

**FOLLOW-UP-WORKSHOP TO TP
NEXT HORIZON: WORKSHOP ON OPEN PROBLEMS IN
HARMONIC
ANALYSIS AND ANALYTIC NUMBER THEORY**

ABSTRACT. Unedited notes from talks at the Hausdorff Research Institute for “Follow-Up-Workshop to TP Next Horizon: Workshop on Open Problems in Harmonic Analysis and Analytic Number Theory”. These notes are incomplete, have not been proofread, and should be considered only a crude approximation to what happened in the lectures, filtered through my own misunderstandings and distractions. Any errors should be assumed to be due to the note-taker.

CONTENTS

1. James Maynard, <i>Large values of Dirichlet polynomials and the zeta function</i>	2
2. Maksym Radziwill, <i>Where we get stuck on the 12th moment of the Riemann zeta function</i>	4
3. Dominique Malagüe, <i>Difficulty of using decoupling ideas for generalized Dirichlet series and cubic Weyl sums</i>	9
4. Ben Green, <i>Properly understanding why nilpotent groups are important in additive combinatorial problems</i>	13
5. Sarah Peluse, <i>Inverse theorems for multidimensional Gowers norms</i>	17
6. Emmanuel Kowalski, <i>Short sums of trace functions: examples, conjectures and applications</i>	20
6.1. What are trace functions?	20
6.2. Examples	20
6.3. Riemann Hypothesis	21
6.4. Short sums	22
6.5. Short sums conjecture	22
7. Philippe Michel, continuation of previous lecture	23
8. Hong Wang, <i>Danzer’s problem and quantitative Besikovitch projection theorem</i>	25
9. Damaris Schindler, <i>Density of rational points near manifolds</i>	27
10. Trevor Wooley, <i>The paucity of knowledge concerning Weyl sums and their mean values</i>	29
10.1. Introduction	29
10.2. How little we know	31
10.3. Mean values	32

https://www.mathematics.uni-bonn.de/him/assets/2025/2025_schedule_next_horizon_update.pdf

1. JAMES MAYNARD, *Large values of Dirichlet polynomials and the zeta function*

We will pitch some of our favorite problems coming from analytic number theory that can be phrased in terms of large values of Dirichlet polynomials, which can in turn be thought of purely in terms of harmonic analysis. (A similar pitch, made over lunch to Guth in January 2020, led to our result on large values of the zeta function.)

Given large real parameters $N \leq T$ and a number $\sigma \in (\frac{1}{2}, 1)$ and a 1-bounded sequence $(a_n)_{n=N}^{2N}$ of complex numbers, let $W \subseteq [T, 2T]$ be a set on which the corresponding Dirichlet polynomial has magnitude at least N^σ , i.e.,

$$\left| \sum_{n=N}^{2N} a_n n^{it} \right| = \left| \sum_{n=N}^{2N} a_n e^{it \log n} \right| \geq N^\sigma \quad \text{for all } t \in W.$$

Note that $\sigma = 1$ means “as large as possible”.

Question 1.1. What is the best upper bound for $|W|$ in terms of N, T, σ .

Hopefully the question is clear. We’re just asking for large values of these sorts of objects. We’re expecting that as you look for larger and larger values, these should occur less and less frequently. Most of the time we won’t care about specific 1-bounded sequences – we’ll try to work with general ones.

Conjecture 1.2 (Montgomery). *We have $|W| \leq T^{o(1)} N^{2-2\sigma}$.*

Conjecture 1.3 (Montgomery). $\int_T^{2T} \left| \sum_{n=N}^{2N} a_n n^{it} \right|^k dt \leq T^{o(1)} (TN^{k/2} + N^k).$

- Conjecture 1.2 implies Conjecture 1.3 which in turn implies the slightly weaker upper bound $|W| \leq T^{2-2\sigma+o(1)}$.
- Both of these conjectures are hard in general – we’re not expecting anyone to completely solve them.
- Bourgain showed that Conjecture 1.3 implies the Kakeya conjecture.
- Conjecture 1.3 implies the density hypothesis for $\zeta(s)$:

$$\# \{ \rho : \zeta(\rho) = 0, \Re(\rho) \geq \sigma, |\Im(\rho)| \leq T \} \leq T^{2-2\sigma+o(1)}.$$

- Lindelöf hypothesis $\leftrightarrow \left| \sum_N^{2N} n^{it} \right| \leq T^{o(1)} N^{1/2}$.
- Riemann hypothesis $\leftrightarrow \left| \sum_N^{2N} (\Lambda(n) - 1) n^{it} \right| \leq T^{o(1)} N^{1/2}$.

Remark 1.4. Is more known about Conjecture 1.2 when $a_n = 1$? Well, in that case, we can focus on the range $N \approx T^{1/2}$, but then outside the case $k \leq 4$, we don’t know much more than for a general 1-bounded sequence,

The “trivial bound” reads

$$\begin{aligned}
& \int_T^{2T} \left| \sum a_n n^{it} \right|^2 dt \\
& \leq \int \omega\left(\frac{t}{T}\right) \left| \sum_{n=N}^{2N} a_n n^{it} \right|^2 dt \\
& = \sum_{n,m} a_n \bar{a}_m \underbrace{\int \omega\left(\frac{t}{T}\right) \left(\frac{n}{m}\right)^{it} dt}_{\approx \begin{cases} 0 & \text{if } \left|\frac{n}{m} - 1\right| \geq \frac{c}{T}, \\ O(T) & \text{otherwise} \end{cases}} \\
& \leq \sum_{n=m(1+O(\frac{1}{T}))} T \leq TN + N^2.
\end{aligned}$$

Note that $|\sum a_n|^2$ is also a Dirichlet polynomial of length N^k , which implies Montgomery’s Conjecture 1.3 for k an even integer. More generally, one obtains

$$|W| \leq TN^{k(1-2\sigma)} + N^{2k(1-\sigma)} \quad \text{for any integer } k.$$

Key point: for many arithmetic applications, one just needs to beat trivial bounds for specific parameters.

Challenge 1: zero density estimates and primes in short intervals. This was the specific question motivating our work with Guth. As a consequence of our work, there is a new limiting case

$$\sigma = 7/10, \quad N = T^{5/13}, \quad (a_n)_N^{2N} = (1).$$

In this case, the trivial bounds come from looking at the *fourth* moment, which gives $TN^{2-4\frac{7}{10}}$. **Can you beat this?** Our work with Guth *does* beat this if $\sigma > 7/10$, but doesn’t use the following features:

- That $a_n = 1$.
- That $N = T^{5/13}$ is quite small. We instead square to get a Dirichlet polynomial of length $T^{10/13}$. It should, in principle, be possible to do better using that this polynomial is a square.
- We also can improve except when $|\{t_1 + t_2 - t_3 - t_4 = O(1), t \in W\}| = \frac{|W|}{T}$.

Challenge 2: moments of $\zeta(s)$. We know that $\int_T^{2T} |\zeta(\frac{1}{2} + it)|^4 dt \leq T^{1+o(1)}$, and expect the same for all moments, but the next interesting result is $\int_T^{2T} |\zeta(\frac{1}{2} + it)|^{12} dt \leq T^{2+o(1)}$. This implies in particular that $\int_T^{2T} |\zeta(\frac{1}{2} + it)|^\kappa dt \leq T^{\frac{1}{2} + \frac{\kappa}{8} + o(1)}$ for $4 \leq \kappa \leq 12$. This is the best we know for these intermediary exponents. The critical case is when $N = T^{1/2}$, $\sigma = 3/4$ (which corresponds to the zeta function taking values of size $T^{1/8}$), and you want to beat the trivial bound $T^{1/2}$ coming from the fourth moment, say for $(a_n)_N^{2N} = (1)$. (Our work with Guth was designed for $\sigma = 3/4$ but gives nothing in this case. We’re not using that you’re squaring a shorter polynomial.)

Challenge 3: progress towards the Lindelöf hypothesis. Bourgain: $|\zeta(\frac{1}{2} + it)| \leq T^{13/84+o(1)}$. To improve this bound, you typically want to show that

$$\left| \sum_{n=N}^{2N} n^{it} \right| \leq N^{\frac{1}{2}} T^{\frac{13}{84}+o(1)} \quad \text{for } N \leq T^{\frac{1}{2}}.$$

The hardest case again is morally when $N = T^{1/2}$, corresponding (if we have done our arithmetic correctly) to $\sigma = 17/21$. Want to show in that case that $|W| = 0$ (shoehorning this L^∞ bound into our original question).

Challenge 4: progress on the zero free region. This is somewhat more speculative. For the zero-free region, there is a method of Vinogradov–Korobov that gives

$$\left| \sum_{n=N}^{2N} n^{it} \right| \leq N^{1-\sigma_0}, \quad \sigma_0 \simeq \frac{(\log N)^2}{(\log T)^2}.$$

This is what's behind the best-known zero-free region. In particular, this is non-trivial for $N \geq \exp((\log T)^{2/3+o(1)})$. The question is now essentially, how small can we take N to get a nontrivial bound here? Essentially, nothing nontrivial is known for $N \leq \exp((\log T)^{2/3})$, so the challenge is to prove something nontrivial when N is of that size.

2. MAKSYM RADZIWILL, *Where we get stuck on the 12th moment of the Riemann zeta function*

$\zeta(s) = \sum_{n \geq 1} 1/n^s$ for $\Re(s) > 1$, analytic continuation to (e.g.) $\Re(s) > 0$, functional equation:

$$\xi(s) := \pi^{-s/2} \Gamma(s/2) \zeta(s) = \xi(1-s).$$

In view of this, we can think of $\Re(s) = 1/2$ as a “boundary set” for this analytic function: if you understand it there, you understand it everywhere else. Coincidentally, the Riemann Hypothesis asserts that all non-real zeros are on the line $\Re(s) = 1/2$.

One can use these properties to show that

$$\zeta(\frac{1}{2} + it) \approx \sum_{n < t} \frac{1}{n^{1/2+it}}$$

inside the critical strip. Riemann had the idea of turning this into a more efficient formula. The observation is based on Poisson summation, which gives

$$\sum_{n \sim N} \frac{1}{n^{1/2+it}} \approx e^{it \log t} \sum_{n \sim t/N} \frac{1}{n^{1/2-it}}.$$

There's thus a duality between N and t/N . Thus, writing $\sum_{n < t} = \sum_{n < \sqrt{t}} + \sum_{\sqrt{t} < n < t}$, splitting the sum into dyadic intervals, applying Poisson and recombining, we get

$$\sum_{\sqrt{n} < n < t} \frac{1}{n^{1/2+it}} \approx e^{it \log t} \sum_{n < \sqrt{t}} \frac{1}{n^{1/2-it}}.$$

The resulting formula

$$\zeta(s) = \sum_{n < \sqrt{t}} \frac{1}{n^{1/2+it}} + e^{it \log t} \sum_{n < \sqrt{t}} \frac{1}{n^{1/2-it}}$$

is computationally more efficient than the previous formula, and we expect it to be theoretically more efficient, too. (One can actually do better via Diophantine approximation. Start with the Annals paper of Ghaith–Hiary.)

The Riemann Hypothesis implies the Lindelöf Hypothesis which says that $\zeta(\frac{1}{2} + it) \ll_\varepsilon (1 + |t|)^\varepsilon$ for every $\varepsilon > 0$. Compare this with the trivial bound that we get from the approximate functional equation $\zeta(\frac{1}{2} + it) \ll (1 + |t|)^{1/4}$, called the convexity bound. In many cases, it is useful to know that the Lindelöf Hypothesis holds for “generic points” $T \leq t \leq 2T$. This is quantified by moments $\int_T^{2T} |\zeta(\frac{1}{2} + it)|^{2k} dt$. Since ζ can be approximated by a Dirichlet polynomial containing the constant term 1, we have $\int_T^{2T} |\zeta(\frac{1}{2} + it)|^{2k} dt \gg T$. On the other hand, the Lindelöf Hypothesis implies that $\int_T^{2T} |\zeta(\frac{1}{2} + it)|^{2k} dt \ll_\varepsilon T^{1+\varepsilon}$. It is thus conjectured that for all $k > 0$, we have

$$\int_T^{2T} |\zeta(\frac{1}{2} + it)|^{2k} dt = T^{1+o(1)}. \quad (2.1)$$

Since for any $T \leq t \leq 2T$,

$$\frac{1}{\log T} |\zeta(\frac{1}{2} + it)|^{2k} \ll \int_T^{2T} |\zeta(\frac{1}{2} + it)|^{2k} dt, \quad (2.2)$$

we see that the conjecture for moments implies the Lindelöf Hypothesis. They are thus equivalent. This was the original idea for introducing moments: embed ζ in some family of harmonics, which should make things easier to understand.

What do we know about moments? We get stuck very quickly. The conjecture (2.1) is known only for $k \leq 2$. To see this, we use the approximation functional equation and plug it into the moment, giving something like

$$\int_T^{2T} |\zeta(\frac{1}{2} + it)|^{2k} dt \ll \int_T^{2T} \left| \sum_{n \leq \sqrt{T}} \frac{1}{n^{1/2+it}} \right|^{2k} dt.$$

Pulling the k th power inside, we get $\int_T^{2T} \left| \left(\sum_{n \leq \sqrt{T}} \frac{1}{n^{1/2+it}} \right)^k \right|^2 dt$. This can be

rewritten $\int_T^{2T} \left| \sum_{n \leq T^{k/2}} \frac{c(n)}{n^{1/2+it}} \right|^2 dt$ for some coefficients $c(n) \ll_\varepsilon n^\varepsilon$. At this point, we introduce a smooth kernel $1 \leq \Phi(\frac{t}{T})$, for $0 \leq t \leq T$, say with the property that $\hat{\Phi}$ is compactly-supported. Thus we have to compute

$$\int_{\mathbb{R}} \left| \sum_{n \leq T^{k/2}} \frac{c(n)}{n^{1/2+it+iT}} \right|^2 \Phi\left(\frac{t}{T}\right) dt.$$

(Notice the shift n^{iT} .) Opening the square and integrating over T gives

$$T \sum_{n,m \leq T^{k/2}} \frac{c(n)\overline{c(m)}}{\sqrt{nm}} \left(\frac{n}{m}\right)^{iT} \hat{\Phi}\left(T \log \frac{n}{m}\right).$$

Since $\hat{\Phi}$ is compactly supported, the sum is restricted to values of n and m such that $|\log n/m| \leq 1/T$. These are two types of terms to consider: $n = m$, which we call the diagonal, and $n = m + h$ with $0 < |h| < m/T$, which we call the off-diagonal.

The contribution of the diagonal is $T \sum_{m \leq T^{k/2}} \frac{|c(m)|^2}{m} \ll_\varepsilon T^{1+\varepsilon}$, and is therefore OK. If $T^{k/2} < T$, then the off-diagonal does not exist. And if $T^{k/2} > T$, then the off-diagonal is approximately

$$\sum_{M=2^\ell \leq T^{k/2}} \frac{1}{M/T} \sum_{h \sim M/T} \sum_{m \sim M} c(m) \overline{c(m+h)} \left(1 + \frac{h}{m}\right)^{iT}.$$

(Even having an optimal bound for the shifted convolution problem would only solve the moment problem up to the 8th – to go beyond that, one would need further cancellation in the sum over h .)

When $T^{k/2} < T$, that is $k \leq 2$, there are no off-diagonal terms. This gives the conjecture in such cases. Those are still the only cases in which we know the correct exponent in the moment conjecture. Dealing with off-diagonal terms is hard, though not impossible. Heath–Brown’s 12th moment bound:

$$\int_T^{2T} |\zeta(\tfrac{1}{2} + it)|^{2t} dt \ll T^2.$$

Using (2.2), we deduce then that $\zeta(\tfrac{1}{2} + it) \ll (1 + |t|)^{1/6+o(1)}$. Using Holder’s inequality, we get

$$\int_T^{2T} |\zeta(\tfrac{1}{2} + it)|^{2k} dt \ll T^{1+(2k-4)/8}.$$

Those are the best results currently known for all moments with $k \leq 6$. There are results for higher k generalizing Heath–Brown’s result, but they are of little interest, and interpolation against them leads to weaker exponents for $k \leq 6$. I will explain shortly why they are of little interest. The distribution content of Heath–Brown’s result is that

$$\text{vol}\{|\zeta(\tfrac{1}{2} + it)| > V\} \ll \frac{1}{V^{12}} \int_T^{2T} |\zeta(\tfrac{1}{2} + it)|^{12} dt \ll \frac{T^2}{V^{12}}.$$

Compare it with the 4th moment, that gives

$$\frac{1}{V^4} \int_T^{2T} |\zeta(\tfrac{1}{2} + it)|^4 dt \ll \frac{T^{1+o(1)}}{V^4}.$$

This means that Heath–Brown’s result is stronger than the fourth moment for $V > T^{1/8}$ and weaker for $V < T^{1/8}$.

Thus the key point to improve on Heath–Brown’s result is to show that $\text{vol}\{|\zeta(\tfrac{1}{2} + it)| > T^{1/8}\} \ll T^{1/2-\delta}$ for some $\delta > 0$. This would improve the exponent in all moments with $4 < 2k < 12$. Improving the exponent in the eighth moment would have application to understand the variance of partial sums of $d_4(n)$, for instance.

What is the meaning of this level set? If we use the approximate functional equation, we need to understand when

$$\left| \sum_{n < \sqrt{T}} \frac{1}{n^{1/2+it}} \right| > T^{1/8}.$$

Imagine that this localized at $n \sim \sqrt{T}$ and $n^{1/2} \sim T^{1/4}$. We then need to understand the level set

$$\left| \sum_{n \sim \sqrt{T}} n^{it} \right| > T^{3/8}.$$

This means that we have a Dirichlet polynomial of length N and we need to understand the frequency of values that are larger than $N^{3/4}$. (This is the same range that causes trouble in the Huxley–Ingham bound, so maybe the result of Guth–Maynard is applicable? Unfortunately, it does not improve, because even Montgomery’s conjecture isn’t sufficient – one needs to use something more about the a_n .)

Perhaps looking at the value distribution is the wrong move here. Let’s try an attack through coefficients. To make things simple, let’s focus on the lowest moment that we can. The twelfth moment implies $\int_T^{2T} |\zeta(\frac{1}{2} + it)|^6 dt \ll T^{5/4}$. Let’s try to understand what it means in terms of coefficients. It is customary to think of $\zeta(s)^3$ as a GL_3 L -function corresponding to an Eisenstein series. Thus the sixth moment of $\zeta(s)$ is the same thing as the second moment of $\zeta(s)^3$. This should be similar to the second moment of a general GL_3 L -function, say $L(s, \pi)$. The advantage of looking at this variant is that the coefficients $\lambda_\pi(n)$ of a GL_3 L -function are oscillatory,

$$L(s, \pi) := \sum_{n \geq 1} \frac{\lambda_\pi(n)}{n^s}.$$

Thus we can think of $\lambda_\pi(n)$ as a variant of $d_3(n)$ but with the “main term” taken out. Let’s look, then, at

$$\int_T^{2T} |L(\frac{1}{2} + it, \pi)|^2 dt.$$

Applying the approximate functional equation

$$|L(\frac{1}{2} + it, \pi)| \ll \left| \sum_{n \leq T^{3/2}} \frac{\lambda(n)}{n^{1/2+it}} \right|$$

and opening the square, a piece of the off-diagonal terms looks like

$$\frac{1}{T^{1/2}} \sum_{h \sim T^{1/2}} \sum_{n \sim T^{3/2}} \lambda_\pi(n) \overline{\lambda_\pi(n+h)}.$$

At this stage a reasonable move is to move the sum over h inside and apply Cauchy–Schawrz, giving

$$\ll \frac{1}{T^{1/2}} \left(\sum_{n \sim T^{3/2}} |\lambda(n)|^2 \right)^{1/2} \left(\sum_{n \sim T^{3/2}} \left| \sum_{h \sim T^{1/2}} \lambda(n+h) \right|^2 \right)^{1/2}.$$

The best case scenario here is $\sum_{h \sim T^{1/2}} \lambda(n+h) \ll T^{1/4}$, which then just recovers Heath–Brown.

In other words, getting square-root cancellation on average in most short sums over $\lambda(n)$ would only recover Heath–Brown’s result. Thus, to go beyond $T^{5/4}$ it seems that we need to know genuine cancellation in $\sum_{n \sim T^{3/2}} \lambda(n) \overline{\lambda(n+h)}$. Currently we do not know the existence of a single explicit h such that $\sum_{n \sim N} \lambda(n) \overline{\lambda(n+h)} \ll N^{1-\delta}$, and correspondingly no asymptotics for $d_3(n)$. Heath–Brown’s bound of $T^{5/4}$ is consistent with this. We can have $\sum_{n \sim T^{3/2}} \lambda(n) \lambda(n+h) = \varepsilon_\pi T^{3/4}$.

Thus to go beyond in this approach we need not only square root cancellation in the signs ε_h , but also genuine cancellation in $\sum_{n \sim T^{3/2}} \lambda(n) \lambda(n+h)$ for most h . Part of me is afraid that we won’t make progress on the twelfth moment until we manage to show $\sum_{n \leq N} \lambda(n) \lambda(n+h) \ll N^{1-\delta}$.

Incidentally, the problem of computing the second moment of $L(\frac{1}{2} + it, \pi)$ is genuinely harder than the second moment problem for $\zeta(s)^3$. Only recently was it shown that

$$\int_T^{2T} |L(\frac{1}{2} + it, \pi)|^2 dt \ll T^{3/2-\delta}$$

for some $\delta > 0$.

Heath-Brown's original approach is based on finding a formula for $\int_t^{t+H} |\zeta(\frac{1}{2} + iu)|^2 du$ and then averaging over t using Halasz-Montgomery. Iwaniec developed an approach based on finding a formula for $\int_t^{t+H} |\zeta(\frac{1}{2} + iu)|^4 du$ and then averaging over t . Iwaniec's formula is obtained by using the approximate functional equation, and using Kuznetsov's trace formula on the off-diagonal terms ("Kloostermania"). This leads to a formula roughly of the shape

$$\int_T^{T+T^{2/3}} |\zeta(\frac{1}{2} + it)|^4 dt \approx \sum_{t_j} e^{it_j \log t_j} A_j(\frac{1}{2} + it) B_j(\frac{1}{2} + it),$$

where $\frac{1}{4} + t_j^2$ are eigenvalues of the hyperbolic Laplacian,

$$A_j(s) := \sum_{n \sim M} \frac{\lambda_j(n)}{n^s},$$

and $\lambda_j(n)$ is the n th Fourier coefficient of the automorphic form with eigenvalue $\frac{1}{4} + t_j^2$. Iwaniec then obtains an optimal bound for the above by taking absolute values and ignoring cancellations coming from $e^{it_j \log t_j}$. This approach gives $\zeta(\frac{1}{2} + it) \ll (1 + |t|)^{1/6}$.

Recently a sub-Weyl bound of the form

$$L(\frac{1}{2} + it, \pi) \ll (1 + |t|)^{1/6-\delta}$$

was claimed by Holowinsky-Munshi-Sharma-Streipel.

If correct, this might provide an opening for an improvement to Iwaniec's bound and eventually the 12th moments. The methods do not appear spectral in nature, but rather close to Bombieri-Iwaniec.

Despite quite a bit of work, the best we can offer a bound of the form

$$\int_T^{2T} |\zeta(\frac{1}{2} + it)|^{2k} dt = o\left(T^{1+(2k-4)/8}\right).$$

Maybe saving a small power of a logarithm. This gives a tiny improvement on Heath-Brown. This might be enough for the consequence for the variance of $\sum_{n \leq x} d_4(n)$.

What has not worked?

- Spectral theory and/or Munshi's method
- Circle method followed by Kuznetsov, which gives tautologies of the kind $A = A + o(1)$, where A on the LHS is the moment and A on the RHS is the same moment appearing as part of the Eisenstein spectrum.

The behavior of the $2k$ th moment of ζ is generally speaking tied with the Diophantine properties of $n^{1/k}$. For example, with de Faveri, we had an alternative proof of Heath-Brown's result using the double large sieve and the result of Robert-Sargas, giving an optimal bound for the number of quadruples (n_1, n_2, n_3, n_4) , all

of size N , such that

$$\left| \sqrt{\frac{n_1}{N}} + \sqrt{\frac{n_2}{N}} - \sqrt{\frac{n_3}{N}} - \sqrt{\frac{n_4}{N}} \right| \leq \frac{1}{X}.$$

The optimal bound is $\ll N^\varepsilon \left(\frac{N^4}{X} + N^2 \right)$. The proof uses an iterative method, induction on scales, bounding the next thing by the previous thing, so it doesn't give an asymptotic.

If one could refine the result of Robert–Sargos to an asymptotic with a powersaving, there is probably an opening for improving the twelfth moment.

We need this for all of X , but really X of size N^2 is the main range.

In general, any strong result (i.e., uniform in X) on the number of tuples such that

$$\left| \left(\frac{n_1}{N} \right)^{1/k} + \cdots + \left(\frac{n_{2\ell}}{N} \right)^{1/k} \right| \leq \frac{1}{X}$$

can probably be used to improve the 12th moment. But I do not believe such results (with $\ell > 2$) are within reach of the current technology. I am more agnostic on the asymptotic for Robert–Sargos, but the proof of such a result would have to be completely different from the proof of Robert–Sargos (a form of induction on scales). A radically different idea would be required there.

The new result on sub-Weyl subconvexity for GL_2 L -functions might offer ideas on the way forward.

3. DOMINIQUE MALDAGUE, *Difficulty of using decoupling ideas for generalized Dirichlet series and cubic Weyl sums*

We'll discuss limitations of decoupling techniques applied to cubic Weyl sums.

Fourier striction theory: control the behavior of functions $f(x)$ (usually in L^p -norm) based on properties of the Fourier (or frequency) support, emphasizing geometric properties.

Both examples are going to have inputs $x \in \mathbb{R}^2$. We'll use the notation $e(\bullet) = e^{2\pi i(\bullet)}$.

Example 3.1. $f(x) = \sum_{n=1}^N e(x \cdot (n, 0))$. We have $f(0) = N$, $|f(x)| \sim N$ if $|x_1| < \frac{1}{100N}$. We have

$$\int_{[0,1]^2} |f(x)|^2 dx = \sum_{n,m} \int_{[0,1]^2} e(x \cdot (n-m, 0)) dx = N.$$

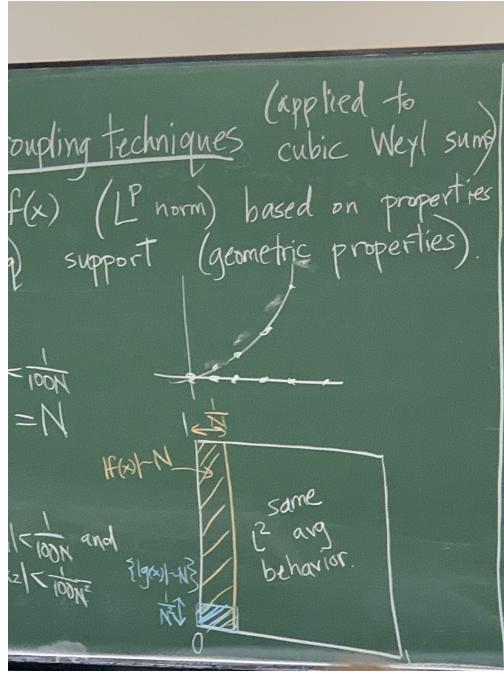
Example 3.2. $g(x) = \sum_{n=1}^N e(x \cdot (n, n^2))$. We have $g(0) = N$, $|g(x)| \sim N$ if $|x_1| < \frac{1}{100N}$ and $|x_2| < \frac{1}{100N^2}$. We have

$$\int_{[0,1]^2} |g(x)|^2 dx = N.$$

The first has frequency support consisting of integers in a line. The second has points on a parabola. We can also draw pictures indicating where the functions f and g are purely constructively intersecting. We get a smaller region for g than for f : same width, but cut off the height. It's a bit trickier to describe their intermediate behavior, but we can at least say that they have the same L^2 average.

Let's now consider $p \geq 2$, and ask: how large are the L^p -norms?

$$N^{p-1} + N^{p/2} \leq \int_{[0,1]^2} |f(x)|^p dx \leq N^{p-2} \int |f(x)|^2 dx = N^{p-1}.$$



$$N^p N^{-\beta} + N^{p/2} \leq \int_{[0,1]^2} |g(x)|^p dx.$$

In this last estimate, $N^{p/2}$ dominates for $2 \leq p \leq 6$, and in the same way, we have the same upper bound

$$\int_{[0,1]^2} |g(x)|^p dx \lesssim N^{p/2}. \quad (3.1)$$

This type of inequality, where you have square root cancellation behavior for higher p : This is exactly of what you call " L^p -decoupling". Easy to check that this is true:

Proof of (3.1). We have

$$\int_{[0,1]^2} |g(x)|^6 dx = \sum_{n_1, \dots, n_6} |g(x)|^6 dx = \sum_{(n_1, \dots, n_6)} \int e(x \cdot \phi(n)) dx,$$

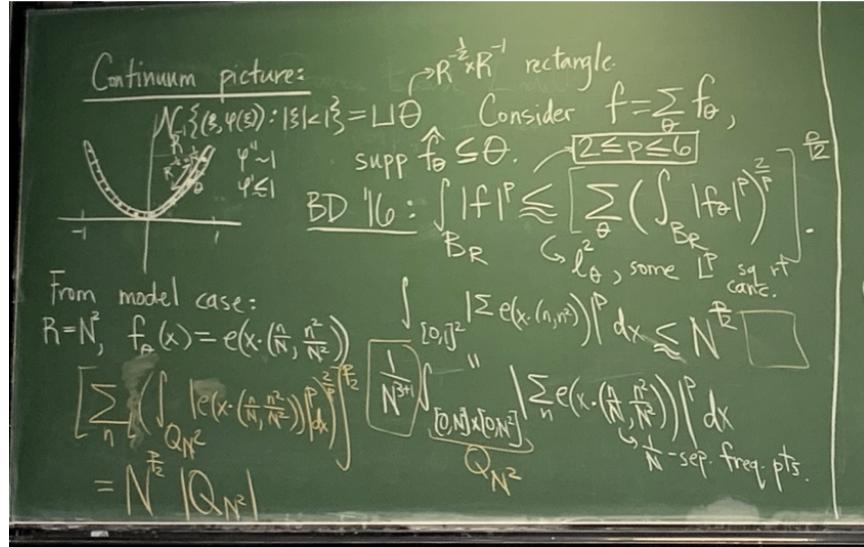
where

$$\phi(n) = (n_1 + n_2 + n_3 - n_4 - n_5 - n_6, n_1^2 + n_2^2 + n_3^2 - n_4^2 - n_5^2 - n_6^2).$$

The integral detects the condition $\phi(n) = 0$. We obtain the bound $\lesssim N^3$ by unique factorization. \square

What is the continuum picture? In the continuum setup, we have

$$\mathcal{N}_{R^{-1}} \{(\xi, \varphi(\xi)) : |\xi| < 1\} = \sqcup \Theta,$$



where each Θ is an $R^{-1/2} \times R^{-1}$ rectangle. Here $\varphi'' \sim 1$ and $\Phi' \lesssim 1$. We consider $f = \sum_\Theta f_\Theta$, where each $\text{supp } f_\Theta \subseteq \Theta$. Bourgain–Demeter '16: for $2 \leq p \leq 6$,

$$\int_{B_R} |f|^p \lesssim \left(\sum_\Theta \left(\int_{B_\Theta} |f_\Theta|^p \right)^{2/p} \right)^{p/12}.$$

From the model case:

$$\int_{[0,1]^2} \left| \sum e(x \cdot (n, n^2)) \right|^p dx \lesssim N^{p/12}.$$

The left hand side of the above is

$$\frac{1}{N^{3+1}} \int_{[0,N] \times [0,N^2]} \left| \sum e \left(x \cdot \left(\frac{n}{N}, \frac{n^2}{N^2} \right) \right) \right|^p dx.$$

We see that $R = N^2$, $f_\Theta(x) = e \left(x \cdot \left(\frac{n}{N}, \frac{n^2}{N^2} \right) \right)$. The frequency points $(\frac{n}{N}, \frac{n^2}{N^2})$ are $1/N$ -separated.

Writing $Q_{N^2} := [0, N] \times [0, N^2]$, we have

$$\left[\sum_n \left(\int_{Q_{N^2}} \left| e \left(x \cdot \left(\frac{n}{N}, \frac{n^2}{N^2} \right) \right) \right|^p dx \right)^{2/p} \right]^{p/2} = N^{p/2} |Q_{N^2}|.$$

Bourgain–Demeter decoupling gives

$$\int_{Q_{N^2}} \left| \sum_{\xi \in \Lambda} e(x \cdot (\xi, \varphi(\xi))) \right|^p dx \lesssim N^{p/2} |Q_{N^2}|$$

for $2 \leq p \leq 6$, $\Lambda \subseteq [-1, 1]$, $\geq \frac{1}{N}$ -separated.

- Sharp (as an inequality) and in the range of $p \in [2, 6]$.
- Very general.
- Small cap decoupling (Demeter–Guth–Wang), level set estimates (FGM), amplitude dependent wave envelope estimates (Guth–Maldaque).

Which exponential sum problems is decoupling useful for?

Conjecture 3.3. $\int |\sum_{n \sim N} b_n e(t \log n)|^p dt \lesssim (TN^{p/2} + N) \|b_n\|_\infty^p.$

The two terms on the right hand side come from square root cancellation and constant interval behavior, respectively.

The frequencies are $\{\log n\}$. This is a concave sequence: the terms are getting closer and closer to one another. What happens when you try to apply decoupling is that you don't really get anywhere. The reason is essentially that the tools of decoupling (wave packets, transversality, locally constant property) apply too generally: they apply to any concave or convex sequence, and some of these sequences have surprising structural properties (e.g., long arithmetic progression subsequences).

There's a related problem: additive energy for convex sequences.

Conjecture 3.4. $\int_{[0, N^2]} |\sum_{n \sim N} e(Ta_n)|^4 \lesssim N^4.$

(a_n) has minimal additive energy. Best bounds not from decoupling!

How about a 2D question? We saw

$$\int_{Q_{N^2}} \left| \sum_{\xi \in \Lambda} e(x \cdot (\xi, \varphi(\xi))) \right|^p dx \lesssim N^{p/2} |Q_{N^2}|$$

for $2 \leq p \leq 6$. What happens when we change the size of Q_{N^2} ? Shrinking \rightsquigarrow small cap decoupling regime. Expanding \rightsquigarrow replace Q_{N^2} by Q_{N^3} . We have

$$\begin{aligned} \int_{Q_{N^3}} |h(x)|^p dx &= \sum_{Q_{N^2} \subseteq Q_{N^3}} \int_{Q_{N^2}} |h(x)|^p dx \lesssim \sum_{Q_{N^2} \subseteq Q_{N^3}} N^{p/2} |Q_{N^2}| \\ &= N^{p/2} |Q_{N^3}|. \end{aligned} \quad (3.2)$$

This lack of sharpness has to do with the range $2 \leq p \leq 6$.

If you take what we saw at the beginning, namely

$$g(x) = \sum_{n \sim N} e\left(x \cdot \left(\frac{n}{N}, \frac{n^2}{N^2}\right)\right),$$

and we study it on this square of side-length N^3 . Let's look at the regions of constructive interference. From this, we see that the estimate in (3.2) is sharp, with a sharp range of exponents. Focus on $g(x) = \sum_{n \sim N} e\left(x \cdot \left(\frac{n}{N}, \frac{n^3}{N^3}\right)\right)$. Then

$$N^p N^2 + N^{p/2} |Q_{N^3}| \leq \int_{Q_{N^3}} |g(x)|^p dx \lesssim N^{p/2} |Q_{N^2}|.$$

Here $2 \leq p \leq 8$.

Current best: '14 (Wooley, verified $p = 9$) and '20 (Hughes–Wooley, verified $p = 10$). Hua 1947: $p = 6, p = 10$ versions.

Take decoupling proof approach to this conjecture (G–M–W).

$$\int_{H \subseteq Q_{N^3}} |g(x)|^8 dx \leq \int_{Q_{N^3}} |g(x)|^8 dx,$$

where $H = \{|g(x)| \sim N\}$.

Our goal is to show that $N^8|H| \lesssim N^4|Q_{N^3}|$. (It turns out that you only have to consider $\log N$ many level sets inside of here, so you don't have to, say, try to show this last estimate with $|H|$ and then add up.) We have

$$g(x) = \sum_{n \sim N} e\left(x \cdot \left(\frac{n}{N}, \frac{n^2}{N^2}\right)\right) = \sum_{m \sim \frac{N}{N^\varepsilon}} \underbrace{\sum_{k < N^\varepsilon} e\left(x \cdot \left(\frac{mN^\varepsilon + k}{N}, \frac{(mN^\varepsilon + k)^3}{N^3}\right)\right)}_{g_m(x)}.$$

Here we write $n = mN^\varepsilon + k$. On H , $|g(x)| \sim N$, $\sum_m |g_m(x)|^2 \sim N^{1+\varepsilon}$, thus

$$\sum_{m \sim N^{1-\varepsilon}} |g_m(x)|^2 = \sum_{m \sim N^{1-\varepsilon}} \sum_{k, k' < N^\varepsilon} e\left(x \cdot \left(\frac{k - k'}{N}, \frac{(mN^\varepsilon + k)^3 - (mN^\varepsilon + k')^3}{N^3}\right)\right).$$

We can reduce to considering the case $k \neq k'$, i.e., $|k - k'| > 0$, because the diagonal case contributes less than the total sum of squares.

Step 1 of “hi-lo” decoupling starts with a local L^4 estimate:

$$N^4 |H| \sim \int_H |g(x)|^4 dx \lesssim \int_{\mathcal{N}(H)} \left| \sum_m |g_m(x)|^2 \right|^2.$$

This says that we only need to consider the high-frequency part of the square function. We said earlier that this was true pointwise on the set H , but it's still true on this small enough neighborhood. So we just need to deal with the “high” part.

The normal decoupling proof notices that this exponential sum $g_m(x)$ satisfies its own L^2 -orthogonality estimate. We thereby obtain that the above satisfies, by the normal “hi-lo” proof of decoupling,

$$\lesssim \int \sum_m |g_m|^4 Q_{N^3},$$

which leads to the sharp L^6 estimate.

Let's write

$$\left[\sum_m |g_m(x)|^2 \right]_{\text{high}} = \sum_{\substack{k, k' < N^\varepsilon \\ |k - k'| > 0}} \sum_{m \sim N^{1-\varepsilon}} e\left(x \cdot \left(\frac{k - k'}{N}, \frac{k - k'}{N^3} \sigma\right)\right),$$

where

$$\sigma := 3m^2 N^{2\varepsilon} + 3mN^\varepsilon k' + k^2 + kk' + (k')^2.$$

argue with a The condition on x is basically that $x_2 \in [0, N^3]$. We then argue that

$$\frac{1}{N^2} \int_{\mathcal{N}(H)} \left| \sum_m \left[|g_m(x)|^2 \right]_{\text{high}} \right|^{2+2} \leq \frac{1}{N^2} \int_{Q_{N^3}} \left| \sum_m |g_m|^4 \right|^2 \sim |Q_{N^3}|.$$

4. BEN GREEN, *Properly understanding why nilpotent groups are important in additive combinatorial problems*

We'll talk today about what's called *higher-order Fourier analysis*. We'll give a slightly speculative discussion of what the state of the art is now, and what we'd like it to be in 10-20 years time.

The central objects of this subject are *Gowers norms*. Let $k \geq 2$ be an integer. For today, we consider functions $f : [N] \rightarrow \mathbb{C}$. The k th Gowers norm of f (which also depends upon N) is defined as follows:

$$\|f\|_{U^k[N]} := \left(\frac{c}{N^{k+1}} \sum_{x, h_1, \dots, h_k} f(x) \overline{f(x+h_1)} \cdots f(x+h_1+\cdots+h_k) \right)^{1/2^k}.$$

The last factor has a conjugate if k is odd. These turn out to define norms. Up to normalization, we have $\|f\|_{U^2} \leq \|f\|_{U^3} \leq \|f\|_{U^4} \leq \dots$

A two-dimensional parallelepiped is just a parallelogram. In k -dimensions, it will be a product of 2^k vertices of a parallelepiped.

Why is this a central object? They control counts of quite general linear configurations, e.g., arithmetic progressions. For instance, if $f_1, f_2, f_3, f_4 : [N] \rightarrow \mathbb{C}$ are functions with $|f_i(x)| \leq 1$ pointwise, then

$$\left| N^{-2} \sum_{x,d} f_1(x) f_2(x+d) f_3(x+2d) f_4(x+3d) \right| \ll \|f_i\|_{U^3[N]} \quad \text{for } i = 1, 2, 3, 4.$$

This result is known as a “generalized von Neumann theorem”. There are many generalizations of this result, to other patterns. This particular estimate is just three applications of Cauchy–Schwarz.

You can use this to reduce questions about arithmetic progressions (e.g., Szemerédi’s theorem, arithmetic progressions in primes) to questions about Gowers norms.

The Gowers norms are themselves a count of linear configurations in a set, weighted by a function. You can think of them as saying that parallelepipeds are universal objects in this theory.

We would like to understand when the Gowers norm of a function is large. Specifically, for $\delta > 0$, what can we say about $f : [N] \rightarrow \mathbb{C}$ with $|f(x)| \leq 1$ and $\|f\|_{U^k[N]} \geq \delta$? This is known as the *inverse problem* for Gowers norms.

Two fairly basic observations.

The first is that for the largest possible size 1 of a Gowers norm, we have the equivalence

$$\|f\|_{U^k[N]} = 1 \iff f(x) = e^{2\pi i \varphi(x)}, \quad \varphi : \text{polynomial of degree } \leq k-1.$$

Indeed, if $\|f\|_{U^k[N]} = 1$, then

$$f(x) \overline{f(x+h_1)} \cdots = e^{2\pi i (\varphi(x)-\varphi(x+h_1)-\cdots)}.$$

We think of the phase as $\partial_{h_1} \partial_{h_2} \cdots \partial_{h_k} \varphi(x)$. This forces φ to be a polynomial of the indicated degree.

The second is that if $\|f\|_{U^2[N]} \geq \delta$, then there exists θ such that

$$\frac{1}{N} \sum_{x \leq N} f(x) e^{-2\pi i \theta x} \gg \delta^2.$$

This exploits a special feature when $k = 2$, that the sum over parallelepipeds is given by the L^4 -norm of the Fourier transform:

$$\sum_{x, h_1, h_2} f(x) \overline{f(x+h_1)} \overline{f(x+h_2)} f(x+h_1+h_2) = \int_0^1 |\hat{f}(\theta)|^4 d\theta,$$

where $\hat{f}(\theta) = \sum_{x \leq N} f(x)e^{-2\pi i \theta x}$. This follows from orthogonality of characters. So the inverse theorem for the U^2 -norm says that it is large iff f correlates with some linear phase function.

The most naive conjecture for the inverse theorem for the U^k -norm would be that if $\|f\|_{U^k[N]} \geq \delta$, then

$$\frac{1}{N} \sum_{x \leq N} f(x)e^{-2\pi i \varphi(x)} \gg_{\delta} 1, \quad \varphi \text{ of degree } \leq k-1.$$

This turns out to be false. It was discovered in ergodic theory in a somewhat related context, then also by Gowers in a related context. There are exotic degree $k-1$ objects that enter the picture, called *nilsequences*. Let's give a rough definition modulo some details.

Definition 4.1. Let G be a simply-connected nilpotent Lie group of step s :

$$G = G_0 \geq G_1 \geq G_2 \geq \cdots \geq G_{s+1} = \{\text{id}\}, \quad G_{i+1} = [G, G_i].$$

Let $\Gamma \leq G$ be a lattice. A *nilsequence* is something of the form $(F(g^n))_{n=1}^{\infty}$, where $g \in G$ is fixed and $F : G \rightarrow \mathbb{C}$ satisfies $F(\gamma x) = F(x)$ for all $\gamma \in \Gamma$.

Example 4.2. $G = \begin{pmatrix} 1 & \mathbb{R} & \mathbb{R} \\ & 1 & \mathbb{R} \\ & & 1 \end{pmatrix}$ has $s = 2$. One can take $\Gamma = \begin{pmatrix} 1 & \mathbb{Z} & \mathbb{Z} \\ & 1 & \mathbb{Z} \\ & & 1 \end{pmatrix}$.

It turns out that nilsequences are natural “higher-degree characters”.

We'll now state the main theorem in the subject, the inverse theorem for the Gowers norms, which roughly states that nilsequences account for Gowers norms being large, but the version we'll state is a recent one, as of last year, that gives strong quantitative information.

Theorem 4.3 (Leng–Sah–Sawhney 2024, “Inverse theorem for the Gowers norms with quasi-polynomial dependencies”). *Let $s \geq 1$ be fixed. Suppose $\|f\|_{U^{s+1}[N]} \geq \delta$, where $f : [N] \rightarrow \mathbb{C}$ is 1-bounded. Then f correlates with one of these nilsequences*

$$\left| \frac{1}{N} \sum_{n \leq N} f(n)F(g^n) \right| \geq c_s(\delta),$$

where $c_s(\delta) \sim \exp(-\log(\frac{1}{\delta})^{C_s})$ and $\dim G \ll \log(\frac{1}{\delta})^{C_s}$, and various complexity parameters, e.g., “smoothness” of F , “rationality” of Γ , are $\ll \exp(\log(\frac{1}{\delta})^{C_s})$.

This is an impressive theorem that I didn't think I'd ever see proven. A non-quantitative version of the theorem was proved by Green–Tao–Ziegler some 15 years ago. Manners had a similar result that was much weaker quantitatively. Applications include improved bounds of Gowers on Szemerédi. Moreover, they've essentially removed any need for sieve theory (“pseudorandom measures”) from the theory of linear equations in primes; rather than putting the primes in a pseudorandom set, they can afford to work directly with the normalized characteristic function of the primes, because the error terms are now strong enough to absorb the logarithms.

Example 4.4. $e^{2\pi i \theta n} = F(g^n)$, $G = \mathbb{R}$, $\Gamma = \mathbb{Z}$, $g = \theta$, $F(x) = e^{2\pi i x}$.

What are the shortcomings?

- Well, currently, the proof of this is approximately 200 pages long (in a 100 page paper that “imports” a lot of additional work), whereas the statement is quite natural.
- Also, there are some weird things about the statement. The most weird thing is that you have to choose this G , Γ and F , and there are many possibilities for these. By comparison, for the U^2 -norm, there are only a few possibilities: we don’t say “take a Lipschitz function”, we just say “take $F(x) = e^{2\pi i x}$ ”. So it would be natural to ask whether there are some “natural” objects (G, Γ, F) that suffice, i.e., a natural refinement of the full class of nilsequences.
- Next, there should be polynomial bounds, not merely quasipolynomial. What is even the correct statement of that strength? We would like a statement that goes from having a large Gowers norm to correlating with a nilsequence:

$$\|f\|_{U^{s+1}} \geq \delta \implies \frac{1}{N} \sum_{n \leq N} f(n) F(g^n \Gamma) \geq c(\delta) \implies \|f\|_{U^{s+1}} \geq \delta^{O(1)}.$$

In the case of the $\|\cdot\|_{U^3}$ -norm, basically equivalent to the polynomial Freiman–Ruzsa conjecture in \mathbb{Z} (recently proved over finite fields).

- The proof is very complicated. I’ll give some brief indication in the moment. I think these nilpotent objects are quite natural, and there should be some way to explain where they really come from. Where do they come from, naturally? In the proof as-is, they arise in a very weird and *ad hoc* way. It’s unsatisfactory to have a natural statement with a very unnatural proof.

Let’s now give some sense of how complicated the proof is. Consider the U^3 -case:

$$\|f\|_{U^3[N]} \geq \delta.$$

We proceed by induction. We get that $\|\Delta_h f\|_{U^2} \gg_\delta 1$ for many h , where $\Delta_h f(x) = f(x)\overline{f(x+h)}$. If we plug in the U^2 -inverse theorem, we get that $\widehat{\Delta_h f}(\varphi(h)) \gg_\delta 1$. By a magical argument of Gowers (using Cauchy–Schwarz), φ has weak additive structure, which means that there are many additive quadruples $h_1 + h_2 = h_3 + h_4$ with $\varphi(h_1) + \varphi(h_2) = \varphi(h_3) + \varphi(h_4) + O(\frac{1}{N})$. Then one needs to understand the structure of these “somewhat additive functions”. The graph of such a function resembles a set of small doubling, but one needs the strongest quantitative forms of this, due to Sanders, which says that φ agrees with $\tilde{\varphi}$, a linear function on a Bohr set,

$$\tilde{\varphi} : \text{Bohr}(\theta_1, \dots, \theta_d, \frac{1}{10}) \rightarrow \mathbb{R}/\mathbb{Z},$$

with $d \ll (\log \frac{1}{\delta})^C$. One then needs to use classic geometry of numbers (Minkowski-type theorems) to deduce that the only such linear functions on a Bohr set are the obvious ones, namely

$$\tilde{\varphi}(x) = \alpha_1 \{\theta_1 x\} + \dots + \alpha_d \{\theta_d x\} + \alpha x + \beta.$$

Then, going back to the first step, what you have is that

$$\frac{1}{N^2} \sum_{x,h} f(x) f(x+h) e^{2\pi i \alpha_1 h \{\theta_1 x\} + \dots + 2\pi i \alpha_d h \{\theta_d x\}} \gg_\delta 1.$$

From this one deduces that

$$\frac{1}{N} \sum_{n \leq N} f(n) e^{2\pi i \alpha n \{\beta n\}} \gg_\delta 1,$$

or something like that. (But this is not a straightforward step – in the world of multi-dimensional maps, the derivative of a quadratic is a symmetric bilinear form, so there needs to be some additional symmetry argument here.) But the final step here, which is somehow the least satisfactory, is to observe in an *ad hoc* manner that such a bracket can be constructed on a Heisenberg group, using the sequence

$$\begin{pmatrix} 1 & -\lfloor \alpha n \rfloor & -\lfloor \alpha n \rfloor \beta n \\ & 1 & -\lfloor \beta n \rfloor \\ & & 1 \end{pmatrix} \begin{pmatrix} 1 & \alpha n & 0 \\ & 1 & \beta n \\ & & 1 \end{pmatrix}.$$

That was a sketch of the U^3 -case. In the higher order case, it gets much worse. You really need to understand how these higher order characters behave. You don't have orthogonality relations for them. You need to import the whole theory for that, which is a separate, substantial body of work.

Hopefully this gives some impression of the fact that the inverse theorem for Gowers norms is a fundamental black box in higher-order Fourier analysis and additive combinatorics, but as things stand, it seems we'll have to live with some enormously complicated proof – maybe there's a better way to do things, which would have other applications, e.g., to polynomial Freiman–Rusza over \mathbb{Z} .

Question 4.5. Could there be something like a natural orthonormal basis of nilsequences that describes the U^k -norm? Not obvious whether there is, and people have differing opinions. Seems like the system is overdetermined, but that's not a proof. The theory feels like we're doing Fourier analysis, but deprived of standard tools like orthogonality and Parseval.

5. SARAH PELUSE, *Inverse theorems for multidimensional Gowers norms*

Theorem 5.1 (Furstenberg–Katznelson, 1978). *Let $S \subseteq \mathbb{Z}^d$ be finite and nonempty. If $A \subseteq [N]^d$ has no nontrivial homothetic copies of S , i.e.,*

$$a + b \cdot S \quad (b \neq 0),$$

then $|A| = o_S(N^d)$.

Example 5.2. Take $d = 1$ and $S = \{0, 1, \dots, k-1\}$. This gives Szemerédi's theorem for length k arithmetic progressions (" k -APs").

The original proof was via ergodic theory and gave no estimates at all for the error term. There are now three effective proofs. From these you can extract a quantitative estimate for the size of A – you save an extremely tiny amount over N^d . For instance, when $|S| = 3$, you save \log_* , and it gets worse for larger S . The fundamental question in this area, stated in full generality by Gowers, is the following:

Problem 1 (Gowers, early 2000s). Prove a quantitative version of the multidimensional Szemerédi theorem with reasonable bounds:

$$|A| \ll_S \frac{N^d}{\log \cdots \log N},$$

where the number of log's is $\ll_S 1$.

Others have stated special cases of this problem, e.g.:

Question 5.3 (Graham, 1997). If $A \subseteq \mathbb{N}^2$ such that

$$\sum_{(n,m) \in A} \frac{1}{n^2 + m^2} = \infty,$$

then must A contain a $k \times k$ grid (so a homothetic copy of $[k] \times [k]$) for all $k \in \mathbb{N}$?

This is a two-dimensional version of a question of Erdős, which we now know in the affirmative.

Both of the above questions, in full generality, seem totally out of reach.

There are some cases where we know reasonable bounds:

- S is collinear. Look at all the cosets of that line. On each line, homothetic copies of S correspond to certain linear configurations in the integers. We know that for any fixed linear configuration, it will be some subset of some k -term arithmetic progression for k sufficiently large. You can now just apply Gowers's bounds (or now the Leng–Sah–Sawhney bounds) in the 1-dimensional case to the cosets to deduce reasonable bounds.
- Corners, i.e., homothetic copies of $(x, y), (x, y+d), (x+d, y)$ where $d \neq 0$, are the only genuinely multidimensional configuration for which we know reasonable bounds. This was first done by Shkredov, 2006.
- The speaker has a result for L -shapes $(x, y), (x, y+d), (x, y+2d), (x+d, y)$ in the finite model setting, and some work being written up in the integer setting.
- Today we'll talk about the case of four vertices of an axis-aligned square, i.e., $(x, y), (x, y+d), (x+d, y), (x+d, y+d)$. We cannot prove bounds for sets lacking such vertices.

Problem 2. Prove reasonable bounds for $A \subseteq [N]^2$ lacking homothetic copies of $\{(0, 0), (0, 1), (1, 0), (1, 1)\}$.

Let G be a finite abelian group. For us, G will always be a cyclic group $\mathbb{Z}/N\mathbb{Z}$ or \mathbb{F}_p^n . Given functions $f_0, f_1, f_2, f_3 : G^2 \rightarrow \mathbb{C}$, we define

$$\Lambda(f_0, f_1, f_2, f_3) := \mathbb{E}_{x,y,d \in G} f_0(x, y) f_1(x, y+d) f_2(x+d, y) f_3(x+d, y+d).$$

Thus

$$\Lambda(1_A, 1_A, 1_A, 1_A) = |G|^{-3} \# \{\text{axis-aligned squares in } A\}.$$

When the functions f_0, \dots, f_3 are 1-bounded, it follows from a couple applications of the Cauchy–Schwarz inequality that $|\Lambda(f_0, f_1, f_2, f_3)| \leq \|f_3\|$, where

$$\|f\|^8 = \mathbb{E}_{x,y,h,k,\ell \in G} \Delta_{(h,0),(0,k),(\ell,\ell)} f(x, y)$$

where

$$\Delta_{(a,b)} f(x, y) := f(x, y) \overline{f(x+a, y+b)}.$$

In order to study axis-aligned squares, we want an inverse theorem for this norm: we want a classification of 1-bounded f such that $\|f\| \geq \delta$ for general $\delta > 0$.

Let's discuss first the easiest situation: proving a 100% inverse theorem. That is, what must f look like if $\|f\| = 1$. If the norm of f is 1, we must have $f(x, y) = e(\phi(x, y))$ for some $\phi : G^2 \rightarrow \mathbb{T}$, where the corresponding additive discrete derivatives of ϕ must vanish: with

$$\partial_{(a,b)} \phi(x, y) := \phi(x, y) - \phi(x+a, y+b),$$

we have

$$\partial_{(h,0),(0,k),(\ell,\ell)}\phi(x,y) \equiv 0.$$

The problem is then to classify ϕ satisfying this last equation. An easier subproblem is to classify functions such that the 2-fold discrete derivative in the directions $(h, 0)$ and $(0, k)$ is identically zero. That is, when is $\partial_{(h,0),(0,k)}\psi(x,y) \equiv 0$? Then,

$$\psi(x,y) = \psi(x+h,y) + \psi(x,y+k) - \psi(x+h,y+k) \quad (5.1)$$

for all $x, y, h, k \in G$. By a change of variables, we get that $\psi(x,y) = \psi(h,y) + \psi(x,k) + \psi(h,k)$. If we now just fix h and k , say we fix both to be zero, then we get $\psi(x,y) = \psi(0,y) + \psi(x,0) + \psi(0,0)$. The upshot is that we can write and ψ satisfying the functional equation (5.1) as

$$\psi(x,y) = a(x) + b(y)$$

for some one-variable functions $a, b : G \rightarrow \mathbb{T}$. To analyze this, write

$$\partial_{(n,0),(0,k),(\ell,\ell)}\phi(x,y) = \partial_{(h,0),(0,k)}[\partial_{(\ell,\ell)}\phi](x,y).$$

Then, $\partial_{(\ell,\ell)}\phi(x,y) = a(\ell,x) + b(\ell,y)$ for all $x, y, \ell \in G$.

The next step is to understand how a and b depend upon ℓ ; if we can do that, then we can say something interesting about ϕ . We have

$$\begin{aligned} \partial_{(\ell,\ell)}\phi(x,y) &= [a(\ell,x) - a(x+\ell,0) + a(x,0)] + [a(x+\ell,0) - a(x,0)] \\ &\quad + [b(\ell,x) - b(x+\ell,0) + b(x,0)] + [b(x+\ell,0) - b(x,0)] \\ &= a'(\ell,x) + a''(\ell,x) + b'(\ell,y) + b''(\ell,y). \end{aligned}$$

We use that $\partial_{(\ell,\ell)}\phi(x,y) = \phi(x,y) - \phi(x+\ell,y+\ell)$, so

$$\partial_{(\ell+\ell',\ell+\ell')}\phi(x,y) = \partial_{(\ell,\ell)}\phi(x,y) + \partial_{(\ell',\ell')}\phi(x+\ell,y+\ell).$$

If we apply this last identity to the right hand side of the previous equality, then we get the corresponding equality for a', a'', b', b'' :

$$\begin{aligned} &a'(\ell+\ell',x) - a'(\ell,x) - a'(\ell',x+\ell) \\ &\quad + a''(\ell+\ell',x) - a''(\ell,x) - a''(\ell',x+\ell) \\ &\quad + b'(\ell+\ell',y) - b'(\ell,y) - b'(\ell',y+\ell) \\ &\quad + b''(\ell+\ell',y) - b''(\ell,y) - b''(\ell',y+\ell) = 0. \end{aligned}$$

We see that the second and fourth lines here vanish, so that

$$a'(\ell+\ell',x) - a'(\ell,x) - a'(\ell',x+\ell) = -b'(\ell+\ell',y) + b'(\ell,y) + b'(\ell',y+\ell).$$

Now the left (resp. right) hand side is independent of y (resp. x), so both sides are equal to $c(\ell, \ell')$, and after a few manipulations, we see that in fact $c(\ell, \ell') = -a'(\ell', \ell)$. The upshot is that

$$a'(\ell+\ell',x) - a'(\ell,x) - a'(\ell',x+\ell) + a'(\ell',\ell) = 0$$

for all $x, \ell, \ell' \in G$. This last equation shows that a' is a 2-cocycle for the trivial action $G \curvearrowright \mathbb{T}$. For $f : G^k \rightarrow \mathbb{T}$, we define $d^{k+1}f : G^{k+1} \rightarrow \mathbb{T}$ by setting

$$\begin{aligned} (d^{k+1}f)(x_1, \dots, x_{k+1}) &= f(x_2, \dots, x_{k+1}) + \sum_{i=1}^k (-1)^i f(x_1, \dots, x_{i-1}, x_i + x_{i+1}, x_{i+2}, \dots, x_{k+1}) \\ &\quad + (-1)^{k+1} f(x_1, \dots, x_k). \end{aligned}$$

We say that f is a k -cocycle if $d^{k+1}f = 0$, and a k -coboundary if it is of the form $d^k g$. Then

$$H^k(G, \mathbb{T}) = \{k\text{-cocycles}\} / \{k\text{-coboundaries}\}.$$

The possibilities for a' depend on $H^2(G, \mathbb{T})$. The answer is different in our two cases.

- $H^2(\mathbb{Z}/N\mathbb{Z}, \mathbb{T}) = 0$. Then we must have $\phi(x, y) = a(x) + b(y) + c(x - y)$.
- $H^2(\mathbb{F}_p^n, \Gamma) \neq 0$ when $n \geq 2$. Then $\phi(x, y) = a(x) + b(y) + c(x - y) + B(x, y)$, where B is a skew-symmetric bilinear form.

Note that we don't know any proofs for this 100% inverse theorem that don't use cohomology, even in disguise. We (Ben and I) can prove a 99% inverse theorem for $\|\cdot\|$ with polynomial bounds. Specifically, we can show that if $\|f\| \geq 1 - \varepsilon$ for some small enough $\varepsilon > 0$, then f satisfies

$$\mathbb{E}_{x,y} f(x, y) a(x) b(y) c(x - y) e(u(y, x - y)) \geq 1 - O(\varepsilon^2).$$

, where u is a 2-cocycle. We would like an inverse theorem for this norm where $1 - \varepsilon$ is replaced by (say) ε that says that f correlates with some structured version.

Question 5.4. Why does group cohomology come up here? We don't understand why these elaborate algebraic manipulations work.

6. EMMANUEL KOWALSKI, *Short sums of trace functions: examples, conjectures and applications*

6.1. What are trace functions? They lie at the intersection of algebraic geometry, harmonic analysis, and number theory. Intuitively, a *trace* function is a function $t : k \rightarrow \mathbb{C}$ for k a finite field $k = \mathbb{F}_p$, or $t : k^d \rightarrow \mathbb{C}$, or $V(k) \rightarrow \mathbb{C}$ where V/k is an algebraic variety, of “algebraic origin”.

To get some intuition, it is best to think of examples. In all cases, the value $t(x)$ at a point is of the form $t(x) = \text{trace } \theta(x)$, where $\theta(x)$ is of some size $r(x)$ with eigenvalues λ such that $|\lambda| = |k|^{w(\lambda)/2}$, with $w(\lambda) \in \mathbb{Z}$.

Moreover, there is a numerical invariant $c(t)$ which measures the arithmetic complexity of t . (Fouvry–Kowalski–Michel for $t : k \rightarrow \mathbb{C}$, Sawin for general case.)

6.2. Examples.

- (1) Let
- $f \in k(X)$,
 - χ a multiplicative character, $\chi \neq 1$,
 - ψ an additive character, $\psi \neq 1$.

Then $x \mapsto \chi(f(x))$ and $x \mapsto \psi(f(x))$ are trace functions (take the value to be 0 at poles or zeroes), with complexity bounded in terms of the degree (numerator and denominator) of f .

- (2) *Formalism.* Easy: if t, t_1, t_2 are trace functions, then so are

- $t_1 + t_2$,
- $t_1 t_2$ and
- \bar{t} .

Moreover, one can easily bound the resulting conductors.

A much deeper example, due to Deligne, is that the class of trace functions is stable by Fourier transform

$$\hat{t}(y) = \frac{1}{|k|^{d/2}} \sum_{x \in k^d} t(x)\psi(x \cdot y),$$

where $\psi \neq 1$ is a fixed additive character. One can bind the conductor of \hat{t} (linearly) in terms of the conductor of t (Fouvry–Kowalski–Michel for $d = 1$, Sawin for $d \geq 2$). Moreover, the Fourier transform preserves the weight. These results contain very general forms of the Riemann Hypothesis over finite fields.

Example 6.1. Taking one of the basic examples $t(x) = e(\frac{\bar{x}}{p})$ (here $x\bar{x} \equiv 1 \pmod{p}$) and computing the Fourier transform, we obtain

$$\frac{1}{\sqrt{p}} \sum_{x \in \mathbb{F}_p^\times} e\left(\frac{xy + \bar{x}}{p}\right).$$

This is a Kloosterman sum. Together with the bounds coming out of the Riemann Hypothesis, one gets “for free” from Deligne’s result concerning the Fourier transform that these are bounded as y and p vary.

6.3. Riemann Hypothesis. Let’s now state the Riemann Hypothesis in a way that is recognizable as a statement in harmonic analysis. The statement requires a couple definitions.

Definition 6.2. We say that t is of *weight 0* if all eigenvalues of $\theta(x)$ have weight $w(\lambda) = 0$ (i.e., $|\lambda| = 1$).

Definition 6.3. A trace function t of weight 0 is *geometrically irreducible* if

$$\sum_{x \in k} |t(x)|^2 \sim |k|.$$

(This definition is not quite rigorous, but is good enough for our purposes.)

Theorem 6.4 (Deligne, 1981). *Let $t_1, t_2 : k \rightarrow \mathbb{C}$ trace functions in one variable of weight 0 and geometrically irreducible. Then either*

- (1) (“Causation”) There exists $\alpha \in \mathbb{C}$, with $|\alpha| = 1$, such that $t_1 = \alpha t_2$.
- (2) (“Non-correlation”) $\sum_{x \in k} t_1(x)\overline{t_2(x)} = O(\sqrt{|k|})$, with implied constant depending only on $c(t_1)$ and $c(t_2)$ (in fact the product of these conductors basically suffices).

Example 6.5. Take $t_2 = 1$ and $t_1(x) = e(\frac{\bar{x}+yx}{p})$, where $y \in \mathbb{F}_p^\times$. Then $\sum t_1(x)\overline{t_2(x)}$ is a Kloosterman sum, hence if $O(\sqrt{p})$.

Example 6.6 (Gowers norms of trace functions (Fouvry–Kowalski–Michel)). Inverse theorem of the form: if $t : \mathbb{F}_p \rightarrow \mathbb{C}$ is a trace function of weight 0 and geometrically irreducible, then either

- $\|t\|_{U^d}$ is as small as it can be (i.e., something like $p^{-1/2^d}$, as if t were a random function), or
- $t(x) = \alpha e(\frac{f(x)}{p})$ for some $\alpha \in \mathbb{C}$ with $|\alpha| = 1$ and some $f \in \mathbb{F}_p[X]$ with $\deg f \leq d - 1$.

This follows by arguments like in Ben Green’s lecture, but each step becomes exact, hence simpler. Let’s now discuss short sums.

6.4. Short sums. In applications, one often has to deal with *short* sums $\sum_{1 \leq x \leq N} t(x)$ for some $N \leq p$, e.g., $N = p^\theta$ for some $\theta < 1$.

- (1) Least quadratic non-residue modulo p : want

$$\left| \sum_{1 \leq n \leq N} \left(\frac{n}{p} \right) \right| < N$$

for N as small as possible.

- (2) (Kowalski–Sawin) Given p and $y \in \mathbb{F}_p^\times$, define a continuous function $K_p(y) : [0, 1] \rightarrow \mathbb{C}$ by requiring that

$$\frac{i}{p-1} \mapsto \frac{1}{\sqrt{p}} \sum_{1 \leq x \leq i} e\left(\frac{xy + \bar{x}}{p}\right).$$

Then ask, does this have a limiting distribution? What we prove (among other things) is the following conditional result.

These functions behave statistically like a specific random Fourier series

$$\text{ST}_0 + \sum_{k \neq 0} \frac{e(ht) - 1}{2\pi ih} \text{ST}_h,$$

where the ST_i are independent random variables on $[-2, 2]$ given by the semi-circle (or Sato–Tate) distribution, *provided* that there exists $\delta > 0$ such that one has nontrivial bounds for

$$\frac{1}{\sqrt{p}} \sum_{x \in I} e\left(\frac{xy + \bar{x}}{p}\right)$$

for $I \subseteq \mathbb{F}_p^\times$, an interval of length $|I| \leq p^{1/2-\delta}$. (Bourgain tried this, and could make it work with a bit of averaging over p , but it seems like a hard problem for a fixed prime.)

6.5. Short sums conjecture.

Conjecture 6.7. Let p be a prime, $d \geq 2$, $\varepsilon > 0$. Then there exists $\delta > 0$ such that for all trace functions t_1, t_2 modulo p , of weight 0 and geometrically irreducible, such that

$$\sum_{1 \leq x \leq N} t_1(x) \overline{t_2(x)} = O_{c(t_1), c(t_2)}(N^{1-\delta})$$

if $N \geq p^{1/d+\varepsilon}$, unless $t_1 \overline{t_2}(x) = \alpha e\left(\frac{f(x)}{p}\right)$ for $f \in \mathbb{F}_p[X]$ of degree $\leq d$.

Example 6.8. Take $t_2 = 1$ and $t_1 = \psi(f)$, f a polynomial of degree at most d .

Example 6.9. (1) Take $d = 2$. Then the conjecture gives $\sum_{1 \leq x \leq N} t(x) = O(N^{1-\delta})$, unless t is an additive character. This is true (Pólya–Vinogradov, Fouvry–Kowalski–Michel). It yields a general form of Pólya–Vinogradov.

- (2) Arbitrary d . Take $t(x) = e\left(\frac{g(x)}{p}\right)$ (Weyl-differencing).
(3) Bourgain–Glibichuk–Konyagin–Chang, $t(x)$ the characteristic function of a multiplicative subgroup (additive combinatorics).

7. PHILIPPE MICHEL, CONTINUATION OF PREVIOUS LECTURE

We focus first on the trace function $t = \chi$, with χ a Dirichlet character. One would like to show that $\sum_{n \sim N} \chi(n) \ll N^{1-\delta}$. One can do so for $N \geq q^{1/4+\eta}$.

Let's discuss the “ $+ab + uv$ ” trick. With $UV = o(N)$, we have

$$\begin{aligned} \sum_n \chi(n) &\approx \frac{1}{UV} \sum_{u,v \sim U \times V} \sum_{n \sim N} \chi(n + uv) \\ &\approx \frac{1}{UV} \sum_{u,n \sim U \times N} \chi(u) \sum_{v \sim V} \chi(\bar{v}n + v). \end{aligned}$$

After Hölder, one gets

$$\frac{1}{2\ell} + 1 - \frac{1}{2\ell} \frac{1}{UV} (UN)^{1-\frac{1}{2\ell}+o(1)} \left(\sum_{\underline{v} \sim V^{2\ell}} \varepsilon_{\underline{v}} \sum_{x \in \mathbb{F}_p} \prod_{i=1}^{\ell} \frac{\chi(x + v_i)}{\chi(x + v_{i+\ell})} \right)^{1/2\ell}.$$

We can rewrite the inner sum as

$$\sum_{\underline{v} \sim V^{2\ell}} \varepsilon_{\underline{v}} \sum_{x \in \mathbb{F}_p} \chi \left(\frac{x + v_i}{x + v_{i+\ell}} \right).$$

The $\chi(\dots)$ term is a trace function, so one would usually expect the x -sum to be $\ll_{\ell} p^{1/2}$, but this is not always true: if your vector \underline{v} is in a kind of diagonal state (for instance, $v_i = v_{i+\ell}$ for every i), then there is no cancellation, but this does not occur very often. One obtains

$$\ll V^{\ell} p + V^{2\ell} p^{1/2}.$$

To balance the two terms, one takes $V = p^{1/2\ell}$.

Instead of bounding a single sum, we will study this question on average and bound bilinear sums. With t a trace function, one would like to show that

$$\sum_{m,n \sim M \times N} \alpha_m \beta_n t(mn) = O((MN)^{1-\delta})$$

for M and N large enough. These are called “Type II” sums. There are also the simpler “Type I” sums, for which one wishes to show

$$\sum_{m,n \sim M \times N} \alpha_m t(mn) = O((MN)^{1-\delta}).$$

Theorem 7.1. *Let $M, N \geq p^{\eta}$, say with $M, N < p$. For “suitable” t , one has*

- the “Type II wish” if $MN \geq p^{3/4+\eta}$ and
- the “Type I wish” if $MN^2 \geq p^{1+\eta}$.

For instance, with $M = N$, these conditions read $N \geq p^{3/8+\eta/2}$ and $N \geq p^{1/3+\eta/3}$. What's important is that the exponent $3/8$ and $1/3$ are both smaller than $1/2$, which is the natural Pólya–Vinogradov limit.

Cases where a theorem of this kind (for Type II sums) was obtained:

- Karatsuba–Vinogradov: $t(x) := \chi(x+2)$. What they proved is that this function oscillates if you sum it along the prime integers in the interval $[1, p]$.
- Friedlander–Iwaniec and Birch–Bombieri: $t(x) = e(\frac{\bar{x}}{p})$, in their work on the ternary divisor function in large arithmetic progressions.

- Fouvry–Michel: $t(x) = e\left(\frac{f(x)}{p}\right)$.
- Kowalski–Michel–Sawin: $t(x) = \text{Kl}_2(x)\text{Kl}_k(x)$, motivated by moments of L -functions.
- Dunn–Zaharescu:

$$t(x) := \text{Salié}(x) = \frac{1}{p^{1/2}} \sum_{y \in \mathbb{F}_p} \left(\frac{y}{p}\right) e\left(\frac{\bar{y} + xy}{p}\right) \doteq 2 \sum_{y^2=x} e\left(\frac{y}{p}\right).$$

- Fouvry–Kowalski–Michel–Sawin: $\text{sym}(\text{Kl}_2)(f(x))$, for any non-constant rational function $f \in \mathbb{F}_p(x) - \mathbb{F}_p$. The proof is less *ad hoc* than the earlier one of Kowalski–Michel–Sawin. It's also more flexible: one can work with $t(m^a n^b)$ rather than $t(mn)$ in the bilinear form.

Example 7.2. We can treat stuff like, for $(a, b, c) = 1$,

$$K_{a,b,c}(x) = \frac{1}{p} \sum_{\substack{x_1^a x_2^b x_3^c = y \\ x_1 x_2 x_3 = y}} e\left(\frac{x_1 + x_2 + x_3}{p}\right).$$

What we're motivated by are cubic moments like

$$\sum_{\chi(p)} L(\chi^a, \frac{1}{2}) L(\chi^b, \frac{1}{2}) L(\chi^c, \frac{1}{2}) = \text{MT} + O(q^{1-\delta}).$$

The crucial step in getting such an asymptotic formula is the Type I sum for $K_{a,b,c}$:

$$\frac{1}{UV} \sum_{u,v \sim U \times V} \sum_{m,n} \alpha_m t(um(\bar{u}n + uv)).$$

Here $UV = o(N)$. After Hölder, one obtains a sum over a tuple

$$\sum_{v \in V^{2\ell}} \varepsilon_{\bar{V}} \sum_{r,s \in \mathbb{F}_p \times \mathbb{F}_p^\times} \prod_{i=1}^{\ell} t(s(r+v_i)) \bar{t}(s(r+v_{i+\ell})).$$

We call this $\sum_{I,\ell}(\underline{v}, \mathbb{F}_p)$. We say that \underline{v} is

- *very bad* if $\sum_I \ll p^2$
- *not too bad* if $\sum_I \ll p^{3/2}$, and
- *good* if $\sum_I \ll p$.

The set of very bad (resp. not too bad) \underline{v} is the intersection of $[V, 2V]^{2\ell}$ with a subvariety $V_{\text{vb}}(\mathbb{F}_p)$ of dimension ℓ (resp. $V_{\text{ntb}}(\mathbb{F}_p)$ of dimension $\leq \ell + \frac{\ell}{2}$).

J. Xu: “seize the moment”.

$$\sum_{\underline{v} \in \mathbb{F}_p^{2\ell}} \left| \sum_{I,\ell} \right|^{2m} \ll p^{2m+2\ell} + p^{4m+\ell}.$$

Given a trace function $t(x)$ on \mathbb{F}_p , one obtains a family of trace functions $t(x; k)$, defined for finite extensions k/\mathbb{F}_p and $x \in k$ by precomposing with the trace: $t(x; k) = t(\text{trace}_{k/\mathbb{F}_p}(x))$. It's important to have more generally

$$\sum_{\underline{v} \in k} \left| \sum_{I,\ell} \right|^{2m} \ll |k|^{2m+2\ell} + |k|^{4m+\ell}.$$

We have

$$\sum_{x \in k} t(x; k) = \sum_{i=0}^2 (-1)^i \operatorname{trace} \left(\operatorname{Frob}_p^d \mid H_c^i \right), \quad d = [k : \mathbb{F}_p].$$

We then deduce bounds for trace functions from bounds for the trace of Frobenius acting on this finite-dimensional space. Suppose that you want to know what is the dimension of the space H_c^i . Suppose for instance that the eigenvalues of $(\operatorname{Frob}_p^d \mid H)$ are of modulus $p^{w/2}$. Then one can show that

$$\dim H_c^1 = \frac{1}{|k|^{w/2}} \overline{\lim}_{d \rightarrow \infty} \operatorname{trace} \left(\operatorname{Frob}_p^d \mid H \right).$$

(Similar in spirit to Turán power sums.)

We open the moment, switch summations and factor into a moment:

$$\sum_{v \in k^{2\ell}} \left| \sum_{r,s} \right|^{2m} = \sum_{(\underline{r}, \underline{s}) \in (k \times k^\times)^{2m}} \left| \sum_{v \in k} \prod_{j=1}^m t(s_j(v + r_j)) \bar{t}(s_{j+m}(v + r_{j+m})) \right|^{2\ell}.$$

Let ρ_t denote the Galois representation associated with G_{geom} , the geometric Galois group inside $\operatorname{GL}(V_\rho)$. If $G_{\text{geom}}^0 \neq \{1\}$ (i.e., G_{geom} is not finite), then we say that t is *suitable* if G_{geom}^0 is simple and acts irreducibly. The simplicity condition is useful because it is very stable and robust.

8. HONG WANG, *Danzer's problem and quantitative Besikovitch projection theorem*

The problem is very simple to state. Let $\mathcal{D} \subseteq \mathbb{R}^2$ be a discrete set of points in the plane. The *density* $d(\mathcal{D})$ of \mathcal{D} is defined to be

$$d(\mathcal{D}) := \lim_{T \rightarrow \infty} \sup \frac{\#(\mathcal{D} \cap B(0, T))}{T^2}.$$

Problem 3 (Danzer's problem, 1965). Does there exist a set of finite density intersecting any convex body of area 1?

Here you can replace “convex body” with “rectangle”. The difficulty is that you need to guarantee that *any* rectangle of area 1 (say of length $1/\delta$ and width δ) must intersect at least one point in your discrete series.

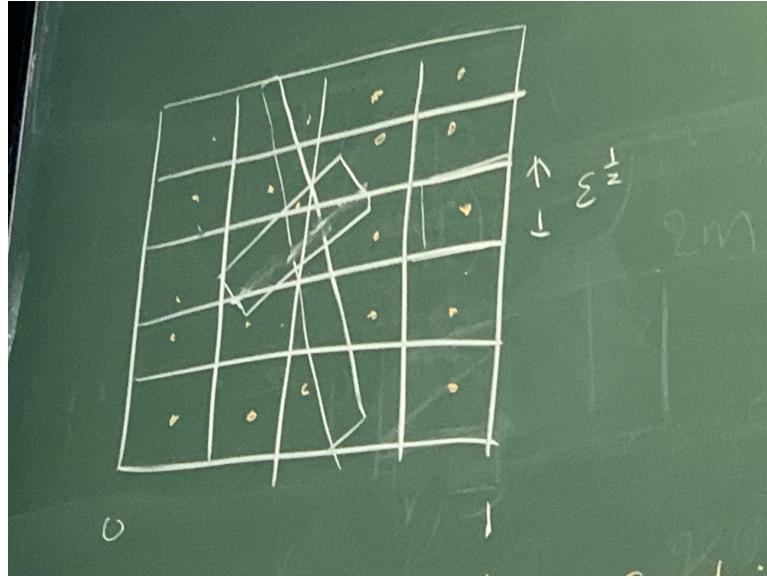
Imagine the elements in your discrete set are particles. Consider their δ -neighborhoods. The particles have some spacing. The question is, if you randomly choose a direction, can you guarantee that you'll hit one of your particles up to error δ within time $1/\delta$.

Many people have independently asked some version of this problem. It's related to Besikovitch projection theorem. We'll write down some equivalent statements and some partial results.

We say that $N_\varepsilon \subseteq [0, 1]^2$ is an ε -net if N_ε intersects any rectangle of area ε . Then Danzer's problem is equivalent to the following finite version:

Problem 4 (Danzer–Rogers problem, 1996). Given $\varepsilon > 0$, does there exist an ε -net N_ε of cardinality $\#N_\varepsilon = O(\frac{1}{\varepsilon})$?

Haussler–Welzl (1987): there exists an ε -net $N_\varepsilon \subseteq [0, 1]^2$ such that $\#N_\varepsilon = O(\varepsilon^{-1} \log(\varepsilon^{-1}))$.



How do we visualize such a problem? Note that by considering squares of width $\varepsilon^{1/2}$, we get a lower bound of $\gg \varepsilon^{-1}$ for such a set. It's known that if you allow the rectangles to be a bit – replacing boxes by *quasi-boxes*, bended by a bounded angle – then the answer to the question is “no”.

Question 8.1 (Gowers's problem). Does there exist a set $\mathcal{D} \subseteq \mathbb{R}^2$ and a constant $m \geq 1$ such that for any box K of area 1,

$$1 \leq \#K \cap \mathcal{D} \leq m?$$

Answer 1 (Solan–Solomon–Weiss, 2017). No.

This is saying that for every fixed set \mathcal{D} , we can find a box K of area 1 so that $K \cap \mathcal{D}$ is either empty or has cardinality $\gg 1$.

An ε -net N_ε contains a *well-spaced set* M_ε of points: any $\varepsilon^{1/2}$ -square contains ≥ 1 points and for any $x, y \in M_\varepsilon$, $|x - y| \gtrsim \varepsilon^{1/2}$.

Problem 5 (Boshernitzan–Conway problem). Does there exist a well-spaced ε -net?

Let $S \subseteq \mathbb{R}^2$. We define

$$H'_\delta(S) := \inf \left\{ \sum_{i=1}^{\infty} \text{diam } U_i : S \subseteq \cup_i U_i, \quad \text{diam } U_i \leq \delta, \right\}$$

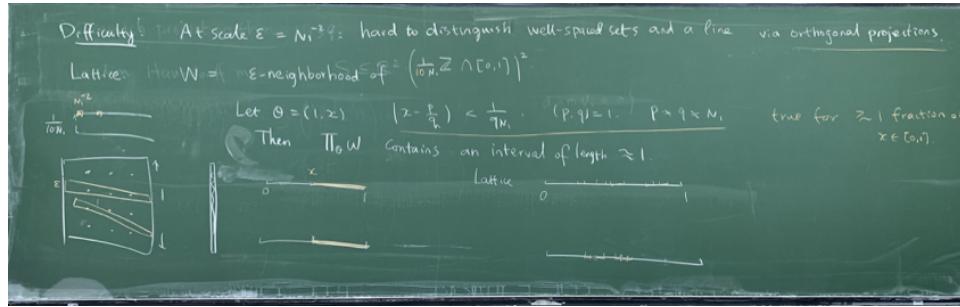
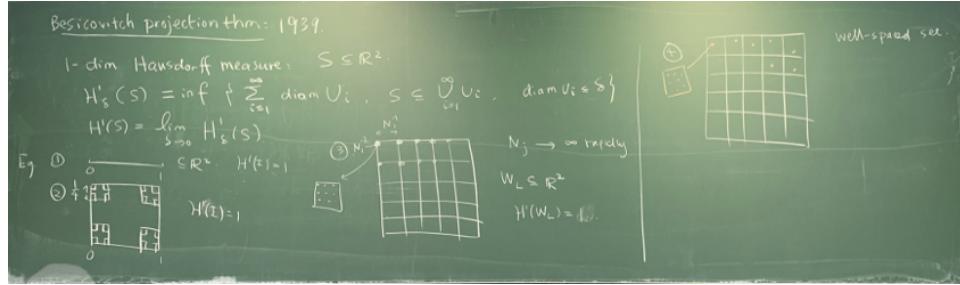
$$H'(S) := \lim_{\delta \rightarrow 0} H'_\delta(S).$$

Suppose given a 1-dimensional Hausdorff measure.

Example 8.2.

Definition 8.3. A set $E \subseteq \mathbb{R}^2$ is *purely 1-unrectifiable* if for every Lipschitz map $f : \mathbb{R} \rightarrow \mathbb{R}^2$, we have $\mathcal{H}^1(E \cap f(\mathbb{R})) = 0$.

In other words, we can't find a “Lipschitz curve” whose intersection with E has positive 1-Hausdorff measure.



Theorem 8.4 (Besicovitch projection theorem). *If $E \subseteq \mathbb{R}^2$ is purely 1-unrectifiable and $0 < H^1(E) < \infty$, then $H^1(\pi_e(E)) = 0$ for H^1 -almost every $e \in \mathbb{S}^1$, where $\pi_e : \mathbb{R}^2 \rightarrow \mathbb{R}$ is given by $x \mapsto x \cdot e$.*

Naive thoughts: if a “quantitative Besicovitch projection theorem” is true, then for most of the ε -separated directions e , we have that $\pi_\ell(W)$ is small. This shows that the ε -net version of the set does not exist. But this turns out to be too wishful thinking – it’s not true, as we’ll explain for the rest of the talk.

The **difficulty** is that at scale $\varepsilon = N_1^{-2}$, it’s hard to distinguish well-spaced sets and a line via orthogonal projections.

As for our lattice: take W to be an ε -neighborhood of $\left(\frac{1}{10N_1}\mathbb{Z} \cap [0,1]\right)^2$. Let $\theta = (1, x)$. Inductively, write

$$\left|x - \frac{p}{q}\right| < \frac{1}{qN_1}, \quad (p, q) = 1, \quad p \approx q \approx N_1.$$

This is true for $\gtrsim 1$ fraction of $x \in [0, 1]$. Then $\pi_\theta W$ contains an interval of length ≈ 1 . The shapes of the bad direction regions can be a couple different ones.

9. DAMARIS SCHINDLER, Density of rational points near manifolds

Let $M \subseteq \mathbb{R}^n$ be a manifold of dimension m . Denote by $R := n - m$ the codimension. For $Q > 1$ and $0 \leq \delta < 1/2$, let $N_M(Q, \delta)$ denote the number of rational vectors $p/q \in \mathbb{Q}^n$, with $p \in \mathbb{Z}^n$ and $1 \leq q \leq Q$, such that

$$\text{dist}\left(M, \frac{p}{q}\right) \leq \frac{\delta}{q}.$$

We have the trivial bound $N_M(Q, \delta) \ll Q^{m+1}$. Heuristically, we expect that

$$N_M(Q, \delta) \sim Q \left(\frac{\delta}{Q}\right)^2 Q^n \sim \delta^R Q^{m+1}.$$

For $\delta > 0$, we have $N_M(Q, \delta) \geq N_M(Q, 0)$.

What is the size of $N_M(Q, 0)$ for “non-flat” M ?

Let $V \subseteq \mathbb{P}_Q^n$ be an irreducible variety of degree $d > 1$. Denote by $H : \mathbb{P}_{\mathbb{Q}}^n(\mathbb{Q}) \rightarrow R_{\geq 0}$ the height function $H(\underline{x}) := \max_{0 \leq i \leq n} |x_i|$, for $\underline{x} = (x_0, \dots, x_n) \in \mathbb{Z}_{\text{prim}}^{n+1}$. Denote by $N_V(Q)$ the number of $x \in V(\mathbb{Q})$ with $H(x) \leq Q$.

Conjecture 9.1 (Weak dimension growth conjecture). *For $d \geq 2$ and $\varepsilon > 0$, we have $N_V(Q) \ll Q^{\dim V + \varepsilon}$.*

Example 9.2. Take $V = \{F = 0\}$, where $F \in \mathbb{Z}[x_0, \dots, x_n]$ with $\deg F \geq 2$. Compare to $M \subseteq V \cap \{x_0 = 0\}$ given by $F(1, x_1/x_0, \dots, x_n/x_0) = 0$.

Conjecture 9.3 (Huang '20, Beresnevich–Kleinbock '22). *If the manifold M is non-flat in some sense, say it satisfies “proper curvature conditions”, then if $\delta \geq Q^{-\frac{1}{k} + \varepsilon}$, we have the asymptotic $N_M(Q, \delta) \sim \delta^R Q^{m+1}$ as in the heuristic argument.*

Note that this implies that $N_M(Q, \delta) \leq Q^{m+\varepsilon} + \delta^R Q^{m+1}$ for all δ .

This conjecture has been proved by Huang '20 for smooth hypersurfaces of non-vanishing Gaussian curvature with $m \geq 2$. This is proved using methods purely from harmonic analysis. It's interesting because we can also run the above argument backwards: if we have an algebraic variety that satisfies this conditions on curvature, then we can stick the affine patches of our variety into the above conjecture and then get back the dimension growth conjecture. We get the weak dimension growth conjecture for hypersurfaces which satisfy the curvature condition.

Let's give a proof sketch. Locally parametrize M as $(\underline{x}, f_1(\underline{x}), \dots, f_R(\underline{x}))$, with $\underline{x} \in \mathbb{R}^m$. Then

$$N(Q, \delta) \sim \sum_{\substack{1 \leq q \leq Q \\ a \in \mathbb{Z}^m}} \omega\left(\frac{a}{q}\right) \prod_{i=1}^R \omega_0\left(\frac{\|qf_i(\frac{a}{q})\|}{\delta}\right).$$

This last factor effectively enforces the condition that $|f_i(\frac{a}{q}) - \frac{b_i}{q}| \leq \frac{\delta}{q}$ for some integer b_i . The above should be

$$\begin{aligned} &\sim \delta^R \sum_{\substack{q \leq Q \\ a \in \mathbb{Z}^m}} \omega\left(\frac{a}{q}\right) + (\text{error}) \\ &\sim \delta^R Q^{m+1} + (\text{error}). \end{aligned}$$

The error term here is

$$\ll \sum_{j \in \mathbb{Z}^R - \{0\}} b_{\delta, j} \sum_{\substack{q \leq Q \\ a \in \mathbb{Z}^m}} \omega\left(\frac{a}{q}\right) e\left(q \sum_{i=1}^R j_i f_i\left(\frac{a}{q}\right)\right).$$

Here the meaningful range is $j \ll Q^\varepsilon/\delta$. Now, maybe you want to apply Poisson summation to the sum over a . This gives

$$\sum_{q \leq Q} q^m \sum_{k \in \mathbb{Z}^m} I(q, j, k),$$

where

$$I(q, j, k) = \int_{\mathbb{R}^m} \omega(x) e\left(q \left(\sum_{i=1}^R j_i f_i(x) - k \cdot x \right)\right) dx.$$

Now we stick to $R = 1$ and assume non-vanishing Gaussian curvature. Then, if there exists a critical point $x_0 = (\nabla f)^{-1} \left(\frac{k}{j} \right)$ (there is only one index j left, so we drop the index), then

$$I(q, j, k) \sim e \left(qj \left(f(x_0) - \frac{k}{j} x_0 \right) \right) (qj)^{-\frac{m}{2}}.$$

We define the “dual” or “Legendre dual” of f to be

$$f^*(y) := x \cdot y - f(x) \text{ if } y = \nabla f(x).$$

Taking the dual of the phase of interest and summing, we obtain

$$\sum_{q \leq Q} e \left(-qj f^* \left(\frac{k}{j} \right) \right),$$

which leads us to asking for the number of $j \ll Q^\varepsilon \delta^{-1}$ and $k \ll Q^\varepsilon \delta^{-1}$ with $\|jf^*(k/j)\|$ small. We then iterate.

Frequently asked questions:

- Full alternative proof for dimension growth conjecture? For example, the cone over an algebraic variety violates the curvature condition.
- Curvature in higher codimension?
- “Proper curvature conditions”? We explain below.

Definition 9.4. Let $M \subseteq \mathbb{R}^n$ be given locally by $(x, f_1(x), \dots, f_R(x))$, again with $x \in \mathbb{R}^m$. Then we say that M is ℓ -non-degenerate at $x_0 \in \mathbb{R}^m$ if all the partitions derivatives in x of order up to ℓ of $(x, f_1(x), \dots, f_R(x))$ span \mathbb{R}^n .

Example 9.5. If $R = 1$ and you have a hypersurface, then to be ℓ -non-degenerate, there should be a partial derivative of order up to ℓ that is nonvanishing.

Conjecture 9.6 (Huang '22). *If M is $(R + 1)$ -non-degenerate and $\delta \geq Q^{-\frac{1}{R} + \varepsilon}$, then we expect the heuristic asymptotic $N_M(Q, \delta) \sim \delta^R Q^{m+1}$ to hold.*

(Consider $(x_1, \dots, x_m, x_1^\ell, \dots, x_R^\ell)$.)

10. TREVOR WOOLEY, *The paucity of knowledge concerning Weyl sums and their mean values*

10.1. **Introduction.** Consider a Vinogradov exponential sum

$$f(\alpha) := \sum_{1 \leq x \leq N} e(\alpha_1 x + \dots + \alpha_k x^k).$$

Here N is a large positive integer, $e(z) := e^{2\pi i z}$. We will specialize to the case $k = 1$:

$$f(\alpha) = \sum_{1 \leq x \leq N} e(\alpha x^k).$$

(Circle method) complexity of α :

$$H(\alpha) := H(\alpha; N^k) = \min_{\substack{a \in \mathbb{Z} \\ q \in \mathbb{N}}} \{q + N^k |q\alpha - a|\}.$$

Informally, this will be small if α has good rational approximations.

Dirichlet: given $\alpha \in \mathbb{R}$, there exists $a \in \mathbb{Z}$, $q \in \mathbb{N}$ with $1 \leq q \leq N^{k/2}$ and $|q\alpha - a| \leq N^{-k/2}$ such that

$$H(\alpha) \leq q + N^k |q\alpha - a| \leq 2N^{k/2}.$$

Thus, if $H(\alpha) = Q$, then there exists $a \in \mathbb{Z}$, $q \in \mathbb{N}$ with $(a, q) = 1$ such that $q \leq Q$ and $|q\alpha - a| \leq QN^{-k}$.

One basic question in the subject is to understand the behavior of an exponential sum like $f(\alpha)$ for $1 \leq H(\alpha) \leq 2N^{k/2}$. Three trivial ranges for $H(\alpha)$:

Range (0). $H(\alpha) \ll 1$. Say $|\alpha - \frac{a}{q}| \leq \frac{1}{100}N^{-k}$, where $q \leq 100$. Then $|f(\alpha)| \approx N$.

Range (1). $H(\alpha) \ll N$. Use mean value theorem. We have $\alpha = \beta + a/q$, say. Then

$$\begin{aligned} f(\alpha) &= \sum_{1 \leq x \leq N} e((\beta + a/q)x^k) = \sum_{r=1}^q \sum_{\substack{\frac{1-r}{q} \leq y \leq \frac{N-r}{q} \\ x=qy+r}} e\left((\beta + \frac{a}{q})(qy+r)^k\right) \\ &= \sum_{r=1}^q e\left(\frac{a}{q}r^k\right) \sum_{\frac{1-r}{q} \leq y \leq \frac{N-r}{q}} e(\beta(qy+r)^k) \\ &= \sum_{r=1}^q e\left(\frac{a}{q}r^k\right) \left(q^{-1} \int_0^N e(\beta\gamma^k) d\gamma + O(1 + N^k |\beta|) \right) \\ &= q^{-1} S(q, a)v(\beta) + \underbrace{O(q + N^k |q\alpha - a|)}_{H(\alpha)}, \end{aligned}$$

where

$$S(q, a) := \sum_{r=1}^q e\left(\frac{a}{q}r^k\right), \quad v(\beta) = \int_0^N e(\beta\gamma^k) d\gamma.$$

Vaughan, 1981: use Poisson summation to refine this estimate, giving

$$f(\alpha) - q^{-1} S(q, a)v(\beta) \ll H(\alpha)^{1/2+\varepsilon}.$$

Note: this gives

$$f(\alpha) - q^{-1} S(q, a)v(\beta) \ll N^{k/4+\varepsilon},$$

but if $k \geq 4$, then this is worse than trivial. That's bad news, but what it does show is that at last when $H(\alpha) \leq N$, the error term is $O(N^{1/2+\varepsilon})$. This gives

$$\int_{H(\alpha) \leq N} |f(\alpha)|^s d\alpha \ll N^{s-k+\varepsilon} \quad \text{when } \begin{cases} s \geq k+1 & \text{if } k \geq 3, \\ s \geq k+2 & \text{if } k = 2. \end{cases}$$

We should mention that Vaughan, 1981, had what he would like to call a "speculation" (not a conjecture): that

$$f(\alpha) - q^{-1} S(q, a)v(\beta) \ll H(\alpha)^{1/k+\varepsilon}.$$

Remember that $H(\alpha)$ can be as large as $N^{k/2}$, so this is \sqrt{N} in the worst case, but usually much better. However, this speculation was shown to be false in work of Brüdern–Daemen published in 2009. There is much more that was proved in this work, but in particular,

$$|f(\alpha) - q^{-1} S(q, a)v(\beta)| \gg N^{1/2}$$

for almost all α .

Conjecture 10.1. *We have $f(\alpha) - q^{-1}S(q, a)v(\beta) \ll N^{1/2+\varepsilon}$ for $H(\alpha) \leq 2N^{k/2}$.*

Problem 6 (Challenge Problem). Take k large, $q \approx N^3$. Estimate

$$\left| \sum_{1 \leq x \leq N} e\left(\frac{a}{q}x^k\right) - \frac{N}{q}S(q, a) \right| \ll N^\theta,$$

where

$$\theta < \min\left\{\frac{k}{4}, \theta_0\right\},$$

with θ_0 independent of k and < 1 .

10.2. How little we know. Let's focus on

$$f\left(\frac{a}{q}\right) = \sum_{1 \leq x \leq N} e\left(\frac{a}{q}x^k\right).$$

The usual approach:

(A) Weyl's inequality. Write

$$|f(\alpha)|^2 = f(\alpha)f(-\alpha) = \sum_h \sum_x e(\alpha((x+h)^k - x^k)) h(kx^{k-1} + \dots).$$

This gives

$$|f(\alpha)|^{2^{k-1}} \ll N^{2^{k-1}-1} + N^{2^{k-1}-k} \left| \sum_{|h_1| \leq N} \dots \sum_{\substack{|h_{k-1}| \leq N \\ h_1 \dots h_{k-1} \neq 0}} e(h_1 \dots h_{k-1} \alpha(k!x + c(h))) \right|,$$

giving

$$\begin{aligned} f\left(\frac{a}{q}\right) &\ll N^{2^{k-1}-1} + N^{2^{k-1}-k+\varepsilon} \sum_{1 \leq h \leq k!N^{k-1}} \min\left\{N, \left\|h\frac{a}{q}\right\|^{-1}\right\} \\ &\ll \left(N^{1+\varepsilon} (q^{-1} + N^{-1} + qN^{-k})^{2^{1-k}}\right)^{2^{k-1}}. \end{aligned}$$

(B) Vinogradov, I. There's various ways of setting up this approach, which more-or-less amount to the same thing. We'll choose the one that is easiest to describe from first principles. It is slightly wasteful, but illustrates the right ideas.

$$|f(\alpha)|^2 = \sum_{|h| \leq N} \sum_{x \in I(h)} e(h\alpha(kx^{k-1} + \dots)).$$

Then

$$\begin{aligned} |f(\alpha)|^{4s} &\leq N^{2s-1} \sum_{|h| \leq N} \left| \sum_{x \in I(h)} e(h\alpha(kx^{k-1} + \dots)) \right|^{2s} \\ &\leq N^{2s-1} \sum_{|h| \leq N} \left| \sum_{x_1, \dots, x_{2s}} e(h\alpha(kv_{k-1}(x) + \dots)) \right|, \end{aligned}$$

where $\sigma_j(x) = x_1^j + \cdots + x_s^j - x_{s+1}^j - \cdots - x_{2s}^j$. Write $\sigma_j(x) = \ell_j$ for $1 \leq j \leq k-1$. Then by fixing the values of the ℓ_j and counting, we obtain

$$|f(\alpha)|^{4s} \leq N^{2s-1} \cdot N^{1+2+\cdots+(k-2)} J_{s,k-1}(N) \cdot \left| \sum_{|h| \leq N} \sum_{|\ell| \leq sN^{k-1}} e(kh\ell\alpha) \right|,$$

where

$$\begin{aligned} J_{s,k-1}(N) &= \# \left\{ \sum_{i=1}^s (x_i^j - y_i^j) = 0 \quad (1 \leq j \leq k-1) : 1 \leq x, y \leq N \right\} \\ &\ll N^{2s-\frac{1}{2}k(k-1)+\varepsilon} + N^{s+\varepsilon}. \end{aligned}$$

(Bourgain–Demeter–Guth, 2016, Wooley, 2016/2019.) Glueing together all of these estimates, we get an upper bound

$$\left| f\left(\frac{a}{q}\right) \right| \ll N^{1+\varepsilon} \left(\frac{1}{q} + N^{-1} + qN^{-k} \right)^{1/k(k-1)}.$$

(C) Vinogradov, II (Bombieri–Korobov). Replace the initial exponential sum with shifts:

$$f(\alpha) = \frac{1}{UV} \sum_{\substack{u \sim U \\ v \sim V}} \sum_{1 \leq x \leq N} e(\alpha(x+uv)^k).$$

One uses that

$$f\left(\frac{a}{q}\right) \ll N^{1+\varepsilon} (q^{-1} + U^{-k} + V^{-k} + q(UV)^{-k})^{1/4s^2}$$

for $s \geq \frac{1}{2}k(k+1)$.

10.3. Mean values. If you can't achieve pointwise bounds, maybe you'll have more success with mean values. The nice thing about mean values is that there are nice arithmetic interpretations for what happens. To illustrate the ideas, we'll again look at the same exponential sum $f(\alpha)$. We consider the integral

$$I_s(N) := \int_0^1 |f(\alpha)|^{2s} d\alpha,$$

which by orthogonality is counting the number of solutions to the equation

$$x_1^k + \cdots + x_s^k = x_{s+1}^k + \cdots + x_{2s}^k \quad (1 \leq x_i \leq N).$$

We conjecture that

$$I_s(N) \sim (CN^{2s-k} + s!N^s)(1 + o(1)).$$

Here C is a product of local densities, $H(\alpha) \leq N$, and the second term comes from the diagonal solutions where $\{x_1, \dots, x_s\} = \{x_{s+1}, \dots, x_{2s}\}$.

The basic question for these mean values from the point of view of the circle method is to understand the source of the various contributions to this mean value $I_s(N)$ arising from points α in the full range $1 \leq H(\alpha) \leq 2N^{k/2}$.

As a challenge, identify the source of the diagonal solutions (arising from linear spaces). The contribution from the small height is

$$\int_{H(\alpha) \leq N} |f(\alpha)|^{2s} \sim CN^{2s-k},$$

where C is a product of local densities.

Conjecture 10.2 (Vaughan and Wooley, 1995).

$$\int_{N < H(\alpha) \leq 2N^{k/2}} |f(\alpha)|^{2s} d\alpha \sim \text{contribution of special subvarieties.}$$

Example 10.3. We were able to show that

$$\int_0^1 \int_0^1 \left| \sum_{1 \leq x \leq N} e(ax^3 + \beta x) \right|^6 d\alpha d\beta = 6N^3 + CN^2(\log N)^5(1 + o(1)).$$

(“C” is due to de la Breteche.)

Example 10.4 (Brüdern–Wooley, 2019).

$$\int_{[0,1]^3} \left| \sum_{|x| \leq N} e(\alpha x^3 + \beta x) \right|^6 \left| \sum_{|x| \leq N} e(\alpha x^3 + \gamma x) \right|^4 d\alpha d\beta d\gamma \sim (45 + \mathcal{C})N^3(1 + o(1)).$$

Here \mathcal{C} is again a product of local densities.

Example 10.5.

$$\begin{aligned} \int_0^1 \left| \sum_{1 \leq x \leq N} e(\alpha x^4) \right|^6 d\alpha &= \# \{x_1^4 + x_2^4 + x_3^4 = x_4^4 + x_5^4 + x_6^4 : 1 \leq x_i \leq N\} \\ &\sim 3!N^3 + \mathcal{C}N^2 + \mathcal{C}'N^2 \log N. \end{aligned}$$

Here

- the first term comes from the diagonal,
- in the second term, \mathcal{C} is a product of local densities, and we have $H(\alpha) \leq N$, and
- the third term comes from the contribution of

$$\begin{aligned} x_3 &= x_1 + x_2, \\ x_6 &= x_4 + x_5, \\ x_1^2 + x_1x_2 + x_2^2 &= x_4^2 + x_4x_5 + x_5^2. \end{aligned}$$

This uses the identity $a^4 + b^4 + (a+b)^4 = 2(a^2 + ab + b^2)^2$.

Let’s now describe a **model problem** to think about, which embodies many of the characteristics we see here and gets back to the observation about pointwise estimates boiling down to understand congruences and glueing variables to make the summation length bigger than the congruence modulus. The model problem is to obtain “good” estimates for the q -analogue problem, by which we mean

$$\{x_1^k + \cdots + x_s^k \equiv x_{s+1}^k + \cdots + x_{2s}^k \pmod{q}, 1 \leq x_i \leq N\} = \frac{1}{q} \sum_{a=1}^q \left| f\left(\frac{a}{q}\right) \right|^{2s},$$

with q fixed and $N^2 < q \leq N^{k/2}$. What’s going on in this problem is that if $q \leq N$, then this problem really looks like the congruence problem, because the variables are big enough that you can just reduce things to congruences – it’s just a $\mathbb{Z}/q\mathbb{Z}$ problem. On the other hand, for $q > N^k$, you have to consider equations.

11. PO LAM YUNG, *Discrete restriction in 2+1 vs 1+1 dimensions*

[Slide talk, notes very partial]

$$\mathcal{I}(a) := \left\| \sum_{n \in \mathbb{Z}} a_n e(nx + n^2 t) \right\|_{L^6([0,1]^2)}.$$

Bourgain–Demeter: if $\text{supp}(a) \subseteq [N]$, then

$$\mathcal{I}(a) \lesssim_\varepsilon N^\varepsilon \|a\|_{\ell_2}.$$

Discrete version of Fourier restriction for the parabola.

Bourgain: for a the indicator function on $[N]$,

$$\mathcal{I}(a) \gtrsim (\log N)^{1/6} \|a\|_{\ell_2}.$$

Circle method shows reverse inequality is true for this a .

Herr–Kwak: if a is supported in a finite set $E \subseteq \mathbb{Z}^2$, then

$$\left\| \sum_{n \in E} a_n e(n \cdot x + |n|^2 t) \right\|_{L^4([0,1]^3)} \lesssim (\log \#E)^{1/4} \|a_n\|_{\ell_2}.$$

Obtaining the sharp power of log allows them to study global well-posedness for the cubic non-linear Schrödinger equation on \mathbb{T}^2 ,

$$i\partial_t u + \Delta_x u = \pm|u|^2 u,$$

which is L^2 -critical.

We have, with $\tilde{n} := (n, |n|^2)$ for $n \in \mathbb{R}^2$,

$$\int_{[0,1]^3} \left| \sum_{n \in E} e(n \cdot x + |n|^2 t) \right|^4 dx dt = \# \{(n_1, n_2, n_3, n_4) \in E^4 : \tilde{n}_1 - \tilde{n}_2 = \tilde{n}_4 - \tilde{n}_3\}$$

is the number of parallelograms in \mathbb{R}^3 with all vertices in \tilde{E} .

Observation: the \tilde{n}_j form a parallelogram in \mathbb{R}^3 if and only if the n_j form a rectangle in \mathbb{R}^2 .

Pach and Sharir (1990s): the number of rectangles with vertices in E is $\lesssim (\#E)^2 \log(\#E)$.