $\alpha \geq 2$. In fact, I would expect it to be true for all real $\alpha \geq 2$. Once it is true for some fixed $\alpha > 0$, it implies that for all smooth, even functions $f : [-1/2, 1/2] \to \mathbb{R}$,

$$\max(\widehat{f}) \geq \frac{\alpha + 1}{\alpha\pi} \int_{-1/2}^{1/2} (1 - |2x|^{\alpha}) f(x) dx,$$

where

$$\max(\widehat{f}) = \max \left\{ \sup_{k \in \mathbb{N}} \left(2k + \frac{1}{2}\right) \widehat{f}\left(2k + \frac{1}{2}\right), - \inf_{k \in \mathbb{N}} \left(2k + \frac{3}{2}\right) \widehat{f}\left(2k + \frac{3}{2}\right) \right\}.$$

Update (Oct. 2021)**.** The sign pattern has been established in

> Yong-Kum Cho and Young Woong Park, The zeros of certain Fourier transforms:Improvements of Pólya's results, arXiv:2110.01885

71. A Refinement of Smale's Conjecture?

Let $f : \mathbb{C} \to \mathbb{C}$ be a polynomial normalized to $f(0) = 0$ and $|f'(0)| = 1$. Smale proved in 1981 that there exists a critical point ($z \in \mathbb{C}$ such that $f'(z) = 0$) satisfying

$$|f(z)| \leq 4|z|.$$

The question is whether 4 can be replaced by 1.

> **A Stronger Conjecture?** Let $g : \mathbb{C} \to \mathbb{C}$ be a polynomial with $|g(0)| = 1$ and consider the subset
>
> $$A = \{z \in \mathbb{C} : |g(z)| < 1\} \subset \mathbb{C}.$$
>
> Let B be the connected component of A whose closure contains 0. Then the polynomial $zg(z)$ contains a critical point in B.

This, if true, would slightly refine Smale's conjecture (which says that there is a critical point in A). In practice, the statement seems to be true – in most cases, the number of roots of $zg(z)$ in B seems to be the same as the number of roots of $g(z)$ in B (which is at least 1). For a while I thought that this stronger statement might be true until Peter Müller constructed a counterexample of degree 5.

The counterexamples are 'barely' counterexamples, so I am naturally still wondering whether something along these lines might be true...

72. A Type of Kantorovich-Rubinstein Inequality?

Let $f : [0, 1]^d \to \mathbb{R}$ and let μ be a probability measure on $[0, 1]^d$. Is there an inequality

$$\left| \int_{[0,1]^d} f(x) dx - \int_{[0,1]^d} f(x) d\mu \right| \leq c \cdot \|\nabla f\|_{L^{d,1}} \cdot W_{\infty}(\mu, dx),$$

where $L^{p,q}$ is the Lorentz space and W_{∞} the $\infty-$Wasserstein distance. This inequality is 'almost' (in a suitable sense) proven in 'On a Kantorovich-Rubinstein inequality' (arXiv: Oct 2020). The most general question is whether there exist inequalities of the type

$$\left| \int_{[0,1]^d} f(x)dx - \int_{[0,1]^d} f(x)d\mu \right| \leq c \cdot \|\nabla f\|_{L^{p,q}} \cdot W_r(\mu, dx),$$

The case $p = q = \infty$ and $r = 1$ is, of course, the famous Kantorovich-Rubinstein inequality that also holds for more general combination of measures (it is not necessary for one of them to be dx).

Update (March 2022). The conjectured inequality has been established by Filippo Santambrogio in the preprint 'Sharp Wasserstein estimates for integral sampling and Lorentz summability of transport densities' (cvgmt: 5463 and Journal of Functional Analysis)

73. A Pretty Inequality involving the Cubic Jacobi Theta Function

Here is a pretty inequality: for all $0 < q < 1$,

$$\sum_{m,n=-\infty}^{\infty} q^{m^2+mn+n^2} \geq \frac{2\pi}{\sqrt{3}\log(1/q)}.$$

It came up naturally in unrelated work ('On the Logarithmic of Points on $\mathbb{S}^2$', arXiv Nov. 2020). The inequality seems to be remarkably accurate as $q \to 1$. I think a way of proving it for $q \in (q_0, 1)$ for some absolute q_0 would be to combine an identity of Borwein & Borwein (1991)

$$\sum_{m,n=-\infty}^{\infty} q^{m^2+mn+n^2} = \theta_3(q)\theta_3(q^3) + \theta_2(q)\theta_2(q^3),$$

where

$$\theta_2(q) = \sum_{k=-\infty}^{\infty} q^{(k+1/2)^2} \qquad \text{and} \qquad \theta_3(q) = \sum_{k=-\infty}^{\infty} q^{k^2}$$

with the identities

$$\theta_2(q) = (q^2, q^2)_\infty \cdot \exp\left(-\frac{1}{\log q}\frac{\pi^2}{12} + \frac{\log q}{12} + \sum_{k=1}^{\infty} \frac{1}{k \sinh\left(\frac{\pi^2 k}{\log q}\right)} \right),$$

and

$$\theta_3(q) = (q^2, q^2)_\infty \cdot \exp\left(-\frac{1}{\log q}\frac{\pi^2}{12} + \frac{\log q}{12} + \sum_{k=1}^{\infty} \frac{(-1)^k}{k \sinh\left(\frac{\pi^2 k}{\log q}\right)} \right),$$

and

$$(q^2; q^2)_\infty = \exp\left(-\frac{\pi^2}{12\log(1/q)} - \frac{1}{2}\log\left(\frac{\log(1/q)}{\pi}\right) + \frac{\log(1/q)}{12} - \sum_{k=1}^{\infty} \frac{1}{k}\frac{\widehat{q}^k}{1-\widehat{q}^k} \right),$$

where $\widehat{q}$ is an abbreviation for

$$\widehat{q} = \exp\left(-\frac{2\pi^2}{\log(1/q)} \right).$$

For $q \in (0, q_0)$, one could probably establish it using a computer.

Question. Is there a 'nice' proof? Is there a more fundamental reason why the inequality is true? What if m^2+mn+n^2 is replaced by another positive-definite quadratic form?

Update (Feb 2023). Nian-Hong Zhou tells me that there is a very simple proof by taking the Poisson Summation Formula since all the Fourier coefficients are positive and gives Equation 5 on page 205 of Schoenberg's *Elliptic Modular Functions* as a reference. So this one really has a nice and simple solution!

74. An Inequality for Eigenvalues of the Graph Laplacian?

Let $G = (V, E)$ be a connected, finite graph and let $L = D - A$ be the (Kirchhoff) Graph Laplacian. It has eigenvalues

$$0 = \lambda_1 < \lambda_2 \leq \lambda_3 \leq \cdots \leq \lambda_n.$$

Question. Is there a universal constant $c > 0$ such that

$$\mathrm{diam}(G)^2 \leq c \sum_{k=2}^{n} \frac{1}{\lambda_k}?$$

A famous inequality of Alon-Milman shows that already the first term of the sum is essentially big enough provided one compensates for vertices of large degree: if we denote the maximal degree by Δ, then

$$\mathrm{diam}(G)^2 \leq \frac{2\Delta(\log \#V)^2}{\lambda_2}.$$

The question is now whether one can replace the dependency on Δ and $\#V$ by summing over the remainder of the spectrum.

Motivation. This inequality would imply that the notion of graph curvature introduced in 'Graph Curvature via Resistance Distance' satisfies a Bonnet-Myers-type inequality: if the curvature is bounded from below by $K > 0$, then

$$\mathrm{diam}(G) \leq c \cdot K^{-1/2}.$$

Update (May 2023)**.** Using an identity of McKay, we can rephrase the question in terms of the average commute time. The commute time between two vertices i, j is the expected time a random walk needs to go from i to j and then back to i. The question is then whether

$$\frac{1}{|V|^2} \sum_{v,w \in V} \mathrm{commute}(v, w) \gtrsim \frac{|E|}{|V|} \mathrm{diam}(G)^2.$$

The inequality is sharp up to constants for cycle graphs C_n where the average commute time is $\sim n^2$. One of the reasons the inequality is interesting is that $|E|$ appears: adding more edges increases the global commute time.

Update (July 2023)**.** I asked the question on mathoverflow and Yuval Peres produced the following counterexample.

Let $\ell = \lfloor \sqrt{n}/2 \rfloor$ and let $G_1, \ldots, G_\ell$ be disjoint cliques of size ℓ. Let K be a clique on the remaining $n - \ell^2 > n/2$ nodes. Connect every node in G_i to every node in G_{i+1} for $i < \ell$ and connect every node in G_ℓ to every node in K. This defines a graph of diameter ℓ. For $i < \ell/2$ the effective resistance between a node v in G_i and a node w in K is $\Theta(1/\ell)$ so the commute time between v and w is $\Theta(\ell^3)$. Consequently, the average commute time between all pairs is $\Theta(\ell^3)$ which contradicts the conjectured inequality.

75. Pasteczka's conjecture

Pasteczka is interested in convex domains $\Omega \subset \mathbb{R}^n$ such that for all convex functions $f : \Omega \to \mathbb{R}$

$$\frac{1}{|\Omega|} \int_{\Omega} f dx \leq \frac{1}{|\partial\Omega|} \int_{\partial\Omega} f d\sigma$$

Pasteczka (Annales Universitatis Paedagogicae Cracoviensis Studia Mathematica, 2018) remarks that by plugging in $f(x) = x_i$ (the $i-$th coordinate function) and $f(x) = -x_i$ (both of which are convex), we can deduce that such a domain Ω needs to satisfy

$$\text{center of mass of } \Omega = \text{center of mass of } \partial\Omega.$$

He conjectures that this condition implies that the convex Hermite-Hadamard inequality holds with constant 1. Or, put differently, the worst case is given by linear functions. This would be *very* nice if it were true – maybe too nice?

Update (June 2023). I asked whether anyone had any thoughts on Pasteczka's conjecture on mathoverflow, no replies.

Update (July 2023). Noah Kravitz and Mitchell Lee ('Hermite–Hadamard inequalities for nearly-spherical domains') just uploaded a short paper to the arXiv that **proves** Pasteczka's conjecture for domains sufficiently close to the ball.

Update (July 2023). Nazarov and Ryagobin prove Pasteczka's conjecture in two dimensions and disprove it in general in dimensions ≥ 3. This resolves Pasteczka's conjecture but, naturally, raises a new question (especially when combining it with the inequality of. Kravitz-Lee): when do domains admit a Hermite-Hadamard inequality?

Department of Mathematics, University of Washington, Seattle
Email address: `steinerb@uw.edu`