Geometric and Fourier analytic questions in Euclidean space

Summer School, Lake Arrowhead *

Sep 10th - Sep 15th 2011

Organizers:

Michael Bateman, University of California, Los Angeles

Ciprian Demeter, Indiana University, Bloomington

Christoph Thiele, University of California, Los Angeles

*supported by NSF grant DMS 1001535

Contents

1 On the Erdős distinct distance problem in the plane **6**

Matthew Bond, UBC 6

1.1 Introduction 6

1.2 Elekes-Sharir reduction steps 7

1.3 The main theorem 8

1.3.1 The polynomial method - the joints problem 8

1.3.2 Polynomial method in the Guth-Katz $k = 2$ case 10

1.3.3 The Szemeredi-Trotter theorem and the Guth-Katz $k \geq 3$ case 10

2 A Bilinear Fourier Extension Theorem and Applications to the Distance Set Problem **12**

Francesco Di Plinio, IU Bloomington 12

2.1 The main results 12

2.2 Theorem 2 and the distance set problem 13

2.3 Sketch of proof of Theorem 2. 16

3 Decay of circular means of Fourier transforms of measures **19**

Taryn C Flock, UC Berkeley 19

3.1 Connecting distance sets and spherical means of Fourier transforms of measures 19

3.2 The result 20

3.3 Reductions 20

3.4 The proof 21

4 Bochner-Riesz conjecture, restriction conjecture and oscillatory integral **25**

Yi Hu, UIUC 25

4.1 Bochner-Riesz conjecture and restriction conjecture 25

4.2 Oscillatory integral 27

5 The endpoint case of the Bennett–Carbery–Tao multilinear Kakeya conjecture **30**

K. Hughes, Princeton University 30

5.1 Introduction 30

5.1.1 The multilinear Kakeya problem 30

5.1.2 Motivation: The linear Kakeya problem 30
5.2 Previous results: Bennett–Carbery–Tao 31
5.3 MLK Jr. 32
5.4 Proof of the endpoint multilinear Kakeya estimate 33

6 An incidence theorem in higher dimensions 37
Mark Lewko, UT-Austin 37
6.1 Introduction 37
6.2 Some topology and algebraic geometry 39
6.3 Proof of the theorem 40
6.4 Contribution from the cells 40
6.5 Contribution from the boundary 41

7 L^p estimates for the Hilbert transform along a one-variable vector field 44
Diogo Oliveira e Silva, UC Berkeley 44
7.1 Introduction 44
7.1.1 History of the problem and related work 45
7.2 The main approach 46
7.3 Construction of the set G 47
7.4 The “end” of the proof 49

8 Maximal Operators for Arbitrary Sets of Directions 52
James Scurry, Georgia Institute of Technology 52
8.1 Introduction 52
8.2 Preliminaries 53
8.3 Sketch of the Proof of the Main Theorem 54
8.3.1 Sketch of the Proofs of Proposition 3 and Theorem 2 . 55
8.3.2 Sketch of the Proof of Lemma 5 56

9 Maximal theorems for the directional Hilbert transform 58
Prabath Silva, IUB 58
9.1 Introduction 58
9.2 Discretization of the operator 59
9.3 Idea of the Proof 60

10 A remark on the maximal function associated to an analytic vector field **62**
Stefan Steinerberger, University Bonn 62
10.1 Introduction 62
10.2 Bourgain's theorem 63
10.3 Sketch of the proof 65

11 The Power Law For The Buffon Needle Probability Of The Four-Corner Cantor Set **69**
Krystal Taylor, Rochester 69
11.1 Introduction 69
11.1.1 The context: analytic capacity, hausdorff measure, and favard length 69
11.1.2 The set up: definitions and notation 70
11.2 Results 71
11.2.1 Statements of background results 71
11.2.2 Statements of the main result and a generalization . . 71
11.2.3 Sketch of the proof of 1 71
11.2.4 The counting function 72
11.2.5 Fourier-analytic part 73
11.2.6 Finishing the proof 74

12 Algebraic methods in discrete Kakeya-type problems **76**
Faruk Temur, UIUC 76
12.1 Introduction 76
12.2 Preliminaries 77
12.3 The joints problem and Bourgain's problem 78
12.4 The finite field Kakeya problem 80

13 Bounds on oscillatory integral operators based on multilinear estimates **82**
Joshua Zahl, UCLA 82
13.1 Introduction 82
13.1.1 Proof Sketch 83
13.2 Background Preliminaries 86
13.2.1 The multilinear restriction estimates of Bennett-Carbery-Tao 86
13.2.2 Localization and ϵ–removal 86

13.2.3 Scaling 87
13.2.4 Wolff's $\mathbb{R}^3$ Kakeya estimates 87
13.3 Proof of Theorem 2: the $p = 10/3$ bound 88
13.3.1 Case 1: transverse interaction 89
13.3.2 Case 2: collinear interaction 89
13.3.3 Case 3: coplanar interaction 89
13.4 Proof of Theorem 1: the $p = 3.3$ bound 90
13.4.1 First Bound: Hölder's inequality 94
13.4.2 Second bound: Wolff's $L^{5/2}$ Kakeya estimate 94

1 On the Erdős distinct distance problem in the plane

after Larry Guth and Nets Hawk Katz [1]
A summary written by Matthew R. Bond

Abstract

We summarize the recent Guth-Katz lower bound in the Erdős distance problem due to Guth and Katz. Namely, for any set of N points in $\mathbb{R}^2$, at least $c\frac{N}{\log N}$ distinct numbers can be obtained as distances between pairs of them.

1.1 Introduction

In 1946, Paul Erdős asked how one might minimize the number of distinct distances arising between N points in $\mathbb{R}^2$. For a square grid, the number of distances is $\approx \frac{N}{\sqrt{\log N}}$, and it remains an open problem whether this example is in fact a sharp lower bound.

Let us call the set of points P, and let $d(P) := \{d(s_1, s_2) : s_1, s_2 \in S)\}$. For any finite set A, let $|A|$ denote the number of elements in A, so that $|P| = N$, for example.

Up until the recent result of [1], the only available estimates in the Erdős distance problem were of the form $|d(P)| \gtrsim N^{1-\epsilon}$, for $0 < \epsilon < 1$ and with ϵ getting closer to 0 with each new paper.

Theorem 1. *(Larry Guth and Nets Hawk Katz)* $|d(P)| \gtrsim \frac{N}{\log N}$

While the above theorem is not necessarily sharp (the square grid is the minimum example so far), the way the bound is proved in [1] is as a corollary of another theorem which, by contrast, is sharp. There are two reduction steps - first, a Cauchy-Schwartz estimate where the sharpness must have been lost, if the above theorem is not in fact sharp. Second, the reduced problem is translated into a problem about rigid motions sending some of P into itself, and then this is translated to a problem about incidences between lines in $\mathbb{R}^3$. These steps are due to Elekes and Sharir. Guth and Katz solved the problem resulting from a slightly simplified variation of the Elekes-Sharir setup.

First, a definition. A **ruled surface** in $\mathbb{R}^3$ is a continuous one-parameter family of lines, and this family of lines is called a **ruling**. A **regulus** is a doubly-ruled surface, that is, a ruled surface which can be described by two genuinely different rulings - that is, for each x in the surface, there is a line through x in the first ruling and a distinct line through x from the second ruling.[(1)] With these definitions, we can state the incidence theorem from which Theorem 1 follows:

Theorem 2. *Let $\mathcal{L}$ be a set of N^2 lines in $\mathbb{R}^3$. Suppose $\mathcal{L}$ contains $\lesssim N$ lines in any plane or regulus. Suppose that $2 \leq k \leq N$. Then the number of points that lie in at least k lines is $\lesssim N^3k^{-2}$.*

1.2 Elekes-Sharir reduction steps

Let **the set of distance quadruples** $Q(P) := \{(p_1, p_2, p_3, p_4) \in P^4 : d(p_1, p_2) = d(p_3, p_4) \neq 0\}$. [(2)]. Clearly $Q(P)$ measures redundancy in representations of distances. Of course $(p_1, p_2, p_1, p_2) \in Q(P)$ regardless, so $|Q(P)| \gtrsim N^2$. In fact, a short Cauchy-Schwartz estimate yields

$$|d(P)| \gtrsim \frac{N^4}{|Q(P)|}, \tag{1}$$

which is sharp in the trivial case where $|d(P)| = \binom{N}{2}$.

In fact, Guth and Katz prove that

$$|Q(P)| \lesssim N^3 \log N \tag{2}$$

Together, these clearly imply Theorem 1. Now on to the second reduction, the reduction to rigid motion groups.

Let G be the group of rigid motions on $\mathbb{R}^2$, that is, the group of all possible maps definable as a translation followed by a rotation. Then it is not hard to see that the elements (p_1, p_2, p_3, p_4) define a unique $g \in G$ such that $g(p_1) = p_3$ and $g(p_2) = p_4$. Let the mapping $(p_1, p_2, p_3, p_4) \to g$ be denoted E (for Elekes).

A lemma shows that if $|P \cap g(P)| = k$, then $|E^{-1}(g)| = 2\binom{k}{2}$. This is a big hint toward how to further organize the problem - i.e., around k. If

[1]A two-sided cone is a (singly)-ruled surface, for example, and $z = xy$ is a doubly-ruled surface, as one can see by slicing along $y = y_0$ and $x = x_0$ and parameterizing the rulings by x and y, respectively. By way of convention, a plane is not a regulus.

[2]The condition $\neq 0$ is convenient for organizing the argument around a small technicality

$G_k(P) := \{g \in G : |P \cap g(P)| \geq k\}$, then it can be shown that $|Q(P)| = \sum_{k=2}^{N}(2k-2)|G_k(P)| \approx \sum_{k=2}^{N} k|G_k(P)|$.

To get (2), one wants $|G_k(P)| \lesssim N^3 k^{-2}$. This is what Guth and Katz prove. [3]

Now let's reduce the rigid motion group problem to incidence geometry. We would like to exploit the unique fixed point of g, which means handling pure translations as a separate case (they have no fixed points, but translations are by far the easier case anyway). Call the resulting set of non-translations $G' \subset G$. If the fixed point of g is (x, y) and the rotation is θ, then define $\rho(g_{x,y,\theta}) := \rho(x, y, \theta) := (x, y, \cot(\theta))$. Then if we consider **all** $g \in G'$ such that $g(p) = q$, then the image of all such g is exactly a line whose projection to the (x, y) plane is the perpendicular bisector of the line segment $\bar{pq}$.

So now we can interpret the following set S_{pq} as a line: $L_{pq} = \rho(S_{pq})$, where $S_{pq} = \{g \in G' : g(p) = q\}$. Then let $\mathcal{L} := \cup_{p,q \in P} L_{pq}$. Some facts:

- $\rho(G_k(P))$ is the set of k-fold (or greater) incidences of lines from $\mathcal{L}$
- No more than N lines of $\mathcal{L}$ lie in a single plane, and at most $O(N)$ lines of $\mathcal{L}$ lie in a single regulus.

The new problem is[4]:

Theorem 3. *Let $\mathcal{L}$ be any set of $N = M^2$ lines such that at most M lie in any common plane and at most $O(M)$ lie in any common regulus. Then the number of k-fold incidences between lines of $\mathcal{L}$ is at most CN^3k^{-2}.*

1.3 The main theorem

1.3.1 The polynomial method - the joints problem

A main ingredient of 3 is the polynomial method. The following **joints problem**, long thought to be intrinsically difficult, actually turns out to be

[3]In particular, it should come as no surprise that the square grid of N integer points saturates this last inequality for $2 \leq k \leq N/2000$. Guth and Katz prove this in the appendix.

[4]The condition about reguli is not needed when $k \geq 3$. $k = 2$ is somewhat special, in that if two lines lie in an algebraic surface and cross at a point, that point can still be somewhat generic, while if three such lines meet at a point, then it is "special" in that the point is either "critical" or "flat" (depending respectively on whether the directions are or are not an independent set).

an ideal model problem for this method. It is proved in [2], which the reader is encouraged to peruse at (its rather short) length.

Theorem 4. *Let $\mathcal{L}$ be a set of lines in $\mathbb{R}^n$. A **joint** of $\mathcal{L}$ is a point lying in some n lines of $\mathcal{L}$ having linearly independent directions. Then* $|\{\textit{joints of } \mathcal{L}\}| \leq C_n|\mathcal{L}|^{\frac{n}{n-1}}$

Proof. This will be just a sketch. Steps 1-4 are generalizeable to other situations, with the generalization given in bold:

1) **Assume a conterexample** – Assume there are a lot of joints - that is, $\geq 2C_n|\mathcal{L}|^{\frac{n}{n-1}}$. We want a contradiction.

2) **Cull the slackers** – A proportion of the lines contain a proportion of as many joints as they are supposed to, even after you remove all lines and joints associated with the "below average" subset of lines. The total number of joints is unharmed by this culling.

3) **Construct a low-degree polynomial** – A non-trivial polynomial p on $\mathbb{R}^n$ can be constructed to be of degree $d \lesssim N^{1/n}$ and to vanish at N points.

4) **Find lines along which the polynomial vanishes** – If a line contains more than d points where p vanishes, then p vanishes on the entire line.[5]

5) If p vanishes on an entire line, then so does its directional derivative in that direction.

6) At a joint, such directions span $\mathbb{R}^n$. So for example, $\frac{\partial}{\partial x_1}p$ vanishes at this point.

7) Induct; for example, replace p with $\frac{\partial}{\partial x_1}p$ at step 3. In particular, $\frac{\partial}{\partial x_1}p$ and its partial derivatives vanish along the same set of lines.

8) **Find a contradiction** – Of course, all mixed partial derivatives of p of all orders must vanish; but p is non-trivial, contradiction. □

In other problems, the polynomial in step 3 will be different. The other example of immediate interest is a discrete corollary of **the polynomial ham sandwich theorem** which lets you construct an even lower-degree polynomial p which bisects M finite point sets S_j simultaneously according to $p > 0$ and $p < 0$. More precisely, these two sets each contain at most half the points of S_j, and some or even all the points may lie on the surface $p = 0$. Working in $\mathbb{R}^n$, one may take $M = \binom{n+d}{n} - 1$.

[5] The exponent in the theorem corresponds to the critical exponent beyond which this step is invalid.

1.3.2 Polynomial method in the Guth-Katz $k = 2$ case

Consider Theorem 3 for $k = 2$, arguing by contradiction. Assume that some choice of less than N^2 lines (of which no CN of them lie in any single plane or regulus) have at least QN^3 incidences for Q larger than the implied constant in the theorem, and also assume that N is the smallest example for this fixed Q. First, one removes from the line set $\mathcal{L}$ those lines which are "below average," calling the reduced set $\mathcal{L}'$. Generically, a randomly-chosen subset $\mathcal{L}''$ contains the structure of $\mathcal{L}'$: the lines of $\mathcal{L}'$ each meet N or more lines of $\mathcal{L}''$. By sampling points from each line of $\mathcal{L}''$, we can construct a polynomial p which has degree $d < N$ which vanishes on these points. p must then vanish along each line of $\mathcal{L}''$ because there are more than d zeroes of p on one such line by construction. Then the larger set $\mathcal{L}'$ meets more than $N > d$ lines of $\mathcal{L}''$, so again p vanishes on all of these lines.

Consider the irreducible factors of p. Some have zero sets which are ruled, and others are unruled. The Elekes-Sharir setup excludes the possibility of having many incidences within any plane or regulus, leaving only unruled surfaces and strictly singly-ruled surfaces to consider. Singly-ruled surfaces have a structure used in one argument, while the unruled surfaces are controled by a degree bound resulting from consideration of the **flecnode polynomial**. Let us turn to this now.

Given a level set $Z := \{(x, y, z) : q(x, y, z) = 0\}$ of a polynomial $q : \mathbb{R}^3 \to \mathbb{R}$, a flecnode is a point at which some directional derivative of p vanishes to order 3. If the degree of q is m, there is polynomial Fl_q of degree $11m - 24$ which vanishes exactly at flecnodes. A theorem of Cayley states that Z is a ruled surface if and only if $Fl(q) \equiv 0$ on Z. Bezout-type lemma implies that any degree-m hypersurface in $\mathbb{R}^3$ having $11m^2 - 24m + 1$ lines must contain a ruled surface as a subset, which in turn is the zero set of a factor of q. In other words, if a surface of degree m is **unruled**, then it can contain at most $O(m^2)$ lines.

1.3.3 The Szemeredi-Trotter theorem and the Guth-Katz $k \geq 3$ case

Fix k, and let $\mathcal{G}$ be the set of points where $\mathcal{L}$ has k-fold or greater incidences. Now that $k \geq 3$, polynomials vanishing along k lines through a point must be either "critical" or "flat" at such a point, allowing the cases $k \geq 3$ to have a lot in common.

The Szemeredi-Trotter theorem states that for $L := |\mathcal{L}|$, $|\mathcal{G}| \lesssim L^2k^{-3} + Lk^{-1}$, regardless of the dimension of the space $\mathbb{R}^n$. Guth-Katz looks at the case $n = 3$ and introduces the assumption that at most B lines lie in any particular plane. In this case,[6]

$$|\mathcal{G}| \lesssim L^{3/2}k^{-2} + LBk^{-3} + Lk^{-1}. \tag{3}$$

A sketch of the proof: We can actually make simplifying uniformity assumptions and assume that the lines are all comparable to average as in the joints problem. Namely, at least one percent of them contain at least one percent of as many points as they should. Then one proves (3) for such a set of lines. Let us suppose that we attempt to find a counterexample to (3) of the type $|\mathcal{G}| \geq AL^{3/2}k^{-2} + CLk^{-1}$, where C is a bit bigger than the Szemeredi-Trotter constant and A is a very large number that has to be chosen later. If it turns out that $B \gtrsim |\mathcal{G}|L^{-1}k^3$, then all is well.

First, one uses the polynomial ham sandwich method to repeatedly bisect $\mathcal{G}$. Then one shows that if A is large, then there must be a way to catch a fixed large majority of $\mathcal{G}$ in the low-degree surface itself rather than in its interior components. Combining Szemeredi-Trotter with the highly uniform behavior of the lines, the vast majority of points end up in the walls of Z; but then it can be shown that Z contains a lot of lines of $\mathcal{L}$, and finally, a plane containing many such lines.

References

[1] Guth, L. and Katz, N. H., *On the Erdős distinct distance problem in the plane* arXiv:1011.4105 ;

[2] Kaplan, Sharir, and Shustin *On lines and Joints* arXiv:0906.0558

[3] Iosevich, Roche-Newton, and Rudnev, *On an application of Guth-Katz Theorem* arXiv:1103.1354 ;

Matthew R. Bond, UBC
email: **bondmatt@math.ubc.ca**

[6]Note that as B becomes small, this becomes a marked improvement over Szemeredi-Trotter. By considering "typical" projections to $\mathbb{R}^2$, Szemeredi-Trotter reduces to the case $B = L$ anyway, where it is a corollary of the Gutz-Katz estimate.

2 A Bilinear Fourier Extension Theorem and Applications to the Distance Set Problem

after M.B. Erdogan [E1]
A summary written by Francesco Di Plinio

Abstract

We summarize Erdogan's article [E1], which contains a weighted version of Tao's bilinear Fourier restriction estimate for paraboloids [T] with improved range. The weighted estimate is used to extend the validity of Falconer's distance problem conjecture in $d \geq 3$ to sets with Hausdorff dimension greater than $\frac{d}{2} + \frac{1}{3}$.

2.1 The main results

Throughout the summary $\hat{\mu}$ will denote the Fourier transform of the measure μ on $\mathbb{R}^d$:

$$\hat{\mu}(\xi) = \int_{\mathbb{R}^d} \mathrm{e}^{-2\pi i x\cdot\xi} \mathrm{d}\mu(x), \qquad \xi \in \mathbb{R}^d.$$

In [T], Tao proved the following bilinear Fourier restriction for the d-dimensional paraboloid

$$S = \{x \in \mathbb{R}^d : x_d = x_1^2 + \cdots + x_{d-1}^2\}.$$

Theorem 1. *Let $d \geq 2$ and S_1, S_2 be compact subsets of S with $\operatorname{dist}(S_1, S_2) > 1$. Then*

$$\|\widehat{f_1\mathrm{d}\sigma}\widehat{f_2\mathrm{d}\sigma}\|_{L^q(\mathbb{R}^d)} \lesssim_{q,d} \|f_1\|_{L^2(\mathrm{d}\sigma)}\|f_2\|_{L^2(\mathrm{d}\sigma)}, \qquad q > \frac{d+2}{d}, \tag{1}$$

for all $f_i \in L^2(\mathrm{d}\sigma)$, supported in S_j, $j = 1, 2$. Here $\mathrm{d}\sigma$ stands for the Lebesgue surface measure on S.

The main result of Erdogan's paper [E1] is a weighted version of Theorem 1, where the weight is (some appropriate smoothing of) an α-dimensional measure in $\mathbb{R}^d$.

Theorem 2. *Let $d \geq 3$ and $\alpha \in (0, d)$. Suppose the function $H : \mathbb{R}^d \to \mathbb{R}$ satisfies*

$$\|H\|_{L^\infty(\mathbb{R}^d)} \leq 1, \tag{2}$$

$$\int_{B(x,r)} |H(y)| \,\mathrm{d}y \leq r^\alpha, \qquad x \in \mathbb{R}, r > 0. \tag{3}$$

Then, under the assumptions of Theorem 1,

$$\|\widehat{f_1 \mathrm{d}\sigma}\widehat{f_2 \mathrm{d}\sigma}\|_{L^q(H\mathrm{d}\xi)} \lesssim_{q,d} \|f_1\|_{L^2(\mathrm{d}\sigma)}\|f_2\|_{L^2(\mathrm{d}\sigma)}, \qquad q > q_0(\alpha, d), \tag{4}$$

where $q_0(\alpha, d) = \max\big\{1, \min\{\frac{4\alpha}{d-2+2\alpha}\}, \frac{d+2}{d}\big\}$.

Note that, in view of (2), (4) is a consequence of Theorem 1 in the range $q > \frac{d+2}{d}$; Theorem 2 is an actual improvement whenever $\alpha < \frac{d+2}{2}$. Motivation for considering the weighted version comes from applications to the following *distance set problem.* For $E \subset \mathbb{R}^d$, the *distance set* is defined as

$$\Delta(E) = \{|x - y| : x, y \in E\}.$$

In [F], Falconer formulated the following conjecture: whenever $d \geq 2$ and E is a compact subset of $\mathbb{R}^d$ of Hausdorff dimension $\dim(E) > \frac{d}{2}$, then $\Delta(E)$ has positive one-dimensional Lebesgue measure.

The best known result in dimension $d = 2$ is due to Wolff [W]: $\dim(E) > 4/3$ implies $|\Delta(E)| > 0$. Regarding $d \geq 3$, in [E2], the author describes how to use Tao's Theorem 1 to show that $|\Delta(E)| > 0$ whenever $\dim(E) > \frac{d(d+2)}{2(d+1)}$. Theorem 2 allows for a modification of the proof in [E2] which gives the following improvement, which is currently the best result for all $d \geq 3$.

Theorem 3. *Let* $d \geq 3$ *and* $E \subset \mathbb{R}^d$ *be a compact subset with*

$$\dim(E) > \frac{d}{2} + \frac{1}{3}.$$

Then $|\Delta(E) > 0|$.

Plan of the summary. In §2, we describe the general strategy (initially due to Mattila, [M]) to attack the distance problem using Fourier transforms of measures, and explain how Theorem 2 enters the proof of Theorem 3. The proof of Theorem 2 is a simple modification of Tao's proof of Theorem 1. We sketch it in §3.

2.2 Theorem 2 and the distance set problem

Before entering the details, the general idea of the Fourier analytic approach to the distance set problem is the following: given a compact set E with

$\dim(E) = \alpha$ and a measure μ supported on E, one considers the measure δ_μ on $\Delta(E)$ given by pushing forward $\mu \times \mu$ by the distance map, i.e.

$$\int f(s)\,\mathrm{d}\delta_\mu(s) := \int_{E\times E} f(|x-y|)\,\mathrm{d}\mu(x)\mathrm{d}\mu(y).$$

and tries to show that δ_μ has an L^2 Fourier transform (so that necessarily $|\Delta(E)| > 0$). The connection between dimensionality of E and L^2 energy of the Fourier transform on μ can be made precise. If $\dim(E) > \alpha$, there exists an α-dimensional probability measure μ supported on E, that is, there holds

$$\mu(D(x,r)) \leq C_\mu r^\alpha, \qquad x \in \mathbb{R}^d, r > 0,$$

and for each $\beta < \alpha$, $I_\beta(\mu) < \infty$, where

$$I_\beta(\mu) := \int s^{-\beta}\,\mathrm{d}\delta_\mu(s) \sim_\beta \int_{\mathbb{R}^d} |\hat{\mu}(\xi)|^2 |\xi|^{-(n-\beta)}\,\mathrm{d}\xi.$$

The core of Mattila's argument is the next theorem [M].

Theorem 4. *Suppose that $\alpha \geq 1$ is a number with the property that*

$$\sup_{R>0} \left(R^{d-\alpha} \int_{S^{d-1}} |\hat{\mu}(R\mathrm{e}^{i\theta})|^2\,\mathrm{d}\theta \right) \lesssim C_\alpha I_\alpha(\mu) \tag{5}$$

for each compactly supported positive measure μ with $I_\alpha(\mu) < \infty$. Then $|\Delta(E)| > 0$ whenever E is a compact subset of $\mathbb{R}^d$ with $\dim(E) > \alpha$..

In the case $d = 2$ Wolff [W] has shown that (5) holds for $\alpha > \frac{4}{3}$, and that $\alpha = \frac{4}{3}$ is sharp. Therefore Mattila's method is structurally unable to decide Falconer's conjecture in $d = 2$, for dimensions below $\frac{4}{3}$.

For $d \geq 3$, estimate (5) for all $\alpha > \frac{d}{2} + \frac{1}{3}$ is an easy consequence of the proposition below. As a corollary, Theorem (3) follows through Theorem 4 above.

Proposition 5. *Let $\alpha \in (0,d), q > q_0(\alpha,d)$. For all α-dimensional measures μ, all $R > 1$ and all functions f supported in $A_1(R) := \{|\xi| \in (R-1,R+1)\}$, we have the estimate*

$$\left| \int f^\vee(x)\,\mathrm{d}\mu(x) \right|^2 \leq C_{q,\alpha} I_\alpha(\mu) R^{(d-1)-\frac{\alpha}{q}} \|f\|_2^2. \tag{6}$$

The proof uses the weighted bilinear estimate of Theorem 2, in the form of the next corollary. The corollary actually follows from the generalization of Theorem 2 to (M,ε)-elliptic surfaces of $\mathbb{R}^d$ (see [T] for details).

Corollary 6. *Let μ be an α-dimensional probability measure. Let $\varepsilon > 0$ and $\varepsilon R^{-1/2} \lesssim \eta \lesssim 1$. Let I_1, I_2 be subsets of $A_\varepsilon(R) := \{|\xi| \in (R-\varepsilon, R+\varepsilon)\}$, with $\operatorname{dist}(I_1, I_2) \sim R\eta \sim \operatorname{diam}(I_j)$. Then, for all functions f_j supported in I_j, $j = 1,2$, and $q > q_0(\alpha, d)$,*

$$\|\widehat{f_1 \mathrm{d}\sigma}\widehat{f_2 \mathrm{d}\sigma}\|_{L^q(\mathrm{d}\mu)} \lesssim_{q,d} \varepsilon (R\eta)^{(d-1)-\frac{\alpha}{q}} \eta^{-\frac{1}{q}} \|f_1\|_{L^2(\mathrm{d}\sigma)} \|f_2\|_{L^2(\mathrm{d}\sigma)}. \tag{7}$$

Sketch of proof of Proposition 5. For $R^{\frac{1}{2}} \lesssim 2^n \lesssim R$, we decompose $A_1(R)$ into spherical caps I with dimensions $2 \times 2^n \times \cdots \times 2^n$ and say that such I has length $\ell(I) = 2^n$. Let $f_I = f\mathbf{1}_I$ and decompose

$$\|f^\vee\|^2_{L^2(\mathrm{d}\mu)} \leq \sum_{R^{\frac{1}{2}} \lesssim 2^n \lesssim R} \sum_{\ell(I)=\ell(J)=2^n, I\sim J} \|f_I^\vee f_J^\vee\|_{L^1(\mathrm{d}\mu)} + \sum_{I \in \mathcal{I}} \|f_I^\vee\|^2_{L^2(\mathrm{d}\mu)}$$

where $I \sim J$ if they are not adjacent but their parents are, and $\mathcal{I}$ is a collection of caps of sidelength $R^{\frac{1}{2}}$ and bounded overlap. The second summand is bounded elementarily by $R^{\frac{d-\alpha}{2}}$, which is harmless for (6). Regarding the first summand, since there are $O(\log R)$-many values of n, and $\{f_I : \ell(I) = 2^n\}$ are almost orthogonal, it suffices to show that, for any given n, any $I \sim J$, $\ell(I) = \ell(J) = 2^n$,

$$\|f_I^\vee f_J^\vee\|_{L^1(\mathrm{d}\mu)} \lesssim_{\alpha,q,d} R^{(d-1)-\frac{\alpha}{q}} \|f_I\|_2 \|f_J\|_2, \qquad q > q_0(\alpha, d). \tag{8}$$

Let us tile $\mathbb{R}^d$ with rectangles P of dimensions $C \times C(2^n/R) \times \cdots \times C(2^n/R)$, with the long axis in the direction of the center of mass of $I \cup J$. Let $f_{I,P} = \widehat{f_I^\vee \phi_P}$, where ϕ_P is the L^∞-normalized rescaling of a positive Schwartz function ϕ with compact support in frequency. The main estimate is the following:

$$\begin{aligned}
\|f_I^\vee f_J^\vee\|_{L^1(\mathrm{d}\mu)} &\lesssim \sum_P \int |f_{I,P}^\vee(x) f_{J,P}^\vee(x)| \phi_p(x)\,\mathrm{d}\mu(x) \\
&\lesssim \sum_P \|f_{I,P}^\vee f_{J,P}^\vee\|_{L^q(\mathrm{d}\mu)} \|\phi_P\|^{q'}_{L^1(\mu)}
\end{aligned} \tag{9}$$

Now, it is easy to see that the sets $I_P = \mathrm{supp}\, f_{I,P}$, $J_P = \mathrm{supp}\, f_{J,P}$ lie in the slightly bigger annulus $A_{10}(R)$, and $\mathrm{dist}(I_P, J_P) \sim 2^n \sim \mathrm{diam}(I_P), \mathrm{diam}(J_P)$. Hence, we can apply Corollary 6, and estimate

$$\|f_{I,P}^{\vee} f_{J,P}^{\vee}\|_{L^q(\mathrm{d}\mu)} \lesssim R^{1/q} 2^{n(d-1-(\alpha+1)/q)} \|f_{I,P}\|_2 \|f_{J,P}\|_2.$$

The Schwartz decay of ϕ and the α-dimensionality of μ easily give

$$\|\phi_P\|_{L^1(\mu)} \lesssim 2^{n\alpha - n} R^{1-\alpha}.$$

The proof is concluded by collecting the last two estimates, applying Cauchy-Schwarz in P and using $\sum_P \|f_{I,P}\|_2^2 \lesssim \|f_I\|_2$, and noting that the exponent of 2^n is nonnegative, so that one may replace 2^n with R. $\square$

2.3 Sketch of proof of Theorem 2.

By standard ε-removal, it suffices to prove the localized version of (4)

$$\|\widehat{f_1 \mathrm{d}\sigma}\widehat{f_2 \mathrm{d}\sigma}\|_{L_0^q(B(0,R), H\mathrm{d}\xi)} \lesssim_{\varepsilon,\alpha,d} R^{\varepsilon} \|f_1\|_{L^2(\mathrm{d}\sigma)} \|f_2\|_{L^2(\mathrm{d}\sigma)}, \qquad R > 1. \qquad (\star)_\varepsilon$$

It is easy to see that $(\star)_\varepsilon$ holds for each $\varepsilon > \frac{\alpha}{q_0}$. We proceed inductively and show that $(\star)_\varepsilon$ implies $(\star)_{\tilde{\varepsilon}}$, $\tilde{\varepsilon} := \max\{(1-\delta_1)\varepsilon, C\delta_1\} + C\delta_2$ for each $0 < \delta_1, \delta_2 < 1$. The first step is a wave packet decomposition of $\widehat{f_i \mathrm{d}\sigma}$ at scale R, namely, for $j = 1, 2$,

$$\widehat{f_j \mathrm{d}\sigma}(\xi) = \sum_{T_j} c_{T_j} \phi_{T_j}(\xi), \qquad \sum_{T_j} |c_{T_j}|^2 \lesssim \|f_j\|_{L^2(\mathrm{d}\sigma)}^2.$$

Here the tubes T_j come from a collection of $R^{\frac{1}{2}}$-separated $R^{\frac{1}{2}} \times \cdot \times R^{\frac{1}{2}} \times R$ tubes, ϕ_{T_j} are bumps adapted to T_j, with $\phi_{T_j}^{\vee}$ supported in a dual rectangle to T_j contained in an $O(R^{-1})$ neighborhood of the surface S, and normalized as $\|\phi_{T_j}\|_2 \sim R^{\frac{1}{2}}$. Note also that $\{\phi_{T_j}\}_{T_j}$ is an almost orthogonal collection of functions. By dyadic pigeonholing and normalization, we can reduce to the case where c_j is either 0 or 1 and, assuming $(\star)_\varepsilon$ holds, prove that, with $\tilde{\varepsilon}$ as above,

$$\Big\| \sum_{T_1 \in \mathbf{d}_1} \sum_{T_2 \in \mathbf{d}_2} \phi_{T_1} \phi_{T_2} \Big\|_{L^{q_0}(B, H\mathrm{d}\xi)} \lesssim R^{\tilde{\varepsilon}} (\#\mathbf{d}_1)^{\frac{1}{2}} (\#\mathbf{d}_2)^{\frac{1}{2}} \qquad (10)$$

The next step is covering B with finitely overlapping balls $Q \in \mathcal{B}$ of radius $R^{1-\delta_1}$. As in [T], a relation $\sim$ between balls $Q \in \mathcal{B}$ and tubes in $\mathbf{d}_1 \cup \mathbf{d}_2$ is introduced so that

$$\#\{Q \in \mathcal{B} : Q \sim T\} \lesssim R^{\delta_2}, \quad \forall T \in \mathbf{d}_1 \cup \mathbf{d}_2 \tag{11}$$

$$\Big\| \sum_{(T_1,T_2)\in Q^{\not\sim}} \phi_{T_1}\phi_{T_2} \Big\|_{L^2(Q)} \lesssim R^{C(\delta_1+\delta_2)} R^{-\frac{d-2}{4}} (\#\mathbf{d}_1)^{\frac{1}{2}} (\#\mathbf{d}_2)^{\frac{1}{2}} \tag{12}$$

where $Q^{\not\sim}$ is the collection of pairs (T_1, T_2) for which T_i are not both related to Q. Now, split the lhs of (10) by applying triangle inequality in $Q \in \mathcal{B}$. Applying the bilinear estimate $(\star)_\varepsilon$ (Q is a $R^{1-\delta_1}$-ball) and Cauchy-Schwarz,

$$\begin{aligned}
\sum_{Q\in\mathcal{B}} \Big\| \sum_{T_1,T_2\sim Q} \phi_{T_1}\phi_{T_2} \Big\|_{L^{q_0}(Q,H\mathrm{d}\xi)} &\lesssim R^{(1-\delta_1)\varepsilon-1} \sum_{Q\in\mathcal{B}} \Big(\sum_{T_1\sim Q} \|\phi_{T_1}\|_2^2 \sum_{T_2\sim Q} \|\phi_{T_2}\|_2^2 \Big)^{\frac{1}{2}} \\
&\lesssim R^{(1-\delta_1)\varepsilon} \sum_{Q\in\mathcal{B}} \big(\#\{T_1\sim Q\}\#\{T_2\sim Q\}\big)^{\frac{1}{2}} \\
&\lesssim R^{(1-\delta_1)\varepsilon + C\delta_2} \big(\#\mathbf{d}_1 \#\mathbf{d}_2\big)^{\frac{1}{2}} \qquad (13)
\end{aligned}$$

Here, we used the normalization of ϕ_{T_i} in going from the first to the second line and (11) to get to the last line. Let us now estimate the "$Q^{\not\sim}$"-part (this is the only part where Erdogan's proof differs from the one in [T]). Set

$$F_Q = \sum_{(T_1,T_2)\in Q^{\not\sim}} \phi_{T_1}\phi_{T_2}.$$

Then, using the assumptions (2)-(3),

$$\begin{aligned}
\sum_{Q\in\mathcal{B}} \big\|F_Q\big\|_{L^{q_0}(Q,H\mathrm{d}\xi)} &\lesssim \sum_{Q\in\mathcal{B}} \|F_Q\|_{L^2(Q,\mathrm{d}\xi)} \Big(\int_Q |H(\xi)|^{\frac{2}{2-q_0}} \Big)^{\frac{1}{q_0}-\frac{1}{2}} \\
&\lesssim \sum_{Q\in\mathcal{B}} \|F_Q\|_{L^2(Q,\mathrm{d}\xi)} \Big(\int_Q |H(\xi)| \Big)^{\frac{1}{q_0}-\frac{1}{2}} \\
&\lesssim R^{\frac{\alpha}{q_0}-\frac{\alpha}{2}} \sum_{Q\in\mathcal{B}} \|F_Q\|_{L^2(Q,\mathrm{d}\xi)} \qquad (14)
\end{aligned}$$

Now, we apply estimate (12), and by the definition of q_0 we conclude that

$$\begin{aligned}
(14) &\lesssim R^{C(\delta_1+\delta2)} R^{-\frac{d-2}{4}} R^{\frac{\alpha}{q_0}-\frac{\alpha}{2}} \big(\#\mathbf{d}_1\#\mathbf{d}_2\big)^{\frac{1}{2}} \\
&\lesssim R^{C(\delta_1+\delta_2)} \big(\#\mathbf{d}_1\#\mathbf{d}_2\big)^{\frac{1}{2}}. \qquad (15)
\end{aligned}$$

Putting together (13) and (15) completes the proof.

References

[E1] Erdogan, M. Burak, *A Bilinear Fourier Extension Theorem and Applications to the Distance Set Problem,* Int. Math. Res. Notices **23** (2005), 1411–1425.

[E2] Erdogan, M. Burak, *A note on the Fourier transform of fractal measures,* Math. Res. Lett **11** (2004), 299–313.

[F] Falconer, Kenneth J., *On the Hausdorff dimensions of distance sets,* Mathematika **32** (1985), 206–212.

[M] Mattila, Pertti, *Spherical averages of Fourier transforms of measures with finite energy; dimension of intersections and distance sets,* Mathematika **34** (1987), 207–228.

[T] Tao, Terence *A sharp bilinear restriction estimate for paraboloids,* GAFA **13** (2003), 1359–1384.

[W] Wolff, Thomas *Decay of circular means of Fourier transforms of measures,* Int. Math. Res. Notices **10** (1999), 547–567.

Francesco Di Plinio, IU Bloomington
email: **fradipli@indiana.edu**

3 Decay of circular means of Fourier transforms of measures

after Thomas Wolff [5]
A summary written by Taryn Flock

Abstract

We summarize the proof in [5] of a sharp decay estimate for the L^2 circular means of the Fourier transform of a measure in $\mathbb{R}^2$ with a given fractal dimension.

3.1 Connecting distance sets and spherical means of Fourier transforms of measures

Let E be a compact set in $\mathbb{R}^2$. The corresponding distance set $\Delta(E)$ is defined as $\Delta(E) = \{|x-y| : x, y \in E\}$. It is conjectured that if $\dim E > 1$ then $|\Delta(E)| > 0$. This is one statement of Falconer's distance set problem. In [3] Mattila connects this problem to that of determining the rate of decay of circular means of the Fourier transform of a measure with specified fractal dimmension. The basic idea (as stated in [6]) is as follows. To show that $|\Delta(E)| > 0$, it suffices to show that it supports a measure that is absolutely continous with respect to Lebesque measure, which may be reduced to showing that $\Delta(E)$ supports a measure with an L^2 Fourier transform. The first step is to transform the statement E has (Hausdorff) dimmension α into a statement about measures. We introduce α-dimensional energy:

$$I_\alpha(\mu) = \int \frac{d\mu(x)d\mu(y)}{|x-y|^\alpha}$$

It is known (see for instance [6]) that if E is a compact set with $\dim E > \alpha$ then E supports a measure μ with $I_\alpha(\mu) < \infty$.

Mattila [3] shows the following

Proposition 1. *Assume that $I_\alpha(\mu) < \infty$ for some $\alpha > 1$. Let ν_0 be the pushforward of $\mu\times\mu$ by $\Delta(x,y) \mapsto |x-y|$ and define $d\nu(t) = e^{i\pi/4}t^{-1/2}d\nu_0(t) + e^{-i\pi/4}|t|^{-1/2}d\nu_0(-t)$ supported on $\Delta(E) \cup -\Delta(E)$. Then $\hat{\nu} \in L^2$ if and only if $\int_{R=1}^{\infty}(\int |\hat{\mu}(Re^{i\theta})|^2)^2 d\mu R dR < \infty$.*

At the heart of the proof is that $\hat{\nu}(k) = |k|^{1/2} \int |\hat{\mu}(ke^{i\theta})|^2 d\theta + \mathcal{O}(|k|^{1/2-\alpha} I_\alpha(\mu))$ which can be shown using the asymptotics of $\widehat{\sigma_R}$.

Having introduced this connection, Mattila's paper [3] begins the systematic study of estimates for the L^2 spherical means of Fourier tranforms of measures with finite energy. In [3] it is shown that for $\alpha < 1$ the best exponent is $\min(\alpha, 1/2)$. The $\alpha \geq 1$ case is addressed by Mattila [3], Sjölin [4], who obtained $\alpha - 1$, Bourgain [1] who obtained $\alpha/2 - (1/6) - \epsilon$, and Wolff in the paper summarized here [5] who obtains the bound $\alpha/2 - \epsilon$, which is sharp for $\alpha > 1$ ([3],[4]). Wolff's method is closely related to that of Bourgain, in particular both make use of the techniques introduced by Fefferman and Cordoba in the context of the restriction and the Bochner-Riesz problems.

3.2 The result

Theorem 2. *Fix $\alpha \in (0, 2)$. Then for any $\epsilon > 0$, there is a constant C_ϵ such that the following is true. Let μ be a positive measure in $\mathbb{R}^2$ supported in the unit disc and with α-dimensional energy*

$$I_\alpha(\mu) = \int \frac{d\mu(x)d\mu(y)}{|x-y|^\alpha} = 1.$$

Then for any $R \geq 1$

$$\int_{-\pi}^{\pi} |\hat{\mu}(Re^{i\theta})|^2 d\theta \lesssim C_\epsilon R^{-((\alpha/2)-\epsilon)}.$$

Notation: Let ϕ be a radial function and T a rectangle then ϕ_T is ϕ composed with the affine map taking T to the unit square. Suppose $t > 0$ then tT is the dilation of T around its center by a factor of t. The *dual* rectangle to T is the rectangle that shares axis directions but has reciprocal side lengths. By "the axis"' we mean the longer axis unless both sides are comparable. Length is measured along the longer axis, width along the shorter. Indicator functions are denoted χ_E and Lebesque measure as $|E|$. R is always assumed to be sufficiently large.

3.3 Reductions

Our first step is to transform the given condition $I_\alpha(\mu) = 1$ into a form that is easier to manipulate. This is handled by a lemma whose proof we omit.

Lemma 3. *Under the assumption that μ is a positive measure supported on the unit the disc with $I_\alpha(\mu) = 1$, we may decompose μ as a sum of $O(logR)$ measures μ_j with disjoint support such that*

$$\mu_j(\mathbb{R}^2) \sup_x \sup_{r \geq R^{-1}} \frac{\mu_j(D(x,r))}{r^\alpha} \leq 1.$$

Using the lemma, it suffices to show that given a measure μ supported on the unit disk such that $\sup_x \sup_{r \geq R^{-1}} \frac{\mu(D(x,r))}{r^\alpha} = B_1$ and $\mu(\mathbb{R}^2) = B_2$, we have that $\int_{-\pi}^{\pi} |\hat{\mu}(R_e^{i\theta})|^2 d\theta \leq B_1 B_2 R^{-(\alpha/2)+\epsilon}$.

From here we further reduce using the uncertainty principle to showing that for $A_R = \{x \in \mathbb{R}^2 : R-1 < |x| < R+1\}$ we have

$$\frac{1}{R} \int_{A_R} |\hat{\mu}(x)|^2 dx \lesssim B_1 B_2 R^{-(\alpha/2)+\epsilon}.$$

We next use duality to further reduce the problem. Fix a function f supported in A_R such that $\int |f|^2 dx = 1$, and let G be its inverse Fourier transform. Then $\|\widehat{\mu}\chi_{A_R}\|_2 = \sup_{f:\|f\|_2=1} |\int \widehat{\mu}\chi_{A_R} f d\xi| = |\int G d\mu| = J$. Thus we must prove

$$J \lesssim B_1 B_2 R^{1/2-(\alpha/4)+\epsilon}.$$

3.4 The proof

Decompose A_R into disjoint circular rectangles β of angular length $R^{-1/2}$. $\beta = \{\xi : |\xi| \in (R-1, R+1), \frac{\xi}{|\xi|} \in \gamma\}$ where γ is an arc on the unit circle of appropriate length. For each β fix ρ_β a standard rectangle $C \times CR^{1/2}$ which contains β. Then $G = \sum_\rho G_\rho$ where G_ρ the inverse Fourier transform of $\chi_\beta f$. For fixed ρ tile $\mathbb{R}^2$ with rectangles dual to ρ, say ρ^*. Recall that b_{ρ^*} is b composed with the affine map taking ρ^* to the unit square. Let $G_\rho^{\rho^*} = b_{\rho^*} G_\rho$, for b a radial Schwartz function with compact support on the Fourier side, such that the $\mathbb{Z}^2$ translations of b form a partion of unity. Thus we have that $\sum_{\rho^*} G_\rho^{\rho*} = G$. Note that by convention the sum varies over both ρ and the dual family ρ^*.

The next stage of the arguement is to find a small family of the ρ^* which contain a resonable fraction of the mass of our integral. To this end we introduce a lower bound for J. By the definitions shows that $B_2 \leq B_1$, thus we may assume that $J \geq B_2 R^{-10}$ as otherwise there is nothing to prove.

Fix M sufficiently large and $\phi = \min(1, |x|^{-M})$. Let $\mathcal{F}_0 = \{\rho^* : \|\phi_{\rho^*}^{-1} G_\rho^{\rho^*}\|_\infty \in (h/2, h]\}$.[7] We wish to choose h such $\|\int \sum_{\rho \in \mathcal{F}_0} G_\rho^{\rho^*} d\mu\| \gtrsim (\log R)^{-1} J$. Note that $[a, b]$ where $0 < a, b$ are powers of R contains $C \log R$ dyadic points. Hence, by the pigeon hole principle, it suffices to find upper and lower bounds for h that are powers of R.

The upper bound $h \lesssim R^{1/4}$ comes from the following

Lemma 4. *We have* $\sum_\rho \sum_{\rho^*} \|\phi_{\rho^*}^{-1} G_\rho^{\rho^*}\|_\infty^2 \le C_M R^{1/2}$.

The lower bound $h > R^{-100}$ comes from the observation that

$$|\int \sum_{\rho^* : \|\phi_{\rho^*}^{-1} G_\rho^{\rho^*}\|_\infty \le R^{-100}} G_\rho^{\rho^*} d\mu| \lesssim B_2 R^{-100}.$$

Thus there is an h such that for $\mathcal{F}_0 = \{\rho^* : \|\phi_{\rho^*}^{-1} G_\rho^{\rho^*}\|_\infty \in (h/2, h]\}$ we have

$$\|\int \sum_{\rho \in \mathcal{F}_0} G_\rho^{\rho^*} d\mu\| \gtrsim (\log R)^{-1} J.$$

Lemma 4 gives that $Nh^2 \lesssim R^{1/2}$ where N is the cardinality of $\mathcal{F}_0$. Next we choose a fixed square with side 10, Q, such that for $\mathcal{F}_1 = \mathcal{F}_0 \cup Q$ we have $\|\int \sum_{\rho \in \mathcal{F}_1} G_\rho^{\rho^*} d\mu\| \gtrsim R^{-\epsilon/2} J$. This is possible (again by the pigeon hole priciple) because μ is supported on the unit disk and $|G_\rho^{\rho^*}| \le h\phi_{\rho^*}$.

Bourgain's proof in [1] tackles integrals over rectangles analogous to β using the Kakeya maximal function discussed in [2]. Wolff replaces that arguement with this careful accounting of rectangles and the following lemma:

Lemma 5. *Let C_0 be a large constant and let Q_0 be a square with side length C_0. Let $\mathcal{F}$ be a finite set of rectangles in $\mathbb{R}^2$ with length 1 and width $\delta = R^{-1/2}$. Further assume that all the rectablge in $\mathcal{F}$ are contained in Q_0 and that the cardinality of $\mathcal{F}$ is less than δ^{-100}. Then it is possible to partion $\mathcal{F}$ into at most $C(\log(1/\delta))^2$ families $\mathcal{F}_{ij}$ so that for each fixed i and j we have numbers p and θ and a family of rectangles $\mathcal{G}_{ij} = \{\tau_k\}$ with length between 1 and C_2 and width between θ and 2θ satisfying*

1. *If $T \in \mathcal{F}_{ij}$ then $T \subset \tau_k \in \mathcal{G}_{ij}$.*

2. *If $T \in \mathcal{F}$, then T is contained in at most a bounded number of τ_k's.*

[7]Note that ρ is varying as is ρ^*.

3. *Each $\tau_k \in \mathcal{G}_{ij}$ contains roughly $p(\theta/\delta)$ T's $T \in \mathcal{F}$.*

4. $\| \sum_{\tau_k \in \mathcal{G}_{ij}} \chi_{\tau_k} \|_2^2 \leq C \log(1/\delta) \sum_{\tau_k \in \mathcal{G}_{ij}} |\tau_k|$.

5. $\| \sum_{\tau_k \in \mathcal{F}_{ij}} \chi_T \|_2^2 \leq C \log(1/\delta) \sum_{\tau_k \in \mathcal{G}_{ij}} |\tau_k|$.

We note that (4) and (5) can be adjusted to apply to the approximate cut-off functions.

The families $\mathcal{F}_{ij}$ and $\mathcal{G}_{ij}$ are constructed as follows. Define for each $T \in \mathcal{F}$ Π_T a rectangle with the same center and axes as T that maximizes $d(\Pi) = \delta \frac{\operatorname{card}(|T' \in \mathcal{F} : T' \in \Pi)}{|\Pi|}$. Π has width, $\theta_T \in [\delta, 2C_0]$ and length$\in [1/2, 2C_0]$. Note $d(\Pi) \in [1, \delta^{-101}]$. Define $\mathcal{F}_{ij} = \{T : \theta_T \in (2^{-i+1}, 2^{-i}] d(\Pi_T) \in (2^j, 2^{j+1}]\}$. For each $\mathcal{F}_{ij}$ θ is defined to be $C_1 2^{-i}$ for C_1 a large constant (to be chosen such that (1) holds) and $p = 2^j$. The bounds on the θ_T and Π_T restricts the number of families. Take $\{\tau_k\} = \{C_1 \Pi_T : T\mathcal{G}\}$ where $\mathcal{G} \subset \mathcal{F}_{ij}$ is the maximal subset such that $T_1, T_2 \in \mathcal{G}$ implies $\Pi_{T_1} \not\subseteq 2\Pi_{T_2}$. (1-3) Follow readily from the construction, (4) and (5) require more calculation (see [5] also [2]).

As $\mathcal{F}_1$ is contained in a square of sidelength 10. It suffices to verify that the cardinality of $\mathcal{F}_1$ is less than δ^{-100}. This condition is satisfied as the rectangles T belong to a δ separated set on the circle and the rectangles with a given axis direction are disjoint. As there are at most logarithmically many $\mathcal{F}_{ij}$ we may fix one family such that the following hold

1. If $\rho^* \in \mathcal{F}$, then $|G_\rho^{\rho^*}| \leq h\phi_{\rho^*}$

2. There are numbers p and $\theta \geq R^{-1/2}$ and a set $\mathcal{G}$ of rectangles of width about θ so that each rectangle in $\mathcal{F}$ is contained in at least one rectangle of $\mathcal{G}$; each $\tau \in \mathcal{G}$ contains roughly $p\theta\sqrt{R}$ rectangles from $\mathcal{F}$ and each rectangle in $\mathcal{F}_1$ is contained in at most a bounded number of rectangles from $\mathcal{G}$. As the cardinality of $\mathcal{F} \leq N$, $\operatorname{card}(G) \lesssim \frac{N}{p\theta\sqrt{R}}$

3. $\| \int \sum_{\rho^* \in \mathcal{F}} G_\rho^{\rho^*} d\mu \| \geq R^{-\epsilon} J$

Futher decompose the circle into arcs Θ of angular length roughly θ. Define $\mathcal{F}(\rho) = \mathcal{F} \cup \rho^*$ and $\mathcal{G}(\Theta)$ to be $\tau \in \mathcal{G}$ whose axis directions are in a suitable fixed dilate of Θ. Define $H_\rho = \sum_{\rho^* \in \mathcal{F}(\rho)} G_\rho^{\rho^*}$, $H_\Theta = \sum_{\rho \in \Theta} H_\rho$, $P = (\sum_\Theta |H_\Theta|^2)^{1/2}$, and $\Phi_\Theta = \sum_{\tau \in \mathcal{G}(\Theta)} \phi_\tau$.

Decompose space into squares of side θ. If $\max_Q \sum_\Theta \Phi_\Theta \in [A, 2A]$, call Q and A-square. Note that $\max_Q \sum_\Theta \Phi_\Theta \sim \min_Q \sum_\Theta \Phi_\Theta$. We again find that at most $\log R$ dyadic values of A such that the A-squares contribute to the integral. Thus the final computation will be to integrate over a family of A-squares.

Lemma 6. *There is a fixed radial function q such that $|q(x)| \leq C_N(1+|x|)^{-N}$ for any given N such that the following holds. Let $t = (\theta R)^{-1}$, $q^t(x) = t^{-2}q(t^{-1}x)$ and $\bar{\mu} = q^t * \mu$. Then for any Θ and any square Q of side length $\geq t$ $\int_Q |H_\Theta| d\mu \lesssim \sum_{j\geq 0} 2^{-Mj} \int_{2^jQ} |H_\Theta| d\bar{\mu}$.*

This lemma, holders inequality, and estimates for $||P||_4$ and $\|\frac{d\bar{\mu}}{dx}\|_L^{\frac{4}{3}}(E_A^j)$ where E_A^j are the 2^j dilates of A squares, completes the proof. Our careful decompostion of space and tight control of supremums of $|H_\Theta|$ provide good estimates for the square function P. Estimates for $\bar{\mu}$ use that $\mu(\mathbb{R}^2) = B_2$, $\mu(D_r) \lesssim B_1 r^\alpha$, and q's rapid decay we have $\mu(\bar{\mathbb{R}}^2) \lesssim B_2 \|\frac{d\bar{\mu}}{dx}\|_\infty \lesssim B_1(\theta R)^{2-\alpha}$. The final estimates are $\|P\|_4 \lesssim h(\log R)^{1/4}(pNR^{-1/2})^{1/4}$ and $\|\frac{d\bar{\mu}}{dx}\|_L^{\frac{4}{3}}(E_A^j) \lesssim B_1(2\theta R)^{2-\alpha}$ which prove the theorem .

References

[1] Bourgain, J., *Hausdorff dimension and distance sets*, Isreal J. Math, 1994, no. 87, 193-201

[2] Cordoba, A. *The Kakeya maximal function and the spherical summation multipliers* Amer. J. Math, 1977, no .99, 1-22

[3] Matilla, P., *Spherical averages of Fourier transforms of measures with finite energy; dimensions of intersections and distance sets* Mathematika, 1987, no. 34, 207-228

[4] Sjölin,*Estimates of spherical averages of the Fourier transforms and dimensions of sets*, 1993, Mathematika, no. 40, 322-330

[5] Wolff, T., *Decay of circular means of Fourier transforms of measures.* Internat. Math. Res. Notices 1999, no. 10, 547 - 567

[6] Wolff, T., *Lectures in Harmonic Analysis*, http://www.math.ubc.ca/ ilaba/wolff/notes_march2002.pdf

Taryn Flock, UC Berkeley
email: **tflock@math.berkeley.edu**

4 Bochner-Riesz conjecture, restriction conjecture and oscillatory integral

after T. Tao [1] and J. Bourgain and L. Guth [2]
A summary written by Yi Hu

Abstract

We state a theorem that the Bochner-Riesz conjecture implies the restriction conjecture. Also we discuss some examples to illustrate the optimality of the exponent in certain oscillatory integral.

4.1 Bochner-Riesz conjecture and restriction conjecture

This section is devoted to the discussion on the relation between the Bochner-Riesz conjecture and the restriction conjecture, discovered by T. Tao [1]. Let $n \geq 2$ be the dimension. We use $BR(p, \alpha)$ to denote that $S^{\delta(p)+\alpha}$ is bounded on L^p, where

$$\delta(p) = \max\left(n\left|\frac{1}{p} - \frac{1}{2}\right| - \frac{1}{2}, 0\right)$$

and S^δ is the Bochner-Riesz multiplier

$$\widehat{S^\delta f}(\xi) = \left(1 - |\xi|^2\right)_+^\delta \widehat{f}(\xi).$$

A well-known conjecture is

Conjecture 1 (Bochner-Riesz conjecture). *$BR(p, \varepsilon)$ holds for all $1 \leq p \leq \infty$ and $\varepsilon > 0$.*

Similarly, we use $R(p, \alpha)$ to denote the localized restriction estimate

$$\|\mathfrak{R}f\|_{L^p(S^{n-1})} \lesssim R^\alpha \|f\|_{L^p(B(0,R))}$$

for f supported in $B(0, R)$, where $\mathfrak{R}f = \hat{f}\Big|_{S^{n-1}}$ is the sphere restriction operator. The restriction conjecture states

Conjecture 2 (Restriction conjecture). *$R(p, 0)$ holds for all $1 \leq p < \frac{2n}{n+1}$.*

These two conjectures are widely believed to be at least heuristically equivalent. Here, the main conclusion is, under some condition, the Bochner-Riesz conjecture implies the restriction conjecture (see Corollary 5). This is an immediate consequence of the following two results.

Theorem 3. *If $1 \leq p \leq \frac{2n}{n+1}$, then $BR(p,\alpha)$ implies $R(p,2\alpha)$.*

Theorem 4. *If $1 < p < 2$ and $0 < \alpha \ll 1$, then $R(p,\alpha)$ implies $R(q,0)$ whenever*

$$\frac{1}{q} > \frac{1}{p} + \frac{C}{\log \frac{1}{\alpha}} .$$

Therefore we could easily get

Corollary 5. *If $1 < p \leq \frac{2n}{n+1}$ and $BR(p,\varepsilon)$ holds for all $\varepsilon > 0$, then $R(q,0)$ holds for all $1 \leq q < p$.*

Since, as we know, that $R(p,0)$ is false for any $p \geq \frac{2n}{n+1}$, so the exponent in the Corollary 5 is optimal.

Theorem 3 is based on the observation that the Bochner-Riesz operator resembles the restriction operator when evaluated at points far away from the support of the function. The basic idea of the proof is to exploit the heuristic approximation (it is accurate when x is far away from the support of f)

$$S^{\delta(p)+\alpha} f(x) \sim |x|^{-\alpha} \frac{e^{\pm 2\pi i |x|}}{|x|^{\frac{n}{p}}} \widehat{f}\left(\pm \frac{x}{|x|}\right).$$

Theorem 4 can be viewed as, if one could control the (L^p, L^p)-norm of $\mathfrak{R}$ reasonably well on large balls, then one could control the (L^q, L^q)-norm of the same operator on all of $\mathbb{R}^n$, where q is less (or worse) than p. The proof of this theorem relies on two ideas. First we bootstrap the localized restriction estimate so that it applies to functions supported on a "sparse" union of balls of constant radius, and then we need a Calderón-Zygmund-type decomposition which covers a set E by some small number of sparse collections of balls whose size could be well controlled.

To be precise, we describe the "sparse" phenomenon as follows. A collection $\{B(x_i, R)\}_{i=1}^{N}$ of balls is sparse if the centers x_i are $R^C N^C$ separated. Here C is a constant changing from line to line. Then the two main steps of proving Theorem 4 could be written as

Lemma 6. *Suppose $R(p,\alpha)$ holds. Then*

$$\|\mathfrak{R}f\|_p \lesssim R^{\alpha}\|f\|_p$$

whenever f is supported on $\cup_i B(x_i,R)$ and $\{B(x_i,R)\}$ is a sparse collection of balls.

Lemma 7. *Suppose E is the union of cubes of size c ($c \sim 1$) and $N \geq 1$. Then there exist $O(N|E|^{\frac{1}{N}})$ sparse collections of balls that cover E, such that the balls in each collection have radius $O\left(|E|^{C^N}\right)$.*

In both Theorem 3 and 4 the uncertainty principle plays a minor but recurring role. We introduce a spatial uncertainty of $O(1)$, and conversely we use a spatial localization at scale R to introduce a frequency uncertainty of $O(1/R)$. In fact, we have the following lemma.

Lemma 8. *For each $R \gg 1$, let $K_R(x,y)$ be a bounded, compactly supported function, where the bounds and support are independent of R, and let $1 \leq p,q \leq \infty$. Suppose $b(x,y,\tau)$ is a bounded function that is C^{∞} in x and τ, for $0 \leq \tau \lesssim 1$, and x,y in the support of K. Then, if the operators A_R and B_R are defined by*

$$A_R f(x) = \int K_R(x,y)f(y)dy,$$
$$B_R f(x) = \int K_R(x,y)b(x,y,1/R)f(y)dy,$$

then for all $N > 0$ we have

$$\|B_R\|_{L^p\to L^q} \lesssim \|A_R\|_{L^p\to L^q} + O(R^{-N}).$$

Lemma 8 shows that one can freely modify the amplitude or phase of a compactly supported oscillatory integral by a smooth $O(1)$ factor without affecting its regularity properties.

4.2 Oscillatory integral

In this section we present an example that will illustrate the optimality of the exponent of certain oscillatory integral. Let the oscillatory integral operator be defined as

$$(T_\lambda f)(x) = \int e^{i\lambda\psi(x,y)}f(y)dy \qquad (\|f\|_\infty \leq 1) \tag{1}$$

Here $x \in \mathbb{R}^n$, $y \in \mathbb{R}^{n-1}$ are restricted to a neighborhood of 0, ψ is a real analytic phase function with

$$\psi(x, y) = x_1 y_1 + \cdots + x_{n-1} y_{n-1} + x_n \langle Ay, y \rangle + O(|x||y|^3) + O(|x|^2 |y|^2)$$

and A is a non-degenerate square matrix. The following theorem in [2] describes an estimate of $T_\lambda f$.

Theorem 9. *Let T_λ be as above with A positive or negative definite. Then*

$$\|T_\lambda f\|_p \leq C_p \lambda^{-\frac{n}{p}} \|f\|_\infty \tag{2}$$

holds for p satisfying

$$\begin{cases} p > \frac{2(4n+3)}{4n-3} & \text{if } \; n \equiv 0 \pmod 3 \\ p > \frac{2n+1}{n-1} & \text{if } \; n \equiv 1 \pmod 3 \\ p > \frac{4(n+1)}{2n-1} & \text{if } \; n \equiv 2 \pmod 3 \end{cases}.$$

For $n = 3$, the exponent $\frac{10}{3}$ in Theorem 9 is optimal. Here we will give two examples showing the optimality and also explaining the differences between the elliptic and hyperbolic cases.

To illustrate the hyperbolic case, let the phase function be

$$\psi(x, y) = -x_1 y_1 - x_2 y_2 + 2 x_3 y_1 y_2 + x_3^2 y_2^2.$$

Then for the oscillatory operator (1), take

$$f(y) = e^{i \lambda y_1^2}.$$

The key point is, by restricting x to a $\frac{1}{\lambda}$-neighborhood of the surface

$$S : x_1 x_3 = x_2,$$

we could use stationary phase method to obtain

$$\left| \int e^{i\lambda\psi(x,y)} f(y) dy \right| \sim \frac{1}{\sqrt{\lambda}},$$

which gives Theorem 9 for $p \geq 4$.

For the elliptic case, take in (1)

$$\psi(x, y) = -x_1 y_1 - x_2 y_2 + \frac{1}{2} x_3 y_1^2 + x_3^2 y_1 y_2 + \frac{1}{2}(x_3 + x_3^3) y_2^2$$

and

$$f(y) = \sum_{s<\sqrt{\lambda}} \sigma_s 1_{[\frac{s}{\sqrt{\lambda}}, \frac{s+c}{\sqrt{\lambda}}]}(y_2) e^{i\lambda \frac{s}{\sqrt{\lambda}} y_1},$$

where $\sigma_s = \pm 1$ and $c > 0$ is a small constant. Restricting x to the region

$$R = \left\{ x_3 \sim 1,\ |x_2 - x_1 x_3| = o\left(\frac{1}{\sqrt{\lambda}}\right) \right\},$$

we get

$$\|T_\lambda f\|_{L^p(R)} \gtrsim \left(\frac{1}{\lambda}\right)^{\frac{3}{4}+\frac{1}{2p}},$$

which implies $p \geq \frac{10}{3}$ is necessary. Also, this is where the exponent $\frac{10}{3}$ comes from.

References

[1] T. Tao, *The Bochner-Riesz conjecture implies the restriction conjecture.* Duke Math Journal **96** (1999), no. 2, 363–375.

[2] J. Bourgain and L. Guth, *Bounds on oscillatory integral operators based on multilinear estimates.* Preprint (March 28, 2011).

Yi Hu, UIUC
email: **yihu1@illinois.edu**

5 The endpoint case of the Bennett–Carbery–Tao multilinear Kakeya conjecture

after L. Guth [1]
A summary written by K. Hughes

Abstract

We formulate and motivate the multilinear Kakeya problem of Bennett–Carbery–Tao and mention it's relation with the linear Kakeya conjecture. Then we give an exposition of L. Guth's endpoint estimate for the multilinear Kakeya conjecture.

5.1 Introduction

5.1.1 The multilinear Kakeya problem

We will sketch of proof of the endpoint case of the Bennett–Carbery–Tao multilinear Kakeya conjecture, which is

Theorem 1. *In $\mathbb{R}^d$, we consider collections of tubes $\{T_{j,a}\}$ such that a tube $T_{j,a}$ has infinite length in the direction $v_{j,a}$ and unit width in all perpendicular directions. The tubes are indexed so that for $1 \leq j \leq d$, there are $A(j)$ tubes. If the tubes satisfy the transversality condition $|det(v_{1,a_1}, \ldots, v_{d,a_d})| > \theta$ for any choice $1 \leq a_j \leq A(j), 1 \leq j \leq d$ then*

$$\int_{\mathbb{R}^d} \prod_{j=1}^{d} \left(\sum_{a=1}^{A(j)} T_{j,a}(x) \right)^{\frac{1}{d-1}} \ll_d \theta^{\frac{1}{1-d}} \prod_{j=1}^{d} A(j)^{\frac{1}{d-1}}.$$

Remark 2. *Throughout this summary, we use the notation $f \ll g$ to mean $|f| \leq C|g|$ and all implicit constants can depend on dimension.*

5.1.2 Motivation: The linear Kakeya problem

The multilinear version of the Kakeya conjecture came about from first considering a bilinear Kakeya operator (see [3]); there, they established an equivalence between the linear and bilinear Kakeya estimates and improved on the linear Kakeya conjecture by adapting previous arguments to the bilinear setting. The linear Kakeya problem can be formulated in a variety of ways; we choose the following formulation.

Conjecture 3 (Kakeya)**.** *Suppose that Ω is a δ-separated subset of the sphere S^{d-1} in $\mathbb{R}^d$ and $\{T_e\}_{e\in S^{d-1}}$ is a collection of tubes with width δ and length 1, e.g. $T_e^\delta(a) := \{x \in \mathbb{R}^d : |(x-a)\cdot e| \le 1/2; |(x-a)^\perp| < \delta\}$ where $\perp$ refers to the direction perpendicular to e; a is the center of the tube T_e while e is its direction. Then*

$$\left\|\sum_{e\in\Omega} \mathbf{1}_{T_e}\right\|_{\frac{d}{d-1}} \ll_\epsilon \delta^{-\epsilon}\left(\sum_{e\in\Omega}|T_e|\right)^{\frac{d}{d-1}}$$

Remark 4. *A solution to this conjecture implies the more classical Kakeya problem of showing that any Kakeya set (a set containing a unit line segment in every direction) must have dimension 0. The factor $\delta^{-\epsilon}$ is necessary since there exist Kakeya sets of measure 0.*

5.2 Previous results: Bennett–Carbery–Tao

In [2], Bennett–Carbery–Tao formulated the multilinear Kakeya problem and proved the optimal estimates except for the endpoint. More precisely, they proved:

Theorem 5 (Bennett–Carbery–Tao)**.** *In $\mathbb{R}^d$, we consider collections of tubes $\{T_{j,a}\}$ such that a tube $T_{j,a}$ has infinite length in the direction $v_{j,a} \in \mathbb{S}^{d-1}$ and unit width in all perpendicular directions. The tubes are indexed so that for $1 \le j \le d$, there are $A(j)$ tubes. If the tubes satisfy the transversality condition: the directions $v_{j,a}$ are assumed to lie in a sufficiently small neighborhood of the axial direction e_j, say within $(100d)^{-1}$ of a degree. Then for $\frac{d}{d-1} < q \le \infty$,*

$$\left\|\prod_{j=1}^{d}\sum_{a=1}^{A(j)} T_{j,a}(x)\right\|_{L^{\frac{q}{d}}(\mathbb{R}^d)} \ll_d \prod_{j=1}^{d} A(j).$$

Remark 6. *Guth's transversality condition is more quantitative than the one here and the factor $\theta^{\frac{1}{1-d}}$ comes from considering affine transformations.*

Remark 7. *Bennett–Carbery–Tao's methods were very different from Guth's; they used multiscale analysis, monoonicity estimates and heat flows to achieve all but the endpoint space $L^{\frac{d}{d-1}}$.*

5.3 MLK Jr.

We warm up to proving Guth's theorem by proving a weaker result. This was originally done in [1] because the proof is simple and it helps in understanding MLK. For the same reasons we include it here.

Proposition 8 (MLK Jr.)**.** *In $\mathbb{R}^d$, we consider collections of tubes $\{T_{j,a}\}_{j=1,a=1}^{d,A}$ such that a tube $T_{j,a}$ has infinite length in the direction $v_{j,a}$ and unit width in all perpendicular directions. If the tubes satisfy the transversality condition: the direction $v_{j,a}$ of the tube $T_{j,a}$ is within a $1/d$ of a degree of the j-axis, then we have the following bound for the set of points contained in a tube in each direction (taking a slightly perturbative definition of direction)*

$$I := \left|\cap_{j=1}^{d}\left(\cup_{a=1}^{A} T_{j,a}\right)\right| \ll A^{\frac{d}{d-1}}.$$

The three main ideas in the proof are the Polynomial Ham Sandwich Theorem, directed volumes of hypersurfaces and the pigeon-hole principle.

Definition 9. *We write $Vol_{d-1}(S)$ for the $d-1$ dimensional volume of a compact subset S of a hypersurface Z in $\mathbb{R}^d$.*

Definition 10. *For a set S, a hypersurface Z and direction $v \in \mathbb{S}^{d-1}$, we write $V_{S\cap Z}(v)$ for the $d-1$-dimensional volume of $S \cap Z$ in the v direction.*

Lemma 11 (Guth's cylinder estimate)**.** *If T is a cylinder with unit width and direction v, and Z is an algebraic hypersurface in $\mathbb{R}^d$, then $V_{Z\cap T}(v) \ll deg(Z)$.*

Lemma 12 (Polynomial Ham Sandwich Theorem)**.** *If $U_1, \ldots, U_N$ are open subsets of $\mathbb{R}^d$ with finte volume, then there exists a hypersurface Z with $deg(Z) \le N^{1/d}$ bisecting each U_i.*

Proof. Impose the unit lattice on $\mathbb{R}^d$ to find a collection of cubes Q_n for $1 \le n \le N$ (Q will always denote such a unit cube and we drop the subscripts unless necessary). By the Polynomial Ham Sandwich Theorem, there is a hypersurface Z of degree $D \ll N^{1/d}$ bisecting the cubes. For each cube Q, since the hypersurface bisects it, $Vol_{d-1}(Q \cap Z) \gg 1$. Thus in one of the standard orthogonal directions $Q \cap Z$ has a big projection i.e. the directed volume of $Q \cap Z$ in the direction of v, $V_{Q\cap Z}(v) \gg 1$. Furthermore, since there is a tube $T_{j(Q)}$ in each direction passing through Q; we associate the tube who's direction has a big projection with Q; we call this tube T_Q. $Vol_{d-1}(T_Q\cap$

Z) maybe small as T_Q could be shifted to the side of a cube. However the cubes have unit width, so we know that the tube $\tilde{T}_Q$ dilated by $1+\sqrt{n}$ covers the tube. Then the $d-1$-volume $Vol_{d-1}(\tilde{T}_Q \cap Q \cap Z) \gg V_{Q\cap Z}(v) \gg 1$ where v is the tube's direction.

There are $d \cdot A$ tubes and we have a map from cubes to tubes. Therefore, by the pigeonhole principle there exists a tube intersecting a lot of cubes, say $\gg N/A$ with the property that $Vol_{d-1}(\tilde{T}_Q \cap Q \cap Z) \gg 1$ for each of these cubes. So $d-1$-volume of the tube in its direction is $Vol_{d-1}(\tilde{T} \cap Z) \gg N/A$. However, by the cylinder estimate $Vol_{d-1}(\tilde{T} \cap Z) \ll N^{1/d}$. Combining these two estimates, we're done. □

5.4 Proof of the endpoint multilinear Kakeya estimate

We would like to imitate the proof for MLK Jr but there are added difficulties. The most obvious one is that our directions are varying more than before. Also, the number of tubes in the "j-direction", by which we mean the collection of tubes with j fixed, is varying. These issues require more cohomological machinery (which we'll take for granted) than the PHS Theorem, but the basic game is the same. Since our transversality condition is more general we'll need to consider more than the projection of a hypersurface in only a few directions; this leads to studying the visibility of a hypersurface, $Vis(Z)$ which combines the directional volume of every direction. The visibility is big when a surface has a large directional volume in a few directions or a moderate directional volume in many directions. Geometrically, the surface defines a norm and the surface's visibility is the volume of the unit ball under its norm. Due to analytic difficulties, Guth actually uses a mollified version which makes Vis continuous. We ignore these difficulties.

Before we begin, we make same reduction to cubes as before. Decompose $\mathbb{R}^d$ into unit cubes given by the standard unit lattice. For such a cube Q, let $m_j(Q)$ be the number of tubes $T_{j,a}$ intersecting Q and $M(Q) = \prod_j m_j(Q)$. This is our multiplicity function. Since there are only finitely many tubes, there are only finitely many intersctions and thus only finitely many cubes to consider, say Q_n, $1 \leq n \leq N$. Our problem reduces to

Theorem 13. *In $\mathbb{R}^d$, we consider collections of tubes $\{T_{j,a}\}$ such that a tube $T_{j,a}$ has infinite length in the direction $v_{j,a}$ and unit width in all perpendicular directions. The tubes are indexed so that for $1 \leq j \leq d$, there are $A(j)$ tubes. If the tubes satisfy the transversality condition $|det(v_{1,a_1}, \ldots, v_{d,a_d})| > \theta$ for* *any choice* $1 \leq a_j \leq A(j), 1 \leq j \leq d$*, then*

$$\sum_{n=1}^{N} M(Q_n)^{\frac{1}{d-1}} \ll \theta^{\frac{1}{1-d}} \prod_{j=1}^{d} A(j)^{\frac{1}{d-1}}$$

Proof. Now that we're thinking of Vis like the projected volumes above, we want estimates for visibility. For the proof of MLK Jr., we used the Polynomial Ham Sandwich theorem to find a hypersruface with bounded degree that bisected the cubes which gave a us a lower bound for the volume. Instead, we now find a hypersurface of bounded degree that has large visibility on the cubes giving us a lower bound for the visibility.

Lemma 14 (Large visibility lemma)**.** *Denote the collection of unit cubes* $\mathcal{Q}$ *as given by decomposing* $\mathbb{R}^d$ *with respect to the lattice. If* $f : \mathcal{Q} \to \mathbb{Z}_{\geq 0}$*, then there exists a hypersurface* Z *such that* $Vis(Z \cap Q) \geq f(Q)$ *for all cubes* $Q \in \mathcal{Q}$ *and with* $deg(Z) \ll |\sum_{q \in \mathcal{Q}} f(Q)|^{1/d}$*.*

Remark 15. *All the necessary algebraic topology is hidden in this statement. For more background, see [1].*

For the time being we won't choose our function f and collect some basic facts involving the hypersurface Z with degree D. If our hypersurface contains hyperplanes in d transverse directions, then we can bound its visibility by the projected volumes.

Lemma 16 (Upper bound for visibility of a hypersurface in a cube)**.** *Given a hypersurface* Z *and cube* Q *such that* Z *contaings a hyperplane intersecting the cube in each of the standard directions, we have*

$$Vis(Z \cap Q) \ll \theta^{-1} \prod_{j=1}^{d} V_{Z \cap Q}(v_j).$$

We redefine our hypersurface by adding hyperplanes for each cube as in the above lemma. This doesn't change the degree of the hypersurface and helps us since we now have upper and lower bounds for visibility. We see that projected volumes are related to visibility, again we can use Guth's cylinder estimate to relate this back to the degree of the hypersurface. For any unit cube Q and tube T in our collection with direction v, we have $V_{T \cap Z}(v) \ll D$. Therefore,

$$\sum_{Q \cap T \neq \emptyset} V_{Q \cap Z}(v) \ll D$$

since we only need to consider cubes that intersect the tube which is what our sum is over. Choose a tube in the j-direction, say T_j with direction v_j, then $\sum_{a=1}^{A(j)} \sum_{Q \cap T \neq \emptyset} V_{Q \cap Z}(v_j) \ll D \cdot A(j)$. By Fubini's theorem, this double sum is just $\sum_{Q \in \mathcal{Q}} m_j(Q) V_{Q \cap Z}(v_j)$ so that

$$\sum_{Q \in \mathcal{Q}} m_j(Q) V_{Q \cap Z}(v_j) \ll D \cdot A(j).$$

Now we take the geometric average of this over each j-direction to get

$$\prod_{j=1}^{d} \left(\sum_{Q \in \mathcal{Q}} m_j(Q) V_{Q \cap Z}(v_j) \right)^{1/d} \ll D \prod A(j)^{1/d}.$$

Using Holder's inequality to bound this below, we get

$$\sum_{Q \in \mathcal{Q}} \left(\prod_{j=1}^{d} m_j(Q) V_{Q \cap Z}(v_j) \right)^{1/d} \ll D \prod A(j)^{1/d}.$$

The product is just $M(Q) \prod_{j=1}^{d} V_{Q \cap Z}(v_j)$. But $Vis(Q \cap Z) \ll \theta^{-1} \prod_{j=1}^{d} V_{Q \cap Z}(v_j)$ meaning that $\sum_Q (\theta M(Q) Vis(Q \cap Z))^{1/d} \ll D \prod A(j)^{1/d}$. By the large visiility lemma $Vis(Q \cap Z) \geq f(Q)$ and $D \ll |\sum_Q f(Q)|^{1/d}$ implying

$$\sum_Q (M(Q) f(Q))^{1/d} \ll \theta^{-1/d} |\sum_Q f(Q)|^{1/d} \prod A(j)^{1/d}.$$

We want the summand $(M(Q)f(Q))^{1/d}$ to be $M(Q)^{\frac{1}{d-1}}$; setting these equal and caculating, we get $f(Q) = M(Q)^{\frac{1}{d-1}}$. Plugging this into the above equation and rearranging, we get the desired inequality.

$\square$

References

[1] L. Guth, *The endpoint case of the Bennett–Carbery–Tao multilinear Kakeya conjecture.* Arxiv 2009;

[2] J. Bennett, A. Carbery and T. Tao, *On the multilinear restriction and Kakeya conjectures.* Acta Mathematica, 196 (2006), pp. 261-302;

[3] T. Tao, A Vargas and L. Vega, *A bilinear approach to the restriction and Kakeya conjectures.* Journal of the American Mathematical Society, Vol. 11, No. 4, 1998, pp 967-1000.

[4] T. Wolff, I. Laba and C. Shubin, *Lecture notes on Harmonic Analysis.* American Mathematical Society, University Lecture Series, vol. 29, October 2003;

Kevin Hughes, Princeton
email: **kjhughes@princeton.edu**

6 An incidence theorem in higher dimensions

after J. Solymosi and T. Tao [1]
A summary written by Mark Lewko

Abstract

We give an exposition of Solymosi and Tao's recent high dimensional analog of the Szemerédi-Trotter incidence theorem.

6.1 Introduction

We start by recalling the classical Szemerédi-Trotter incidence bound in the plane.

Theorem 1. *(Szemerédi-Trotter) Let $\mathcal{P}$ denote a set of points and $\mathcal{L}$ a set of lines in $\mathbb{R}^2$. Define the set of incidences $I(\mathcal{P},\mathcal{L}) := \{(p,l) \in \mathcal{P}\times\mathcal{L} : p \in l\}$, then*

$$|I(\mathcal{P},\mathcal{L})| \ll |P|^{2/3}|L|^{2/3} + |P| + |L|.$$

This inequality has found many applications, and it is natural to seek higher dimensional analogs. Perhaps the most obvious potential generalization would be to place the points and lines in $\mathbb{R}^d$ for $d > 2$ and ask for the same conclusion. This is true, and is easily deduced from the $d = 2$ case stated above. Indeed, consider a projection $\pi : \mathbb{R}^d \to \mathbb{R}^2$. With the exception of a finite number degenerate choices of π (such as those that send lines in $\mathcal{L}$ to points, or multiple points in $\mathcal{P}$ to the same point), we have that $|\pi(\mathcal{P})| = |\mathcal{P}|$, $|\pi(\mathcal{L})| = |\mathcal{L}|$ and $|I(\pi(\mathcal{P}),\pi(\mathcal{L}))| \leq |I(\mathcal{P},\mathcal{L})|$. Now the $\mathbb{R}^d$ version of the bound follows from applying the $\mathbb{R}^2$ bound to the sets $\pi(\mathcal{P})$ and $\pi(\mathcal{L})$. We will apply this 'generic projection trick' again below.

A more interesting, and less trivial, analog would be to replace lines with higher dimensional objects such as hyper-planes or algebraic varieties. In this direction, Toth has made the following conjecture where lines are replaced by k-dimensional affine subspaces (or *k-flats*). In order to avoid degenerate cases, we will require that no two k-flats intersect in more than one point (in which case we will say that the k-flats are *nearly disjoint*).

Conjecture 2. *(Toth) Let $d \geq 2k$. Then for a set of points, $\mathcal{P}$, and nearly disjoint k-flats, $\mathcal{L}$, in $\mathbb{R}^d$ we have that*

$$|I(P,L)| \ll_k |P|^{2/3}|L|^{2/3} + |P| + |L|.$$

Our main objective will be to prove a weak form of this conjectured inequality, namely $|I(\mathcal{P},\mathcal{L})| \leq C_{k,\epsilon}|\mathcal{P}|^{2/3+\epsilon}|\mathcal{L}|^{2/3} + \frac{3}{2}|\mathcal{P}| + \frac{3}{2}|\mathcal{L}|$ (under the above hypotheses). We'll actually prove a generalization of this with k-flats replaced by certain collections of algebraic varieties. In fact, we must work with these more general objects in order to run the inductive step of the proof. We let $\mathcal{L}$ be a set of real algebraic varieties in $\mathbb{R}^d$ and $\mathcal{P}$ a set of points in $\mathbb{R}^d$. Given a subset of incidences $\mathcal{I} \subseteq I(\mathcal{P},\mathcal{L})$, we say the tuple $\mathcal{P},\mathcal{L},\mathcal{I}$ is an *admissible* (with parameters k, d, C_0) if it satisfies the following 'pseudoline' axioms:

1. Each $\ell \in \mathcal{L}$ is a k-dimensional ($2k \geq d$) real algebraic variety of degree at most C_0 in $\mathbb{R}^d$.

2. If $j, \ell \in \mathcal{L}$ are distinct, then P contains at most C_0 points such that $(p,j), (p,\ell) \in \mathcal{I}$.

3. If $p, q \in \mathcal{P}$ are distinct, then there are at most C_0 elements $\ell \in \mathcal{L}$ such that $(p,\ell), (q,\ell) \in \mathcal{I}$.

4. If $(p,\ell) \in I(\mathcal{P},\mathcal{L})$, then p is a real smooth point of l.

5. If $p \in j, l$ (j and l distinct), then the respective tangent spaces $T_p l$ and $T_p j$ intersect only at p (that is they are transverse).

Theorem 3. *(Solymosi-Tao) Let $d \geq 2k \geq 0$ and $\epsilon, C_0 \geq 0$, and $\mathcal{P}$, $\mathcal{L}$, $\mathcal{I} \subseteq I(\mathcal{P},\mathcal{L})$ be an admissible configuration. Then*

$$|I(\mathcal{P},\mathcal{L})| \leq C_{k,\epsilon,C_0}|\mathcal{P}|^{2/3+\epsilon}|\mathcal{L}|^{2/3} + \frac{3}{2}|\mathcal{P}| + \frac{3}{2}|\mathcal{L}|.$$

First we notice that the constant in this inequality is allowed to depend on the parameters k, ϵ, and C_0. There are several steps in the proof where the choice of constant will need to be refined. To aid with the presentation, we let C_i ($i \geq 0$) denote a a constant that is allowed to depend on $k, \epsilon, C_0, C_1, C_2, \ldots, C_{i-1}$.

It suffices to consider the case $d = 2k$ by the generic projection trick. To see this, one must verify that if a configuration $\mathcal{P},\mathcal{L},\mathcal{I}$ is admissible, then the projection $\pi(\mathcal{P}), \pi(\mathcal{L}), \pi(\mathcal{I})$ is also admissible for a generic projection $\pi : \mathbb{R}^d \to \mathbb{R}^{2k}$ with $d > 2k$. We leave this easy verification to the reader. We also need π to map distinct points/varieties to distinct points/varieties, but this follows for generic π, as remarked above.

Next we record some trivial bounds on $|\mathcal{I}|$, which are easily proved via Cauchy-Schwarz.

Lemma 4. *(Trivial Bound) For an admissible configuration $(\mathcal{P}, \mathcal{L})$ we have*

$$|\mathcal{I}| \leq C_0^{1/2}|\mathcal{P}||\mathcal{L}|^{1/2} + |\mathcal{L}|, |\mathcal{I}| \leq C_0^{1/2}|\mathcal{L}||\mathcal{P}|^{1/2} + |\mathcal{P}|.$$

Thus the conclusion of Theorem 3 follows unless we have that

$$C_2|\mathcal{P}|^{1/2} \leq |\mathcal{L}| \leq C_2^{-1}|\mathcal{P}|^2. \tag{1}$$

We now summarize the proof. We will induct on the number of points in $\mathcal{P}$ and the dimension of the varieties k, with the base cases being trivial. We will find a low dimensional polynomial Q such that the variety $\{Q = 0\} := \{x \in \mathbb{R}^d : Q(x) = 0\}$ induces a "cell decomposition", such that (i) the connected components ("cells") $\{\Omega_i\}$ cutout by $\{Q = 0\}$, partition P into M sets of size at most $O(|\mathcal{P}|/M)$, and (ii) a variety $l \in \mathcal{L}$ can intersect only a small number of the cells $\{\Omega_i\}$. The second hypothesis will enable us to efficiently recover a bound on the total number of incidences from bounds on the number of incidences in each cell (which will then follow from the induction hypothesis). The first hypothesis ensures that each cell has less than $|\mathcal{P}|$ points, enabling us to apply the induction hypothesis to bound the number of incidences in each cell. Finally, we will need to handle the contribution from the points on the cell boundaries $\{Q = 0\}$. This case will be reduced to the case of varieties of a smaller dimension (and then handled by the induction hypothesis).

6.2 Some topology and algebraic geometry

We collect here several results from topology and algebraic geometry that will be used in the proof. These results can be found as Theorem 3.4, B.2, Corollary A.5, and Lemma A.3, respectively, in [1]. The reader should refer there for attribution and proofs.

Theorem 5. *(Cell Decomposition) Let $\mathcal{P}$ be a finite set of points in $\mathbb{R}^d$. Then there exists a polynomial of degree at most D and open sets $\{\Omega_i\}_{i=1}^M$, $M = O_d(D^d)$ such that $\mathbb{R}^d = \{Q = 0\} \cup \Omega_1 \cup \ldots \cup \Omega_M$ and $|\mathcal{P} \cap \Omega_i| \leq O_d(|\mathcal{P}|/D^d)$.*

Theorem 6. *Let ℓ be a k-dimensional variety of degree D, and $\{Q=0\}$ a hyper-surface of degree at most C_1. Then either ℓ is contained in $\{Q=0\}$ or $\ell \setminus \{Q=0\}$ has at most $O_D(C_1^k)$ connected components.*

Theorem 7. *Let $\mathcal{S}$ be a k-dimensional algebraic set in $\mathbb{C}^d$ of degree at most D. One can then find varieties $\ell_{\alpha,\beta}$ of degree at most $O_{D,d}(1)$ such that*

$$\mathcal{S} = \mathcal{S}^{smooth} + \bigcup_{\alpha=0}^{k-1} \bigcup_{\beta=1}^{M} \ell_{\alpha,\beta}^{smooth}, \tag{2}$$

where $\ell_{\alpha,\beta}^{smooth}$ denotes the smooth points of $\ell_{\alpha,\beta}$, and these varieties have dimension α. Moreover, $M = O_{D,d}(1)$.

Theorem 8. *Let $V = \{x \in \mathbb{C}^d : P_1(x) = \ldots = P_m(x) = 0\}$ for $m \geq 0$ and where $P_1, \ldots, P_m$ are polynomials of degree at most D. Then V is the union of $O_{m,D,d}(1)$ varieties of degree $O_{m,D,d}(1)$.*

6.3 Proof of the theorem

We wish to prove that $|\mathcal{I}| \leq C_3|\mathcal{P}|^{2/3+\epsilon}|\mathcal{L}|^{2/3} + \frac{3}{2}|\mathcal{P}| + \frac{3}{2}|\mathcal{L}|$ holds for $C_3 = O_{D,k,C_0,\epsilon}(1)$. We will proceed by induction on $|\mathcal{P}|$ and k, with the base cases being trivial.

We apply Theorem 5 with $M = O(C_1^{2k})$, to obtain open sets $\{\Omega_i\}_{i=1}^M$ such that $|\Omega_i \cap \mathcal{P}| = O(|P|/C_1^{2k})$. Write $\mathcal{L}_i$ for the subset of varieties in $\mathcal{L}$ that are incident to the points in $\mathcal{P} \cap \Omega_i$. We then have that

$$|\mathcal{I}| = |\mathcal{I} \cap I(\mathcal{P} \cap \{Q=0\}, L)| + \sum_{i=1}^{M} |\mathcal{I} \cap I(\mathcal{P} \cap \Omega_i, L_i)|. \tag{3}$$

We consider the contribution from the boundary and the cells separately.

6.4 Contribution from the cells

By the inductive hypothesis we have that

$$|\mathcal{I} \cap I(\mathcal{P} \cap \Omega_i, \mathcal{L}_i)| \leq C_1^{-\frac{4k}{3}-2k\epsilon} C_3 |\mathcal{P}|^{2/3+\epsilon}|\mathcal{L}_i|^{2/3} + \frac{3}{2}C_1^{-2k}|\mathcal{P}| + \frac{3}{2}|\mathcal{L}_i|. \tag{4}$$

By Theorem 6, we have that $\sum_{i=1}^{M}|\mathcal{L}_i| \ll C_1^k|\mathcal{L}|$. It follows (by Hölder's inequality) that $\sum_{i=1}^{M}|L_i|^{2/3} \ll C_1^{4k/3}|L|^{2/3}$. Combining this with (4), we have that

$$\sum_{i=1}^{M}|\mathcal{I}\cap I(\mathcal{P}\cap\Omega_i,\mathcal{L}_i)| \ll C_1^{-2k\epsilon}C_3|\mathcal{P}|^{2/3+\epsilon}|\mathcal{L}|^{2/3}+\frac{3}{2}|\mathcal{P}|+\frac{3}{2}C_1^k|\mathcal{L}|.$$

From the assumption (1) we may assume $\frac{3}{2}|P| \leq \frac{3}{2}C_2^{-2}|\mathcal{P}|^{2/3}|\mathcal{L}|^{2/3}$ and $\frac{3}{2}C_1^k|\mathcal{L}| \leq \frac{3}{2}C_1^kC_2^{-1/3}|\mathcal{P}|^{2/3}|\mathcal{L}|^{2/3}$. Using this and taking C_1, C_2 sufficiently large, we may conclude

$$\sum_{i=1}^{M}|\mathcal{I}\cap I(\mathcal{P}\cap\Omega_i,\mathcal{L}_i)| \leq \frac{1}{2}C_3|\mathcal{P}|^{2/3+\epsilon}|\mathcal{L}|^{2/3}.$$

6.5 Contribution from the boundary

To complete the proof we must handle the points on the boundary of the cell decomposition, that is $\mathcal{P}\cap\{Q=0\}$. We will prove

$$|\mathcal{I}\cap I(\mathcal{P}\cap\{Q=0\})| \ll \frac{1}{2}C_3|\mathcal{P}|^{2/3+\epsilon}|\mathcal{L}|^{2/3}+|\mathcal{P}|+|\mathcal{L}|.$$

We prove the following more general statement (our case follows by choosing $r=2k-1$, $D=C_0$ and using a covering of $\mathcal{K}=\{Q=0\}$) by $O_{1,D,d}(1)$ varieties via theorem 8):

Lemma 9. *Let $0 \leq r < 2k$, and $\mathcal{K}$ a r-dimensional variety in $\mathbb{R}^{2k}$ of degree at most D. Then*

$$|\mathcal{I}\cap I(\mathcal{P}\cap\mathcal{K},L)| \ll_{C_0,D,\epsilon} |\mathcal{P}|^{2/3+\epsilon}|\mathcal{L}|^{2/3}+|\mathcal{P}|+|\mathcal{L}|. \tag{5}$$

The proof will proceed by induction, assuming that the statement has been verified for all smaller values of r, the case of $r=0$ being trivial since $\mathcal{K}$ will consist of a single point (thus each $\ell\in\mathcal{L}$ has at most one incidence, and $|\mathcal{I}|\leq|\mathcal{L}|$). Next we argue that we may assume that all of the points in $\mathcal{P}$ are smooth. This is immediate from Theorem 7 and the induction hypothesis since we may cover the singular points of $\mathcal{K}$ by a bounded number $(O_{d,D}(1))$ of lower dimensional varieties of acceptable degree.

Each $\ell \in \mathcal{L}$ either is contained in $\mathcal{K}$ or intersects on an algebraic set of lower dimension. We claim that at most one element of $\mathcal{L}$ will be contained in $\mathcal{K}$. To see this, consider a point in distinct varieties $p \in l, l'$. It follows that $T_p l, T_p l' \subseteq T_p \mathcal{K}$, where $T_p l, T_p l'$ are k-dimensional and $T_p \mathcal{K}$ is r-dimensional with $r < 2k$. Thus it would follow that $T_p l$ and $T_p l'$ are not be transverse, contradicting axiom 4. Thus each point in $\mathcal{P}$ is incident to at most one $l \in \mathcal{L}$. This contributes at most $|\mathcal{L}|$ to the right side of (5), which is acceptable.

We may now assume that each variety in $\mathcal{L}$ intersects $\mathcal{K}$ in an algebraic set of dimension less than k. Applying Theorem 7, we may take $\ell \cap \mathcal{K}$ to be the union of the smooth points of $M = O_{C_0,D}(1)$ varieties of dimension at most $k-1$. Or, $\ell \cap \mathcal{K} = \bigcup_{\alpha=0}^{k-1} \bigcup_{\beta=1}^{M} \ell_{\alpha,\beta}^{\text{smooth}}$ (the $\ell_{\alpha,\beta}^{\text{smooth}}$ are defined in Theorem 7). We denote by $\mathcal{I}_{\alpha,\beta} \subseteq \mathcal{I}$ the incidences of $(p,\ell) \in \mathcal{I}$ such that $p \in \ell_{\alpha,\beta}^{\text{smooth}}$. Thus, $|\mathcal{I} \cap I(\mathcal{P} \cap \mathcal{K}, \mathcal{L})| \leq \sum_{\alpha=0}^{k-1} \sum_{\beta=1}^{M} |\mathcal{I}_{\alpha,\beta}|$.

Note that there are at most $M = O_{C_0,D}(1)$ non-zero terms on the right of this inequality, and it is acceptable to lose a factor of this size in the inequality we wish to prove. Thus if we let $\mathcal{L}_{\alpha,\beta}$ denote the set of all $\ell_{\alpha,\beta}$ (so $|\mathcal{L}_{\alpha,\beta}| \leq \mathcal{L}$), it will suffice to prove that

$$|\mathcal{I}_{\alpha,\beta}| \ll_{C_0,D,\epsilon} |\mathcal{P}|^{2/3+\epsilon} |\mathcal{L}_{\alpha,\beta}|^{2/3} + |\mathcal{P}| + |\mathcal{L}_{\alpha,\beta}| \tag{6}$$

Since all of the sets $\mathcal{L}_{\alpha,\beta}$ contain varieties of dimension at most $\alpha < k$, this will follow from the inductive hypothesis once we establish that the elements of $\mathcal{L}_{\alpha,\beta}$ are distinct and that the configuration $\mathcal{P}$, $\mathcal{L}_{\alpha,\beta}$, $\mathcal{I}_{\alpha,\beta}$ is admissible. The admissibility is easily seen to be inherited from that of $\mathcal{L}$.

Lastly, we argue that we may assume that the elements of $\mathcal{L}_{\alpha,\beta}$ are distinct. First, consider the case where $|\mathcal{L}_{\alpha,\beta}|$ is incident to $\leq C_0$ points. These varieties will contribute at most $C_0|\mathcal{L}|$ to (6), which is acceptable. Now if the same variety, containing more than C_0 points, is included in $\mathcal{L}_{\alpha,\beta}$ from the decompositions of two distinct $\ell, \ell' \in \mathcal{L}$, then it follows that ℓ and ℓ' are incident to more than C_0 identical points. However, this contradicts axiom 2.

References

[1] Solymosi, J. and Tao, T., *An incidence theorem in higher dimensions* . Preprint, arXiv:1103.2926;

Mark Lewko, UT-Austin
email: **mlewko@math.utexas.edu**

7 L^p estimates for the Hilbert transform along a one-variable vector field

after M. Bateman and C. Thiele [2]
A summary written by Diogo Oliveira e Silva

Abstract

The authors of [2] prove L^p estimates for the full Hilbert transform along a measurable, non-vanishing, one-variable vector field in the plane. We summarize their results.

7.1 Introduction

We are interested in singular integral operators on functions of two variables, which act by performing a one-dimensional transform along a particular line in the plane. The choice of lines is to be variable. Thus, for a non-vanishing measurable vector field $v : \mathbb{R}^2 \to \mathbb{R}^2 \setminus \{0\}$ and a measurable function $f : \mathbb{R}^2 \to \mathbb{R}$, we define the directional Hilbert transform

$$H_v f(x, y) := p.v. \int_{\mathbb{R}} \frac{f((x, y) - tv(x, y))}{t} dt$$

Specializing to vector fields which depend on one variable only, the authors of [2] are able to prove the following result:

Theorem 1. *Let $p \in (\frac{3}{2}, \infty)$, and let v be a non-vanishing, measurable vector field in the plane such that for all $x, y \in \mathbb{R}$, $v(x, y) = v(x, 0)$. Then*

$$\|H_v f\|_p \lesssim \|f\|_p$$

for every $f \in L^p(\mathbb{R}^2)$.

A few remarks may help to further orient the reader:

1. The case of a constant vector field follows from the classical L^p estimates for the one-dimensional Hilbert transform.

2. The class of vector fields depending on the first variable only is invariant under linear transformations which preserve the vertical direction. These symmetries, together with those of the Hilbert transform, allow

us to assume without loss of generality that $v(x,y) = (1, u(x))$ for some real-valued measurable function u satisfying

$$\|u\|_\infty \le 10^{-2}. \tag{1}$$

3. Sharpness of the endpoint exponent $p = \frac{3}{2}$ is an open problem. It is known however that the exponent in Theorem 1 can be improved to $p = \frac{4}{3}$ under the additional assumption that the function f is an elementary tensor. See [2].

4. The case $p = 2$ of Theorem 1 is actually equivalent to the celebrated Carleson-Hunt theorem on pointwise convergence of Fourier series. See [3].

7.1.1 History of the problem and related work

The discovery of the Besicovitch set in the 1920s inspired Zygmund to ask if integrals of $L^2(\mathbb{R}^2)$ functions could be differentiated in a Lipschitz choice of directions. Much later, Stein raised the singular integral variant of this conjecture: if v is Lipschitz, is (a truncated version of) H_v a bounded operator on $L^2(\mathbb{R}^2)$? For a fuller history of these conjectures, see [3].

In a somewhat different direction, the Hilbert transform along a one variable vector field has been previously studied by Carbery, Seeger, Wainger and Wright, who proved L^p boundedness for $p > 1$ under additional smoothness assumptions on the vector field. On the other hand, Christ, Nagel, Stein and Wainger proved similar estimates under the additional geometric hypothesis that no integral curve of the vector field forms a straight line. For further references, see [2].

Finally, the companion paper [1] proves L^p estimates for $p \in (1, \infty)$ for the Hilbert transform along a one-variable vector field v acting on functions with frequency supported in an annulus. Since their main result will be of importance to us already in the next section, we state it here in a form invariant under the linear transformation group mentioned in remark 2 above:

Theorem 2. *[1] Let $p \in (1, \infty)$, and assume $\widehat{f}(\xi, \eta)$ is supported in a horizontal pair of strips $A < |\eta| < 2A$ for some $A > 0$. Then*

$$\|H_v f\|_p \lesssim \|f\|_p.$$

7.2 The main approach

It is a common theme to reduce L^p estimates for a given operator to restricted weak-type estimates for a model operator. In this spirit, instead of trying to estimate H_v directly, we start by defining the closely related operator

$$H_k := P_k H_v P_c. \quad (k \in \mathbb{Z}/100)$$

Here, by P_c we mean the restriction to a cone in the Fourier plane (ξ, η)

$$\widehat{P_c f}(\xi, \eta) = 1_{\{10|\xi| \le |\eta|\}}(\xi, \eta) \widehat{f}(\xi, \eta),$$

whereas P_k denotes the Fourier multiplier given by $\widehat{P_k f} = 1_{B_k} \widehat{f}$, where B_k is the horizontal pair of bands given by

$$B_k := \{(\xi, \eta) \in \mathbb{R}^2 : |\eta| \in [2^k, 2^{k+.01})\}.$$

By Littlewood-Paley theory and a limiting argument, it will be enough to prove that, for all $k_0 > 0$,

$$\Big\| \Big(\sum_{|k| \le k_0} |H_k f_k|^2 \Big)^{\frac{1}{2}} \Big\|_p \lesssim \Big\| \Big(\sum_{|k| \le k_0} |f_k|^2 \Big)^{\frac{1}{2}} \Big\|_p \tag{2}$$

holds[8] for any sequence of functions $f_k \in L^2$, with implicit constant independent of k_0.

Note that H_k is bounded on L^p for $1 < p < \infty$ for each k, by Theorem 2. In particular, (2) is true for $p = 2$. For other values of p, we recall the notion of restricted weak-type estimates in the spirit of ([4], chapter 3) and observe that it suffices to show, for bounded $G, H \subseteq \mathbb{R}^2$ and $\sum_k |f_k|^2 \le 1_H$, that

$$\Big| \Big\langle \Big(\sum_{|k| \le k_0} |H_k f_k|^2 \Big)^{\frac{1}{2}}, 1_G \Big\rangle \Big| \lesssim |H|^{\frac{1}{p}} |G|^{1-\frac{1}{p}}. \tag{3}$$

In what follows we restrict our attention to the case $\frac{3}{2} < p \le 2$. Since we already have (3) for $p = 2$, we immediately obtain this estimate for $p < 2$ provided $|G| \lesssim |H|$. By a standard inductive procedure, it will suffice to prove the following result:

[8] It is understood that the index k runs through elements of $\mathbb{Z}/100$, that is, multiples of $1/100$.

Lemma 3. *Let $G', H' \subset [-N,N]^2$ be measurable, and let $\frac{3}{2} < p < 2$. If $|H'| < \frac{1}{10}|G'|$, then there exists a subset $G \subset G'$ (depending only on p, G' and H') with $|G| \geq |G'|/2$ such that* (3) *holds with $H = H'$, for any sequence of functions f_k with $\sum_k |f_k|^2 \leq 1_H$.*

In the next section we present the construction of the set G of Lemma 3 and sketch the proof of the size estimate $|G| \geq |G'|/2$. In the last section we outline very briefly how time-frequency analysis comes into play to prove strong L^2 bounds for the sets G and H, from which (3) follows.

7.3 Construction of the set G

Following general principles of wave packet analysis [4], it is natural to decompose the operator H_v into wave packets, which can be visualized by acting with the same group element in the unit square of the plane. The shapes thus obtained are parallelograms with a pair of vertical edges, and because of (1) it is enough to consider parallelograms whose non-vertical edges are close to horizontal.

Given a parallelogram R with two vertical edges, we refer to [2] for the precise definitions of the height $H(R)$, the shadow $I(R)$ and the interval of uncertainty $U(R)$. Given $c > 0$, we denote by cR the parallelogram with the same central line segment as R but height $cH(R)$. We also define

$$E(R) := \{(x,y) \in R : u(x) \in U(R)\}.$$

The following observation will be used several times and its proof is an easy but amusing exercise in elementary geometry which we recommend to the reader:

Lemma 4. *Let R, R' be two parallelograms and assume $R \cap R' \neq \emptyset$, $I(R) = I(R')$, $U(R) \cap U(R') \neq \emptyset$ and $H(R) \leq H(R')$. Then $R \subseteq 7R'$.*

After these preliminaries, we indicate how to construct the set G. Let G' and H' be as in Lemma 3. For $i \in \{1,2\}$, define

$$G_i := \bigcup_{j \in \mathbb{Z}_-} \{R \in \mathcal{R}_i : \frac{|E(R)|}{|R|} \geq 2^j \text{ and } \frac{|H' \cap R|}{|R|} \geq C_\epsilon 2^{-(\frac{1}{2}+\epsilon)j} \Big(\frac{|H'|}{|G'|}\Big)^{\frac{1}{2}}\},$$

where $\mathcal{R}_i$ is a finite set of parallelograms with vertical edges and dyadic shadow adapted to a shifted dyadic grid $\mathcal{I}_i$ on the real line and having some nice properties about which we will not be completely precise. The small parameter $\epsilon > 0$ and the large constant $C_\epsilon < \infty$ will be chosen as a function of p later on in the argument in order to force the set G defined by

$$G' \setminus G := G_1 \cup G_2$$

to satisfy the desired size estimate $|G| \geq |G'|/2$. That this is indeed possible is a consequence of the following result, which holds for parallelograms of *arbitrary* height:

Lemma 5. *Let $\delta, \sigma \in [0,1]$, let H be a measurable set, and let $\mathcal{R}$ be a finite collection of parallelograms with vertical edges and dyadic shadow such that for each $R \in \mathcal{R}$ we have*

$$|E(R)| \geq \delta |R| \text{ and } |H \cap R| \geq \sigma |R|.$$

Then

$$\Big| \bigcup_{R \in \mathcal{R}} R \Big| \lesssim \delta^{-1} \sigma^{-2} |H|.$$

Proof. Adopting the covering lemma approach of Córdoba and Fefferman (see [3]), it will be enough to find a "good" subset $\mathcal{G} \subset \mathcal{R}$ such that

$$\Big| \bigcup_{R \in \mathcal{R}} R \Big| \lesssim \sum_{R \in \mathcal{G}} |R| \quad \text{and} \quad \int (\sum_{R \in \mathcal{G}} 1_R)^2 \lesssim \delta^{-1} \sum_{R \in \mathcal{G}} |R|. \tag{4}$$

We accomplish this by a recursive procedure, which we initialize by setting $\mathcal{G} := \emptyset$ and $STOCK := \mathcal{R}$. As long as we have a nonempty $STOCK$ of parallelograms, we may choose $R \in STOCK$ with maximal $|I(R)|$, and update[9]:

$$\begin{array}{rcl} \mathcal{G} & \leftarrow & \mathcal{G} \cup \{R\} \\ \mathcal{B} & \leftarrow & \{R' \in STOCK : R' \subset \{x : M_V(\sum_{R \in \mathcal{G}} 1_R)(x) \geq 10^{-3}\}\} \\ STOCK & \leftarrow & STOCK \setminus \mathcal{B}. \end{array}$$

The first inequality in (4) is then a trivial consequence of the Hardy-Littlewood weak-type (1,1) bound. For the second one, we organize the set $\mathcal{P}$ of pairs $(R, R') \in \mathcal{G} \times \mathcal{G}$ such that $R \cap R' \neq \emptyset$ and R is chosen prior to R' into two sets, according to whether the two rectangles are "well-aligned" or not. Define

$$\mathcal{P}' := \{(R, R') \in \mathcal{P} : U(R) \subset 100 U(R')\}$$

[9]Here, M_V denotes the Hardy-Littlewood maximal operator in the vertical direction.

and $\mathcal{P}'' := \mathcal{P} \setminus \mathcal{P}'$. It will be enough to show that, for fixed $R' \in \mathcal{G}$ we have

$$\sum_{R \in \mathcal{R}:(R,R') \in \mathcal{P}'} |R \cap R'| \lesssim |R'|, \tag{5}$$

and that for fixed $R \in \mathcal{G}$ we have

$$\sum_{R' \in \mathcal{R}:(R,R') \in \mathcal{P}''} |R \cap R'| \lesssim \delta^{-1}|R|. \tag{6}$$

The proof of (5) is based on the following observation: if $R' \subset cR$, then $R' \subset \{M_V 1_R > c^{-1}\}$. We use this together with Lemma 4 to show that $H(R) \leq H(R')$ for every $(R, R') \in \mathcal{P}'$, and then the same lemma again to conclude that

$$R \cap (I(R') \times \mathbb{R}) \subset 700R'.$$

It follows that, for some point $(x, y) \in R'$,

$$10^{-3} \geq M_V\Big(\sum_{R:(R,R') \in \mathcal{P}'} 1_R\Big)(x, y) \geq \frac{1}{700} \sum_{R:(R,R') \in \mathcal{P}'} \frac{|R \cap R'|}{|R'|}.$$

We omit the proof of (6), hoping to say something about it at the Summer School. $\square$

7.4 The "end" of the proof

In what follows, we omit almost all details.

Let $p \in (\frac{3}{2}, 2)$, and let $G', H' \subseteq \mathbb{R}^2$ be as in Lemma 3. Once again by restricted weak-type interpolation, it will be enough to establish the following single frequency band estimate: for any measurable sets $E, F \subseteq \mathbb{R}^2$ and each $|k| \leq k_0$, we have that

$$|\langle H_{k,G,H} 1_F, 1_E \rangle| \lesssim \left(\frac{|G|}{|H|}\right)^{\frac{1}{2}-\frac{1}{p}} |F|^{\frac{1}{2}} |E|^{\frac{1}{2}}, \tag{7}$$

where $G \subset G'$ is the set constructed in the last section, $H := H'$, and the operator $H_{k,G,H}$ is defined by

$$H_{k,G,H} f := 1_G H_k(1_H f).$$

Assuming without loss of generality that $E \subset G$ and $F \subset H$, we have that $\langle H_{k,G,H} 1_F, 1_E\rangle = \langle H_k 1_F, 1_E\rangle$. Following [1], we write the latter form as a linear combination of a bounded number of model forms

$$\langle H_k 1_F, 1_E\rangle = \sum_{s\in\mathcal{U}_k} \langle C_{s,k} 1_F, 1_E\rangle, \tag{8}$$

where $\mathcal{U}_k$ is a set of parallelograms with vertical edges and height depending on k only. To estimate the sum in (8), one starts by proving estimates for the sum over certain subsets of $\mathcal{U}_k$ called trees. Each tree T is assigned a parallelogram $\text{top}(T)$, a density $\delta(T)$ and a size $\sigma(T)$. One obtains for each tree T:

$$\sum_{s\in T} |\langle C_s 1_F, 1_E\rangle| \lesssim \delta\sigma|\text{top}(T)|.$$

Denoting by $\mathcal{T}_{\delta,\sigma}$ the collection of trees with density at most δ and size at most σ, it remains to estimate $\sum_{\delta,\sigma} S_{\delta,\sigma}$ with

$$S_{\delta,\sigma} := \sum_{T\in\mathcal{T}_{\delta,\sigma}} \delta\sigma|\text{top}(T)|.$$

The desired estimate (7) follows from the estimates for $S_{\delta,\sigma}$ proved in [1] (and presented in J. Jung's summary) together with one new maximal estimate.

I am indebted to Michael Bateman for a very useful discussion of some parts of [1] and [2].

References

[1] Bateman, M., *Single Annulus L^p estimates for Hilbert transforms along vector fields.* Preprint (2011).

[2] Bateman, M. and Thiele, C., *L^p estimates for the Hilbert transform along a one-variable vector field.* Preprint (2011).

[3] Lacey, M. and Li, X., *On a Conjecture of E.M. Stein of the Hilbert Transform on Vector Fields*, Memoirs of the AMS **205** (2010), no. 965.

[4] C. Thiele, *Wave Packet Analysis,* CBMS Regional Conference Series in Mathematics, 105. American Mathematical Society, Providence, RI, 2006.

Diogo Oliveira e Silva, UC Berkeley
email: **dosilva@math.berkeley.edu**

8 Maximal Operators for Arbitrary Sets of Directions

after Nets Hawk Katz [4]
A summary written by James Scurry

Abstract

We consider a maximal operator defined using an arbitrary collection of N unit vectors as an operator acting on L^2 and obtain a sharp bound for its operator norm.

8.1 Introduction

Let $\Omega \subset S^{\mathbf{1}}$ be a set consisting of N elements. For each $v \in \Omega$, we define a directional maximal operator M_v

$$M_v f(x,y) = \sup_{r>0} \frac{1}{2r} \int_{-r}^{r} |f((x,y)+tv)|dt.$$

and a maximal operator $M_\Omega f(x,y) = \sup\limits_{v\in\Omega} M_v f(x,y)$, where $f \in L^1_{\text{loc}}(\mathbb{R}^2)$. In [4], Katz was mainly interested in providing a sharp bound for the operator norm of M_Ω when acting from $L^2(\mathbb{R}^2)$ to $L^2(\mathbb{R}^2)$. Stated more precisely, his result is the following:

Theorem 1. *For $f \in L^2(\mathbb{R}^2)$ we have*

$$\|M_\Omega f\|_{L^2} \;\; \lesssim \;\; \log N \|f\|_{L^2},$$

with with one the sharp exponent on $\log N$*.*

Estimating the operator norm of M_Ω and similar types of maximal operators had been considered previously. Restricting to the case when Ω consists of equidistributed directions, Stömberg showed $\|M_\Omega\|_{L^2\to L^2} \lesssim \log N$. Further, when Ω is taken to be a lacunary sequence of directions,we have $\|M_\Omega\|_{L^2\to L^2} \lesssim 1$ [5]. For Ω an unrestricted collection of N directions, the best possible bounds prior to [4] had been obtained by Barrionuevo in [2], wherein he demonstrated $\|M_\Omega\|_{L^2\to L^2} \lesssim N^{\frac{2}{\log N}}$.

There have also been subsequent papers which studied similar problems. Notably, in [1] the authors obtain a weak $(2,2)$ inequality which they use to show $\|M_\Omega\|_{L^2\to L^2} \lesssim \log N^\alpha$ for some α; however, they are not able to obtain the sharp value of α. In [3] an argument analogous to that presented in [4] for Theorem 1 is used in a slightly more general context.

8.2 Preliminaries

In this section we introduce some notation and ideas which will be useful throughout. First, take $\mathcal{D}$ to be the collection of all dyadic intervals. For $I \subset \mathbb{R}$ and $s \in [0,1]$, define

$$m_{s,I} f(y) \;=\; \frac{1}{|I|}\int_I f(x, y+sx)dx$$

with $f \in L^1_{\mathrm{loc}}(\mathbb{R}^2)$ and for an interval $L \subset \mathbb{R}$ take

$$O_{s,I,L} \;=\; \{(x,z) : x \in I, z = y + sx \text{ for some } y \in 3L\}$$

where $3L$ is the interval with the same center as L but three times its length.

Suppose $\Sigma \subset [0,1]$ and $\alpha = (\alpha_\Sigma, \alpha_{\mathcal{D}}) : \mathbb{R}^2 \to \Sigma \times \mathcal{D}$. Given $E \subset \mathbb{R}^2$ we define

$$\begin{aligned} M_\alpha f(x,y) &= (m_{\alpha(x,y)} f)(y - x(\alpha_\Sigma(x,y))) \\ M_{\alpha_E} f(x,y) &= \mathbf{1}_E (m_{\alpha(x,y)} f)(y - x(\alpha_\Sigma(x,y))) \end{aligned}$$

for $f \in L^1_{\mathrm{loc}}(\mathbb{R}^2)$ and $b^E_{s,I}(y) = |\{x : \alpha(x, y+sx) = (s,I),\ (x, y+sx) \in E\}|^{\frac{1}{2}}$. We note in passing that $b^E_{s,I}$ satisfies a Carleson property: if $O \subset \mathbb{R}^2$ is measurable with finite measure and

$$P(O) = \{(s,I,y) \in \Sigma \times \mathcal{D} \times \mathbb{R} : \{(x,z) : x \in I, z = y + sx\} \subset O\}$$

then $\beta^E(P(O)) \le C|O|$, where

$$\beta^E(F) \;=\; \sum_{s\in\Sigma}\sum_{I\in\mathcal{D}} \int b^E_{s,I}(y)^2 \mathbf{1}_F(s,I,y)dy$$

for $F \subset \Sigma \times \mathcal{D} \times \mathbb{R}$. For fixed $s \in \Sigma$ and $I \in \mathcal{D}$ we may define

$$B^{\text{bajo}}_{s,I}(y) \;=\; m_{s,I}\left(\sum_{t\in\Sigma}\sum_{J\subset I} m^*_{t,J}(b^E_{t,J})^2\right)(y)$$

and

$$\begin{aligned} B^{\text{bajo}}_{s,I,F}(y) = &\sum_{J\subset I} \frac{b^E_{s,J}(y)^2}{|I|}\mathbf{1}_F(s,J,y) + \\ &\sum_{s\neq t} \frac{1}{|I||J||s-t|}\int_J \mathbf{1}_F(t,J,y+(s-t)x) b^E_{t,J}(y+s-tx)^2 dx \end{aligned}$$

for $F \subset \Sigma \times \mathcal{D} \times \mathbb{R}$.

8.3 Sketch of the Proof of the Main Theorem

The proof of Theorem 1 hinges on a strong-type estimate for a particular class of operators and a technical proposition. Namely, we have

Theorem 2. *Let $E \subset \mathbb{R}^2$ be any set. Then there is a universal constant C such that for any linearization $\mathcal{M}$ of M_Ω,*

$$\|\mathcal{M}^*(\mathbf{1}_E)\|_{L^2} \leq C\sqrt{\log N}|E|^{\frac{1}{2}}.$$

Proposition 3. *Suppose X is a measure space and L is a linear operator acting on $L^2(X)$ having a positive kernel. Suppose there is a fixed constant $A > 0$ so that for any $E \subset X$ one has*

$$\|L^*(\mathbf{1}_E)\|_{L^2} \leq A|E|.$$

Then for any $f \in L^2(X)$ and any $\lambda > 0$, one has

$$|\{x : |Lf(x)| > \lambda\}| \leq \frac{A\|f\|_{L^2}^2}{\lambda^2}.$$

Conceding the proofs of Theorem 2 and Proposition 3, the proof of our main theorem follows easily. Using the theorem and proposition in concert, we may obtain the following estimate

$$|\{(x, y) : M_\Omega f(x, y) > \lambda\}| \leq \frac{C\log N\|f\|_{L^2}^2}{\lambda^2}.$$

Well-known bounds for the Hardy-Littlewood maximal operator give

$$\begin{aligned} \|M_\Omega f\|_{L^\infty} &\leq \|f\|_{L^\infty} \\ |\{(x, y) : M_\Omega f(x, y) > \lambda\}| &\leq \frac{CN\|f\|_{L^1}}{\lambda}. \end{aligned}$$

Now if we take $f_{1,\lambda}(x, y) = f(x, y)$ for (x, y) satisfying $|f(x, y)| \geq \frac{N\lambda}{3}$ and zero otherwise; $f_{2,\lambda}(x, y) = f(x, y)$ for x satisfying $\frac{\lambda}{3} \leq |f| \leq \frac{N\lambda}{3}$ and zero otherwise; and $f_{3,\lambda}(x, y)$ such that for all (x, y), $f(x, y) = f_{1,\lambda}(x, y) + f_{2,\lambda}(x, y) + f_{3,\lambda}(x, y)$, then we easily have

$$\begin{aligned} \|M_\Omega f\|_{L^2(\mathbb{R}^2)}^2 &= 2\int_0^\infty \lambda|\{(x, y) : M_\Omega f(x, y) \geq \lambda\}|d\lambda \\ &\leq \int_0^\infty CN\|f_{1,\lambda}\|_{L^1}d\lambda + \int_0^\infty \frac{C\log N}{\lambda}\|f_{2,\lambda}\|_{L^2}^2 d\lambda \\ &\leq C\|f\|_{L^2}^2 + C(\log N)^2\|f\|_{L^2}^2, \end{aligned}$$

which gives $\|M_\Omega\|_{L^2\to L^2} \lesssim \log N$. To conclude the proof of Theorem 1, it only remains to notice our estimate for $\|M_\Omega\|_{L^2\to L^2}$ must be sharp by Stömberg's estimates in [6].

8.3.1 Sketch of the Proofs of Proposition 3 and Theorem 2

With the proof of Theorem 1 established under the assumption of Proposition 3 and Theorem 2, we restrict our attention to their respective proofs. Proposition 3 can be shown via a simple proof by contradiction, but the argument of Theorem 2 is lengthy, occupying the bulk of Katz' paper and consisting of several different stages.

The first part of the proof is largely a series of reductions which serve to streamline the exposition. In totem, these simplifications indicate the following implies Theorem 2:

Theorem 4. *Let $\Sigma \subset [0,1]$ and $E \subset \mathbb{R}^2$ be any set. Then there exists $C > 0$ so that for any $\alpha = (\alpha_\Sigma, \alpha_\mathcal{D}) : \mathbb{R}^2 \to \Sigma \times \mathcal{D}$, we have*

$$\|M^*_\alpha(\mathbf{1}_E)\|_{L^2(\mathbb{R}^2)} \quad \leq \quad C(\log N)^{\frac{1}{2}}|E|^{\frac{1}{2}}.$$

Hence, we focus on the proof of Theorem 4, from which Theorem 2 may be deduced. To this end, we fix $\Sigma \subset [0,1]$, $\alpha : \mathbb{R}^2 \to \Sigma \times \mathcal{D}$, and $E \subset \mathbb{R}^2$. Katz defines an attendant paraproduct operator

$$\pi_{b^E} f(s, I, y) = b^E_{s,I}(y) m_{s,I} f(y)$$

for $f \in L^1_{\text{loc}}(\mathbb{R}^2)$. A simple computation gives $(M^*_\alpha \mathbf{1}_E)(x,y) = (\pi^*_{b^E} b^E)(x,y)$ for almost all $(x,y) \in \mathbb{R}^2$, so that it suffices to show

$$\|\pi^*_{b^E} b^E\|^2_{L^2(\Sigma\times\mathcal{D}\times\mathbb{R})} \lesssim \log N|E|. \tag{1}$$

The remainder of Theorem 4's proof largely consists of proving (1) through the following John-Nirenberg type lemma:

Lemma 5. *Let*

$$F_\mu = \{(s, I, y) : B^{\text{bajo}}_{s,I}(y) \geq \mu \log N\}.$$

Then there exists $C, c > 0$ such that $\beta^E(F_\mu) \leq C|E|e^{-c\mu}$.

Indeed, the conclusion of Theorem 4 follows almost immediately from the lemma since

$$\begin{aligned}
\int_{\Sigma\times\mathcal{D}\times\mathbb{R}} (\pi^*_{b^E} b^E)^2 &\leq \sum_{s\in\Sigma}\sum_{I\in\mathcal{D}} \int_{\mathbb{R}} b^E_{s,I}(y)^2 B^{\text{bajo}}_{s,I}(y)dy \\
&\leq \int_0^\infty \log N |F_\mu| d\mu \\
&\leq C|E|\log N.
\end{aligned}$$

8.3.2 Sketch of the Proof of Lemma 5

The proof of Lemma 5 is a stopping time argument. Specifically, for each $s \in \Sigma$ and $I \in \mathcal{D}$, Katz considers the largest dyadic intervals $K^j_{s,I}$ such that

$$\frac{1}{|K^j_{s,I}|}\int_{K^j_{s,I}} \sum_{t\neq s, J\subset I, |s-t||J|\geq |K^j_{s,I}|} \frac{1}{|I||J|}\int_{27J} b^E_{t,J}(y+(s-t)x)^2 dxdy \geq \lambda,$$

where λ is a sufficiently large constant. Straightforward computations imply for $y \in 3K^j_{s,I}$

$$\lambda \;\lesssim\; \sum_{t\neq s, J\subset I, |s-t||J|\geq |K^j_{s,I}|} \frac{1}{|I||J|}\int_{3J} b^E_{t,J}(y+(s-t)x)^2.$$

Defining

$$O^1_s = \bigcup_{I\in\mathcal{D}}\bigcup_{j\in\mathbb{N}} O_{s,I,K^j_{s,I}}$$

for $s \in \Sigma$ and using basic properties of the Hardy-Littlewood maximal operator, Katz obtains $|O^1_s| \leq \frac{C|E|}{\lambda}$ and observes that the support of $B^{\text{bajo}}_{s,IP(O^1_s)}(y)$ is contained in $\tilde{O}^1_s = \cup_{I\in\mathcal{D}}\cup_j O_{s,I,10K^j_{s,I}}$. Now an iterative process is obtained by application of the previous steps to $\mathbf{1}_{P(O^j_s)} b^E$ instead of b^E; this gives rise to a collection $\{\tilde{O}^j_s\}_{j\in\mathbb{N}}$ of sets satisfying

$$|\tilde{O}^j_s| \leq |E|\left(\frac{C}{\lambda}\right)^j$$

for each j and such that for $j \sim \log N$ we have

$$|\tilde{O}^j_s| \;\leq\; \frac{C|E|e^{-c\mu}}{N}$$

with $F_\lambda \subset \cup_{s\in\Sigma} P(\tilde{O}^j)$. Immediately, this implies $\beta^E(F_\lambda) \leq C|E|e^{-c\mu}$ and concludes the proof of the lemma.

References

[1] A. Alfonseca, F. Soria, and A. Vargas, *A Remark on Maximal Operators Along Directions in* $\mathbb{R}^2$, Mathematical Research Letters **10** (2003), 41-49.

[2] J. Barrionuevo, *A Note on the Kakeya Maximal Operator*, Mathematical Research Letters **3** (1996), no. 1, 61-66.

[3] M. Bateman, *Maximal averages along a planar vector field depending on one variable*, http://arxiv.org/abs/1104.2859.

[4] N. Katz, *Maximal Operators Over Arbitrary Sets* Duke Math. J. **97** (1999), no. 1, 67-79.

[5] A. Nagel, E. M. Stein, and S. Wainger, *Differentiation in Lacunary Directions*, Proc. Natl. Acad. Sci. U.S.A. **75** (1978), no. 3, 1060?1062.

[6] J. Strömberg, *Maximal functions associate3d to rectangles with uniformly distributed directions*, Ann. of Math. **107** (1978), no. 3, 399-402.

James Scurry, Georgia Institute of Technology
email: **jscurry3@math.gatech.edu**

9 Maximal theorems for the directional Hilbert transform

after Michael T. Lacey and Xiaochun Li [4]
A summary written by Prabath Silva

Abstract

We show that following operator maps L^p into L^p for $p > 2$ and L^2 into $L^{2,\infty}$.

$$Tf(x) = \sup_{v \in \mathbb{R}^2 \setminus \{0\}} p.v. \int_{\mathbb{R}} \boldsymbol{\zeta} f(x - vy) \frac{dy}{y}$$

Here f is defined on $\mathbb{R}^2$ and $\boldsymbol{\zeta} f = \zeta * f$, where ζ is a Schwartz function with frequency support in the annulus $1 \leq |\xi| \leq 2$.

9.1 Introduction

First consider the directional Hilbert transform H_v for $v \in \mathbb{R} \setminus \{0\}$ defined by

$$H_v f(x) = p.v. \int_{\mathbb{R}} f(x - vy) \frac{dy}{y};$$

here f is a Schwartz function on $\mathbb{R}^2$.

In this paper we look at the maximal operator $H^*(f) = \sup_{v \in \mathbb{R} \setminus \{0\}} H_v(\boldsymbol{\zeta} f)$. $\boldsymbol{\zeta} f$ is the frequency restriction of f to a single frequency annulus, i.e. $\boldsymbol{\zeta} f = \zeta * f$, where ζ is a Schwartz function with frequency support on $1 \leq |\xi| \leq 2$.

Theorem 1. *The maximal operator H^* maps L^p into L^p for $p > 2$ and L^2 into weak L^2.*

A counterexample by M. Christ shows that the L^2 estimate is sharp, in the sense that H^* does not map L^2 into L^2.

We use time frequency analysis to prove this theorem. This operator is very similar to the Carleson operator, as we can see in next section, and in fact the proof follows the same form of argument as in [1].

9.2 Discretization of the operator

The objective of this section is to get the model sum for the operator. We will go through the main steps of this without going into details.

First we linearize the operator; the idea is to think of $v : \mathbb{R}^2 \to \mathbb{T}$ as the choice of the direction where the supremum is achieved. The linearized operator is

$$T_v f(x) = \int_{\mathbb{R}} \boldsymbol{\zeta} f(x - v(x)y)\frac{dy}{y}.$$

It is enough to show the boundedness of T with bound independent of the choice of v.

Next we decompose the kernel $\frac{1}{y} = \sum_{l \in \mathbb{Z}} \psi_l$, where ψ_l has frequency support in $2^l \leq |\eta| \leq 2^{l+1}$.

Next we look at one term in the above sum $\int_{\mathbb{R}} \boldsymbol{\zeta} f(x - v(x)y)\psi_l(y)dy$. $\boldsymbol{\zeta} f$ supported in frequency in $1 \leq |\xi| \leq 2$ we only need to decompose this annulus. we do it as in the below picture with the small side of the rectangle ω_s having the length $2^l = scl(s)$.

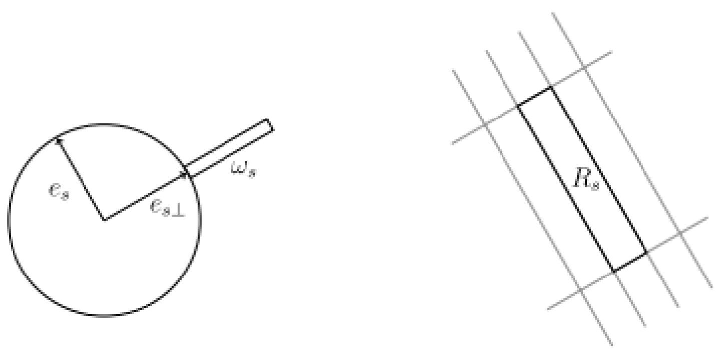


FIGURE 2. The two rectangles ω_s and R_s whose product is a tile. The gray rectangles are other possible locations for the rectangle R_s.

So we get a $\boldsymbol{\zeta} f$ as a sum of functions each having frequency support in a rectangle ω_s. Next, writing windowed fourier series in 2D for each of those functions, we get the decomposition

$$\boldsymbol{\zeta} f = \sum_{s \in AT, scl(s) = 2^l} < f, \varphi_s > \varphi_s.$$

Here φ_s is time frequency adapted to $s = \omega_s \times R_s$, and AT denotes the complete collection of such s.

Now we get

$$T_v f(x) = \sum_{s \in AT} < f, \varphi_s > \phi_s.$$

Here $\phi_s(x) = \int \varphi_s(x - yv(x))\psi_s(y)dy$. Now the important observation is that the frequancy localization of ψ_s makes it possible to get

$$\phi_s(x) = 1_{\omega_{s_2}}(v(x)) \int \varphi_s(x - yv(x))\psi_s(y)dy;$$

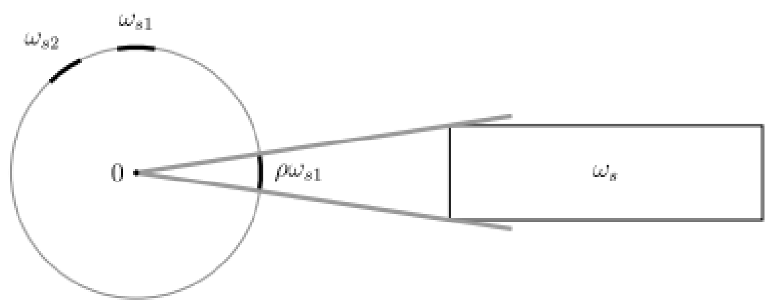


FIGURE 3. An annular rectangular ω_s, and three associated subintervals of $\rho\omega_{s1}$, ω_{s1}, and ω_{s2}.

here ω_{s_2} and ω_{s_2} can be viewed as chilren of a same dyadic interval in $\mathbb{T}$. ω_{s_2} gives the location of $v(x)$ and ω_{s_1} gives the location of the frequency support of φ_s.

So we have the model sum

$$T_v f(x) = \sum_{s \in AT} < f, \varphi_s > \phi_s 1_{\omega_{s_2}}(v(x)).$$

9.3 Idea of the Proof

In the proof of Carleson's theorem in [1] we have similar type of model sum. In there φ_s was supported in frequency in an interval, but in here it is supported in rectangle; but the good news is all those rectangles have fixed length for one side. In fact we will be able to define a Fefferman type [2] order relation on these tiles, which is impossible to do if we allow both sides of the rectangles to vary.

Having the relation between tiles in this case allows us to define trees. We prove the main theorem by proving a size lemma, mass lemma, and a tree lemma as in [1] in this setting, with additional complications.

References

[1] Michael Lacey and Christoph Thiele., *A proof of boundedness of the Carleson operator* Math. Res. Lett. 7 (2000), 361370. MR 2001m:42009

[2] Charles Fefferman., *Pointwise convergence of Fourier series* Ann. of Math. (2) 98 (1973), 551 571. MR 49 #5676

[3] J. Bourgain., *A remark on the maximal function associated to an analytic vector eld* Analysis at Urbana, Vol. I (Urbana, IL, 19861987), London Math. Soc. Lecture Note Ser., vol. 137, Cambridge Univ. Press, Cambridge, 1989, pp. 111132. MR 90h:42028

[4] Michael T. Lacey and Xiaochun Li., *Maximal Theorems for The Directional Hilbert transform* Transactions of the American Mathematical Society Volume 358, Number 9, September 2006, Pages 40994117

Prabath Silva, Indiana University Bloomington
email: **pssilva@indiana.edu**

10 A remark on the maximal function associated to an analytic vector field

after Jean Bourgain [1]
A summary written by Stefan Steinerberger

Abstract

We describe an argument by Jean Bourgain showing the boundedness of a maximal operator parametrized by a real-analytic vectorfield in $L^2(\mathbb{R}^2)$. This argument generalizes earlier results by Nagel, Stein & Wainger by replacing the condition of non-vanishing curvature with a much weaker condition and applies to not too slowly turning C^1 vectorfields. It is suspected that a C^1 result should hold unconditionally.

10.1 Introduction

Consider the following averaging operator on $L^2(\mathbb{R}^2)$. For a given vector $v \in \mathbb{R}^2$ with $|v| = 1$ and given $\varepsilon > 0$, we take

$$A_\varepsilon f = \frac{1}{\varepsilon} \int_0^\varepsilon f(x + tv)dt \quad \text{with} \quad f \in L^2(\mathbb{R}^2).$$

The naturally associated maximal operator

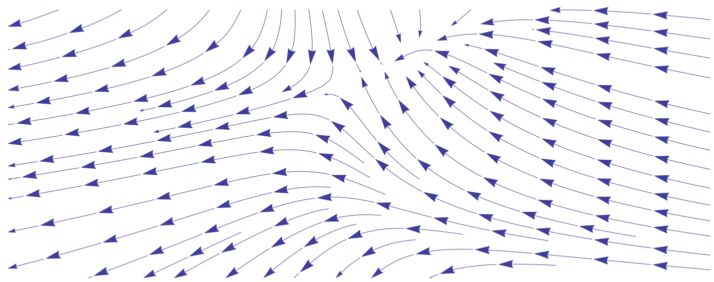
$$\mathcal{M}f = \sup_{\varepsilon > 0} |A_\varepsilon f|$$

is essentially a one-dimensional object and L^2-bounds follow immediately from the boundedness of the one-dimensional Hardy-Littlewood maximal operator. A natural generalization is given by allowing the vector v to depend on the point x. Under what conditions on the vectorfield $v(x)$ is it still possible to prove L^2-boundedness?

Problem statement. Let $\Omega \subset \mathbb{R}^2$ be a bounded open set and $v : \Omega' \to \mathbb{R}^2$ be a vectorfield defined on the neighbourhood Ω' of the closure $\overline{\Omega}$ of Ω. Under which conditions on the vector field v exists $\varepsilon_0 > 0$ such that the maximal function

$$\mathcal{M}_v f = \sup_{0 < \varepsilon < \varepsilon_0} \left| \frac{1}{\varepsilon} \int_0^\varepsilon f(x + tv(x))dt \right|$$

is a bounded operator on $L^2(\mathbb{R}^2)$?



An earlier result by Nagel, Stein & Wainger [2] from the 1970s shows that this is the case if the vector field has non-vanishing curvature on Ω. More precisely: if the vectorfield is normalized, $|v(x)| = 1$, and satisfies

$$\det\left((Dv)(x)v(x), v(x)\right) > 0,$$

then one has L^2-boundedness with a constant that can be formulated in terms of the quantity

$$\sup_{x,y\in\Omega} \frac{\det\left((Dv)(x)v(x), v(x)\right)}{\det\left((Dv)(x)v(y), v(y)\right)}.$$

There are several negative results for vector fields of not sufficient regularity: a Nikodym set $N \subset [0,1]^2$

1. has Lebesgue measure 1 and
2. for every $x \in N$ there is an entire line $\ell \ni x$ with $\ell \cap N = \{x\}$.

The existence of this seemingly paradoxical set shows that some assumptions on the vector field $v(x)$ are necessary. More elaborate constructions based on the Besicovitch set show that Hölder regularity $v \in C^{0,\alpha}$ with $\alpha < 1$ is not sufficient. It is an old conjecture of Zygmund that the Lipschitz class $v \in C^{0,1}$ is sufficient, however, the best result in this direction is Bourgain's theorem for real-analytic vectorfields - even C^∞ is not known.

10.2 Bourgain's theorem

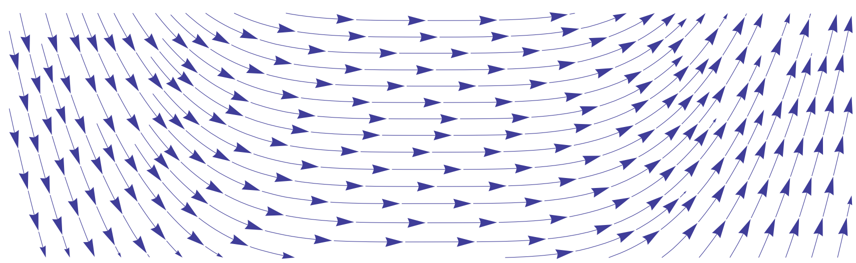
The condition of non-vanishing curvature implies that the vector field changes its direction when moving along the flow. Bourgain realized that it is possible to replace curvature with a more flexible condition on the geometry. Fix $x \in \Omega$ and consider for t small

$$w_x(t) = |\det(v(x + tv(x)), v(x))|.$$

Note that always $w_x(0) = 0$ and that its growth properties indicate how fast the vector field changes. The condition substituting curvature now reads as follows: There are constants $0 < c, C < \infty$ such that

$$\left| \left\{ -\varepsilon \leq t \leq \varepsilon : w_x(t) < \tau \sup_{-\varepsilon \leq t \leq \varepsilon} w_x(t) \right\} \right| \leq C\tau^c \varepsilon$$

holds for all $0 < \tau < 1$, all $0 < \varepsilon < \varepsilon_0$ and all $x \in \Omega$. Non-vanishing curvature implies that the condition holds for $c = 1$. Smaller c allow for



more general vectorfields - even some with zero curvature at certain points.

Theorem 1. *Let $v \in C^1$ satisfy the turning condition. Then the maximal operator M_v is bounded in L^2.*

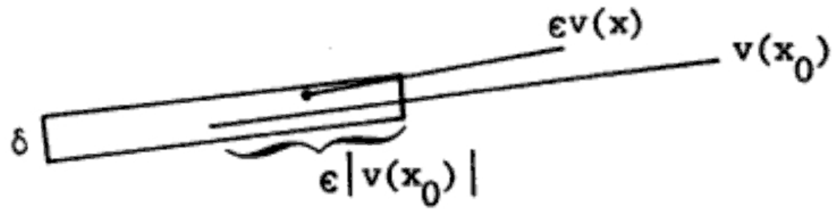
The unconditional result for real-analytic vectorfields follows from showing that the turning condition is automatically fulfilled.

Theorem 2. *Let v be real analytic. Then, for $\varepsilon_0 > 0$ small enough, the maximal operator M_v is bounded in L^2.*

It is not difficult to see that the turning condition is fulfilled for $w_x(t)$ being a polynomial $p(t)$ of degree d in t: factorization of the polynomial in combination with Hölder shows that if $0 < \rho < 1/d$

$$\left(\frac{1}{\varepsilon} \int_{-\varepsilon}^{\varepsilon} \frac{1}{|p(t)|^\rho} dt \right)^{1/\rho} \sup_{|t|<\varepsilon} |p(t)| < C < \infty$$

for some constant C depending only on ρ and the polynomial. The statement for real-analytic vectorfields is shown by algebraic means.



10.3 Sketch of the proof

The proof of Theorem 1 can be decomposed into three main steps.

- Associating with each element $(x_0, \varepsilon) \subset \Omega \times \mathbb{R}_+$ a rectangle in the plane centered at x_0. The ε-parameter will be used to take into account different frequencies. After this, we derive a local estimate for an averaging operator associated to a particular rectangle.

- Studying the geometry of the arising rectangle systems, especially with respect to the possibility of reducing them to a simple subset containing all relevant information.

- Combining the previous steps and using the above rectangle system for the right frequencies to estimate the maximal function by the maximal function for the rectangle systems.

Associating rectangles and local estimates. For given $(x_0, \varepsilon) \subset \Omega \times \mathbb{R}_+$ consider the rectangle R with center x_0, orientation $v(x_0)$, length $\varepsilon|v(x_0)|$ (in direction $v(x_0)$) and width

$$\delta = \varepsilon \cdot \sup_{|t| \leq \varepsilon} \left| \det \left(v(x_0 + tv(x_0)), \frac{v(x_0)}{|v(x_0)|} \right) \right|,$$

which is assumed to be non-zero. A local estimate for $\|A_\varepsilon f\|_{L^2(R)}$ is given for functions f whose support is contained in a neighbourhood of the rectangle (large enough so that the averaging operator never leaves the support) and whose Fourier support is localized to a Littlewood-Paley band. After applying Plancherel, duality and Schur's inequality, one is left with an integral containing purely geometric terms, which can be simplified and does yield terms containing $w_x(t)$ in a natural fashion. Assuming Bourgain's condition and additional arguments, one can simplify the integral further to get

$$\|A_\varepsilon f\|_{L^2(R)} \leq \frac{C}{(T\delta)^c} \left\| f \left(\chi_{R'} * \psi_{\frac{1}{T}} \right) \right\|_{L^2(\mathbb{R}^2)},$$

where c, C come from Bourgain's condition, T is a frequency, δ is the width of the rectangle (depending on ε) and R' is the doubled rectangle R.

Geometric properties of associated rectangles. The crucial argument is that intersecting rectangles of comparable length have comparable width and are indeed exchangable for all relevant purposes. We will again use R' to denote the doubled rectangle R.

Lemma 3. *Let $x' \in R'_{x,\varepsilon}$. Then*

1. $|v(x)| \sim |v(x')|$
2. $\|w_x\|_{L^\infty(-\varepsilon,\varepsilon)} \sim \|w_{x'}\|_{L^\infty(-\varepsilon,\varepsilon)}$
3. $\delta(R_{x,\varepsilon}) \sim \delta(R_{x',\varepsilon})$
4. *$R_{x,\varepsilon}$ is contained in a multiple of $R_{x',\varepsilon}$ and vice versa.*

An inductive application of this Lemma then yields the statement that was hinted at above: each subset $\mathcal{D}_0$ of the set of associated rectangles $\mathcal{D}$ can be further refined to yield a subset $\mathcal{D}_1 \subset \mathcal{D}_0$ with comparable total area

$$\left| \bigcup_{R \in \mathcal{D}_0} R' \right| \leq C \sum_{R \in \mathcal{D}_1} |R|$$

such that the new subset doesn't overlap too much

$$\left\| \sum_{R \in \mathcal{D}_1} \chi_{R'} \right\|_{L^\infty} \leq C.$$

This Besicovitch-type covering argument allows one to conclude that the with $\mathcal{D}$ associated maximal function

$$\mathcal{M}_{\mathcal{D}} f = \sup_{R \in \mathcal{D}, R \ni x} \frac{1}{|R|} \int_R |f|$$

is bounded in L^2. This fact will be of the utmost importance in the conclusion of the proof as it will replace $\mathcal{M}_v$.

Estimation of the maximal function. For $0 < \varepsilon < \varepsilon_0$ and $s > 0$, we define

$$\Omega_{\varepsilon,s} := \left\{x \in \Omega \mid 2^{-s-1} \leq \delta(R_{x,\varepsilon}) \leq 2^{-s}\right\}.$$

We can restrict ourselves to considering only dyadic values 2^{-j} for ε (with $j > j_0$ to account for $\varepsilon < \varepsilon_0$). From the above considerations, it follows that for a fixed value of s the different $\Omega_{\varepsilon,s}$ are essentially disjoint in the sense that

$$\left\|\sum_{\varepsilon} \chi_{\Omega'_{\varepsilon,s}}\right\|_{L^\infty} < C.$$

Let now $f \in L^2(\mathbb{R}^2)$ and perform a Littlewood-Paley decomposition $f = \sum_{\text{dyadic}} f_T$. Then $A_\varepsilon f = \sum_{\text{dyadic}} A_\varepsilon f_T$ and for fixed $x \in \Omega_{\varepsilon,s}$ we can estimate

$$|A_\varepsilon f(x)| \leq \left|\sum_{T \leq 2^s} A_\varepsilon f_T(x)\right| + \sum_{T > 2^s} |A_\varepsilon f_T(x)|.$$

Note that the second expression contains terms which oscillate too fast to be really noticed by the rectangles and can be bounded by $c\|f\|_{L^2}$ independent of ε - the maximal function is not acting on this part of the problem. The same result has to be shown for the first part and this is done in a very nice manner: the key lemma is that for a function g with supp $\hat{g} \subset B(0,T)$ one can estimate $|g(x)| \leq Cg^*(y)$, where g^* is the ordinary Hardy-Littlewood maximal function and y is close enough to x to not be really affected by oscillations, that is $|x-y| < 1/T$. This, however, allows for the maximal function (a line integral) to be estimated by integrating the maximal function on the associated rectangle. For the conclusion, it thus suffices to note that both the maximal function for associated rectangles and the Hardy-Littlewood maximal function are bounded on L^2. This gives the result.

References

[1] Bourgain, J., *A remark on the maximal function associated to an analytic vector field.* Analysis at Urbana, Vol. I, Urbana (1986–1987), pp. 111–132 & London Math. Soc. Lecture Note Ser. **137**, Cambridge Univ. Press, Cambridge (1989).

[2] Nagel, A., Stein, E. M. and Wainger, S., *Differentiation in lacunary directions.* Proc. Nat. Acad. Sci. U.S.A. **75** (1978), pp. 1060–1062.

[3] Nagel, A., Stein, E. M. and Wainger, S., *Hilbert transforms and maximal functions related to variable curves.* Proc. Sympos. Pure Math. **35**, (1979).

Stefan Steinerberger, University Bonn
email: **steinerb@math.uni-bonn.de**

11 The Power Law For The Buffon Needle Probability Of The Four-Corner Cantor Set

after F. Nazarov, Y. Peres, and A. Volberg [5]
A summary written by Krystal Taylor

Abstract

Given a set E in the plane, one may consider the probability that "Buffon's needle," a long line segment dropped at random, intersects the set. In the case that E is a neighborhood of the four-corner Cantor set, a theorem of Besicovitch implies that this probability tends to zero. We discuss a result due to Nazarov, Peres, and Volberg, which gives an explicit upper bound for the rate of this decay in terms of the Favard length [5]. In a sequential paper, Laba and Zhai improve on the techniques in [5] to include a larger class of Cantor Sets [2]. We state these and several other result, and look at some of the key steps in the proof.

11.1 Introduction

11.1.1 The context: analytic capacity, hausdorff measure, and favard length

The Painleve problem asks one to find a geometric characterization of certain removable sets for bounded analytic functions. In 1947, Ahlfors introduced the notion of the analytic capacity of a compact set $E \subset \mathbb{C}$. He showed that the analytic capacity of E, denoted $\gamma(E)$, is zero if and only if E is removable for bounded analytic functions [6]. There have been steps forward in an attempt to understand the situation geometrically, and it was conjectured that the one dimensional Hausdorff measure of E, denoted by $\mathcal{H}^1(E)$, equals zero if and only if $\gamma(E) = 0$. While an argument using complex analysis shows that the forward direction of this conjecture is true, there is a complicated counter example due to Vitushkin for the converse. Later, Garnett found a counter example which is simpler to describe. He observed that when $E = \mathcal{K}$, the four-corner Cantor set, then $\mathcal{H}^1(\mathcal{K}) > 0$ while $\gamma(\mathcal{K}) = 0$ [6] [4].

In light of these counter examples, the conjecture was then re-stated with a new notion of length. Vitushkin conjectured that the Favard length of a set E is zero if and only if $\gamma(E) = 0$. Although this conjecture turns out to not always be true, the four-corner Cantor set, $\mathcal{K}$, has zero Favard length. In this summary, we state several results which give either upper or lower bounds on the rate of decay of the Favard length of the $n-th$ iteration in the construction of $\mathcal{K}$ with our emphasis being on the result of Nazarov, Peres, and Volberg.

11.1.2 The set up: definitions and notation

To construct the four-corner Cantor set, $\mathcal{K}$, begin by replacing the unit square with four sub-squares of side length $\frac{1}{4}$ located at the corners of the unit square. Then, repeat this process indefinitely within each sub-square in a self-similar manner with a scaling factor of $\frac{1}{4}$. Let $\mathcal{K}_n$ denote the set which comes from the $n-th$ iteration of this process; $\mathcal{K}_n$ is a union of 4^n squares of side length $\frac{1}{4^n}$. Then, $\mathcal{K} = \cap \mathcal{K}_n$.

To study the probability that the "Buffon's needle" (an infinite line with direction chosen uniformly at random and then located in a uniformly chosen position in that direction, at a distance at most, say, $\sqrt{2}$ from the origin) intersects a neighborhood of the 4-corner Cantor set, mainly $\mathcal{K}_n$ for some fixed n, we will use the notion of Favard length.

The ***Favard length*** of a set $E \subset \mathbb{R}^2$ is defined by

$$Fav(E) = \frac{1}{\pi}\int_0^{\pi} |ProjR_\theta E| d\theta, \tag{1}$$

where Proj denotes the orthogonal projection from $\mathbb{R}^2$ to the horizontal axis, and where R_θ is the counterclockwise rotation by an angle θ.

The Favard length of a set E is sometimes called the ***Buffon needle Probability*** of the set because, up to a constant factor, it is the probability that Buffon's needle will intersect E. To see this, notice that a line l intersects E if and only if l intersects the orthogonal projection of E onto any line perpendicular to l.

A theorem of Besicovitch implies that the projection of $\mathcal{K}$ to almost every line through the origin has zero length. This means that $Fav(\mathcal{K}) = 0$. A recent result of Nazaroz, Peres, and Volberg reveals that $Fav(\mathcal{K}_n) = O(n^{-\frac{1}{6}})$. [5].

11.2 Results

11.2.1 Statements of background results

Let $A_n \lesssim B_n$ mean that there exists a constant C, which is independent of n, so that $A_n \leq CB_n$. If $A_n \lesssim B_n$ and $B_n \lesssim A_n$, then we write $A_n \sim B_n$.

In 1990, Mattila showed that $Fav(E_n) \gtrsim \frac{1}{n}$. In 2002, Peres and Solomyak proved that $Fav(E_n) \lesssim e^{-c\log^* n}$, where $\log^* n$ denotes the number of times that one must iterate the log function to get from n to 1. Next, in 2008, Bateman and Volberg showed that $Fav(\mathcal{K}_n) \gtrsim \frac{\log n}{n}$.

11.2.2 Statements of the main result and a generalization

The main result of this summary follows.

Theorem 1. *(Nazarov, Peres, and Volberg 2008) For every $\delta > 0$, there exists $C > 0$ such that*

$$Fav(\mathcal{K}_n) \leq Cn^{\delta - 1/6}, \quad \text{for all } n \in \mathbb{N}. \tag{2}$$

A generalization of the previous theorem follows.

Theorem 2. *(Laba and Zhai, 2009) Let E_∞ be a generalized Cantor set. Assume that for some direction θ_0, $|Proj(R_{\theta_0}(E_\infty))| > 0$. Then there exists a $p \in (6, \infty)$, which depends on the choice of K, A, and B, so that*

$$Fav(E_n) \leq Cn^{-1/p}.$$

The explicit range of p is described in [2].

11.2.3 Sketch of the proof of 1

The proof of (2) is is divided into a two sections; the first section deals with a harmonic-analytic estimate, and the second section takes advantage of self-similarity in a combinatorial estimate. Both sections rely on a function which is, up to some minor re-scaling, the sum of the characteristic functions of the projections of $\mathcal{K}_n$ for a fixed rotation and a fixed value of n.

11.2.4 The counting function

Finding the right notation to write down a function which is the sum of the characteristic functions of the projections of $R_\theta \mathcal{K}_n$ can be a bit tedious. This section shows a simple way to write the desired function, with some minor re-scaling, in terms of convolutions.

To begin, we re-center $\mathcal{K}_n$ by replacing it with the set $\mathcal{K}_n - (\frac{1}{2}, \frac{1}{2})$. Due to symmetries, it is enough to average over $\theta \in (0, \frac{\pi}{4})$ in defintion (1) of $Fav(\mathcal{K}_n)$. Now, the projection of $R_\theta(\mathcal{K}_n - (\frac{1}{2}, \frac{1}{2}))$ to the horizontal axis is the union of 4^n intervals of length $4^{-n}(\cos\theta + \sin\theta)$ centered at the points $\sum_{k=0}^{n-1} 4^{-k}\xi_k$, where $\xi_k \in \{\pm\frac{3\sqrt{2}}{8}\cos\theta, \pm 3\frac{\sqrt{2}}{8}\sin\theta\}$. For the sake of notation, we notice that

$$\begin{aligned}
&\Big|Proj(R_\theta(\mathcal{K}_n - (\frac{1}{2}, \frac{1}{2}))\Big| \\
&= \Big| \cup \Big[\sum_{k=0}^{n-1} 4^{-k}\xi_k - \frac{4^{-n}(\cos\theta + \sin\theta)}{2}, \sum_{k=0}^{n-1} 4^{-k}\xi_k + \frac{4^{-n}(\cos\theta + \sin\theta)}{2} \Big] \Big| \\
&= \cos(\frac{\pi}{4} - \theta)\frac{3\sqrt{2}}{8} \Big| \cup \Big[\sum_{k=0}^{n-1} 4^{-k}\eta_k - \frac{4^{-n}\rho}{2}, \sum_{k=0}^{n-1} 4^{-k}\eta_k + \frac{4^{-n}\rho}{2} \Big] \Big|,
\end{aligned}$$

where $\eta_k \in \{\pm 1, \pm\tan(\frac{\pi}{4} - \theta)\}$, $\rho = \frac{8}{3\sqrt{2}}\left(\frac{\cos\theta + \sin\theta}{\cos(\frac{\pi}{4} - \theta)}\right)$, and ξ_k is defined above. This shows that the length of the projection is comparable to the union of 4^n intervals of length $4^{-n}\rho$ centered at the points $\sum_{k=0}^{n-1} 4^{-k}\eta_k$. We are now ready to define f_n. Let $t = \tan(\frac{\pi}{4} - \theta)$. For $\theta \in [0, \frac{\pi}{4})$, define

$$f_n = \sum_{\eta \in \{\pm 1, \pm t\}} \chi_{\left[\sum_{k=0}^{n-1} 4^{-k}\eta_k - \frac{4^{-n}\rho}{2}, \sum_{k=0}^{n-1} 4^{-k}\eta_k + \frac{4^{-n}\rho}{2}\right]}.$$

Since $\chi_{[c-r,c+r]} = \delta_c * \chi_{[-r,r]}$,

$$f_n = \nu^{(n)} * \frac{4^n}{\rho}\chi_{[-\frac{\rho}{2}4^{-n}, \frac{\rho}{2}4^{-n}]},$$

where $\nu^{(n)} = *_{k=0}^{n-1}\nu_k$ and $\nu_k = \frac{1}{4}[\delta_{-4^{-k}} + \delta_{-4^{-k}t} + \delta_{4^{-k}t} + \delta_{-4^{-k}}]$.

It will be useful to observe that $\int_{\mathbb{R}} |\widehat{f_n}(y)| \gtrsim \int_1^{4^{n/2}} |\widehat{\nu}^{(n)}(y)|$ for n sufficiently large, because $\psi = \frac{4^n}{p}\chi_{[-\frac{\rho}{2}4^{-n}, \frac{\rho}{2}4^{-n}]}$ satisfies $\widehat{\psi}(y) \gtrsim 1$ for all $|y| < 4^{n/2}$.

11.2.5 Fourier-analytic part

Let K and S be large positive numbers. Then for $q > 4$, the set

$$E = \{t \in [0,1] : \max_{1 \leq n \leq (KS)^q} \int f_n^2 \leq K\} \tag{3}$$

satisfies $|E| \leq \frac{1}{S}$.

One of the key estimate in the proof of (3) which sets this argument apart from that of the general case is that one can identify the zero set of $\prod_{k=0}^{m} \frac{\cos 4^k y + \cos 4^k ty}{2}$, where this product arises upon re-writing $\widehat{\nu}$ in terms of cosines.

One of the main estimates in the proof of (3), is to show that

$$\int_{4^{-m}}^{1} | \prod_{k=m+1}^{n} \frac{\cos 4^k y + \cos 4^k ty}{2} |^2 dy \gtrsim 4^{m-n} \tag{4}$$

where $m \leq n < \frac{1}{2}(KS)^q$ is choosen so that 4^m is a large multiple of K.

Showing (4) reduces to showing that

$$\int_{4^{-m}}^{1} | \sum_{j=1}^{4^{n-m}} e^{i\lambda_j y} |^2 dy \gtrsim 4^{n-m}. \tag{5}$$

Next, the idea is to introduce a function g on $\mathbb{R}$ with the following properties:
-g is even
-$\int_{\mathbb{R}} g \geq \frac{1}{2}$
-g is supported on $[-1,1] \backslash [-4^{-m}, 4^{-m}]$
-$g \leq 1$
-$\widehat{g} \geq -c\frac{L}{\lambda^2 + L^2}$ with some constant $c > 0$ and $L = 4^m$. This function is given explicitly in [5].

Now

$$\int_{[-1,1]\backslash[-4^{-m},4^{-m}]} | \sum_{j=1}^{4^{n-m}} e^{i\lambda_j y} |^2 dy \geq \int_{\mathbb{R}} g(y) | \sum_{j=1}^{4^{n-m}} e^{i\lambda_j y} |^2 dy$$

To bound the above quantity above, we will take advantage of the fact that for $t \in E$, $\int_{\mathbb{R}} f_{n-m}^2 \leq K$ and, with a change of variable, this is equivalent to

$$\int_{\mathbb{R}} (\sum_j \chi_{[\lambda_j - \frac{\rho}{2}4^m, \lambda_j + \frac{\rho}{2}4^m]})^2 \leq K4^n. \tag{6}$$

Properties of the Poisson kernel play a role.

11.2.6 Finishing the proof

To finish the proof the theorem (1), it suffices to show for $t > 0$ that

$$|\{\theta \in (0, \frac{\pi}{4}) : |Proj R_\theta \mathcal{K}_N| \geq t\}| \lesssim (N^{-1}t^{-p})^{1/q}, \tag{7}$$

where $p > 6$ and $q > 4$.

Indeed, (7) implies that

$$\begin{aligned}
\int_0^{\frac{\pi}{4}} |Proj R_\theta \mathcal{K}_n| d\theta &= \int_0^\infty |\{\theta \in (0, \frac{\pi}{4}) : |Proj R_\theta \mathcal{K}_n| \geq t\}| dt \\
&= \int_0^{cN^{-1/p}} 1 dt + \int_{cN^{-1/p}}^{\sqrt{2}} (N^{-1}t^{-p})^{1/q} dt \\
&\lesssim N^{-1/p}.
\end{aligned}$$

We prove that if K and S large enough and $N \geq K^p S^q$, with $p > 6$ and $q > 4$, then

$$|\{\theta \in (0, \frac{\pi}{4}) : |Proj R_\theta \mathcal{K}_N| \geq \frac{C}{K}\}| \lesssim \frac{1}{S}. \tag{8}$$

The idea behind the proof of (8), is to first fix θ, and then consider the set of x where much stacking occurs for that choice of θ. When this set is small, we show that θ is in a set of small measure. When this set is larger, we show that, beyond a large generation, the size of the projection, $|Proj R_\theta \mathcal{K}_N|$ is small. We combine these observations with the Fourier-analytic part to conclude that the set on the left-hand-side of (8) is small. It is within this proof that the condition that $p > 6$ arises.

In more detail, fix θ, K, and N. Set $F^*(x) = \max_{1 \leq n \leq N} f_n(x)$, and let $\nu = |\{F^* \geq K\}|$. When θ is such that $\nu \lesssim \frac{1}{K^3}$, it can be shown that $\theta \in E$; here we will need $N \geq (KS)^q$ where $q > 4$. When θ is such that $\nu \gtrsim \frac{1}{K^3}$, it can be shown that $|ProjR_\theta\mathcal{K}_N| \lesssim \frac{1}{K}$; here we will need $N \geq K^p S^q$ where $p > 6$ and $q > 4$. We conclude that

$$\{\theta \in (0, \frac{\pi}{4}) : |ProjR_\theta\mathcal{K}_N| \geq \frac{C}{K}\} \subset E.$$

Since the Fourier-analytic part of the argument shows that $|E| \leq \frac{1}{S}$, this concludes (8).

References

[1] M. Bond, *Combinatorial and fourier analytic L^2 methods for buffon's needle problem,* http://bondmatt.files.wordpress.com/2011/05/thesis-7-0.pdf (2011).

[2] I. Laba and K. Zhai, *The favard length of product cantor sets,* Bull. London Math. Soc., doi: 10.1112/blms/bdq059(2010).

[3] J. Mateu, X. Tolsa, J. Verdera *The Planar Cantor Sets of Zero Analytic Capacity and the Local T(b)-Theorem* http://mat.uab.es/ xtolsa/cantset12.pdf

[4] P. Mattila, M. Melnikov, and J. Verdera *The cauchy integral, analytic capacity, and uniform rectifiability* Annals of Math. 144 (1996) 127-136

[5] F. Navarov, Y. Peres, and A. Volberg, *The power law for the buffon needle probability of the four-corner cantor set,* St. Petersburg Math. J. **22** (2010).

[6] H. Pajot *Analytic Capacity, Rectifiability, Menger Curvature, and the Cauchy Integral,* Springer: lecture notes in mathematics 1799 (2000).

Krystal Taylor, University Rochester
email: **taylor@math.rochester.edu**

12 Algebraic methods in discrete Kakeya-type problems

after L. Guth and N. Katz [3]
and Z. Dvir [2]
A summary written by Faruk Temur

Abstract

We prove three discrete analogs of the Kakeya problem using the polynomial method.

12.1 Introduction

The aim of this note is to give an exposition of papers [2],[3]. The common point of these papers is that they answer discrete analogs of the Kakeya conjecture in positive. Let's first take a look at these problems. Here is the first problem:

Conjecture 1. *Let $L \subset \mathbb{R}^3$ be a set of N lines. Let a joint be a point where at least three lines intersect non-coplanarly. Let J be the set of joints. Then $|J| \lesssim N^{3/2}$.*

This conjecture was first considered in [1], where it was proved that $|J| \lesssim N^{7/4}$. Guth and Katz proved this conjecture in [3]. The connection between this problem and the Kakeya problem was uncovered by Wolff and Schlag; see [4]. Now we turn to the second problem:

Conjecture 2. *Let $L \subset \mathbb{R}^3$ be a set of lines with cardinality N^2 and $P \subset \mathbb{R}^3$ be a set of points. Let each line in L be incident to at least N points of P, and no more than N of them lie on the same plane. Then $|P| \gtrsim N^3$.*

This problem was posed by Bourgain in 2004. The analogy with the Kakeya conjecture is quite obvious. It was proved by Guth and Katz in [3]. The third problem is of a somewhat different nature:

Conjecture 3. *Let $\mathbb{F}$ be a finite field with q elements. A Kakeya set in this context is a set $K \subset \mathbb{F}^n$ such that for every $x \in \mathbb{F}^n$ there exist a $y \in \mathbb{F}^n$ with*

$$\{y + a \cdot x | a \in \mathbb{F}\} \subset K.$$

Then $|K| \gtrsim q^n$.

This problem was posed by Wolff in [4]. Dvir proved it in [2]. In his proof he introduced the polynomial method. L. Guth and N. Katz, in [3], proved the first and second conjectures by adapting this method to the Euclidean setting.

We start our exposition by reviewing algebra and geometry facts required for the proofs of these three problems. Then we prove the joints problem and Bourgain's problem. Finally we turn to the finite field Kakeya problem.

12.2 Preliminaries

Usefulness of the polynomial method rests on certain algebraic and geometric facts about polynomials. We, in this section, will review these facts. The first is as follows

Proposition 4. *Let f, g be elements of $\mathbb{R}[x_1, x_2, x_3]$ with positive degrees l and m respectively. If f and g vanish identically on more that lm lines in $\mathbb{R}^3$, then they have a common factor.*

The proof of this relies on the celebrated theorem of Bezout, which has a very similar statement except that f, g are elements of $\mathbb{C}[x_1, x_2]$ and lines are replaced by points in $\mathbb{C}^2$. Though we will not get into Bezout's theorem's proof we remark that a simple proof is given in [3]. The second fact is the following:

Proposition 5. *Let $X \subset \mathbb{R}^3$ be a set of N points. Then there is a non-trivial polynomial in $\mathbb{R}[x_1, x_2, x_3]$ with degree $\lesssim N^{1/3}$ vanishing on each point of X.*

The proof uses only basic linear algebra: A degree d polynomial of three variables has roughly d^3 coefficients, thus with at most this many points we have an undetermined system of equations, which will always have a non-trivial solution. The next proposition is the only tool required to solve the finite field Kakeya problem and is called the Schwartz-Zippel lemma.

Proposition 6. *Let $f \in \mathbb{F}[x_1, \dots, x_n]$ be a non-zero polynomial with degree at most d. Then the zero set of f has size at most $d \cdot q^{n-1}$.*

After a review of necessary algebra, we now turn to some geometric facts we will need . Let p be a non-trivial polynomial and S its zero set. We call a point a of S critical if $\nabla p(a)$ is zero, and regular if it is not critical. If every point of a line is critical we call it a critical line. Appliying Proposition 1 to

p and a non-trivial component of ∇p, and bearing in mind the irreducibility of p we get the following:

Proposition 7. *The zero set S of p cannot contain more than d(d-1) critical lines.*

Define algebraic second fundamental form of a point to be

$$\mathbf{II}(p)(a) = \{(\nabla_{\nabla p \times e_j} \nabla p) \times \nabla p\}_{j=1,2,3}.$$

This is a set of three vectors, and thus it has nine components of degree at most $3d - 4$. We call a regular point flat if algebraic fundamental form vanishes at that point. If a line contained in S is not critical and all of its regular points are flat then we call it a flat line. We give a criterion for flatness of points.

Proposition 8. *Let $a \in S$ be a regular point. If S contains three distinct lines passing through a, then a is a flat point.*

Proof of this proposition relies on the observation that algebraic second fundamental form is closely related to the geometric second fundamental form, and on an application of proposition 1. From this proposition and Proposition 1 the corollary below follows immediately.

Corollary 9. *If S contains more than $3d^2 - 4d$ flat lines, then S is a plane.*

Now we are ready to prove Joints problem and Bourgain's problem.

12.3 The joints problem and Bourgain's problem

On a heuristic level both proofs are very similar. Let's view them first at this level.

For the joints problem if the number of joints is too big then we are able to find a polynomial vanishing on every line with degree $\ll N^{1/2}$. Then taking care that each contain enough number of joints, by non-coplanarity, these lines will all be critical. Thus we have N critical lines, which is not possible for a polynomial of degree $\ll N^{1/2}$. Here the ingenuity is in the set up that gives a lower degree polynomial as number of joints increases.

For Bourgain's problem things are simpler. If number of points is $\ll N^3$ then there is a polynomial vanishing on all of them with degree $\ll N$.

This polynomial then will include all lines, as they each intersect the zero set at least N times. Then it is easy to show that a proportion of these lines comparable to N^2 is either flat or critical, but neither is possible for a polynomial of degree $\ll N$.

Now let's get into some detail for the joints problem. Let $|J| \geq KN^{3/2}$ for some K large. Since at each joint there are at least three incidences, color one red, one blue, and one green. Let $L_R \subset L$ consist of lines with at least $\frac{K}{1000}N^{1/2}$ red incidences and define L_G, L_B similarly. The set of joints J' that has a red incidence with L_R, a blue with L_B, and a green with L_G then satisfy $|J'| \geq \frac{99}{100}|J|$. We say a line l in L_G or L_B meets a line l' of L_R if $l \cap l'$ is a joint in J'. Then each line of L_G or L_B meets at least $\frac{K}{1000}N^{1/2}$ line of L_R.

Here instead of directly taking a polynomial vanishing on all of the joints, we take a random subset $L'_R \subset L_R$, picking each line with probability $\frac{1}{K}$. Then with positive probability $|L'_R| \leq \frac{2N}{K}$ and each line of L_G or L_B meets at least $\frac{1}{2000}N^{1/2}$ lines of L'_R. Take a set S by choosing $\frac{1}{2}N^{1/2}$ points on each line of L'_R. Since $|S| \leq \frac{N^{3/2}}{K}$, there exists a polynomial p of degree $\lesssim \frac{N^{1/2}}{K^{1/3}}$ vanishing on S. Due to its low degree it vanishes on each line of L'_R, and since any line of L_B or L_G meets lines of L'_R at least $\frac{1}{2000}N^{1/2}$ times, p vanishes on all of L_G and L_B. So p vanish on all of J'.

Since p need not be irreducible, we factorize it into irreducibles: $p = \prod_{j=1}^{m} p_j$. Let d_j denote the degree of p_j and J_j be the subset of J' where p_j vanishes. Then by pigeonholing we can find a value j with $|J_j| \gtrsim K^{4/3}Nd_j$.

Let $L_{R,j}, L_{G,j}, L_{B,j}$ be those lines in L_R, L_G, L_B which are incident to at least $d_j + 1$ lines in J_j. This ensures that they lie in the zero set of p_j. Let $J'_j \subset J_j$ be the set of points with respectively red, green, blue incidences with $L_{R,j}, L_{G,j}, L_{B,j}$. Every element of this set is a critical point. Then define $L'_{R,j}, L'_{G,j}, L'_{B,j}$ as sets with at least $d_j + 1$ incidences with J'_j. So these are all critical lines. Finally let J''_j be set of joints defined by these lines. This set satisfy $|J''_j| \geq \frac{99}{100}|J_j|$. Now we have two cases. If none of $L'_{R,j}, L'_{G,j}, L'_{B,j}$ has more than d_j^2 lines then we have a problem with fewer lines and better exponents than the original. If one of them does have more than d_j^2 lines, then clearly we have more critical lines than allowed. Thus we are done.

We turn to Bourgain's problem in detail now. We may assume each line incident to exactly N points. Assume $|P| = \frac{N^3}{K}$. Let $v(x)$ denote the number of incidences the point x has. Let $x \in P_v$ if $v(x) \geq \frac{K}{1000}$. By dyadic pigeonholing we can find $P_j \subset P_v$ such that $\frac{2^{j-1}K}{1000} \leq v(x) < \frac{2^jK}{1000}$ whenever

$x \in P_j$, and $\sum_{x \in P_j} v(x) \geq \frac{999N^3}{2000j^2}$. The set P_j then satisfies $\frac{N^3}{K2^j j^2} \lesssim |P_j| \lesssim \frac{N^3}{K2^j}$. Hence we can find a polynomial p with degree $d \lesssim \frac{N}{K^{1/3}2^{j/3}}$ vanishing on all of P_j. As above we factorize this polynomial into irreducibles, and the zero set $P_{j,k}$ of one factor p_k has the property $|P_{j,k}| \gtrsim \frac{N^2 d_k}{K^{2/3}2^{2j/3}j^2}$ where d_k is the degree of p_k.

Let $Y = P_{j,k}$. Let L' be the set of lines incident to more than $100d_k$ points of Y. On all of these lines p_k vanish. Let Y' be set of points in Y that are incident to more than 3 lines in L'. Thus each point in Y' is either a critical point or a flat point of p_k. Finally let L'' be lines in L' that are incident to at least $10d_k$ points of Y'. Any line in L'', then, is either critical or flat. Incidences I between L'' and Y' satisfy $|I| \gtrsim N^2 d_k K^{1/3} 2^{j/3} j^{-2}$. Since each point in Y' is either critical or flat, any incidence in I is either with a critical point of with a flat point. So number of either critical incidences or flat incidences satisfy $\gtrsim N^2 d_k K^{1/3} 2^{j/3} j^{-2}$. Neither of these is possible for our surface can have at most d_k^2 critical lines and $3d_k^2$ flat lines.

12.4 The finite field Kakeya problem

Using the Schwartz-Zippel lemma we shall first prove that $|K| \gtrsim q^{n-1}$, then improve it to $\gtrsim q^n$. The main idea will be that if a polynomial vanishes on a Kakeya set, then it vanishes everywhere, so Kakeya sets have to be large. Now suppose to the contrary that there exist a Kakeya set K with

$$|K| < \binom{q+n-3}{n-1}$$

Then by basic linear algebra there exist a homogenous polynomial g of degree $q-2$ that vanishes on the entire set K. Then by homogeneity it vanishes on the set K' where

$$K' = \{c \cdot x | x \in K, c \in \mathbb{F}\}.$$

Now we claim that g vanishes on all of $\mathbb{F}^n$. To see this let $x \in \mathbb{F}^n$. Then there exist $y \in \mathbb{F}^n$ such that $y + a \cdot x \in K$ for all $a \in \mathbb{F}$. Except 0 all these elements a have inverses, so multiplying by them we see that $x + a \cdot y \in K'$ for all $a \in \mathbb{F}$ except for 0. Since g vanishes on K', this means that it vanishes on a line at $q-1$ points. But it has degree at most $q-2$, thus it should vanish on all points of this line, hence $g(x) = 0$. So the zero set of the polynomial g has size q^n. But this contradicts the Schwartz-Zippel lemma.

We can upgrade this result as follows. Assume to the contrary that there exist a Kakeya set K with

$$|K| < \binom{q+n-2}{n}.$$

Then we can find a non-zero polynomial g of degree $q-1$ in $\mathbb{F}[x_1, \ldots, x_n]$, vanishing on all of K. We can write g as $\sum_i^{q-1} g_i$ where g_i is the homogenous part of g of degree i. For any fixed $x \in \mathbb{F}^n$ there exist a $y \in \mathbb{F}^n$ such that $y + a \cdot x \in K$ for all $a \in \mathbb{F}$. Thus $g(y + a \cdot x) = 0$ for all such a. But this is a polynomial of degree $q-1$ in a vanishing at q points, so it should identically vanish. Hence its coefficients are zero. The coefficient of a^{n-1} is g_{q-1}, so actually $g = \sum_i^{q-2} g_i$. Iterating the process gives in the end that g is a constant. But since g vanishes at some points it should be the zero polynomial, which is a contradiction.

References

[1] B. Chazelle, H. Edelsbrunner, L. Guibas, R. Pollack, R. Seidel, M.Sharir and J. Snoeyink, Counting and cutting cycles of lines and rods in space, Computational Geometry: Theory and Applications, 1(6) 305-323 (1992)

[2] Z. Dvir, On the size of Kakeya sets in finite fields, J. Amer. Math. Soc. 22, 1093-1097 (2009)

[3] L. Guth and N.H. Katz, Algebraic methods in discrete analogues of the Kakeya problem, Adv. in Math., 225, 2828-2839 (2010)

[4] T. Wolff, Recent work connected with Kakeya problem, Prospects in Mathematics, 129-162 (1999)

Faruk Temur, UIUC
email: **temur1@illinois.edu**

13 Bounds on oscillatory integral operators based on multilinear estimates

after J. Bourgain and L. Guth
A summary written by Joshua Zahl

Abstract

We establish new L^p bounds for the restriction operator associated to a paraboloid in $\mathbb{R}^3$.

13.1 Introduction

In [2], Bourgain and Guth present several new results on the L^p boundedness of the restriction operator associated to surface in $\mathbb{R}^n$, $n \geq 3$. They also examine variable-coefficient analogues of the restriction operator and a Kakeya-type maximal function associated to curved Kakeya sets.

In this summary we shall only consider restriction estimates for paraboloids in $\mathbb{R}^3$. In this direction, Bourgain and Guth prove the following result:

Theorem 1. *Let $S \subset \mathbb{R}^3$ be a compact subset of the paraboloid $\{x_3 = x_1^2 + x_2^2\}$, and let σ be surface measure on S. Let $p_0 = 3.3$. Then for all measures $\mu \ll \sigma$ such that $\frac{d\mu}{d\sigma} \in L^\infty(S, \sigma)$,*

$$\|\hat{\mu}\|_{L^{p_0}} \lesssim \Big\|\frac{d\mu}{d\sigma}\Big\|_\infty. \tag{1}$$

Since (1) with $p_0 = \infty$ is trivial, establishing (1) for one value of p_0 immediately establishes it for all larger values; thus the goal is to prove (1) for as small a value of p_0 as possible. The (conjectured) optimal value of p_0 is 3 (see [4]). We shall not survey previous progress on the restriction conjecture here; see [7, Lecture 1] for a survey of progress up to 2003. We will merely note that in [2], Bourgain and Guth obtain the best known restriction results in dimensions 3 and all dimensions $d > 4$ not divisible by 3. In dimension 2, a sharp result was proved by Cordoba in [3], while in dimension 4, Bourgain and Guth's result is the same as Tao's from [6]. For dimensions $d = 3k,\ k > 1$, Bourgain and Guth's results have recently (8/19/2011) been beaten by the work of Temur in [8], who extended the techniques used to prove Theorem 1 to all dimensions divisible by 3.

Before proving (1) with $p_0 = 3.3$, Bourgain and Guth first give a simpler argument that establishes a weaker version of Theorem 1:

Theorem 2. *Equation* (1) *holds for* $p_0 = \frac{10}{3}$.

While Theorem 2 was already known (indeed, this result is due to Tao in [6]), Bourgain and Guth's arguments are simpler than Tao's. In section 13.3 below we shall first prove the $p_0 = \frac{10}{3}$ result, and then in section 13.4 we shall establish the stronger result with $p_0 = 3.3$.

13.1.1 Proof Sketch

Rather than considering measures μ on S it is easier to work with functions $f \colon \Omega \to \mathbb{R}$, where $\Omega \subset \mathbb{R}^2$ is a small ball, and instead of considering $\|\hat{\mu}\|$, we shall obtain bounds on the operator

$$Tf(x) = \int_{\Omega} e^{i\phi(x,y)} f(y)dy, \tag{2}$$

where

$$\phi(x,y) = x_1 y_1 + x_2 y_2 + x_3(y_1^2 + y_2^2). \tag{3}$$

By standard limiting arguments, a bound of the form $\|Tf\|_p \lesssim \|f\|_\infty$ yields (1) for the same value of p.

Each point $y \in \Omega$ is associated to the point $(y_1, y_2, y_1^2 + y_2^2) \in S$ by the embedding

$$\iota \colon \Omega \hookrightarrow S.$$

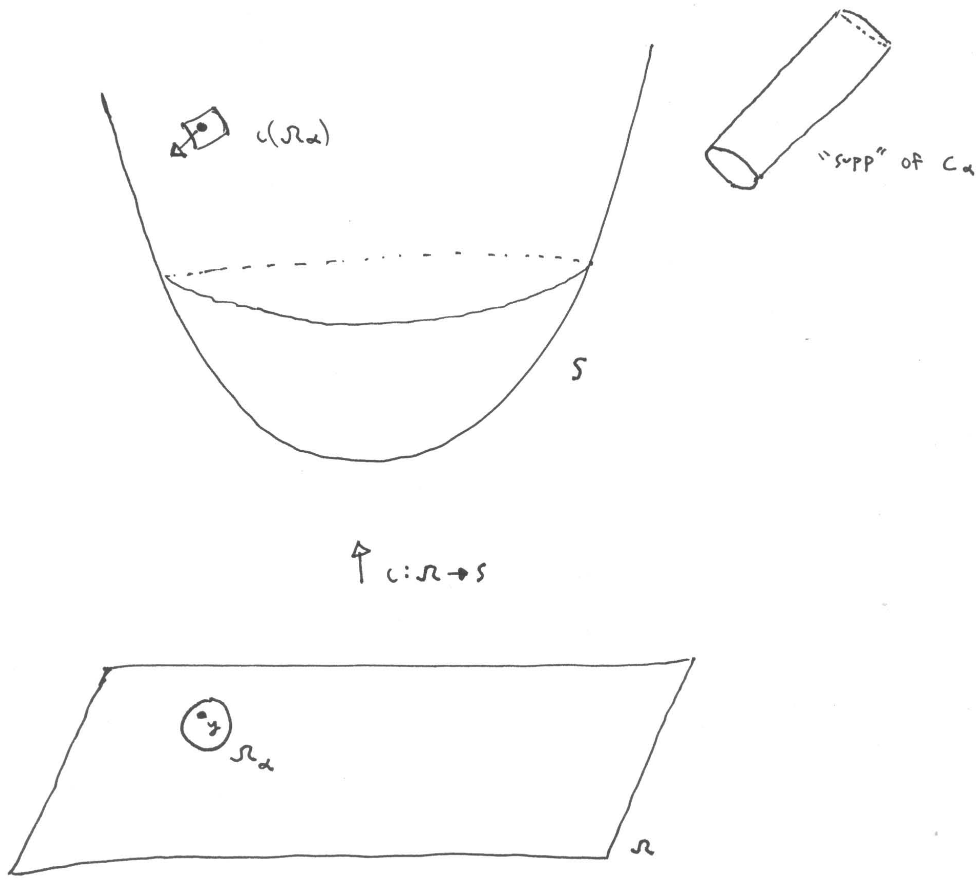
Thus, we shall frequently identify points in Ω with their images in S.

Bourgain and Guth begin with the standard step of writing $\Omega = \bigcup \Omega_\alpha$ as a finitely overlapping union of balls Ω_α of size K^{-1} for K a large constant to be determined later. Let y_α be the center of Ω_α. There are $\sim K^2$ values of α. We have

$$\begin{aligned} Tf(x) &\lesssim \sum_\alpha e^{i\phi(x,y_\alpha)} \Big[\int_{\Omega_\alpha} e^{i[\phi(x,y)-\phi(x,y_\alpha)]} f(y)dy \Big] \\ &= \sum_\alpha e^{i\phi(x,y_\alpha)} (T_{\Omega_\alpha} f)(x) \qquad (4) \\ &= \sum_\alpha c_\alpha(x). \qquad (5) \end{aligned}$$

We have that $\iota(\Omega_\alpha)$ is contained in a "plate"; the K^{-2} neighborhood of a disk in $\mathbb{R}^3$ of radius K^{-1}. Thus c_α is essentially supported on the (Fourier) dual of this plate, which is a "tube," i.e. the K–neighborhood of a line segment of length K^2. Furthermore, each tube points in the same direction as the corresponding plate, but the tube could be centered at any point of $\mathbb{R}^3$ (see figure 13.1.1). See [7] for an excellent discussion of these ideas.

Figure 1: The domain Ω, its image $S = \iota(\Omega)$, the cap Ω_α, its image, and the "support" of c_α (really just the set where c_α is large).



Our goal is now to control the interactions between the various functions c_α. Roughly one of two things can occur:

Case 1: At every point $x \in \mathbb{R}^3$, we can find three disks $\Omega_1, \Omega_2, \Omega_3$ whose corresponding tubes point in "transverse" directions (i.e. the tubes' directions span $\mathbb{R}^3$ in some quantitative way that we will elaborate upon below). Then, the behavior of $\sum c_\alpha$ is already well understood; Bennett, Carbery, and Tao in [1] established a sharp restriction theorem for this "transverse interaction" situation, and this restriction theorem is more than powerful enough to establish (1) with $p = \frac{10}{3}$.

Case 2: There exists some 2–dimensional vector space $V \subset \mathbb{R}^3$ so that "most" of the contributions to $\sum c_\alpha$ come from disks Ω_α such that the normal vectors of $\iota(\Omega_\alpha)$ lie in a (say) $\frac{100}{K}$ neighborhood V. Now, however, we have essentially reduced the problem from $\mathbb{R}^3$ to $\mathbb{R}^2$, and the restriction operator in $\mathbb{R}^2$ is well understood.

Bourgain and Guth carefully balance the "cutoff" of when we are in Case 1 or Case 2, and doing this allows them to establish the $p_0 = \frac{10}{3}$ bound.

To establish (1) with $p = 3.3$, an additional argument is needed. If we are lucky enough to be in Case 1, then (1) holds for a very good value of p (even better than 3.3); it is Case 2 that presents difficulties. If we are in Case 2, then we consider the disks $\Omega_\alpha \subset \Omega$ whose image under ι have normal direction almost lying in V. We then re-scale so that these disks becomes a balls of radius 1. The same dichotomy now applies in this rescaled situation, so again we fall into either Case 1 or Case 2. If we are in Case 1, then we stop. Otherwise, we continue until we either find ourselves in Case 1, or we have iterated a certain number of times (at which point we stop and essentially treat whatever we have left as an error term). We now find ourself juggling a large collection of tubes at different scales. However by pigeonholing, we can assume that "most" of the tubes exist at a single scale. Now, we need to prevent this tubes from all overlapping with each other too much. If they were to do so, then the L^p norm of their sum would be large, which we are trying to show is not the case. However, since the tubes are the duals of disjoint caps (which thus have different normal vectors), the tubes all point in different directions. At this point we can appeal to the Kakeya maximal function, which controls the extent to which tubes pointing in different directions can overlap. Using a result due to Wolff about the behavior of the Kakeya maximal function, we can show that the tubes don't overlap too much, and this allows us to finish the argument.

13.2 Background Preliminaries

13.2.1 The multilinear restriction estimates of Bennett-Carbery-Tao

Bourgain and Guth make fundamental use of the multi-linear restriction estimates developed by Bennett, Carbery and Tao (henceforth BCT) in [1]. BCT develop restriction estimates that apply in the special case where the sum (5) is "dominated by transverse interactions," in a manner we shall make precise below.

Theorem 3. *Let $S \subset \mathbb{R}^3$ be a paraboloid, and for $x \in S$, let $x' \in S^2$ be the unit normal vector of S at x. Let $U_1, U_2, U_3 \subset S$ be small surface patches on S, with*

$$|x_1' \wedge x_2' \wedge x_3'| > c. \tag{6}$$

for all $x_i \in U_i,\ i = 1, 2, 3$.

Let $B_R \subset \mathbb{R}^3$ be a ball of radius R. Let $g_i \in L^\infty(U_i)$, and $q = 3$. Then with certain technical caveats which we will gloss over,

$$\left\| \prod_{j=1}^{3} \int_{U_j} g(x) e^{i\phi(x,\xi)} dx \right\|_{L^1(B_R)} \lesssim R^\epsilon \prod_{j=1}^{3} \|g_j\|_{L^2}. \tag{7}$$

13.2.2 Localization and ϵ–removal

Bourgain and Guth make use of a technical tool called the ϵ–removal lemma, which was first developed by Tao in [5] . The ϵ–removal lemma allows one to obtain global restriction estimates from local ones. The version here is slightly different from Tao's original theorem, and a proof can be found in the appendix of [2].

Lemma 4 (ϵ–removal lemma)**.** *Let $S \subset \mathbb{R}^3$ be a paraboloid. Suppose that for every $\gamma > 0$, there exists a constant C_γ so that*

$$\left\| \hat{f}|_S \right\|_{L^1(d\sigma)} \leq C_\gamma R^\gamma \|f\|_{L^p(B_R)}. \tag{8}$$

Then for any $\tilde{p} < p$, we have

$$\left\| \hat{f}|_S \right\|_{L^1(d\sigma)} \lesssim \|f\|_{L^{\tilde{p}}(\mathbb{R}^3)}. \tag{9}$$

To simplify our notation if A and B are functions of R, we say that $A \lessapprox B$ if for all $\epsilon > 0$, there exists a constant C_ϵ such that $A \le C_\epsilon R^\epsilon B$. Often the dependence on R will be implicit. To avoid confusion, the symbol R will always denote the radius of the ball on which we are obtaining a restriction estimate.

Thus, using this ϵ–removal lemma, in order to prove Theorem 1 it suffices to establish the estimate

$$\|Tf\|_{L^{p_0}(B_R)} \lessapprox \|f\|_\infty. \tag{10}$$

13.2.3 Scaling

Let $S \subset \mathbb{R}^3$ be a compact surface with positive-definite second fundamental form (we think of S as being a piece of a paraboloid), and define

$$Q_R^{(q)} = \sup \left\| \int_S g(\xi) e^{ix\cdot\xi} \sigma(d\xi) \right\|_{L^q(B_R)}, \tag{11}$$

where the supremum is taken over all $g \in L^\infty(S)$ with $|g| \le 1$. Let $U_\rho \subset S$ be a cap of radius ρ. By a change of variables, we have

$$\left\| \int_{U_\rho} g(\xi) e^{ix\cdot\xi} \sigma(d\xi) \right\|_{L^q(B_R)} \lesssim \rho^{2-4/q} Q_{\rho R}^{(q)}. \tag{12}$$

We shall make essential use of this scaling property in the arguments below.

13.2.4 Wolff's $\mathbb{R}^3$ Kakeya estimates

Bourgain and Guth prove Theorem 1 by combining the arguments used to prove Theorem 2 with an L^p estimate on the Kakeya maximal operator. Bounds on the Kakeya maximal operator can be interpreted as bounds on the extent to which collections of tubes can overlap if the tubes point in different directions. In [9], Wolff established the following estimate:

Theorem 5 (Wolff 5/2–Kakeya estimate). *For all $\epsilon > 0$, there exists a constant C_ϵ such that if $\mathcal{T}$ is a collection of tubes of length 1 and radius δ which point in δ–separated directions, then*

$$\Big\| \sum_{\mathcal{T}} \chi_T \Big\|_{5/3} \le C_\epsilon \delta^{-1/5-\epsilon}. \tag{13}$$

13.3 Proof of Theorem 2: the $p = 10/3$ bound

Recall the decomposition $\Omega = \bigcup \Omega_\alpha$ and $Tf(x) \lesssim \sum c_\alpha(x)$ from Section 13.1.1. Note that $\hat{c}_\alpha$ is supported in a ball of radius K^{-1}, and thus by an application of the uncertainty principle, $c_\alpha(x)$ can be controlled by an average of c_α over a suitable ball centered at x. We will make use of this fact in Equation (17) when we are applying the multi-linear theory of BCT.

Now, define

$$c_*(x) = \max_\alpha c_\alpha(x) = c_{\alpha_*}(x). \tag{14}$$

Of course the choice of α that achieves this maximum will depend on x (i.e. α_* is a function of x), but we shall make the following simplifying assumption:

Simplifying Assumption 1. $\alpha_*(x)$ is constant, i.e. it does not depend on x.

This assumption is fairly harmless, since there are only K^2 choices of α_*, and we don't care if we loose a constant depending on K in our final estimate. At each point $x \in B_R$, one of three things can occur:

Case 1: **Transverse interaction**: There exist three indices $\beta_1, \beta_2, \beta_3$ such that $c_{\beta_1}(x),\ c_{\beta_2}(x),\ c_{\beta_3}(x) > K^{-4} c_*(x)$, and

y_{β_1} is distance at least $1000/K$ from the line passing through $y_{\beta_2 1}$ and y_{β_3}, and similarly for the roles of $y_{\beta_1}, y_{\beta_2}, y_{\beta_3}$ permuted. (15)

Case 2: **Collinear interaction**: If $|y_\alpha - y_{\alpha_*}| > K^{-1}$, then $c_\alpha(x) < K^4 c_*$.

Case 3: **Coplanar interaction**: Case 1 does not hold, and there exists α_{**} such that $c_{\alpha_{**}} > K^{-4} c_*$ and $|y_{\alpha_*} - y_{\alpha_{**}}| > K^{-1}$.

Of course, we must choose one of the above three cases for each $x \in B_R$. However, we shall make another simplifying assumption:

Simplifying Assumption 2. Each $x \in B_R$ falls into the same case. If each x lies in Case 1, then the three indices $\beta_1, \beta_2, \beta_3$ are the same for each $x \in B_R$, and if each x lies in Case 3, then the choice of α_{**} is the same for each $x \in B_R$.

Again, since there are only $O(K^2)$ indices α and we are free to loose constants depending on K, these assumptions are fairly harmless.

13.3.1 Case 1: transverse interaction

For $q \geq 3$, we have

$$\int_{B_R} |Tf(x)|^q \leq \int_{B_R} |Tf(x)|^3 \leq K^{18} \int |c_{\beta_1}(x) c_{\beta_2}(x) c_{\beta_3}(x)| dx. \tag{16}$$

Simplifying Assumption 3. By an application of the uncertainly principle which was alluded to at the beginning of Section 13.3, we have that

$$c_{\beta_j}(x) \text{``} = \text{''} \int_{\Omega_{\beta_j}} f(x) e^{i\phi(y,x)} dy. \tag{17}$$

The quotation marks " " around the equals sign indicate that this equality is not true as stated, but it conveys the correct idea and can be made rigorous (often by introducing additional intermediate quantities and technical steps)

Thus, combining (16) and (17) we obtain

$$\|Tf(x)\|_{L^1(B_R)} \lesssim \left\| \prod_{j=1}^{3} \int_{\Omega_{\beta_j}} f(x) e^{i\phi(y,x)} dy \right\|_{L^1(B_R)}, \tag{18}$$

but by the BCT bound from Theorem 3, (18) $\lessapprox 1$.

13.3.2 Case 2: collinear interaction

Case 2 is handled by the same types of scaling arguments discussed in the next section, so for the sake of brevity we shall omit it.

13.3.3 Case 3: coplanar interaction

Let $\mathcal{L} = \mathcal{L}(y_{\alpha_*}, y_{\alpha_{**}})$ be the line passing through y_{α_*} and $y_{\alpha_{**}}$. If y_α is far from this line, then $c_\alpha(x)$ is small. We shall pretend that it is actually 0:

Simplifying Assumption 4. If $\text{dist}(y_\alpha, \mathcal{L}) > 1000\frac{1}{K}$, then $c_\alpha(x) \equiv 0$.

Thus we must estimate the contribution from those c_α with y_α close to $\mathcal{L}$. Again, one of two cases can occur; either one (or a collection of closely clustered) cap dominates $Tf(x)$, or there exist at least two caps near $\mathcal{L}$ that both contribute significantly. We shall assume the latter:

Simplifying Assumption 5. We can find segments $\mathcal{L}_1, \mathcal{L}_2 \subset \mathcal{L}$ with $\mathrm{dist}(\mathcal{L}_1, \mathcal{L}_2) > \frac{10^6}{K}$ such that

$$|Tf(x)|^4 \lesssim \Big| \sum_{\substack{\alpha \\ \mathrm{dist}(y_\alpha, \mathcal{L}_1) < K^{-1}}} c_\alpha(x) \Big|^2 \cdot \Big| \sum_{\substack{\alpha \\ \mathrm{dist}(y_\alpha, \mathcal{L}_2) < K^{-1}}} c_\alpha(x) \Big|^2. \tag{19}$$

Now, by an argument similar to the one Cordoba used in [3] to establish the restriction theorem in $\mathbb{R}^2$, we have the estimate

$$(19) \text{ ``} \lesssim \text{''} \Big(\sum_{\substack{\alpha \\ \mathrm{dist}(y_\alpha, \mathcal{L}) < K^{-1}}} |c_\alpha(x)|^2 \Big)^{1/2}. \tag{20}$$

Since the sum is over $O(K)$ indices y_α, by Hölder's inequality we have

$$\|Tf(x)\|^q_{L^q(B_R)} \lesssim K^{q/2-1} \Big(\sum_\alpha \|c_\alpha(x)\|^q_{L^q(B_R)} \Big). \tag{21}$$

Here, however, the second part of our Simplifying Assumption 2 does not hold, so we must sum over all $O(K^2)$ indices α. Thus

$$\|Tf(x)\|^q_{L^q(B_R)} \lesssim K^{q/2-1} K^2 (Q^{(q)}_{R/K})^q. \tag{22}$$

Now, we apply the rescaling argument from Section 13.2.3 to obtain

$$\|Tf(x)\|_{L^q(B_R)} \lesssim K^{5/q-3/2} Q_{R/K}. \tag{23}$$

If we assume that $q > \frac{10}{3}$ and that $Q_{R/K} \lessapprox 1$ (the latter is our induction on scales hypothesis), then $\|Tf(x)\|_{L^q(B_R)} \lessapprox 1$.

13.4 Proof of Theorem 1: the $p = 3.3$ bound

Recall the decomposition of $Tf(x)$ from Section 13.3:

$$\begin{aligned}
|Tf| &= \text{``transverse interactions''} + \text{``coplanar interactions''} \\
&\quad + \text{``collinear interactions''} \\
&\lesssim \max_{\substack{\beta_1,\beta_2,\beta_3 \\ \text{``transverse''}}} (|c_{\beta_1}| \cdot |c_{\beta_2}| \cdot |c_{\beta_3}|)^{1/3} + \max_{\substack{A \\ A', A''}} \Big| \sum_{\substack{\alpha \\ y_\alpha \in A'}} c_\alpha(x) \Big|^{1/2} \Big| \sum_{\substack{\alpha \\ y_\alpha \in A'}} c_\alpha(x) \Big|^{1/2} \\
&\quad + \max_a |T(f|_{B(a,K^{-1})})| \qquad (24)\\
&= (25) + (26) + (27).
\end{aligned}$$

Here, $A \subset \mathbb{R}^2$ is a rectangle of dimensions $1 \times K^{-1}$, and A', A'' are $\frac{1}{100}$–separated sub-rectangles which satisfy an estimate similar to (19). Each function c_α is obtained by applying to f a restriction operator associated with the "cap" $\iota(\Omega_\alpha)$, (see figure 13.1.1), which has diameter $\sim K^{-1}$. It will be convenient to think of the entire surface S (a compact subset of a paraboloid) as being a "cap" of diameter ~ 1. We shall use the symbol τ to denote a cap (of some diameter—Later, we will consider diameters much smaller than K^{-1}).

If A is the rectangle achieving the supremum in (26), then by Cauchy-Schwartz we have

$$\left\| \max_{\substack{A \\ A',A''}} \Big| \sum_{\substack{\alpha \\ y_\alpha \in A'}} c_\alpha(x) \Big|^{1/2} \Big| \sum_{\substack{\alpha \\ y_\alpha \in A'}} c_\alpha(x) \Big|^{1/2} \right\|_{L^4(B(a,K))} \leq C(K) \Big(\sum_{\substack{\alpha \\ y_\alpha \in A}} |c_\alpha|^2 \Big)^{1/2},$$

so for $x \in B(a, K) \subset \mathbb{R}^3$, we can write

$$\max_{\substack{A \\ A',A''}} \Big| \sum_{\substack{\alpha \\ y_\alpha \in A'}} c_\alpha(x) \Big|^{1/2} \Big| \sum_{\substack{\alpha \\ y_\alpha \in A'}} c_\alpha(x) \Big|^{1/2} = \phi_S \sum_{\substack{\alpha \\ y_\alpha \in A}} |c_\alpha|^2, \tag{28}$$

where in this case $\tau = S$ is the "cap" of radius ~ 1, but in future, we will consider functions ϕ_τ where τ is a smaller camp. Note as well that $\phi = \phi(x)$ is a function of x. However, we shall make the following simplifying assumption:

Simplifying Assumption 6. We shall pretend that if ϕ_τ is the function satisfying the analogue of (28) on the cap τ, then

$$\phi_\tau = \mathrm{diam}(\tau)^{-1/12}. \tag{29}$$

Thus in (28), $\phi_\tau = 1^{1/12} = 1$. In reality, ϕ_τ could be any function which has suitable L^4 averages on certain sets, and in order to deal with the contribution from ϕ_τ we need to prove several estimates and interpolate between them to remove the dependence on ϕ_τ. However, we will gloss over these details; the function ϕ_τ from (29) captures the worst-case behavior of the "actual" ϕ_τ.

Using (28), and letting A be the rectangle that achieves the supremum in (26), (24) becomes

$$\begin{aligned} |Tf| \leq C & \max_{\substack{\beta_1,\beta_2,\beta_3 \\ \text{“transverse''}}} (|c_{\beta_1}| \cdot |c_{\beta_2}| \cdot |c_{\beta_3}|)^{1/3} + \phi_S \sum_{\substack{\alpha \\ y_\alpha \in A}} |c_\alpha|^2. \\ & + \max_a |T(f|_{B(a,K^{-1})})| \qquad (30) \\ & = (31) \ + \ (32) \ + \ (33). \end{aligned}$$

Now, as in the previous section, one of the three terms of (31)–(33) will be dominant.

Simplifying Assumption 7. For all $x \in B_R$, the same term is dominant. Furthermore, this is term is either (31) or (32).

The first part of this simplification is fairly harmless. The second part merely shortens our proof; if the third term is dominant then we use arguments similar to those if the second term is dominant.

If (31) dominates then we can apply the multilinear theory from BCT in a fashion similar to Section 13.3.1. Now, lets suppose (32) is dominant. Instead of using an induction hypothesis (that a suitable estimate holds at smaller scales) to control the second term, we shall apply the decomposition from (24) to each of the terms $\{c_\alpha\}$ appearing in (32). Recall that each term c_α is obtained by applying to f the restriction operator associated with the surface patch Ω_α. If A is the rectangle A that achieves the supremum in (32), then A contains $\sim K$ caps $\{\Omega_\alpha\}$, each of radius K^{-2}. Now, consider in turn each cap Ω_α lying in A. Write $\Omega_\alpha = \bigcup \Omega_{\alpha_i}$ as a finitely overlapping union of disks of radius K^{-2}. We can apply the decomposition from (30) on this cap to obtain

$$\begin{aligned} |c_\alpha| \lesssim & \max_{\substack{i_1,i_2,i_3 \\ \text{“transverse}'}} (|T_{\Omega_{\alpha_{i_1}}} f| \cdot |T_{\Omega_{\alpha_{i_2}}} f| \cdot |T_{\Omega_{\alpha_{i_3}}} f|)^{1/3} \\ & + \phi_{\Omega_\alpha} \max_{A_1 \subset \Omega_\alpha} \Big(\sum_{\substack{i \\ \Omega_{\alpha_i} \in A_1}} |T_{\Omega_{\alpha_i}} f|^2 \Big)^{1/2} + \max_{a \in \Omega_\alpha} |T(f|_{B(a,K^{-2})})|, \end{aligned} \tag{34}$$

where we are performing the same decomposition as in (24), except the entire decomposition is performed inside Ω_α (so for the time being, we have $\tau = \Omega_\alpha$), and the decomposition is at scale K^{-2}. Thus $T_{\Omega_{\alpha_i}}$ is the restriction operator associated to the cap Ω_{α_i}, and A_1 is a rectangle of dimensions $K^{-1} \times K^{-2}$ (which therefore contains $\sim K^{-1}$ caps of diameter K^{-2}). By Assumption 6, $\phi_{\Omega_\alpha} = K^{-1/12}$.

Now, if we apply the decomposition (34) to every term in the sum from

(32), we get

$$
\begin{aligned}
(32) \lesssim \sum_{\alpha} \Bigg(& \max_{\substack{i_1,i_2,i_3 \\ \text{“transverse}'}} (|T_{\Omega_{\alpha_{i_1}}} f| \cdot |T_{\Omega_{\alpha_{i_2}}} f| \cdot |T_{\Omega_{\alpha_{i_3}}} f|)^{2/3} + \max_{A_1 \subset \Omega_\alpha} \sum_{\substack{i \\ \Omega_{\alpha_i} \subset A_1}} |T|_{\Omega_{\alpha_i}} f|^2 \\
& + \max_{a \in \Omega_\alpha} |T(f|_{B(a,K^{-2})})| \Bigg)^{1/2} \qquad (35) \\
= {} & (36) \ + \ (37) \ + \ (38).
\end{aligned}
$$

Let us suppose again that for each α, either (36) or (37) is dominant. If (36) is dominant, then keep it. Note that since there are $\sim K$ caps of diameter K^{-1} in A, there are $O(K)$ terms of the form (36). If (37) is dominant, then apply the same decomposition to each of the $T|_{\Omega_i} f$ appearing in (37). Note that there are $O(K^2)$ such terms (for all possible choices α and i). We keep repeating this process: After decomposing at scale $K^{-\ell}$, keep all terms for which the transverse interactions dominate, and further decompose those for which coplanar interactions dominate (we are continuing to pretend that collinear interactions never dominate). We shall do this until either there are no more terms of the form (37) (i.e. every dominant contribution comes from transverse interactions) or until we have reached length scale $K^{-\ell} = R^{-1/2}$, and we shall then keep any remaining terms.

Note that there are $O(\rho^{-1})$ terms at length scale ρ, for $1 \leq \rho \leq R^{-1/2}$. There are only $\log_K(R^{-1/2})$ different length scales, so by pigeonholing, one of these scales must dominate the contribution to $|Tf(x)|$. Call this length scale δ. Thus we have:

$$
\begin{aligned}
|Tf| \lesssim & \max_{1 > \delta > R^{-1/2}} \delta^{-1/12} \max_{\mathcal{E}_\delta} \Big[\sum_{\tau \in \mathcal{E}_\delta} (|Tf_{\tau_1}| \cdot |Tf_{\tau_2}| \cdot |Tf_{\tau_3}|)^{2/3} \Big]^{1/2} \\
& + \max_{\mathcal{E}_{R^{-1/2}}} \Big[\sum_{\tau \in \mathcal{E}} |Tf_\tau|^2 \Big]^{1/2} \\
= {} & (39) \ + \ (40),
\end{aligned}
$$

Where $\mathcal{E}_\delta$ is a collection of $O(\delta^{-1})$ disjoint δ caps. The $\delta^{-1/12}$ factor comes from (29), where in this case the quantity τ appearing in (29) is a cap of diameter δ.

In the previous section, we considered the case where we had three transverse caps which dominated the contribution to $|Tf(x)|$. Here we have many caps, and in place of transversality we have the condition that the caps are disjoint and their centers are distance $\geq \delta$ apart, and thus their normal vectors point in δ–separated directions.

Bourgain and Guth bound various L^p norms of (39) and (40). We shall gloss over many details and discuss only two of these bounds on (39). Similar arguments can be used to bound (40), but for the sake of brevity we shall not mention them here.

13.4.1 First Bound: Hölder's inequality

$$\|(39)\|_{L^3(B_R)} \lessapprox \delta^{-1/6}. \tag{41}$$

Note that the (full) restriction conjecture in $\mathbb{R}^3$ would imply that (41) $\lessapprox 1$. The bound (41) is proved using standard techniques (essentially Hölder's inequality plus the observation that $|\mathcal{E}_\delta| \lesssim \delta^{-1}$), and we shall not discuss it here.

13.4.2 Second bound: Wolff's $L^{5/2}$ Kakeya estimate

Equation (41) gives us a "bad" (i.e. large) bound on a "good" L^p norm of (39). In this section, we shall obtain a very "good" bound on a "bad" L^p norm of (39) (specifically $p = 10/3$). By interpolating between these two bounds, we can show that $\|(39)\|_{L^p(B_R)} \lessapprox 1$ for some $3 < p < \frac{10}{3}$; it turns out that $p = 3.3$ is the magic value.

On each cap τ (which has diameter $\sim \delta$), let $\mathring{\tau} \subset \mathbb{R}^3$ be the tube which is dual to the cap $\iota(\tau)$, and is located on the "support" of $T_\tau f$ (of course $T_\tau f$ is likely supported on all of $\mathbb{R}^3$, but we can choose the dual cap so that "most" of $T_\tau f$ is located on the dual cap). $\mathring{\tau}$ has dimensions $\delta^{-1} \times \delta^{-1} \times \delta^{-2}$, it points in the normal direction of the cap $\iota(\tau)$, and its center is an arbitrary point of $\mathbb{R}^3$ (i.e. we have no control over where the tube is located). We have that

$$|T_{\tau_{i_1}} f(x)|^{2/3} |T_{\tau_{i_2}} f(x)|^{2/3} |T_{\tau_{i_3}} f(x)|^{2/3} \sim \delta^2 \text{ for } x \in \mathring{\tau} \tag{42}$$

(the δ^2 factor is just due to normalization considerations), and the RHS of (42) decays off of $\mathring{\tau}$. Thus we are justified in making the following simplification:

Simplifying Assumption 8.

$$|T_{\tau_{i_1}} f(x)|^{2/3} |T_{\tau_{i_2}} f(x)|^{2/3} |T_{\tau_{i_3}} f|^{2/3} = \delta^2 \chi_{\mathring{\tau}}(x). \tag{43}$$

Now, $\{\overset{\circ}{\tau}\colon \tau \in \mathcal{E}_\delta\}$ is a collection of $O(\delta^{-1})$ "tubes" of dimension $\delta^{-1} \times \delta^{-1} \times \delta^{-2}$ which point in δ–separated directions. Thus we may apply Wolff's $L^{5/2}$ bound on the Kakeya maximal function to conclude

$$\Big\| \sum_{\tau \in \mathcal{E}_\delta} \chi_{\overset{\circ}{\tau}} \Big\|_{L^{5/2}(B_R)} \lessapprox \delta^{-19/5}. \tag{44}$$

Remark 6. *The "usual" formulation of Wolff's bound is for tubes of radius $\sim \delta$ and length ~ 1. Since our tubes have radius $\sim \delta^{-2}$ and length $\sim \delta^{-1}$, (44) has a different exponent of δ from the usual bound.*

From (44) we obtain

$$\begin{aligned}
&\|(39)\|_{L^{10/3}(B_R)} \\
&\lessapprox \delta^{-1/12} \Big(\int_{B_R} \Big(\Big[\sum_{\tau \in \mathcal{E}_\delta} (|Tf_{\tau_1}| \cdot |Tf_{\tau_2}| \cdot |Tf_{\tau_3}|)^{2/3} \Big]^{1/2} \Big)^{10/3} \Big)^{3/10} \\
&= \delta^{-1/12} \delta^2 \Big\| \sum_{\tau \in \mathcal{E}_\delta} \chi_{\overset{\circ}{\tau}} \Big\|^{1/2}_{L^{5/2}(B_R)} \\
&\lessapprox \delta^{-1/12} \delta^2 \delta^{-19/10} \\
&\lesssim \delta^{-1/60}.
\end{aligned} \tag{45}$$

If we interpolate (41) and (45), we obtain

$$\|(39)\|_{L^{3.3}(B_R)} \lessapprox 1, \tag{46}$$

which is the desired bound.

References

[1] J. Bennett, T. Carbery, T. Tao. On the multilinear restriction and Kakeya conjectures. *Acta Math.* 196:261–302. 2006.

[2] J. Bourgain, L. Guth. Bounds on oscillatory integral operators based on multilinear estimates. arXiv:1012.3760v3. 2011.

[3] A. Cordoba. Geometric Fourier analysis. *Ann. Inst. Fourier.* 32(3):215–226. 1982.

[4] E. Stein. Some problems in harmonic analysis. *Harmonic analysis in Euclidean spaces*. Proc. Sympos. Pure Math. Part 1, 3–20. 1978.

[5] T. Tao. The Bochner-Riesz conjecture implies the restriction conjecture. *Duke Math. J.* 96:363–376. 1999.

[6] T. Tao. A sharp bilinear restriction estimate for paraboloids. *GAFA*. 13:1359–1384. 2003.

[7] T. Tao. Recent progress on the restriction conjecture. To appear in *Park City Pro.* 2003.

[8] F. Temur. An Improved Restriction Estimate for $\mathbb{R}^6$. arXiv:1108.4017v1. 2011.

[9] T. Wolff, An improved bound for Kakeya type maximal functions. *Rev. Mat. Iberoamericana* 11(3):651–674. 1995.

Joshua Zahl, UCLA
email: **jzahl@math.ucla.edu**