Restriction and Kakeya Phenomena

These are notes from a Cambridge Part III course I gave in Michaelmas 2002. The notes have sat undisturbed for over 10 years, and so cannot be expected to take any account of modern developments. For example, Dvir has completely solved the finite field Kakeya conjecture with a very short argument, rendering one or two of the sections somewhat pointless. Furthermore, I hadn't completed my PhD when I gave this course so certain things which appeared mysterious to me then might seem less so now (and vice versa). Perhaps one day I will bring the notes up to date. However, I am informed that the notes have been used now and then over the years. I am making them available now for the first time in a single file.

Ben Green, Cambridge, February 2013.

Contents

1. Besicovitch Sets

In this set of notes we are going to prove the following celebrated result of Besicovitch.

Theorem 1 (Besicovitch)**.** *There is a closed and bounded subset of the plane which has measure zero, yet contains a unit line segment in every direction.*

We will deduce this result by a applying a limiting argument to a discrete analogue of it. We say that a triangle $T \subseteq \mathbb{R}^2$ is *decent* if its base lies on the x-axis and it has height 1. Let T be the decent triangle bounded by points $(0,0), (1,0)$ and $(0,1)$. For an integer N consider a subdivision of T into triangles T_i bounded by $(i/N, 0), ((i+1)/N, 0)$ and $(0,1)$, $i = 0, 1, \ldots, N-1$.

Proposition 1.1 (Discrete Besicovitch)**.** *Let $\epsilon > 0$, and let V be an open set containing T. Then there exists $N = N(\epsilon)$ and real numbers $e_1, \ldots, e_N$ such that*

(i) *The union of the translated triangles $T_i + e_i$ has area at most ϵ and*
(ii) $\bigcup (T_i + e_i) \subseteq V$.

Remarks. There is nothing special about the particular triangle T; by an affine transformation, a similar result holds for any decent triangle.

Now let $N = 2^k$, where k is to be chosen later. Let $\delta > 0$ be a real number, also to be chosen later. We have a subdivision of T into triangles $T_1, \ldots, T_{2^k}$, and we wish to translate the triangles T_i so as to make the total area of the translated copies smaller than ϵ. Let us begin by moving T_2 a distance $\delta 2^{-k}$ to the left. The translated copy of T_2 overlaps with T_1, and this creates a figure that resembles the one below.

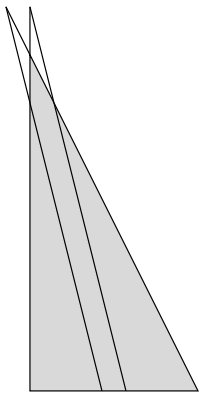

Figure 1. overlapping triangles

The shaded triangle is similar to $T_1 \cup T_2$ but is smaller by a scale factor $(1 - \delta)$. The unshaded part of the figure, B, is called a "bowtie". We need an estimate for its area.

Lemma 1 (Bowtie Lemma)**.** *The area of B is $2\delta^2 |T_1 \cup T_2|$.*

Proof. We may assume, by subjecting the whole figure to an affine transformation, that $T_1 \cup T_2$ is an *isoceles* right triangle with sidelengths 1 as shown in Figure 2. Elementary trigonometry confirms that $|CD| = |BC| = \delta$. Thus

$$\text{Area}(ABC) \;=\; \text{Area}(ADB) - \text{Area}(ACD) \;=\; \tfrac{1}{2}\delta^2.$$

Now by comparing angles and the sides BC, CD we see that triangle ACB is congruent

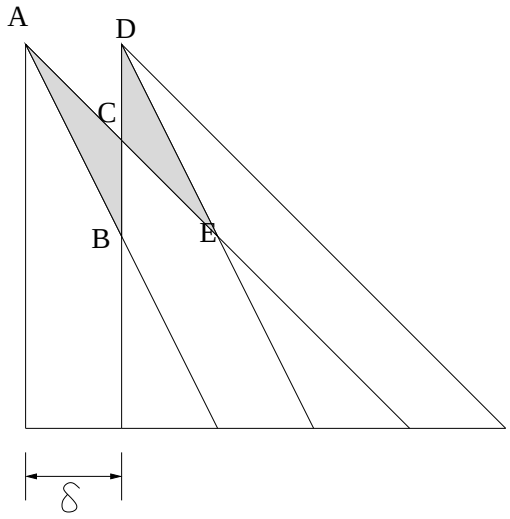

FIGURE 2. The bowtie lemma

to ECD. Thus the area of B is exactly δ^2, which is $2\delta^2$ times the area of $T_1 \cup T_2$. □

Remark. In fact, all we need is that $|B|$ is at most a constant multiple of δ^2. We will not always be so careful as we were in proving this lemma.

We perform analagous translations on the adjacent pairs of triangles (T_3, T_4), (T_5, T_6) and so on as illustrated in Figure 3. The areas of the resulting bowties are all at most $2\delta^2|T_{2i+1} \cup T_{2i+2}|$, because we can always apply an affine transformation so that $T_{2i+1} \cup T_{2i+2}$ becomes an isoceles right triangle.
One could now shift all these pieces along so that the union of the shaded areas is a shaded triangle which is similar to T (that is, isoceles and right-angled) but smaller by a factor $(1-\delta)$. See Figure 3. The union of all the unshaded parts is a union of bowties, and hence has area no more than $2\delta^2|T| = \delta^2$.
The 2^{k-1} shaded areas in Figure 2, then, form a subdivision of an isoceles right-triangle into 2^{k-1} smaller triangles which is very similar to the partition $T = \bigcup T_i$ that we started with. We can therefore iterate our construction, making pairs of adjacent shaded areas overlap as in Figure 4. The bowties from the previous stage get moved around automatically, and there are now some new bowties like the one shaded dark in Figure 4, and some new triangles like the one shaded light grey. The total area of the new

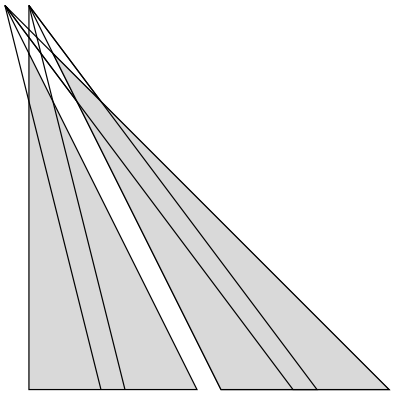

FIGURE 3. translating adjacent pairs ($k = 2$)

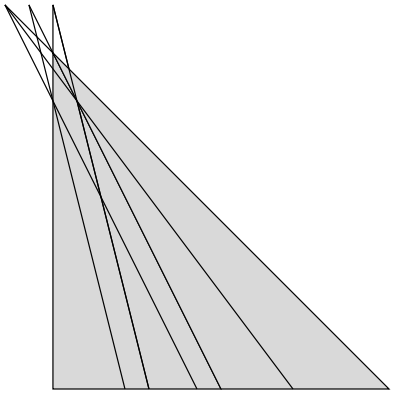

FIGURE 4. a new shaded triangle

triangles is $(1-\delta)^4|T|$, and they fit together as before to form an isoceles right triangle similar to T but smaller by a factor $(1-\delta)^2$. The total area of the new bowties is no more than $\delta^2(1-\delta)^2$.

The construction has now got to the stage where drawing pictures is difficult, but it is clear that we can iterate it k times. This will leave a single shaded triangle of area $\frac{1}{2}(1-\delta)^{2k}$ plus a union of bowties from each stage having area no more than

$$\delta^2\left(1+(1-\delta)^2+\cdots+(1-\delta)^{2k-2}\right) \leqslant \delta.$$

Thus we have a union of translates of the original triangles T_i with total area no more than

$$\tfrac{1}{2}(1-\delta)^{2k}+\delta \leqslant \tfrac{1}{2}e^{-2k\delta}+\delta.$$

Setting $\delta = \log k/k$ we see that this is at most $2\log k/k$ for $k \geqslant 3$. By choosing k sufficiently large we can certainly ensure that this is no more than ϵ, thereby satisfying condition (i) of Proposition B.1. How do we satisfy condition (ii)? Well, since T is compact there is some neighbourhood $N_\eta(T)$ which lies entirely in V. So we can begin

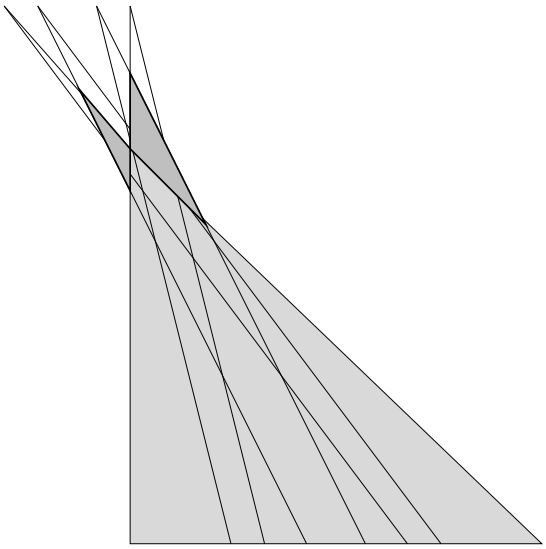

FIGURE 5. the second iteration

by splitting T into triangles with height 1 and base at most η, and then applying the above construction to each of these thin triangles in turn. Note that, by construction, no point in T then gets translated by more than η and so all the points of our union $\bigcup(T_i + e_i)$ lie in V. This completes the proof of Proposition B.1. □

Proof of Theorem 11. A rather obvious observation concerning Proposition B.1 is that the set $\bigcup(T_i + e_i)$ contains a unit line segment in every direction d_θ making an angle $\theta \in [0, \pi/4]$ with the y-axis. Such sets can, therefore, have arbitrarily small area. We will now produce a limiting argument which shows that they can have zero area. Putting 8 rotated copies of such a set next to one another gives a set of the form found by Besicovitch.

Start with the triangle T, and let V_0 be an open set containing T with $|V_0| = 1$. Construct a new set T_1 using Proposition B.1, so that T_1 is a union of N_1 decent triangles, $|T_1| \leqslant 1/4$ and $T_1 \subseteq V_0$. Let V_1 be an open set containing T_1, such that $\overline{V_1} \subseteq V_0$ and $|V_1| = 1/2$. For each of the triangles comprising T_1 apply Proposition B.1 to get a new union of decent triangles which lies in V_1 and has area at most $1/8N_1$. Putting all N_1 of these together gives a set T_2 with $|T_2| \leqslant 1/8$, $T_2 \subseteq V_1$ and T_2 a union of N_2 decent triangles. By construction, of course, T_2 contains a line segment for every direction with $\theta \in [0, \pi/4]$.

Continuing in this vein gives a nested sequence

$$V_0 \supset V_1 \supset V_2 \supset \dots$$

of open sets such that $|\overline{V_i}| \to 0$. Furthermore each V_i contains T_i, and so has a line segment in every direction $d_\theta \in [0, \pi/4]$. Set $F = \bigcap_{i=1}^{\infty} \overline{V_i}$. Then F is closed and bounded and has measure zero. We claim that F contains a line segment in every direction d_θ, $\theta \in [0, \pi/4]$. Indeed, for fixed θ the set V_i contains a unit line segment $x_i + [0,1]d_\theta$. The sequence $\{x_i\}_{i=1}^{\infty}$ is bounded and hence has a convergent subsequence $x_i' \to x$. Suppose that $x \notin F$. Since F is closed, this means that there is some ball $B = N_\delta(x)$ whose closure is disjoint from F. Since $x_i' \to x$, B meets each V_i. The sequence $\overline{B \cap V_i}$ consists of nested, non-empty closed sets whose intersection is therefore non-empty. But this intersection lies in $\bigcap \overline{V_i}$ and hence in F, a contradiction. A similar argument shows that, in fact, the whole line segment $x + [0,1]d_\theta$ lies in F, and this completes the proof of Theorem 11. □

2. The Kakeya Problem I

We begin with a word on notation. In this set of notes the parameter δ will be a real number smaller than 1, and n will be a fixed positive integer, regarded as the dimension of the problem we are considering. We use the symbols $\ll$ and $\gg$ in what might seem a rather cavalier manner. A statement such as

$$|A| \ll \delta^{n-1} \tag{2.1}$$

means that there is some constant c such that

$$|A| \leqslant c\delta^{n-1} \tag{2.2}$$

for all $\delta \in [0, 1)$. The constant c might depend on n, but we do not indicate such dependence explicitly. An even more convenient notation, which seems even more confusing at first sight, is something like

$$|A| \ll \delta^{n-1-\epsilon}. \tag{2.3}$$

The letter ϵ will generally be reserved for such an expression, which means that for any choice of $\epsilon > 0$ there is a constant c_ϵ (which might also depend on n) such that

$$|A| \leqslant c_\epsilon \delta^{n-1-\epsilon}.$$

In words, this means that $|A|$ grows more slowly than any power of δ with exponent smaller than $n-1$ (remember that, as $\delta < 1$, δ^3 is bigger than δ^5, and so on).

Minkowski dimension. Let $E \subseteq \mathbb{R}^n$, and for $\delta > 0$ let $N_\delta(E)$ be the δ-neighbourhood of E. How does the volume $|N_\delta(E)|$ vary as $\delta \to 0$? If E is a line then $|N_\delta(E)|$ is comparable to δ^{n-1}, whilst if E is the unit ball in $\mathbb{R}^n$ then $|N_\delta(E)|$ is of the order of 1. The line and the unit ball differ in that one would expect the line to have "dimension" 1

whereas the ball should have "dimension" n. These considerations lead to the following precise definition of dimension.

Definition 2.1. Let $E \subseteq \mathbb{R}^n$. Define the *lower Minkowski dimension* $\underline{d}(E)$ by

$$\underline{d}(E) \;=\; \inf\left\{d : \liminf_{\delta\to 0} |N_\delta(E)|\delta^{d-n} \;=\; 0\right\},$$

and the *upper Minkowski dimension* $\overline{d}(E)$ by

$$\overline{d}(E) \;=\; \inf\left\{d : \limsup_{\delta\to 0} |N_\delta(E)|\delta^{d-n} \;=\; 0\right\}.$$

These definitions take some unravelling. Observe that if $d > \underline{d}(E)$ then there is a sequence of δ_i's tending to 0 for which $|N_{\delta_i}(E)|$ is eventually smaller than any constant multiple of δ_i^{n-d}. If $d > \overline{d}(E)$ then $|N_{\delta_i}(E)|$ is smaller than $C\delta^{n-d}$ for all sufficiently small δ and any fixed C.

In this course we will not use these notions of dimension a huge amount. The reader eager to know more may consult Mattila's *Geometry of sets and measures in Euclidean spaces*, CUP.

Example. Define C, the *Cantor middle thirds set*, to be the set of all real numbers in $[0,1]$ whose base 3 expansion consists entirely of 0s and 1s. We will show that $\underline{d}(C) = \overline{d}(C) = \frac{\log 2}{\log 3}$. Take some $\delta > 0$, and consider the neighbourhood $N_\delta(C)$. Let $k = \lfloor \log_3(1/\delta) \rfloor$, and set $\eta = 3^{-k}$. It is not hard to give an upper bound for $|N_\eta(C)|$, because if $x \in N_\eta(C)$ then x is at distance at most 2.3^{-k} from some $y = 0.a_1a_2\ldots a_k$, where $a_i \in \{0,1\}$. Thus $|N_\eta(C)| \leqslant 4(2/3)^k$, and

$$|N_\delta(C)| \;\leqslant\; |N_\eta(C)| \;\leqslant\; 4\left(\frac{2}{3}\right)^{\log_3(1/\delta)-1} \;\leqslant\; 6.\delta^{1-\frac{\log 2}{\log 3}}. \tag{2.4}$$

To get a lower bound, set $\kappa = 3^{-k-1}$ and consider $N_\kappa(C)$. This certainly contains an interval of length κ about each point of the form $z = 0.a_1a_2\ldots a_{k+1}$, and for different choices of z these intervals are disjoint. Thus

$$|N_\delta(C)| \;\geqslant\; |N_\kappa(C)| \;\geqslant\; (2/3)^{k+1} \;\geqslant\; \tfrac{2}{3}\cdot\delta^{1-\frac{\log 2}{\log 3}}. \tag{2.5}$$

It is clear from inequalities (5.3) and (2.5), together with the definitions of upper and lower Minkowski dimension, that indeed $\underline{d}(C) = \overline{d}(C) = \frac{\log 2}{\log 3}$.

We now state the Kakeya problem.

Problem 2.2. What is $d(n)$, the infimum over all Besicovitch sets $B \subseteq \mathbb{R}^n$ of $\overline{d}(B)$?

It is conjectured that $d(n) = n$, which is as large as it possibly could be. This is known as the Kakeya conjecture; it is true (and we will prove it) for $n = 2$. For higher dimensions only partial results have been established. We will prove some of those too.

Conjecture 2.3 (Kakeya conjecture)**.** *If $B \subseteq \mathbb{R}^n$ is Besicovitch then $\overline{d}(B) = n$.*

Kakeya in 2 dimensions. In this section we will show that $d(2) = 2$. Roughly speaking, the main peculiarity of two dimensions that makes this possible is the fact that just about any pair of lines in the plane intersect. Sadly, such a principle fails rather dramatically in higher dimensions.

Theorem 2. *$d(2) = 2$, that is all Besicovitch sets in the plane $\mathbb{R}^2$ have Minkowski dimension 2.*

Proof. Let $\delta \in (0, 1]$, and let $k = \lfloor 1/\delta \rfloor$. Let $B \subseteq \mathbb{R}^2$ be a Besicovitch set, and consider the neighbourhood $N_\delta(B)$. We wish to prove an estimate of the form $|N_\delta(B)| \gg \delta^\epsilon$. Observe that for each $i = 1, \ldots, k$ the neighbourhood $N_\delta(B)$ contains a $\delta \times 1$ rectangle whose long axis makes an angle $\pi i/2k$ with the positive x-axis. Let these rectangles be $R_1, \ldots, R_k$; we will show that

$$\left| \bigcup_{i=1}^{k} R_i \right| \gg \delta^\epsilon, \tag{2.6}$$

which clearly suffices to prove Theorem 11. Write χ_i for the characteristic function of R_i, and let $A = \bigcup R_i$. We use Cauchy-Schwarz, which gives

$$\begin{aligned} \left(\int (\chi_1 + \cdots + \chi_k)(x)\, dx \right)^2 &\leqslant |A| \int (\chi_1 + \cdots + \chi_k)(x)^2\, dx \\ &= |A| \sum_{i,j} |R_i \cap R_j|. \end{aligned} \tag{2.7}$$

The left hand side is simply $k^2|R_1|^2$, which is comparable to 1. To estimate the right-hand side of (2.7) we need to know something about how rectangles intersect. This depends on the angle between R_i and R_j. The angle between R_i and R_j is $\theta = |i-j|\pi/2k$ and their intersection is contained within a rhombus, as shown. Each side of this rhombus has length $\delta/\sin\theta$, so its area is precisely $\delta^2/\sin\theta$. This is at most $2\delta^2/\theta$, because one has the inequality $\sin t \geqslant 2t/\pi$ when $t \in [0, \pi/2]$. Since $k \leqslant 1/\delta$, this is at most $2\delta/|i-j|$. Thus for fixed i we have

$$\begin{aligned} \sum_j |R_i \cap R_j| &\leqslant \delta + 2 \sum_{l=1}^{k} \frac{2\delta}{l} \\ &\ll \delta \log\left(\frac{1}{\delta}\right). \end{aligned}$$

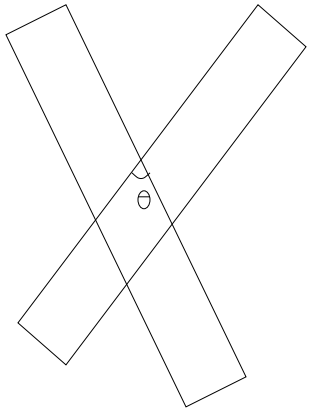

FIGURE 6. intersecting rectangles

Summing over the k values of i and substituting into (2.7) gives the inequality

$$|A| \gg \left(\log\left(\frac{1}{\delta}\right)\right)^{-1}.$$

This confirms the Kakeya conjecture in two dimensions (and in fact it gives a rather strong lower bound on $|N_\delta(B)|$). □

The finite field Kakeya problem. The finite field Kakeya problem, which we shall state presently, may be regarded as a toy version of the Euclidean problem. Although it lacks some of the features of the Euclidean problem it retains the most important aspect - trying to understand the incidence and intersection properties of possibly skew lines.

Let p be a prime and consider the vector space $\mathbb{F}_p^n$. We define a *line* to be a set of the form $\{x_0 + tx : t = 0, 1, \ldots, p-1\}$. The direction of the line is x, and it is uniquely defined up to projective equivalence (that is, up to multiplication by non-zero elements of $\mathbb{F}_p$). A Besicovitch set in $\mathbb{F}_p^n$ is simply a set which contains a line in every direction.

Problem 2.4 (Finite field Kakeya)**.** What is the minimum cardinality of a Besicovitch set in $\mathbb{F}_p^n$?

In this problem we think (of course) of n as a fixed dimension and let p become large. To that end define $d_F(n)$ to be the infimum of all d for which there is a constant $C = C(d)$ and, for all primes p, a Besicovitch subset of $\mathbb{F}_p^n$ with cardinality no more than Cp^d.

Conjecture 2.5 (Finite field Kakeya conjecture)**.** *We have* $d_F(n) = n$.

Although the finite field Kakeya problem is, in many ways, much easier to think about than the Euclidean version, there are some important respects in which it is different.

By modifying the proof of Theorem 11, for example, it is possible to show that there is no proper analogue of the Besicovitch construction we saw in the previous set of notes.

Theorem 3. *Any Besicovitch subset of $\mathbb{F}_p^2$ has cardinality at least $\frac{1}{2}p(p+1)$ (and so, a fortiori, $d_F(2) = 2$).*

Proof. Consider a collection of $p+1$ lines $L_1, \ldots, L_{p+1}$ in $\mathbb{F}_p^2$, one in each direction. Let $A = \bigcup L_i$, and suppose for a contradiction that A is small. In what follows we write χ_i for the characteristic function of the line L_i; that is, $\chi_i(x) = 1$ if $x \in L_i$ and 0 otherwise. Using the Cauchy-Schwarz inequality gives

$$\begin{aligned} p^2(p+1)^2 &= \left(\sum_x (\chi_1 + \cdots + \chi_{p+1})(x) \right)^2 \\ &\leqslant |A| \sum_x (\chi_1 + \cdots + \chi_{p+1})(x)^2 \\ &= |A| \sum_{i,j} |L_i \cap L_j| \\ &= 2|A|p(p+1), \end{aligned}$$

where we have used the fact that $|L_i \cap L_j| = 1$ unless $i = j$, when it equals p. The result follows immediately. □

For completeness, we give a construction which shows that the constant $\frac{1}{2}$ in this theorem is tight. It looks rather different to the "shifting triangles" construction from the previous set of notes!

Theorem 4. *Suppose that $p > 2$. Then there is a Besicovitch subset $B \subseteq \mathbb{F}_p^2$ of cardinality at most $p(p+3)/2$.*

Proof. Consider the set of pairs

$$S = \left\{ (x,t) \in \mathbb{F}_p^2 : x + t^2 \text{ is a square in } \mathbb{F}_p \right\}.$$

For a fixed choice of t there are $(p+1)/2$ choices for x, and so $|S| = p(p+1)/2$. Furthermore S contains the line $(a^2, 0) + \lambda(2a, 1)$ for any a, which gives a line in every direction except the direction $(1, 0)$. Set $B = S \cup \{(\lambda, 0) : \lambda \in \mathbb{F}_p\}$. □

3. The Kakeya Problem II

In the last set of notes we acquired a decent understanding of the Kakeya problem in dimension 2, both in Euclidean space and over finite fields. Our knowledge of the problems in dimensions 3 and higher is considerably less complete, and this set of notes is devoted to proving some partial results. We begin with

Theorem 5. $d(n) \geqslant (n+1)/2$.

Proof. Let $B \subseteq \mathbb{R}^n$ be a Besicovitch set with lower Minkowski dimension d. Although B is (by definition) closed and bounded there is nothing to stop it being fairly spread out. This turns out to be rather inconvenient.

Lemma 2. *Let Γ be the "north cap" of the sphere S^{n-1}, which we define to be the set of vectors γ with $\langle \gamma, e_n \rangle \geqslant 3/4$ (where $e_n = (0, 0, \ldots, 1)$). Then there is a set $B' \subseteq \mathbb{R}^n$ with Minkowski dimension at most d such that $B' \subseteq \mathbb{R}^{n-1} \times [0, 1]$, and for every $\gamma \in \Gamma$ there is a line segment $l_\gamma \in B'$ in direction γ and meeting both hyperplanes $x_n = 0$ and $x_n = 1$.*

Proof. Chop B into slices

$$B_i \;=\; B \cap \left(\mathbb{R}^{n-1} \times [i/4, (i+1)/4) \right),$$

$i \in \mathbb{Z}$, and translate these slices so that they lie on top of one another to form a set B^*. If l_γ is a line in B with direction $\gamma \in \Gamma$ then it certainly intersects at least two of the hyperplanes $x_n = i/4$. (A projection of) this statement is illustrated in the figure. Thus B^* contains a line segment in every direction $\gamma \in \Gamma$ meeting both the hyperplanes

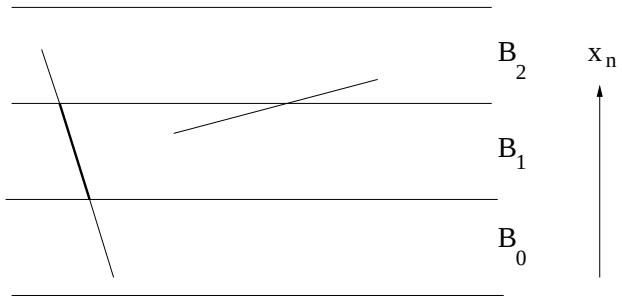

FIGURE 7. a preliminary slicing-up

$x_n = 0$ and $x_n = 1/4$ (and thus having length at least $1/4$). Let B' be the set obtained by applying a scale factor of 4 to B^*. This set has the properties required by the lemma (I am not interested in proving that $\underline{d}(B') \leqslant \underline{d}(B)$ rigorously, but you may care to think about it). □

From now on we drop the dash: that is, we pretend that B had the property of Lemma 29 from the outset. Let $\delta > 0$, and consider the neighbourhood $N_\delta(B) \cap (\mathbb{R}^{n-1} \times [0, 1])$. This contains a *skew tube* (an object whose intersection with every hyperplane $\{x_n = \lambda\}$ is a δ-ball) in every direction $\gamma \in \Gamma$.

In two dimensions, it was very convenient to pass to a subset of directions. On the circle we looked at the explicit directions $\pi i/2k$, $i = 1, \ldots, k$, where $k = \lfloor 1/\delta \rfloor$, but there is no obvious explicit set of directions in Γ.

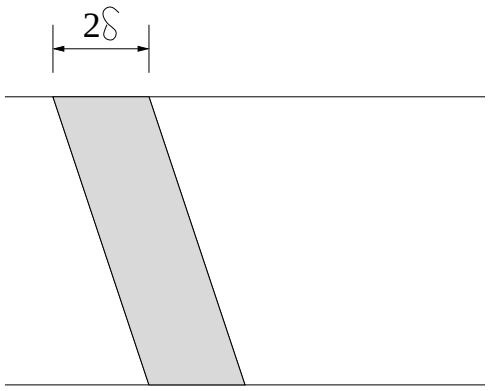

FIGURE 8. a skew tube

We say that a (finite) set $\Omega \subseteq S^{n-1}$ is η-separated if $|\omega - \omega'| \geqslant \eta$ whenever ω, ω' are distinct elements of Ω.

Lemma 3. *There is a 200δ-separated subset of Γ with cardinality $\gg \delta^{1-n}$.*

Proof. Let Ω be a 200δ-separated subset of Γ which is maximal with respect to inclusion. Then the balls $N_{400\delta}(\omega)$, $\omega \in \Omega$, cover the whole of Γ, as if they did not a point in their complement could be added to Ω. Now the area of $N_{400\delta}(\omega)$ is proportional to δ^{n-1}, whereas the area of Γ is a constant depending only on n. It follows that $|\Omega| \gg \delta^{1-n}$. □

Pick such a set Ω, and let A be the union of the skew tubes $T_\omega \in N_\delta(B)$ in directions $\omega \in \Omega$. To prove Theorem 11, it suffices to prove a lower bound $|A| \geqslant C_\epsilon \delta^{(n-1)/2+\epsilon}$.

Now for any $\lambda \in [0,1]$ we may consider the *slice* $S_\lambda = \{\mathbf{x} \in A | x_n = \lambda\}$. If $|A|$ is small then most of the slices must be small too. Indeed if Λ is the set of fat slices, that is slices with $|S_\lambda| \geqslant 100|A|$, then $|\Lambda| \leqslant 1/100$. For if not then we would have

$$\begin{aligned} |A| &= \int_0^1 |S_\lambda| \, d\lambda \\ &\geqslant 100|\Lambda||A| \\ &> |A|, \end{aligned}$$

a contradiction. Because there are so few fat slices there must be two thin slices S_{λ_1} and S_{λ_2} with $|\lambda_1 - \lambda_2| \geqslant 1/2$.

Our next job is to combinatorialise our slices. To do this, set $\kappa = \delta/\sqrt{n}$ and define D_λ to be set of all lattice vectors $\mathbf{v} \in \mathbb{Z}^{n-1}$ for which $(\kappa\mathbf{v}, \lambda) \in S_\lambda$. To relate the size of D_λ to that of S_λ, we need a lemma concerning balls.

Lemma 4. *Let $\{x_i\}_{i\in I}$ be an arbitrary collection of points in $\mathbb{R}^n$, let $\kappa > 0$ and let $t > 1$. Then we have*

$$\left|\bigcup_{i\in I} B(x_i, t\kappa)\right| \leqslant (t+2)^n \left|\bigcup_{i\in I} B(x_i, \kappa)\right|.$$

Proof. This lemma would, of course, be trivial (with $t+2$ replaced by t) if the balls $B(x_i, \kappa)$ did not overlap. To this end let $J \subseteq I$ be maximal so that the balls $B(x_j, \kappa)$, $j \in J$, are disjoint. Then the balls $B(x_j, 2\kappa)$ must between them contain all the x_i, as otherwise the set J would not be maximal. Thus

$$\bigcup_{i\in I} B(x_i, t\kappa) \subseteq \bigcup_{j\in J} B(x_j, (t+2)\kappa),$$

leading to the chain of inequalities

$$\begin{aligned}\left|\bigcup_{i\in I} B(x_i, t\kappa)\right| &\leqslant \left|\bigcup_{j\in J} B(x_j, (t+2)\kappa)\right| \\ &\leqslant (t+2)^n \left|\bigcup_{j\in J} B(x_j, \kappa)\right| \\ &\leqslant (t+2)^n \left|\bigcup_{i\in I} B(x_i, \kappa)\right|,\end{aligned}$$

as required. □

Lemma 5. *For any λ we have $|D_\lambda| \leqslant (4n)^n \delta^{1-n} |S_\lambda|$.*

Proof. S_λ, being a union of δ-balls, is a union $\bigcup_{i\in I} B(x_i, \kappa)$ of (uncountably many) κ-balls. The set of κ-cubes with lower corners in D_λ, which has volume $\kappa^{n-1}|D_\lambda|$, is contained in $\bigcup_{i\in I} B(x_i, (1+\sqrt{n-1})\kappa)$. The result now follows from the previous lemma and a short calculation. □

Now to each tube T_ω we may associate a pair $P_\omega = (v_\omega, v'_\omega) \in D_{\lambda_1} \times D_{\lambda_2}$ such that $v_\omega \in T_\omega \cap S_{\lambda_1}$ and $v'_\omega \in T_\omega \cap S_{\lambda_2}$. Indeed by definition the intersection of T_ω with S_λ is a δ-ball, and such a ball must contain a point of the lattice $\kappa\mathbb{Z}^{n-1}$.

(*This is where I erred in the lecture. It is* not true *that any δ-ball in $\mathbb{R}^m$ intersects the lattice $\delta\mathbb{Z}^m$. This is true in dimesnions* 2,3 *and* 4 *but not in dimensions* 5 *and higher. That is what confused me!*)

Lemma 6. *The vectors $v_\omega - v'_\omega$, $\omega \in \Omega$, are all distinct.*

Proof. The reason this is true is that the vector $v_\omega - v'_\omega$ is "roughly parallel" to the direction ω of T_ω, but ω and ω' are far apart for different ω, ω' (in fact, 200δ-separated).

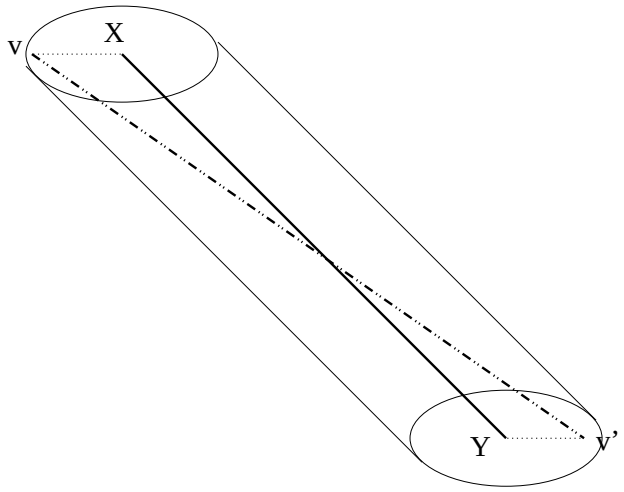

FIGURE 9. 10δ-separated tubes

Actually proving this statement is rather painful, but we have to do it. The picture shows a tube T intersected with the two hyperplanes $H_1 : \{x_n = \lambda_1\}$ and $H_2 : \{x_n = \lambda_2\}$. The line XY is the axis of symmetry of the tube, so XY is in direcion ω. A vector $v - v'$ of the relevant form is also illustrated. Now $v \in B(X, \delta)$ and $v' \in B(Y, \delta)$, so that $|Xv|$ and $|Yv'|$ are at most δ. We may therefore, by translating so that $X = v$, draw the following figure in the plane. Write $|XY| = a$, $|vv'| = a + \epsilon_1$ and $|Yv'| = \epsilon_2$. Thus

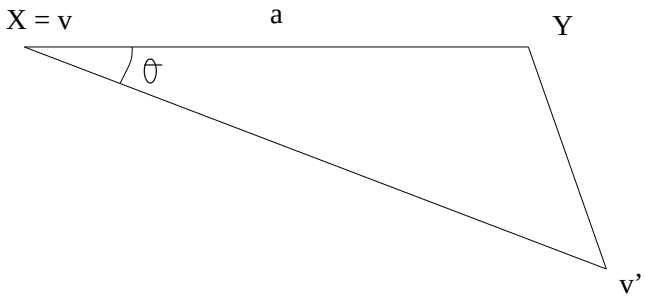

FIGURE 10. Euclidean geometry

$\epsilon_2 \leqslant 2\delta$ and, since we chose λ_1, λ_2 so that $|\lambda_1 - \lambda_2| \geqslant 1/2$, we have $a \geqslant 1/2$.

The cosine rule applies to show that

$$\begin{aligned} \cos\theta &= \frac{a^2 + (a+\epsilon_1)^2 - \epsilon_2^2}{2a(a+\epsilon_1)} \\ &\geqslant 1 - \frac{\epsilon_2^2}{2a(a+\epsilon_1)} \\ &\geqslant 1 - 100\delta^2 \end{aligned}$$

provided that δ is smaller than some absolute consant (which we can assume it is).

Now by considering Taylor series expansions, or by elementary calculus, one can check that $\cos t \leqslant 1 - \frac{1}{4}t^2$ for t sufficiently small. Thus

$$1 - \tfrac{1}{4}\theta^2 \geqslant 1 - 100\delta^2,$$

which implies that $\theta \leqslant 20\delta$. We have shown, then, that $v - v'$ makes an angle at most 20δ with ω. Lemma 6 follows immediately from the fact that the different directions ω are 200δ-separated. □

We may now complete the argument. We chose the slices S_{λ_1} and S_{λ_2} to be thin, that is $|S_{\lambda_1}|$ and $|S_{\lambda_2}|$ are both at most $100|A|$. It follows from Lemma 5 that both $|D_{\lambda_1}|$ and $|D_{\lambda_2}|$ are $\ll \delta^{1-n}|A|$. Lemma 6, however, implies that $|D_{\lambda_1}||D_{\lambda_2}| \geqslant |\Omega|$. Recalling that $|\Omega| \gg \delta^{1-n}$, a short calculation confirms that $|A| \gg \delta^{(n-1)/2}$. By our earlier remarks, this is enough to imply Theorem 11. □

As you can see, technical issues get in the way of even the simplest arguments concerning the Euclidean Kakeya problem. For that reason in all of our subsequent arguments pertaining to the Kakeya problem we think only about the finite field problem. The truly conscientious reader can go through the arguments adapting them to the Euclidean setting, using the argument of Theorem 11 as a model. In what follows we suppose that p is a sufficiently large prime (where "sufficiently large" might depend on the dimension n).

Slice-free arguments. An exercise on the first example sheet asks you to write out the argument of Theorem 11 in the finite field case. You will see that it boils down to something very simple. Here is another very simple way of seeing that $d_F(n) \geqslant (n+1)/2$.

Theorem 6. *Besicovitch sets in $\mathbb{F}_p^n$ have cardinality at least $\frac{1}{4}p^{(n+1)/2}$.*

Proof. Suppose not, and let $A \subseteq \mathbb{F}_p^n$ be a set containing a line in each of the $(p^n-1)/p-1 \geqslant p^{n-1}/2$ possible directions. The number of point-line pairs (p, l) with $p \in l$ is at least $p^n/2$, and so some point x must lie on $L \geqslant p^{(n-1)/2}/2$ lines. These lines are disjoint away from x, and so the union of this point and the L lines through it has cardinality $1 + L(p-1)$, which is at least $p^{(n+1)/2}/4$. □

Wolff's hairbrush argument. The argument of Theorem 6 was incredibly crude, and what is more it fails to recover what we already know in the case $n = 2$. In this section we give an argument, due to Tom Wolff, which proves $d_F(n) \geqslant (n+2)/2$. Amazingly the case $n = 3$ of this result, namely that $d_F(3) \geqslant 5/2$, is the best result currently known in 3 dimensions, at least for the finite field version of the Kakeya problem.

Theorem 7 (Wolff)**.** *Besicovitch sets in $\mathbb{F}_p^n$ have cardinality at least $\frac{1}{8}p^{(n+2)/2}$.*

We start with a lemma, which is a kind of 2-dimensional finite field Kakeya problem in which there need not be very many lines.

Lemma 7. *Suppose that a set $A \subseteq \mathbb{F}_p^2$ is a union of l lines $L_1, \ldots, L_l$, all pointing in different directions. Then $|A| \geqslant pl/2$.*

Proof. We adapt the proof of Theorem 7 from the previous notes. Write $\chi_1, \ldots, \chi_l$ for the characteristic functions of $L_1, \ldots, L_l$. Then we have

$$\begin{aligned} p^2 l^2 &= \left(\sum_x (\chi_1 + \cdots + \chi_l)(x) \right)^2 \\ &\leqslant |A| \sum_x (\chi_1 + \cdots + \chi_l)(x)^2 \\ &= |A| \sum_{i,j} |L_i \cap L_j| \\ &= |A| \left(lp + l(l-1) \right) \\ &\leqslant 2lp|A|. \end{aligned}$$

The lemma follows immediately. □

Now let $A \subseteq \mathbb{F}_p^n$ be Besicovitch, being a union of lines $L_1, \ldots, L_k$ with $p^{n-1} \leqslant k \leqslant 2p^{n-1}$, and suppose that $|A| \leqslant p^{(n+2)/2}$. We can use a Cauchy-Schwarz type argument to prove that some line intersects at least $\frac{1}{4}p^{n/2}$ others. Indeed, we have

$$\begin{aligned} p^2 k^2 &= \left(\sum_x (\chi_1 + \cdots + \chi_k)(x) \right)^2 \\ &\leqslant |A| \sum_x (\chi_1 + \cdots + \chi_k)(x)^2 \\ &= |A| \sum_{i,j} |L_i \cap L_j|, \end{aligned}$$

so that

$$\sum_{i,j} |L_i \cap L_j| \ \geqslant \ p^{(3n-2)/2}.$$

It follows that for some i, say $i = k$, we indeed have

$$\sum_{j \neq i} |L_i \cap L_j| \ \geqslant \ \tfrac{1}{2}p^{n/2} - p \ \geqslant \ \tfrac{1}{4}p^{n/2}.$$

We now restrict attention to the line L_k and the lines $L_1, \ldots, L_m$ intersecting it, where $m \geqslant \frac{1}{4}p^{n/2}$. We call this collection of lines a *hairbrush* H. Each of these lines lies in a unique 2-plane containing L_k. Let $\Pi_1, \ldots, \Pi_t$ be the complete collection of 2-planes containing L_k and some other L_j, and suppose that Π_i contains $n_i \geqslant 1$ of the lines

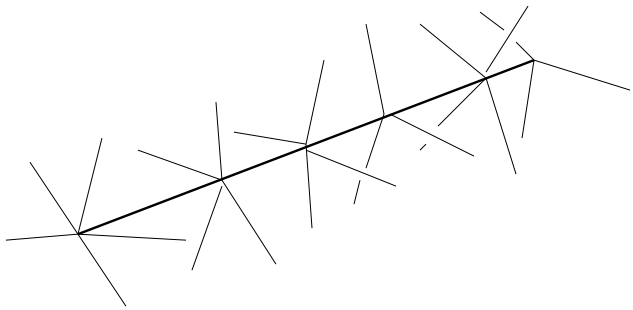

FIGURE 11. a rather schematic depiction of a hairbrush

$L_1, \ldots, L_m$. Since Π_i also contains L_k we have, by Lemma 37, that

$$|\Pi_i \cap H| \geqslant \tfrac{1}{2}(n_i+1)p \geqslant \left(\tfrac{1}{4}n_i + 1\right)p$$

provided that $n_i \geqslant 2$. This inequality clearly also holds when $n_i = 1$. In particular, Π_i contains at least $\frac{1}{4}n_i p$ points of H which do not lie on L_k, and these collections of points are disjoint as i varies. Thus we have the estimate

$$\begin{aligned} |A| &\geqslant |H| \\ &\geqslant \tfrac{1}{4}p\sum_{i=1}^{t} n_i \\ &= \tfrac{1}{4}pm \\ &\geqslant \tfrac{1}{8}p^{(n+2)/2}. \end{aligned}$$

This completes the proof of Theorem 7. □

Adapting the above argument to the Euclidean case is quite a challenge, and Wolff's original paper contained a number of important technical innovations which we do not have time to describe in this course.

More elaborate slicing arguments. Recall that in the previous set of notes we proved that $d_F(n) \geqslant (n+1)/2$ using what we called a slicing argument. By adapting this argument we can get a better bound. Such techniques were introduced to the study of the Kakeya problem by Bourgain, and elaborated by Katz and Tao.

Suppose that $(Z, +)$ is an abelian group and that $G \subseteq Z \times Z$. If r is an integer we write $\pi_r(G)$ for the set of all $a + rb$, where $(a, b) \in G$. We also write $\pi_\infty(G)$ for the set $\{b : (a, b) \in G\}$. The notation is supposed to be suggestive of the fact that the π_i are *projections*. The next proposition shows that if there is a small Besicovitch set in $\mathbb{F}_p^n$ then we can find a group Z and a set G with very strange projection properties.

Proposition 3.1. *Suppose that $A \subseteq \mathbb{F}_p^n$ is a Besicovich set. Then there is an abelian group $(Z,+)$ and a subset $G \subseteq Z \times Z$ so that $|\pi_{-1}(G)| \geqslant \frac{1}{2}p^{n-1}$, but $|\pi_0(G)|$,$|\pi_1(G)|$,$|\pi_2(G)|$ and $|\pi_\infty(G)|$ are all at most $8|A|/p$.*

Proof. Let us look at slices $A_i = A \cap \{x_n = i\}$, $i = 0, 1, \ldots, p-1$. At most $p/8$ of these slices can have cardinality bigger than $8|A|/p$. Of the other "thin" slices, of which there are at least $7p/8$, we can find some seven consecutive ones $A_i, A_{i+1}, \ldots, A_{i+6}$. Now let $Z = \mathbb{F}_p^{n-1}$, and define subsets $B_0, \ldots, B_6 \subseteq Z$ by

$$B_0 \ = \ \{b : (b,i) \in A_i\},$$

$$B_1 \ = \ \{b' : (b', i+1) \in A_{i+1}\}$$

and so on. Now the Besicovitch set A contains lines $L_1, \ldots, L_k$ in different directions, where $k \geqslant p^{n-1}$. At least $\frac{1}{2}p^{n-1}$ of these lines intersect both A_i and A_{i+6}, and they do so in pairs of points (x, x'). Each such pair gives rise to a pair in $B_0 \times B_6$, and we call the set of all such pairs G. Clearly G is a subset of $Z \times Z$, and both $|\pi_0(G)|$ and $|\pi_\infty(G)|$ are at most $8|A|/p$. Furthermore for any pair (x, x') the midpoint $(x+x')/2$ lies on the same line as x and x', and also has $x_n = i+3$. Therefore it lies in the slice A_{i+3}, and this means that $\pi_1(G)$ lies in B_3. Hence $|\pi_1(G)| \leqslant 8|A|/p$ too. Similarly $(x+2x')/3$ lies in A_{i+4}, so that $\pi_2(G) \subseteq B_4$ and $|\pi_2(G)| \leqslant 8|A|/p$. However (as all the lines of A point in different directions) the differences $x - x'$ are all different, and so all of the projections of G in the -1 direction are distinct. Thus $|\pi_{-1}(G)| = |G| \geqslant \frac{1}{2}p^{n-1}$. □

It is all very well finding a set with "strange projection properties", but we have to actually do something with it.

Proposition 3.2. *Let $(Z,+)$ be an abelian group and let $G \subseteq Z \times Z$. Then we have an inequality*

$$|\pi_{-1}(G)| \ \leqslant \ \max_{r \in \{0,1,2,\infty\}} |\pi_r(G)|^{7/4}.$$

Proof. We may clearly suppose that π_{-1} is 1-1 on G, so that $\pi_{-1}(G) = |G|$. Suppose that $\max_{r\in\{0,1,2,\infty\}} |\pi_r(G)| = N$. Let Q be the set of all *suitable quadrilaterals* (g_1, g_2, g_3, g_4), where $g_1, g_2, g_3, g_4 \in G$ and $\pi_0(g_1) = \pi_0(g_2)$, $\pi_0(g_3) = \pi_0(g_4)$, $\pi_\infty(g_1) = \pi_\infty(g_3)$ and $\pi_2(g_2) = \pi_4(g_4)$. A typical such quadrilateral Q is depicted in Figure 2. We're going to give, first of all, a lower bound on the number of such quadrilaterals. First of all we count *vertical line segments* (g_1, g_2) with $\pi_0(g_1) = \pi_0(g_2)$. For each point $x \in \pi_0(G)$ suppose that there are $n(x)$ points $g \in G$ above x, that is with $\pi_0(g) = x$. Then the number of vertical line segments is just $\sum_x n(x)^2$ which, by the Cauchy-Schwarz inequality, is at least $|G|^2/N$.

Now let $m(x,y)$ be the number of vertical line segments (g_1, g_2) with $\pi_\infty(g_1) = x$ and $\pi_2(g_2) = y$. By what we have just shown, $\sum_{x,y} m(x,y) \geqslant |G|^2/N$. The number

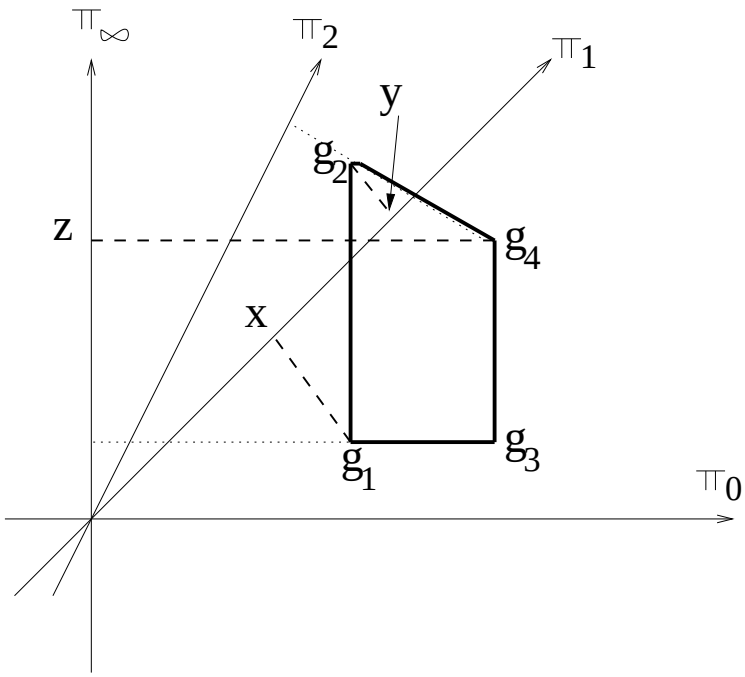

FIGURE 12. a quadrilateral

of suitable quadrilaterals is precisely $\sum_{x,y} m(x,y)^2$ which, by another application of Cauchy-Schwarz, is at least $|G|^4/N^4$.

This gives a lower bound for the number of suitable quadrilaterals. To get an upper bound we show that Q is completely determined by the triple (x,y,z), where $x = \pi_1(g_1)$, $y = \pi_1(g_2)$ and $z = \pi_\infty(g_4)$. Since there are only N^3 such triples the required bound $|G|^4 \leqslant N^7$ will follow immediately. The proof rests on the identity

$$\pi_{-1}(g_3) \;=\; -\pi_1(g_1) + 2\pi_1(g_2) - 2\pi_\infty(g_4).$$

To check this, write $g_i = (a_i, b_i)$ and simplify using the relations $a_1 = a_2$, $a_3 = a_4$, $b_1 = b_3$ and $a_2 + 2b_2 = a_4 + 2b_4$. Thus x, y and z determine $\pi_{-1}(g_3)$ and hence g_3, and therefore a_4 and b_1. This leaves a_1, a_2, b_2, b_4 undetermined, but of these we know that $a_1 = a_2$ plus the three linear relations given by the values of x, y, z. It is a simple matter of linear algebra to check that this 4×4 system is invertible and that all the g_i can be recovered. □

Theorem 8. *$d_F(n) \geqslant (4n+3)/7$.*

Proof. This is just a matter of substituting Proposition 3.1 into Proposition 3.2 and seeing what comes out after doing the algebra. In fact, every Besicovitch subset of $\mathbb{F}_p^n$ has cardinality at least $\frac{1}{16}p^{(4n+3)/7}$. □

The arithmetic Kakeya conjecture. Let us reconsider Proposition 3.2, and in

particular how it might be improved. Suppose that for some $\epsilon > 0$ there are numbers $r_1, \dots, r_k \in \mathbb{Q}_{\geqslant 0} \cup \{\infty\}$ (called a *slice set*) such that we have a bound

$$|\pi_{-1}(G)| \leqslant \max_{i \in \{1,\dots,k\}} |\pi_{r_i}(G)|^{1+\epsilon}.$$

Then, by adapting the argument used to prove Proposition 3.1 in a straightforward way, we can show that

$$d_F(n) \geqslant (n+\epsilon)/(1+\epsilon).$$

We showed, in Proposition 3.2, that one can find a slice set for $\epsilon = 3/4$ (namely the set $\{0, 1, 2, \infty\}$). By more involved arguments of an iterative nature Katz and Tao have shown that there are slice sets for all $\epsilon > 0.67514$.

Conjecture 3.3 (Arithmetic Kakeya)**.** *There are slice sets for all $\epsilon > 0$.*

This conjecture implies the finite field Kakeya conjecture, and in fact the Minkowski dimension Kakeya conjecture in $\mathbb{R}^n$. It seems not to be implied by either of these though; that is, arithmetic Kakeya is stronger than Kakeya.

We do not intend to prove any more Kakeya bounds in this course. The reader interested in persuing the matter further should consult Tao's *Edinburgh lecture notes on the Kakeya problem*, available on his website, and the references therein.

4. The circle, I

One of the aims of this course is supposed to be to achieve an understanding of the Fourier transform, and of how it interacts with geometry. In this set of notes we take our first steps in this direction. It is clear that any understanding of the Fourier transform will be difficult without giving the definition, so we start with that.

A very short introduction to the Fourier transform on Euclidean spaces. We will be dealing with Fourier transforms in various different settings - in Euclidean space, in finite fields, in abelian groups, on the integers. It is well worth understanding (though not essential to the course) that the Fourier transform's natural habitat is a locally compact abelian group. Depending on how much explanation you would like of this statement, I refer you to the following:

(i) Tom Körner's Part III course on topological groups, to be given in the Michaelmas term;
(ii) Katznelson's classic (and inexpensive) text *Harmonic analysis*;
(iii) Some notes that I wrote entitled "Fourier analysis and the zeta function", available at
`http://www.dpmms.cam.ac.uk/~bjg23/papers/zeta.dvi`.
Rather tangential to the things we are discussing in this course.

Now if $f : \mathbb{R}^n \to \mathbb{C}$ is a sufficiently nice function[1] and if $\xi \in \mathbb{R}^n$ then we define the Fourier transform by

$$\widehat{f}(\xi) \;=\; \int_{\mathbb{R}^n} f(x) e^{-2\pi i \xi \cdot x}\, dx$$

Lemma 8 (Basic properties of the FT)**.** *Let $f, g : \mathbb{R}^n \to \mathbb{C}$. Then we have:*

(i) (*Plancherel*) $\|f\|_2 = \|\widehat{f}\|_2$.
(ii) (*Parseval*) $\int f(x)\overline{g(x)}\, dx \;=\; \int \widehat{f}(\xi)\overline{\widehat{g}(\xi)}\, d\xi$;
(iii) (*Inversion*) $f(x) \;=\; \widehat{f}^{\vee}(x)$, *where* $g^{\vee}(x) = \widehat{g}(-x)$;
(iv) (*Convolution*) *Define the convolution of two functions by*

$$(f * g)(x) \;=\; \int f(y)\overline{g(x-y)}\, dy.$$

Then

$$\widehat{f * g}(\xi) \;=\; \widehat{f}(\xi)\overline{\widehat{g}(-\xi)}.$$

(v) (*Scaling*) *Define f_a by $f_a(x) = f(ax)$. Then $\widehat{f_a}(\xi) = a^n \widehat{f}(\xi/a)$.*
(vi) (*Rotational symmetry*) *The Fourier transform of a radially symmetric function is radially symmetric.*

****Remarks****. We shall prove the analagous properties in the finite field case later on. Part (v) and (vi) are straightforward change-of-variables, and (iv) is easy to prove using Fubini's Theorem. To prove (i), I would show that it is true for the *Hermite polynomials* $H_n(x) = p_n(x)e^{-\pi|x|^2}$. These functions ($p_n$ is a polynomial) are eigenvectors of the Fourier transform and are dense in $L^2(\mathbb{R}^n)$. You may care to check that if $f(x) = e^{-\pi x^2}$ (in one dimension) then $\hat{f} = f$, so that (i) is trivial in this case. Property (ii) follows from (i) using a well-known technique called polarization: apply (i) to various linear combinations $\mu f + \nu g$ and perform a linear elimination to get an expression for $\int f\overline{g}$. For more details consult Rudin's *Real and Complex Analysis*.

Our main objective in this set of notes is to look at the Fourier transform of the circle $S^1 \subseteq \mathbb{R}^2$. The circle is endowed with a natural measure. Parametrize S^1 by $(\cos\theta, \sin\theta) : \theta \in [0, 2\pi)$, and consider the arc A from $(\cos\theta_1, \sin\theta_1)$ to $(\cos\theta_2, \sin\theta_2)$, where $0 < \theta_1 <$

[1] I do not wish to address measure-theoretic issues in this course, so I do not intend to say what *sufficiently nice* means. Furthermore I shall not even bother qualifying statements with phrases like "suppose that f and g are sufficiently nice". If this really bothers you, then I will be happy to give a non-examinable talk on measure theory and the Fourier transform. You should also bear in mind that we will be dealing exclusively with discrete phenomenawithin a few lectures.

$\theta_2 < 2\pi$. We define its measure[2] $\sigma(A)$ to be $\theta_2 - \theta_1$. The object of interest to us will be

$$\begin{aligned} \widehat{d\sigma}(\lambda) &= \int_{\mathbb{R}^2} e^{-2\pi i \lambda \cdot x}\, d\sigma(x) \\ &= \int_0^{2\pi} e^{-2\pi i(\lambda_1 \cos\theta + \lambda_2 \sin\theta)}\, d\theta, \end{aligned}$$

defined for any $\lambda \in \mathbb{R}^2$. This[3] is called the Fourier transform of the (surface) measure σ.

We will show that the fact that S^1 is curved has implications for $\widehat{d\sigma}$. To set this work in context, let's look briefly at the Fourier transform of the archetypal non-curved surface, the line $L = \{(x, 0) : 0 \leqslant x \leqslant 1\}$. This has an obvious measure ν, and we have

$$\widehat{d\nu}(\lambda) = \int_0^1 e^{-2\pi i x \lambda_1}\, dx.$$

When $\lambda = (0, \lambda_2)$, this is identically equal to 1. That is, the Fourier transform $\widehat{d\nu}(\lambda)$ does not *decay* as $\lambda \to \infty$. We will see that this is not true of the circle.

A closed form evaluation of $\widehat{d\sigma}(\lambda)$ is not possible, but we can get an asymptotic. This is an example of the *principle of stationary phase* which some of you may have met before. We begin with two lemmas.

Proposition 4.1 (Non-stationary phase)**.** *Let $a \in C_0^\infty(\mathbb{R})$. Then $\hat{a}(\xi) = O(|\xi|^{-N})$ as $\xi \to \infty$ for any positive integer N.*

Proof. We do this by repeated integration by parts. A single such integration, for example, gives

$$\hat{a}(\xi) = \frac{1}{2\pi i \xi} \int_{\mathbb{R}} e^{-2\pi i \xi x} a'(x)\, dx$$

and from this it is clear that $|\hat{a}(\xi)| \ll |\xi|^{-1}$. □

Why give this such a strange name? Well, the integral for $\hat{a}(\xi)$, $\int a(x) \exp(-2\pi i \xi x)\, dx$, contains a phase function (namely x) which is never stationary. This is what gives $\hat{a}(\xi)$ such good decay properties. By contrast one expects substantially less decay with a phase function that can be stationary, a heuristic that the following proposition makes precise.

Proposition 4.2 (Stationary phase lemma)**.** *Let $a \in C_0^\infty(\mathbb{R})$ and for $\lambda \in \mathbb{R}$ let*

$$K(\lambda) = \int_{\mathbb{R}} e^{i\pi\lambda x^2} a(x)\, dx.$$

[2]For the cognoscenti, this is in fact the measure *induced* from Lebesgue measure on $\mathbb{R}^2$.

[3]It is, basically, the classical Bessel function J_0; in fact, we have $\widehat{d\sigma}(|\lambda|) = J_0(-2\pi|\lambda|)$. This is one reason why Bessel functions are actually important in modern mathematics.

Let $\lambda \neq 0$ be a real number. Then

$$K(\lambda) \;=\; 2^{-1/2}|\lambda|^{-1/2}(1 \pm i)a(0) + O\left(\lambda^{-3/2}\right),$$

where we choose the plus sign if $\lambda > 0$ and the minus sign if $\lambda < 0$.

Proof. Let z be a positive real number. We begin by working out the Fourier transform of the function $g(x) = e^{-zx^2}$. We have

$$\begin{aligned}
\hat{g}(\xi) &= \int_{\mathbb{R}} \exp(-zx^2 - 2\pi i\xi x)\, dx \\
&= \exp(-\pi^2\xi^2/z) \int_{-\infty}^{\infty} \exp\left(-z(x + \pi i\xi/z)^2\right)\, dx \\
&= \exp(-\pi^2\xi^2/z) \int_{-\infty}^{\infty} \exp\left(-zx^2\right)\, dx, \qquad (4.1)
\end{aligned}$$

the latter step following by integrating e^{-zx^2} (which is an analytic function of x) around a rectangle with corners $\pm R$ and $\pm R + \pi i\xi/z$ and letting $R \to \infty$. The integral in (4.1) can, however, be easily evaluated as $(\pi/z)^{1/2}$ using the Gaussian integral $\int_{-\infty}^{\infty} \exp(-t^2)\, dt = \sqrt{\pi}$. Thus we have

$$\hat{g}(\xi) \;=\; (\pi/z)^{1/2} \exp(-\pi^2\xi^2/z).$$

An application of Parseval's formula yields

$$\int_{\mathbb{R}} e^{-zx^2} a(x)\, dx \;=\; (\pi/z)^{1/2} \int_{\mathbb{R}} e^{-\pi^2\xi^2/z} \hat{a}(\xi)\, d\xi. \qquad (4.2)$$

At the moment we have only proved this for $z \in (0, \infty)$, but we have not used the latter z in vain.

Lemma 9. *The integral $I_1(z) = \int e^{-zx^2} a(x)\, dx$ converges for all $z \in \mathbb{C}$, and in fact defines an analytic function.*

Proof. The convergence is obvious, because a is a compactly supported function. To prove that $I_1(z)$ is analytic there is really only one obvious thing to do: guess that $I_1'(z)$ is what you get by differentiating under the integral, and then prove this. Write $F(z, x) = e^{-zx^2} a(x)$ and use the inequality

$$\left| F(z + h, x) - F(z, x) - h\frac{\partial F}{\partial z}(z, x) \right| \;\leqslant\; |h|^2 \sup_{w \in B(z, |h|)} \left| \frac{\partial^2 F}{\partial z^2}(w, x) \right|. \qquad (4.3)$$

The second derivative is uniformly bounded by some constant $C_1(z, r)$ on any domain of the form $B(z, r) \times \mathbb{R}$, and so we have

$$\left| I_1(z + h) - I_1(z) - h \int \frac{\partial F}{\partial z}(z, x)\, dx \right| \;\leqslant\; C_1(z, 1)|h|^2$$

for all $|h| \leqslant 1$. This proves that $I_1(z)$ is indeed analytic at z with derivative $\int \frac{\partial F}{\partial z}(z, x)\, dx$. □

Now the right-hand side of (4.2) is more troublesome, particularly at zero. Nonetheless we can show

Lemma 10. *Let S be the set $\{z : \Re z \geqslant 0, z \neq 0\}$. Then the integral*

$$I_2(z) \;=\; (\pi/z)^{1/2} \int_{\mathbb{R}} \exp(-\pi^2\xi^2/z)\hat{a}(\xi)\, d\xi,$$

currently defined for $z \in (0,\infty)$, is in fact continuous on S and analytic on S^o, the interior of S. Here $z^{1/2}$ is that branch of the square root which sends $re^{i\theta}$ to $r^{1/2}e^{i\theta/2}$ when $\theta \in [-\pi/2, \pi/2]$.

Proof. Non-examinable (but see the appendix). □

It follows by the identity principle that (4.1) holds for all $z \in S$, and hence in particular for $z = -i\pi\lambda$, $\lambda \neq 0$. This gives the identity

$$\int_{\mathbb{R}} e^{i\pi\lambda x^2} a(x)\, dx \;=\; 2^{-1/2}|\lambda|^{-1/2}(1 \pm i) \int_{\mathbb{R}} \hat{a}(\xi) e^{i\pi\xi^2/\lambda}\, d\xi, \tag{4.4}$$

the $\pm$ sign depending on the sign of λ. Now observe that

$$\begin{aligned} \left| \int_{\mathbb{R}} \hat{a}(\xi) e^{i\pi\xi^2/\lambda}\, d\xi - \int_{\mathbb{R}} \hat{a}(\xi)\, d\xi \right| &\leqslant \int_{\mathbb{R}} |\hat{a}(\xi)| \left| 1 - e^{i\pi\xi^2/\lambda} \right| d\xi \\ &\leqslant \frac{\pi}{\lambda} \int_{\mathbb{R}} |\hat{a}(\xi)|\, |\xi^2|\, d\xi. \end{aligned} \tag{4.5}$$

This is at most $C\lambda^{-1}$ for some C, because $\hat{a}(\xi)$ is subject to the bounds $|\hat{a}(\xi)| \leqslant \|a\|_1$ and $|\hat{a}(\xi)| \ll |\xi|^{-4}$ (the latter bound follows from the principle of non-stationary phase applied with $N = 4$). Now by the inversion formula

$$\int_{\mathbb{R}} \hat{a}(\xi)\, d\xi \;=\; a(0),$$

and the proposition follows immediately from (10.7) and (10.8). □

Proposition 4.3. $\widehat{d\sigma}(\lambda) = \dfrac{2\cos 2\pi\,(|\lambda| - 1/8)}{|\lambda|^{1/2}} + O(|\lambda|^{-3/2})$.

Proof. Since $\widehat{d\sigma}$ is radially symmetric, it suffices to check this for $\lambda = (\lambda, 0)$, $\lambda > 0$. In this case we have

$$\widehat{d\sigma}(\lambda) \;=\; \int_0^{2\pi} e^{-2\pi i\lambda\cos\theta}\, d\theta. \tag{4.6}$$

Now the function $\cos\theta$ has zero derivative at $\theta = k\pi$, $k \in \mathbb{Z}$. That is to say, the integral in (4.6) has *stationary phase* at these points. To apply Proposition B.2 we must first isolate each of the stationary phase points by introducing a *partition of unity*. This consists of four C^∞ functions $\psi_1, \ldots, \psi_4$ such that $\overline{\mathrm{Supp}\psi_i} \subseteq (i\pi/2, (i+2)\pi/2)$ and $\psi_1 + \cdots + \psi_4 = 1$ (see Figure 3). The existence of such a partition is not something we wish to establish here - it is often done in differential geometry courses, for example

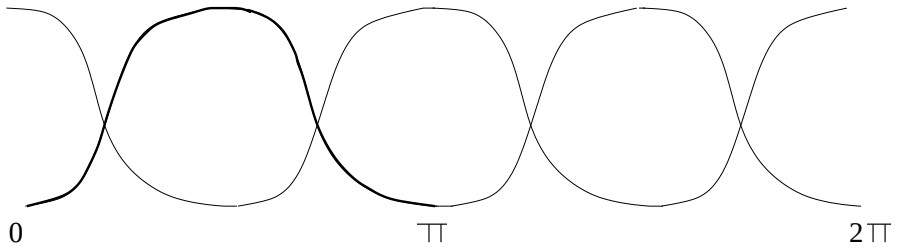

FIGURE 13. Partition of unity. ψ_4 is highlighted.

Madsen and Tornehave's excellent *From Calculus to Cohomology*, CUP.

Now we can evaluate $I_3 = \int \psi_3(\theta) \exp(-2\pi i\lambda \cos\theta)\, d\theta$ by making the substitution $\cos(\theta) = 1 - t^2$. This gives, using Proposition B.2,

$$\begin{aligned} I_3 &= \exp(-2\pi i\lambda) \int_{-1}^{1} \psi_3(t) \exp(2\pi i\lambda t^2) \frac{2dt}{\sqrt{2-t^2}} \\ &= \exp(-2\pi i\lambda)\lambda^{-1/2}(1+i) + O(\lambda^{-3/2}). \end{aligned}$$

Similarly

$$I_1 = \exp(2\pi i\lambda)\lambda^{-1/2}(1-i) + O(\lambda^{-3/2}).$$

Now on the supports of ψ_2 and ψ_4 the phase $\cos\theta$ is non-stationary, and we can apply Proposition B.1. Making the substitution $\cos\theta = t$, it is a simple matter to check that both $I_2(\lambda)$ and $I_4(\lambda)$ are $O(\lambda^{-2})$.

The proposition follows since $\widehat{d\sigma}(\lambda) = I_1(\lambda) + \cdots + I_4(\lambda)$. □

Corollary 4.4. *Let χ be the characteristic function of the unit ball $B(0,1)$ in $\mathbb{R}^2$. Then $|\chi(\lambda)| \ll |\lambda|^{-3/2}$.*

Proof. Writing in polar coordinates we have

$$\begin{aligned} \hat{\chi}(\lambda) &= \int_0^1 r\widehat{d\sigma}(r\lambda)\, dr \\ &= |\lambda|^{-1/2} \int_0^1 r^{1/2} \cos\left(2\pi(|\lambda|r - \tfrac{1}{8})\right)\, dr + O\left(|\lambda|^{-3/2}\right). \end{aligned}$$

To show that the integral here is $O(|\lambda|^{-1})$, write it as

$$\Re e^{-i\pi/4} \int_0^1 r^{1/2} e^{2\pi i|\lambda|r}\, dr$$

and integrate by parts once. □

Gauss circle problem. To conclude this set of notes we are going to use Corollary 4.4 to give an estimate for the number $n(R)$ of lattice points of $\mathbb{Z}^2$ which lie in the

ball $B(0,R)$. This is known as Gauss's circle problem. We begin with an elementary argument.

Proposition 4.5. *We have* $n(R) = \pi R^2 + O(R)$.

Proof. For any $(i_1, i_2) \in \mathbb{Z}^2$ let $S(i_1, i_2)$ be the lattice square with (i_1, i_2) as its bottom left corner. Write $U(R)$ for the set of all (i_1, i_2) for which $S(i_1, i_2)$ is completely contained in $B(0, R+\sqrt{2})$. Then $B(0,R) \cap \mathbb{Z}^2 \subseteq U(R)$, which leads to the bound

$$\left|B(0,R) \cap \mathbb{Z}^2\right| \leqslant |U(R)| \leqslant \pi(R+\sqrt{2})^2 \leqslant \pi R^2 + 10R$$

for large R. To get a bound in the other direction, let $V(R) = B(0,R) \cap \mathbb{Z}^2$. Then the set of squares $S(i_1, i_2)$, $(i_1, i_2) \in V(R)$, completely covers $B(0, R-\sqrt{2})$. Thus

$$\left|B(0,R) \cap \mathbb{Z}^2\right| = |V(R)| \geqslant \pi(R-\sqrt{2})^2 \geqslant \pi R^2 - 10R.$$

This completes the proof. □

If you think carefully about this argument you will realise that pretty much all we have used about the circle is the fact that its circumference grows linearly in R. Using the decay of the Fourier transform gives an improvement, the proof of which will be our main goal for the rest of this set of notes.

Theorem 9. *We have* $n(R) = \pi R^2 + O(R^{2/3})$.

To relate counting lattice points to the Fourier transform, we need a beautiful result called the Poisson Summation Formula. If you want to discover a little more about the natural setting for this result, and about its relation to the Riemann zeta function, I have a set of notes called *Fourier analysis and the ζ-function* on my webpage.

Theorem 10 (Poisson Summation)**.** *Let* $f \in C_0^\infty(\mathbb{R}^k)$ (*this is, in fact, a far more stringent condition on* f *than is necessary*)*. Then*

$$\sum_{n \in \mathbb{Z}^k} f(n) = \sum_{n \in \mathbb{Z}^k} \hat{f}(n). \tag{4.7}$$

Proof. Definitely non-examinable (but see the second appendix). □

Now let ψ be a C_0^∞ function whose support is in $B(0,1)$ and whose integral is equal to 1. For any $\epsilon > 0$ define $\psi_\epsilon(x) = \epsilon^{-2}\psi(x/\epsilon)$, which is then a smooth function supported on the ball $B(0,\epsilon)$ and having integral 1. Write χ_R for the cahracteristic function of the ball $B(0,R)$, and set $\phi(x) = \chi_{R+\epsilon} * \psi_\epsilon$. What on earth is this function? Well, it is a kind of "smoothed out" version of χ_R. In fact we have $\chi_R(x) \leqslant \phi(x)$ for all $x \in \mathbb{R}^2$, which allows us to write

$$n(R) \leqslant \sum_{n \in \mathbb{Z}^2} \phi(n). \tag{4.8}$$

The reason for introducing the smoothing is so that we can apply the Poisson summation formula, and so that certain Fourier transforms decay as we would like them to. You will see what I mean in the course of the proof, but I should add that smoothing in this manner (that is, convolving with a function having small support) is a vital technical tool in many harmonic analysis questions. It has analogues in a discrete setting in which it is not possible to make sense of the word *smooth* in the classical way.

Returning to (4.8), then, Poisson sum together with simple properties of the Fourier transform imply that

$$\begin{aligned} n(R) &\leqslant \sum_{\xi} \widehat{\phi}(\xi) \\ &= \sum_{\xi\in\mathbb{Z}^2} \widehat{\chi_{R+\epsilon}}(\xi)\widehat{\psi}_\epsilon(\xi) \\ &= (R+\epsilon)^2 \sum_{\xi\in\mathbb{Z}^2} \widehat{\chi}((R+\epsilon)\xi)\widehat{\psi}(\epsilon\xi) \\ &= \pi(R+\epsilon)^2 + E, \end{aligned} \tag{4.9}$$

where

$$|E| \;=\; (R+\epsilon)^2 \sum_{\lambda\in\mathbb{Z}^2\setminus\{0\}} \hat{\chi}((R+\epsilon)\xi)\psi(\epsilon\xi).$$

We'll estimate E using Corollary 4.4 together with the non-stationary phase estimate (cf. Proposition B.1)

$$\hat{\psi}(\xi) \;\ll\; \min\left(1, |\xi|^{-N}\right)$$

for any fixed positive integer N. We get

$$\begin{aligned} |E| &\ll R^2 \sum_{\xi\in\mathbb{Z}^2\setminus\{0\}} (R|\xi|)^{-3/2} \min\left(1, |\epsilon\xi|^{-N}\right) \\ &\leqslant R^{1/2} \sum_{0<|\xi|<\epsilon^{-1}} |\xi|^{-3/2} + R^{1/2}\epsilon^{-N} \sum_{|\xi|>\epsilon^{-1}} |\xi|^{-N-3/2} \\ &\ll R^{1/2}\epsilon^{-1/2} \end{aligned}$$

for any $N \geqslant 1$. Observe that if we had taken $N = 0$ (corresponding to no smoothing) then the sum over $|\xi| > \epsilon^{-1}$ would not converge. Returning to (4.9), we see that choosing $\epsilon = R^{-1/3}$ gives an upper bound $n(R) - \pi R^2 \ll R^{2/3}$. A very similar argument gives a corresponding lower bound and thus a proof of Theorem 9. □

You may like to think about a certain aspect of what is going on here, namely the appearance of the *uncertainty principle*. When one tries to localise too much in space (say by making ϵ too small in the above argument) one pays the price on the Fourier side.

It is conjectured that the $R^{2/3}$ in Theorem 9 can be replaced by $R^{1/2+\delta}$ for any $\delta > 0$. Hardy and Littlewood proved a result (called an Ω-result) showing that one cannot hope for more than this. That is, for arbitrarily large R we have $|n(R) - \pi R^2| \geqslant CR^{1/2}$ for every fixed constant C. This problem is one where numerous clever technical and conceptual improvements have led to rather modest advances. I believe that the best results currently known are due to Huxley, who can replace the $R^{2/3}$ by something like $R^{0.63}$.

Appendix: Proof of Lemma 10. Since the product of two analytic (continuous) functions is analytic (continuous) we may safely ignore the factor of $(\pi/z)^{1/2}$. Now observe that when $\Re z > 0$ (and $z \neq 0$) we have

$$\left| e^{-\pi^2\xi^2/z} \right| \leqslant 1,$$

so the integral defining $I_2(z)$ certainly converges if $\hat{a} \in L^1(\mathbb{R})$ (that is, if $\int |\hat{a}(\xi)|\, d\xi < \infty$). We in fact need the slightly stronger inequality

$$\int \max(|\xi|^4, 1)|\hat{a}(\xi)|\, d\xi < \infty.$$

To prove this note that $|\hat{a}(\xi)| \leqslant \|a\|_1$ for all ξ, and that for large $|\xi|$ Proposition B.1 gives the superior estimate $|\hat{a}(\xi)| \ll |\xi|^{-6}$. Thus

$$\int_{-\infty}^{\infty} \max(|\xi|^4, 1)|\hat{a}(\xi)|\, d\xi \ll \|a\|_1 + \int_{|\xi|>1} |\xi|^{-2}\, d\xi < \infty.$$

Now write $F(\xi, z) = \exp(-\pi^2\xi^2/z)\hat{a}(\xi)$, and use (4.3) again. Now $\partial^2 F/\partial z^2(\xi, z)$ is something like

$$\left(\frac{\pi^4\xi^4}{z^4} - \frac{2\pi^2\xi^2}{z^3} \right) e^{-\pi^2\xi^2/z},$$

and on any ball $B(z, r)$ contained in S° this is bounded by $C_2(z, r)\max(|\xi|^4, 1)$. Fix $z \in S^\circ$ and let δ be so small that $B(z, \delta) \subseteq S^\circ$. Then if $|h| < \delta$ we have

$$\begin{aligned} \left| \int F(\xi, z+h)\, d\xi - \int F(\xi, z)\, d\xi - h \int \frac{\partial F}{\partial z}(\xi, z)\, d\xi \right| &\leqslant |h|^2 C_2(z, \delta) \int \max(|\xi|^4, 1)|\hat{a}(\xi)|\, d\xi \\ &\ll C_3|h|^2. \end{aligned}$$

This proves that $\int F(\xi, z)\, d\xi$, and hence $I_2(z)$, is analytic on S°.

To prove that $I_2(z)$ is continuous, we can use the dominated convergence theorem. Fix w, $\Re w = 0$, and let $(z_n)_{n=1}^{\infty} \subseteq S$ be any sequence converging to w. Write $F_n(\xi) = e^{-\pi^2\xi^2/z_n}\hat{a}(\xi)$ and $F(\xi) = e^{-\pi^2\xi^2/w}\hat{a}(\xi)$. Clearly $F_n \to F$ pointwise. Furthermore $|F(\xi) - F_n(\xi)|$ is at most $2|\hat{a}(\xi)|$, which is an integrable function. Thus, by DCT, $\int F_n \to \int F$. $\square$

Appendix: Sketch proof of Poisson summation. We consider two auxilliary functions $F, G : \mathbb{T}^k \to \mathbb{C}$, where $\mathbb{T}^k$ is the torus $[0,1)^k$. These are defined by

$$F(\theta) \;=\; \sum_{n\in\mathbb{Z}^k} \hat{f}(n) e^{2\pi i n\cdot\theta}$$

and

$$G(\theta) \;=\; \sum_{n\in\mathbb{Z}^k} f(n+\theta).$$

We will show that F and G are equal, whereupon (4.7) will follow immediately on setting $\theta = 0$. To do this we look at the Fourier coefficients[4]

$$\widetilde{F}(m) = \int_{\mathbb{T}^k} F(\theta) e^{-2\pi i m\cdot\theta}\, d\theta, \qquad \widetilde{G}(m) = \int_{\mathbb{T}^k} G(\theta) e^{-2\pi i m\cdot\theta}\, d\theta,$$

defined for $m \in \mathbb{Z}^k$. Indeed we have

$$\begin{aligned}
\widetilde{F}(m) &= \int_{\mathbb{T}^k} \sum_{n\in\mathbb{Z}^k} \hat{f}(n) e^{2\pi i n\cdot\theta} e^{-2\pi i m\cdot\theta}\, d\theta \\
&= \sum_{n\in\mathbb{Z}^k} \hat{f}(n) \int_{\mathbb{T}^k} e^{2\pi i (n-m)\cdot\theta}\, d\theta \\
&= \hat{f}(m),
\end{aligned}$$

the interchange of integration and summation being valid because $\sum_{n\in\mathbb{Z}^k} |\hat{f}(n)| < \infty$ (you can prove this using Proposition B.1. Furthermore,

$$\begin{aligned}
\widetilde{G}(m) &= \int_{\mathbb{T}^k} \sum_{n\in\mathbb{Z}^k} f(n+\theta) e^{-2\pi i m\cdot\theta}\, d\theta \\
&= \int_{\mathbb{T}^k} \sum_{n\in\mathbb{Z}^k} f(n+\theta) e^{-2\pi i m\cdot(\theta+n)}\, d\theta \\
&= \sum_{n\in\mathbb{Z}^k} \int_{x\in n+[0,1)^k} f(x) e^{-2\pi i m\cdot x}\, dx \\
&= \hat{f}(m).
\end{aligned}$$

Now it seems reasonable that two continuous functions whose Fourier coefficients agree, such as F and G, must in fact be equal. This is a true and famous result, but it is not trivial to prove. See Theorem 2.7 of Chapter 1 of Katznelson's excellent book *Harmonic analysis* for a proof of this fact (there are numerous other references - another good source are the lecture notes from Tom Körner's 1999 Part III course on Fourier analysis, available from his webpage). □

[4]We use the term "coefficient" hestitantly. Really, $\hat{F}$ is just another type of Fourier transform, but it takes values in $\mathbb{Z}^k$ because that is the dual group of $\mathbb{T}^k$. I suggest going to Dr Körner's course to find out more!

5. The circle, II

The object of this set of notes is to use the information we obtained in *The circle I* concerning the Fourier transform of the unit circle to prove a simple example of what is known as the *restriction phenomenon*. We will use our result to give another proof that Kakeya sets in 2 dimensions have Minkowski dimension 2.

An $L^\infty \to L^4$ local restriction theorem. Let f be a locally measurable function on the unit circle. We are going to study the Fourier transform of $fd\sigma$, where σ is the uniform measure on the circle. This is defined by

$$\widehat{fd\sigma}(\lambda) \ = \ \int_{S^1} f(x) e^{-2\pi i x\cdot\lambda}\, d\sigma(x)$$

for any $\lambda \in \mathbb{R}^2$. When $f = 1$, this is simply the Fourier transform $\widehat{d\sigma}$ for which we derived an asymptotic expression. In particular we know that

$$|\widehat{d\sigma}(\lambda)| \ \ll \ \min(1, |\lambda|^{-1/2}).$$

This means that $\widehat{d\sigma}$ lies in $L^p(\mathbb{R}^2)$ for any $p > 4$. Thus it is "almost" in L^4, a principle that we can quantify by noting that

$$\|\widehat{d\sigma}\|_{L^4((B(0,R))} \ \ll \ (\log R)^{1/4}. \tag{5.1}$$

The point here, of course, is that the logarithm grows rather slowly and in particular more slowly than any power of R. It turns out that an estimate like (11.3) is valid for any bounded function f in place of 1. The following theorem to this effect is known as a local restriction theorem. We'll explain the terminology later on.

Theorem 11. *Let $f : S^1 \to \mathbb{C}$ be measurable. Then $\|\widehat{fd\sigma}\|_{L^4(B(0,R))} \ \ll \ (\log R)^{1/4}\|f\|_\infty$.*

Proof. The proof of this result, which I got from the thesis of Gerd Mockenhaupt, uses a rather nice positivity argument of a type important in harmonic analysis. Write χ for the characteristic function of the unit disc, and set $\psi = \chi * \chi$. Then ψ enjoys the properties described in the following lemma.

Lemma 11. *(i) $\hat{\psi}(\lambda) \geqslant 0$ for all $\lambda \in \mathbb{R}^2$;*
(ii) $\psi(x) \geqslant \chi(x)$ for all $x \in \mathbb{R}^2$;
(iii) $\psi(x) \leqslant \pi$ for all $x \in \mathbb{R}^2$;
(iv) Supp$(\psi) \subseteq B(0, 2)$.

Proof. (i) is immediate, because $\hat{\psi} = |\hat{\chi}|^2$. To see (ii), observe that $\chi(y)$ is precisely the area of the intersection of two unit discs at distance $|y|$ apart. Thus if $y \in \text{Supp}\chi$, so that $|y| \leqslant 1$, then $\psi(y) \geqslant \frac{2\pi}{3} - \frac{\sqrt{3}}{2} \geqslant 1$. (iii) and (iv) follow from the same observation.

□

Write $\psi_R(x) = \psi(x/R)$ and $\chi_R(x) = \chi(x/R)$. Now we have, using properties (i), (ii) and (iii),

$$\begin{aligned}
\|\widehat{fd\sigma}\|^4_{L^4(B(0,R))} &= \int_\lambda \chi_R(\lambda) \left| \int_{S^1} f(x) e^{-2\pi i x\cdot\lambda}\, d\sigma(x) \right|^4 d\lambda \\
&\leqslant \int_\lambda \psi_R(\lambda) \left| \int_{S^1} f(x) e^{-2\pi i x\cdot\lambda}\, d\sigma(x) \right|^4 d\lambda \\
&= \int_{S^1} \dots \int_{S^1} \left(\int_\lambda \psi_R(\lambda) e^{-2\pi i\lambda\cdot(x_1+x_2-x_3-x_4)}\, d\lambda \right) f(x_1) f(x_2) \overline{f(x_3) f(x_4)} \\
&\qquad\qquad \times d\sigma(x_1) d\sigma(x_2) d\sigma(x_3) d\sigma(x_4) \\
&= \int_{S^1} \dots \int_{S^1} \widehat{\psi_R}(x_1+x_2-x_3-x_4) f(x_1) f(x_2) \overline{f(x_3) f(x_4)} \\
&= \qquad\qquad \times d\sigma(x_1) d\sigma(x_2) d\sigma(x_3) d\sigma(x_4) \\
&\leqslant \|f\|^4_\infty \int_{S^1} \dots \int_{S^1} \widehat{\psi_R}(x_1+x_2-x_3-x_4) \prod_{i=1}^4 d\sigma(x_i) \\
&= \|f\|^4_\infty \int_\lambda \psi_R(\lambda) \left| \int_{S^1} e^{-2\pi i x\cdot\lambda}\, dx \right|^4 d\lambda \\
&\leqslant \pi \|f\|^4_\infty \int_\lambda \chi_{2R}(\lambda) \left| \int_{S^1} e^{-2\pi i x\cdot\lambda}\, dx \right|^4 d\lambda \\
&= \pi \|f\|^4_\infty \|\widehat{d\sigma}\|^4_{L^4(B(0,2R))}.
\end{aligned}$$

The theorem follows immediately from (11.3). Look how important it was to find the function ψ, whose Fourier transform is positive, and how simply the result follows once we have thought of this trick. □

Local restriction implies Kakeya. In this section we will show that Theorem 11 leads to another proof that Kakeya sets in $\mathbb{R}^2$ have Minkowski dimension 2.

Lemma 12 (Knapp Example)**.** *Let $\delta > 0$ be a real number, let $\theta \in [0, 2\pi]$, and let A_θ be an arc of the circle of length δ centred on the point $e^{i\theta}$. Write χ_θ for the characteristic function of A_θ. Then we have that $|\widehat{\chi_\theta}(\lambda)| \geqslant \delta/2$ for all $\lambda \in R_\theta$, where R_θ is a $\frac{1}{10}\delta^{-2} \times \frac{1}{10}\delta^{-1}$ rectangle with long axis in direction θ and centre the origin.*

Proof. We suppose that $\theta = 0$, the proof in other cases being virtually identical. Let e_1, e_2 be the standard basic vectors for $\mathbb{R}^2$, so that e_1 is the centre of the arc A.

The basic idea is as follows. Fix $\lambda \in R$. Then we have

$$\widehat{\chi_\theta}(\lambda) \;=\; \int_A e^{-2\pi i x\cdot\lambda}\, d\sigma(x) \;=\; e^{-2\pi i e_1\cdot\lambda} \int_A e^{-2\pi i (x-e_1)\cdot\lambda}\, d\sigma(x),$$

so that

$$|\widehat{\chi_\theta}(\lambda)| \ \geqslant \ \int_A \cos\left(2\pi(x-e_1)\cdot\lambda\right)\, d\sigma(x). \tag{5.2}$$

However for all $x \in A$ and $\lambda \in R$ the vectors $x - e_1$ and λ are almost perpendicular, so that the cosine here is close to 1. It is easy to make this precise – some elementary geometry gives

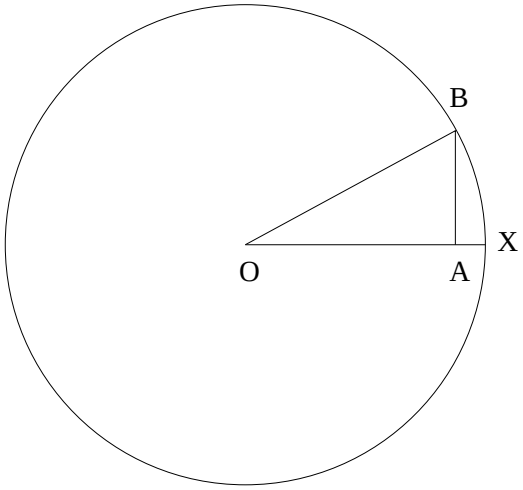

FIGURE 14. Projections.

$$|(x-e_1)\cdot e_1| \ \leqslant \ 1-\cos\delta \ \leqslant \ \delta^2.$$

(In the figure, $e_1 = OX$, $x - e_1 = XB$ and $|(x-e_1)\cdot e_1|$ is just the length $|AX|$.) Thus, writing $\lambda = \lambda_1 e_1 + \lambda_2 e_2$ where $|\lambda_1| \leqslant \frac{1}{20}\delta^{-2}$ and $|\lambda_2| \leqslant \frac{1}{20}\delta^{-1}$, we have

$$\begin{aligned} |(x-e_1)\cdot\lambda| \ &\leqslant \ |\lambda_1||(x-e_1)\cdot e_1| + |\lambda_2||(x-e_1)\cdot e_2| \\ &\leqslant \ 1/10. \end{aligned}$$

The lemma follows immediately from this and (5.2). □

Thus we have a function, namely χ_θ, whose Fourier transform is large on a rectangle in direction θ and centred on the origin. By modifying this function we can easily translate the rectangle. Indeed for any $a \in \mathbb{R}^2$ we see that

$$\widehat{e^{2\pi i a x}\chi_\theta}(\lambda) \ = \ \widehat{\chi_\theta}(\lambda - a).$$

Let's summarise what we have proved.

Proposition 5.1. *Let R be a $\frac{1}{20}\delta^{-2} \times \frac{1}{20}\delta^{-1}$ rectangle in direction θ. Then there is a function f_R with the following properties.*
(i) $\|f_R\|_\infty = 1$;
(ii) $\mathrm{Supp}(f_R) = A_\theta$, where A_θ is an arc of length δ centred on $e^{i\theta}$;
(iii) $|\widehat{f_R d\sigma}(\lambda)| \geqslant \delta/2$ for $\lambda \in R$.

Now let $E \subseteq \mathbb{R}^2$ be a Besicovitch set, that is a compact set containing a line segment in every direction. Consider the collection of all 2×2 squares of the form $\Sigma_{i,j} = [i/2, i/2+2] \times [j/2, j/2+2]$, $(i,j) \in \mathbb{Z}^2$. Let E' be the subset of $\Sigma_{0,0}$ defined by

$$E' \;=\; \bigcup_{(i,j)\in\mathbb{Z}^2} (E \cap \Sigma_{i,j}) - (\tfrac{i}{2}, \tfrac{j}{2}),$$

that is the union of all the intersections of E with squares $\Sigma_{i,j}$, translated to lie one on top of the other in the square $\Sigma_{0,0}$. The set E' still contains a unit line segment in every direction, because all unit line segments are wholly contained in some $\Sigma_{i,j}$. Furthermore $\underline{\dim}(E') \leqslant \underline{\dim}(E)$ (exercise). The most important new feature of E is that it lies in the ball $B(0,4)$. We have no more need of E, so we'll write $E' = E$ for notational simplicity.

Now let $\delta > 0$, and consider the δ-neighbourhood $N_\delta(E)$. This contains T, a union of $1 \times \delta$ rectangles, one in each direction $\theta = j\delta$, $j = 1, \ldots, k$ where $k = \lfloor 1/\delta \rfloor$. Subjecting T to a homothety with scale factor $1/20\delta^2$ gives a collection S of k $\frac{1}{20}\delta^{-2} \times \frac{1}{20}\delta^{-1}$ rectangles. We'll use Proposition 5.1, together with the local restriction estimate of Theorem 11, to get a lower bound on the size of S. This will immediately give a bound for $|T|$, and hence for the Minkowski dimension of E.

Let $R_1, \ldots, R_k$ be the rectangles comprising S. Let $f_i = f_{R_i}$ be the associated functions as described by Proposition 5.1. Finally, for any choice of phases ε_i, $|\varepsilon_i| = 1$ write

$$f_\varepsilon(x) \;=\; \sum_{i=1}^{k} \epsilon_i f_i(x).$$

This introduction of phases ε_i may seem strange at the moment. In a short while we will show that a *random* choice of these phases gives f_ε the properties we would like.

Now observe first of all that for any choice of the phases the L^∞ norm of f_ε is at most 1. Indeed f_i is supported on the arc $A_{i\delta}$, and these arcs are disjoint for different i.

What about $\|\widehat{f_\varepsilon d\sigma}\|_{L^4(B(0,R))}$? Well, Proposition 5.1 tells us that $|\widehat{f_i d\sigma}(\lambda)|$ is at least $\delta/2$ for $\lambda \in R_i$. If the R_i overlap a lot, which is what we would expect if $|S|$ is small, then $\widehat{f_\varepsilon d\sigma}$ ought to be very large at many points, which would cause it to have large L^4 norm. We will show that this can be made precise, at least for a typical choice of the phases ε_i. We have

$$\|\widehat{f_\varepsilon d\sigma}\|^4_{L^4(B(0,R))} \;=\; \sum_{i_1,i_2,i_3,i_4} \varepsilon_{i_1}\varepsilon_{i_2}\overline{\varepsilon_{i_3}\varepsilon_{i_4}} \int_{|\lambda|\leqslant R} \widehat{f_{i_1} d\sigma}(\lambda)\widehat{f_{i_2} d\sigma}(\lambda)\overline{\widehat{f_{i_3} d\sigma}(\lambda)\widehat{f_{i_4} d\sigma}(\lambda)}\, d\lambda \tag{5.3}$$

Now suppose that the ε_i are chosen independently at random. The expected value of $\varepsilon_{i_1}\varepsilon_{i_2}\overline{\varepsilon_{i_3}\varepsilon_{i_4}}$ is zero except when $i_1 = i_3$ and $i_2 = i_4$, or when $i_1 = i_4$ and $i_2 = i_3$, in which case it is 1. In all these cases the integral in (5.3) is non-negative, and so we have

$$\begin{aligned}\mathbb{E}\|\widehat{f_\varepsilon d\sigma}\|^4_{L^4(B(0,R))} &\geqslant \sum_{i_1=i_3, i_2=i_4} \int_{|\lambda|\leqslant R} \widehat{f_{i_1}d\sigma}(\lambda)\widehat{f_{i_2}d\sigma}(\lambda)\overline{\widehat{f_{i_3}d\sigma}(\lambda)\widehat{f_{i_4}d\sigma}(\lambda)}\,d\lambda \\ &= \int_{|\lambda|\leqslant R}\left(\sum_{i=1}^k |\widehat{f_i d\sigma}(\lambda)|^2\right)^2 d\lambda.\end{aligned}$$

Since S is contained in the ball $B(0,\delta^{-2})$, it is natural to set $R = \delta^{-2}$. It then follows from the above that there is at least one specific choice of the phases ε_i for which

$$\|\widehat{f_\varepsilon d\sigma}\|^4_{L^4(B(0,R))} \geqslant \int_S \left(\sum_i |\widehat{f_i d\sigma}(\lambda)|^2\right)^2 d\lambda.$$

However we have, by Proposition 4(iii) and the Cauchy-Schwarz inequality,

$$\begin{aligned}\int_S \left(\sum_i |\widehat{f_i d\sigma}(\lambda)|^2\right)^2 d\lambda &\gg \delta^4 \int_S \left(\sum_i \chi_{R_i}(\lambda)\right)^2 d\lambda \\ &\geqslant \frac{\delta^4}{|S|}\left(\int \sum_i \chi_{R_i}(\lambda)\,d\lambda\right)^2 \\ &\gg \frac{\delta^4 k^2 |R_i|^2}{|S|} \\ &\gg (\delta^4|S|)^{-1}\end{aligned}$$

Thus $\|\widehat{f_\varepsilon d\sigma}\|_{L^4(B(0,R))} \gg \delta^{-1}|S|^{-1/4}$. Substituting into Theorem 11, and recalling that $\|f_\varepsilon\|_\infty \leqslant 1$, one gets that $|S| \gg \delta^{-4}\,(\log(1/\delta))^{-1}$, and hence that $|T| \gg (\log(1/\delta))^{-1}$. Thus we see once again that Besicovitch sets in $\mathbb{R}^2$ have full Minkowski dimension. □

The reader may have noticed some similarities between the argument just presented and that given in *The Kakeya Problem I*. The Cauchy-Schwarz inequality was used in essentially the same way in both cases, but the need to estimate the intersection of pairs of tubes has been replaced by a decay estimate for the Fourier transform of the circle. It might be argued that the new argument is much longer and less intuitive than the old one. Our reason for giving it is that it helps in understanding the higher dimensional situation.

6. Discrete Functional Analysis

In this set of notes we set up the basic languages of harmonic and functional analysis in a discrete setting. We will do our harmonic analysis on $\mathbb{F}^n$, where $\mathbb{F}$ is a finite field (if you like, $\mathbb{F} = \mathbb{Z}/p\mathbb{Z}$). In this setting the basic facts of harmonic analysis, such

as Parseval's identity and the inversion formula, become easy formal exercises. This allows one to understand the combinatorial aspects of the Fourier transform without the distraction of convergence issues.

Vector spaces over finite fields. Fourier transforms. Given a function $f : \mathbb{F}^n \to \mathbb{C}$ we define its integral

$$\int_{\mathbb{F}^n} f(x)\,dx \;=\; \sum_{x\in\mathbb{F}^n} f(x).$$

Fix a non-trivial character $e : \mathbb{F} \to S^1$ (when $\mathbb{F} = \mathbb{Z}/p\mathbb{Z}$, for example, we could take $e(x) = e^{2\pi ix/p} = \omega^x$). Then for $\xi \in \mathbb{F}^n$ we define the Fourier transform

$$\widehat{f}(\xi) \;=\; \int_{\mathbb{F}^n} f(x)e(-x\cdot\xi)\,dx,$$

where the inner product $x \cdot \xi$ is defined by

$$x\cdot\xi \;=\; x_1\xi_1 + \cdots + x_n\xi_n.$$

Now in actual fact what we have said is misleading in one important respect. The Fourier transform is in actual fact defined on the *dual space* $\mathbb{F}_*^n$. This is isomorphic to $\mathbb{F}_*^n$ as an abstract group, but the natural measure on it is different. Thus if $g : \mathbb{F}_*^n \to \mathbb{C}$ is a function then we define the integral

$$\int_{\mathbb{F}_*^n} g(\xi)\,d\xi \;=\; |\mathbb{F}|^{-n}\sum_{\xi} g(\xi).$$

This convention, which may appear strange, emphasises that the spacial and Fourier sides are very different.

Observe that choosing a specific character e amounts to fixing an isomorphism between $\mathbb{F}^n$ and $\mathbb{F}_*^n$. *All* characters $\psi : \mathbb{F}^n \to \mathbb{C}$ are then of the form $x \mapsto e(x\cdot\xi)$ for some $\xi \in \mathbb{F}_*^n$.

The difference between a group and its dual becomes more pronounced for infinite groups. The dual of $\mathbb{R}$ is $\mathbb{R}$ again (which is why Fourier transforms, in the traditional undergraduate sense of the word, are defined on $\mathbb{R}$). However the dual of $\mathbb{Z}$ is the circle group $\mathbb{T} = [0, 1)$, which is why periodic functions are often dicussed in terms of their so-called Fourier coefficients, which are defined at integers. Fourier coefficients are really just another type of Fourier transform, but on the group $\mathbb{T}$. Such matters as this are the topic of a whole Part III course this year by Dr Körner, which may be of interest to those attending this course.

Proposition 6.1 (Fourier facts)**.** *Let $f, g : \mathbb{F}^n \to \mathbb{C}$ be functions. Then we have*
(i) (Parseval) $\int_{\mathbb{F}^n} f(x)\overline{g(x)}\,dx \;=\; \int_{\mathbb{F}_*^n} \widehat{f}(\xi)\overline{\widehat{g}(\xi)}\,d\xi$.
(ii) (Plancherel) $\|f\|_{L^2(\mathbb{F}^n)} = \|\hat{f}\|_{L^2(\mathbb{F}_*^n)}$.

(iii) (Convolution) Define $(f*g)(x) = \int_{\mathbb{F}^n} f(y)\overline{g(x-y)}\,dy$. *Then* $\widehat{f*g}(\xi) = \hat{f}(\xi)\overline{\hat{g}(-\xi)}$.
(iv) (Inversion) For $g : \mathbb{F}_*^n \to \mathbb{C}$ *define*

$$g^\vee(x) \;=\; \widehat{g}(-x) \;=\; \int_{\mathbb{F}_*^n} g(\xi)e(x\cdot\xi)\,d\xi.$$

Then $f(x) = \widehat{f}^\vee(x)$.

Proof. We use the orthogonality relations

$$\int_{\mathbb{F}_*^n} e(x\cdot\xi)\,d\xi \;=\; \delta_0(x).$$

These follow from elementary representation theory or, in the case $\mathbb{F} = \mathbb{Z}/p\mathbb{Z}$, from summing a geometric series.
(i) We have

$$\begin{aligned}\int_{\mathbb{F}_*^n} \widehat{f}(\xi)\overline{\widehat{g}(\xi)}\,d\xi &= \int\int\int f(x)e(-x\cdot\xi)\overline{g(y)}e(y\cdot\xi)\,dxdyd\xi\\ &= \int\int f(x)\overline{g(y)}\left(\int e(-(x-y)\cdot\xi)\,d\xi\right)dxdy\\ &= \int\int f(x)\overline{g(y)}\delta_0(x-y)\,dxdy\\ &= \int f(x)\overline{g(x)}\,dx.\end{aligned}$$

(ii) is just a special case of (i).
(iii) is no more than an application of Fubini's theorem, which in the finite situation is better known as "changing the order of summation":

$$\begin{aligned}\widehat{f*g}(\xi) &= \int\int f(y)\overline{g(x-y)}e(-x\cdot\xi)\,dxdy\\ &= \int f(y)e(-y\cdot\xi)\int \overline{g(x-y)}e(-(x-y)\cdot\xi)\,dxdy\\ &= \hat{f}(\xi)\overline{\hat{g}(-\xi)}.\end{aligned}$$

(iv) Another simple formal exercise. □

Norms, operators and operator norms. In this section we wish to work at the level of a general finite[5] measure space X with measure μ_X. Particular examples are any finite set X together with either counting measure or normailised counting measure. Write $B(X)$ for the vector space of all functions $f : X \to \mathbb{C}$. If $1 \leqslant p < \infty$ then we

[5]I mean that the underlying set is finite – I'm not sure this is standard terminology

define the L^p norm

$$\|f\|_p \;=\; \|f\|_{L^p(X)} \;=\; \left(\int_X |f(x)|^p \, d\mu_X(x)\right)^{1/p}.$$

We also define the L^∞-norm

$$\|f\|_\infty \;=\; \|f\|_{L^\infty(X)} \;=\; \sup_{x\in X} |f(x)|.$$

Proposition 6.2 (L^p facts)**.** *(i) $\|\cdot\|_p$ and $\|\cdot\|_\infty$ are norms.*
(ii) (Hölder's inequality). Define the dual index p' by the equation $1/p + 1/p' = 1$ if $1 < p < \infty$, $\infty' = 1$ and $1' = \infty$. Then for any two functions $f, g : X \to \mathbb{C}$ and any p we have

$$|\langle f, g\rangle_X| \;=\; \left|\int_X f(x)\overline{g(x)}\, d\mu_X(x)\right| \;\leqslant\; \|f\|_p \|g\|_{p'}.$$

(iii) Suppose that $1 \leqslant p_1 \leqslant p_2 \leqslant \infty$ and that μ_X is counting measure. Then $\|f\|_{p_2} \leqslant \|f\|_{p_1}$.
(iv) Suppose that $1 \leqslant p_1 \leqslant p_2 \leqslant \infty$ and that μ_X is normalised counting measure. Then $\|f\|_{p_1} \leqslant \|f\|_{p_2}$.

Proof. (i),(ii) Omitted (see Rudin's red book if you like).
(iii) Suppose, without loss of generality, that $\|f\|_{p_1} = 1$. Then $\|f\|_\infty \leqslant 1$, and so

$$\begin{aligned} \|f\|_{p_2} &\leqslant \|f\|_\infty^{(p_2-p_1)/p_2} \|f\|_{p_1}^{p_1/p_2} \\ &\leqslant 1. \end{aligned}$$

(iv) This is a consequence of Hölder's inequality and the fact that the total measure of X is unity:

$$\begin{aligned} \|f\|_{p_1} &= \left(\int |f(x)|^{p_1}\, d\mu_X\right)^{1/p_1} \\ &\leqslant \left(\int 1^{p_2/(p_2-p_1)}\, d\mu_X\right)^{(p_2-p_1)/p_2} \left(\int |f(x)|^{p_2}\, d\mu_X\right)^{1/p_2}. \end{aligned}$$

The nesting of L^p norms that we see here is peculiar to the counting and normalised counting measures. On $\mathbb{R}$, for example, the function f defined by

$$f(x) \;=\; \begin{cases} 0 & (x = 0) \\ |x|^{-1/3} & (0 < x < 1) \\ |x|^{-1} & (x \geqslant 1) \end{cases}$$

lies in L^2 but in neither L^1 nor L^4.

Now we may use any function $g \in B(X)$ to define a linear map $T_g : B(X) \to \mathbb{C}$ via

$$T_g(f) \;=\; \langle f, g\rangle_X \;=\; \int_X f(x)\overline{g(x)}\, d\mu_X(x).$$

Let $p \in (1,\infty)$. *Any* linear map $T : B(X) \to \mathbb{C}$ has a p-norm $\|T\|_p$ defined to be the supremum of $|Tf|$ over all f with $\|f\|_p = 1$, or equivalently the infimum of all C for which $|Tf| \leqslant C\|f\|_p$. It turns out that the p-norm of T_g depends on g in a very natural way.

Proposition 6.3. $\|T_g\| = \|g\|_{p'}$. *That is,* $\|g\|_{p'} = \sup_{\|f\|_p=1} \langle f, g\rangle_X$.

Proof. It is clear from Hölder's inequality that $\|T_g\| \leqslant \|g\|_{p'}$. For the converse inequality, apply T_g to the function $f(x) = |g(x)|^{p'-1} \cdot e^{i\arg(g(x))}$. It can be checked that $|T_g(f)| = \|g\|_{p'}\|f\|_p$. □

Now let X, Y be two finite measure spaces with measures μ_X, μ_Y. Let $T : B(X) \to B(Y)$ be a linear map. For example, we could take X to be $\mathbb{F}^n$ and Y to be $\mathbb{F}^n_*$, both with their natural measures as discussed above, and T to be the Fourier transform. Given $p, q \in [1,\infty]$ we write $\|T\|_{p\to q}$ for the infimum of all constants C such that $\|Tf\|_{L^q(d\mu_Y)} \leqslant C\|f\|_{L^p(d\mu_X)}$ for all $f \in B(X)$.

Lemma 13. *Let $T : B(X) \to B(Y)$ be a linear operator. Then there is a linear operator $T^* : B(Y) \to B(X)$, called the adjoint of T, which satisfies*

$$\langle f, T^*g\rangle_X \;=\; \langle Tf, g\rangle_Y$$

for all $f \in B(X)$, $g \in B(Y)$.

Proof. This is simple linear algebra. For fixed g the map $\psi_g : B(X) \to \mathbb{C}$ defined by $f \mapsto \langle Tf, g\rangle$ is linear or, in other words, $\psi_g \in B(X)^*$. But every such functional is of the form $f \mapsto \langle f, h\rangle$ for some unique $h \in B(X)$. Define $T^*g = h$, and check that T^* is a linear map. □

The following example of adjoint oprators T and T^* is of great relevance to this course.

Example. Let $S \subseteq \mathbb{F}^n_*$, and endow S with the normalised counting measure σ *(*so $\sigma(x) = |S|^{-1}$ if $x \in S$ and 0 otherwise*)*. let $T : B(\mathbb{F}^n) \to B(S)$ be the *restriction* map

$$f \mapsto \widehat{f}|_S.$$

Then the adjoint, T^*, of T is the *extension* map

$$g \mapsto (gd\sigma)^\vee.$$

Proof. We must check that

$$\int f(x)\overline{(gd\sigma)^\vee(x)}\,dx \;=\; \int \hat{f}(\xi)\overline{gd\sigma(\xi)}.$$

But this is simply Parseval's identity. □

Another important example is the following.

Example. Let $T : B(X) \to B(X)$ be given by convolution by a kernel K, so that $Tf = f * K$. Then $T^*g = g * K'$, where $K'(x) = \overline{K(-x)}$. □

In the examples we will encounter, K will equal K'. This is the case when $K = \hat{h}$, where h is a real-valued function (such as a measure). In these cases, then, $f \mapsto f * K$ is a self-adjoint operator.

For any linear operator T the various norms of T and of T^* are related in the following pleasant way.

Proposition 6.4. *Let $p, q \in (1, \infty)$. Then we have*

$$\|T\|_{p\to q} \;=\; \|T^*\|_{q'\to p'}.$$

Proof. By Proposition B.3 and Hölder we have

$$\begin{aligned}
\|T^*f\|_{p'} &= \sup_{\|g\|_p=1} \langle T^*f, g\rangle \\
&= \sup_{\|g\|_p=1} \langle f, Tg\rangle \\
&\leqslant \sup_{\|g\|_p=1} \|f\|_{q'}\|Tg\|_q \\
&\leqslant \|T\|_{p\to q}\|f\|_{q'}.
\end{aligned}$$

It follows that $\|T^*\|_{q'\to p'} \leqslant \|T\|_{p\to q}$, and the reverse inequality may be demonstrated in an identical manner. □

This result tells us that for any problem requiring us to bound the L^p–L^q norm of some operator $T : B(X) \to B(Y)$, there is a corresponding *dual problem* concerning the operator $T^* : B(Y) \to B(X)$. It is often extremely helpful to think about the dual formulation of a problem. There is a special situation in which one can manufacture a third problem which is equivalent to both the original problem and its dual. This is known as the T^*T technique.

Proposition 6.5. *Suppose that $T : B(X) \to B(Y)$. Then $\|T^*T\|_{p\to p'} = \|T\|^2_{p\to 2} = \|T^*\|^2_{2\to p'}$.*

Proof. The inequality $\|T^*T\|_{p\to p'} \leqslant \|T^*\|_{2\to p'}\|T\|_{p\to 2} = \|T\|^2_{p\to 2}$ is immediate. On the other hand we have, using Hölder, that for any f

$$\begin{aligned}
\|Tf\|^2_{p\to 2} &= \langle Tf, Tf\rangle_Y \\
&= \langle T^*Tf, f\rangle_X \\
&\leqslant \|T^*Tf\|_{L^{p'}(d\mu_X)}\|f\|_{L^p(d\mu_X)} \\
&\leqslant \|T^*T\|_{p\to p'}\|f\|^2_{L^p(d\mu_X)}.
\end{aligned}$$

Thus we have the reverse inequality $\|T\|_{p\to 2}^2 \leqslant \|T^*T\|_{p\to p'}$. □

Thus, in considering L^p–L^2 bounds on an operator $T : B(X) \to B(Y)$, we can if desired consider the L^p–$L^{p'}$ bounds on the operator $T^*T : B(X) \to B(X)$.

****Some remarks on infinite measure spaces.** The classical space $L^p(\mathbb{R})$ is defined to be the vector space of all locally integrable functions f for which $\int |f(x)|^p\,dx < \infty$. Even this simple construct has no analogue in the finite case because *every* function on a finite measure space has bounded L^p norm. Discussion of the L^p-L^q norm of a linear operator T can become difficult because often there is no natural space on which T is *a priori* defined.

To give a classical example, the Plancherel theorem states that

$$\|\hat{f}\|_{L^2(\mathbb{R})} \;=\; \|f\|_{L^2(\mathbb{R})}. \tag{6.1}$$

But what does this mean? As it stands, nothing. The Fourier transform on $\mathbb{R}$ is defined, for functions $f \in L^1(\mathbb{R})$, by the equation

$$\hat{f}(\lambda) \;=\; \int_{-\infty}^{\infty} f(x)e^{-2\pi i x\lambda}\,dx. \tag{6.2}$$

It is not the case that $L^2(\mathbb{R}) \subseteq L^1(\mathbb{R})$, so (6.1) does not make sense even for all $f \in L^2(\mathbb{R})$. The normal way to resolve this issue is to prove (6.1) for the class of Schwarz functions (smooth functions, all of whose derivatives decay at infinity faster than any polynomial). If f is Schwarz then both f and $\hat{f}$ lie in all L^p spaces and so it certainly makes sense to write down (and prove) (6.1). One can then get (6.1) for a wider class of functions via a limiting argument, using the fact that the Schwarz functions are dense in every L^p space. One can even use (6.1) to give a *definition* of the Fourier transform on $L^2(\mathbb{R})$ which agrees with (6.2) on $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$.

This is how the Plancherel theorem is often stated. Namely, that the Fourier transform as defined on $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ by (6.2) extends to an isometry of $L^2(\mathbb{R})$. Let me reiterate that this does *not* mean that one can make sense of the Fourier integral (6.2) for all L^2 functions.

We do not wish to get embroiled in matters such as this in the present course. For that reason on the rare occasions that we have dealt with Euclidean L^p spaces we have tacitly assumed that all functions involved are sufficiently regular that equations such as (6.2) are *a priori* meaningful**.

7. Riesz-Thorin and its consequences

In this set of notes (X, μ_X) and (Y, μ_Y) will be finite vector spaces with measures μ_X, μ_Y.

The classical Riesz-Thorin interpolation theorem is a result concerning the behaviour of general linear operators in Euclidean L^p spaces. Here we prove a version for linear operators on finite measure spaces. We give a proof which is almost identical to the argument for the Euclidean case. However, as we are working in a finite setting we can avoid several measure-theoretic difficulties.

Theorem 12 (Riesz-Thorin)**.** *Let $T : B(X) \to B(Y)$ be linear and suppose that $p_0, p_1, q_0, q_1 \in [1, \infty]$ satisfy $p_0 < p_1$ and $q_0 < q_1$. For any $t \in [0, 1]$ define p_t, q_t by*

$$\frac{1}{p_t} \;=\; \frac{1-t}{p_0} + \frac{t}{p_1}$$

and

$$\frac{1}{q_t} \;=\; \frac{1-t}{q_0} + \frac{t}{q_1}.$$

Then

$$\|T\|_{p_t \to q_t} \;\leqslant\; \|T\|_{p_0 \to q_0}^{1-t} \|T\|_{p_1 \to q_1}^{t}.$$

It turns out that this result is a short, if slightly tricky, deduction from the following result in complex analysis known as the Hadamard three-circles theorem.

Proposition 7.1. *Let D be the strip $0 < \Re(z) < 1$. Suppose that a function f is analytic on D and continuous and bounded on $\overline{D}$. Suppose further that $|f(z)| \leqslant M_0$ on $\Re(z) = 0$ and $|f(z)| \leqslant M_1$ on $\Re(z) = 1$. Then $|f(a + ib)| \leqslant M_0^{1-a} M_1^a$ for any $a \in [0, 1]$ and $b \in \mathbb{R}$.*

Proof Write $M = \min(M_0, M_1)$. For any $\epsilon > 0$ set

$$F_\epsilon(z) \;=\; \frac{e^{-\epsilon(1-z)z} f(z)}{M_0^{1-z} M_1^z}.$$

It is easy to check the bound

$$|F_\epsilon(a + ib)| \;\leqslant\; \frac{\|f\|_\infty \, e^{-\epsilon(a(1-a)+b^2)}}{M_0^{1-a} M_1^a}.$$

In particular we have $|F_\epsilon(a + ib)| \leqslant 1$ if $|b| \geqslant L_\epsilon$, where we can take

$$L_\epsilon \;=\; \epsilon^{-1/2} \left(\log(\|f\|_\infty / M)\right)^{1/2}.$$

It is also clear, straight from the definition, that $|F_\epsilon(z)| \leqslant 1$ when $\Re(z) = 0$ or 1. By applying the maximum principle to very large rectangles of the form $0 \leqslant \Re(z) \leqslant 1$, $-R \leqslant \Im(z) \leqslant R$ (where $R \geqslant L_\epsilon$) we see that F_ϵ is in fact bounded by 1 throughout D.

Thus

$$|f(z)| \leqslant e^{\epsilon(a(1-a)+b^2)} M_0^{1-a} M_1^a$$

for all $z = a + ib \in D$. For any fixed z let $\epsilon \to 0$, and we get Proposition B.1. □

Remarks There is another Hadamard three-circles theorem, which actually involves three circles (and is rather easy to prove). One might be tempted to deduce Proposition B.1 from that result by a transformation such as $z \mapsto e^z$. This map not being one-to-one makes such an undertaking rather difficult, but the map $z \mapsto e^{\epsilon z}$ is one-to-one on an increasingly large portion of D as $\epsilon \to 0$. This perhaps motivates the above proof.

Proof of Theorem 12. Write $M_0 = \|T\|_{p_0 \to q_0}$ and $M_1 = \|T\|_{p_1 \to q_1}$. It suffices to prove that

$$\langle Tf, g \rangle \leqslant M_0^{1-t} M_1^t \tag{7.1}$$

whenever $\|f\|_{p_t} = \|g\|_{q_t'} = 1$. Indeed we have

$$\begin{aligned} \|T\|_{p_t \to q_t} &= \sup_{\|f\|_{p_t}=1} \|Tf\|_{q_t} \\ &= \sup_{\|f\|_{p_t}=1} \sup_{\|g\|_{q_t'}=1} \langle Tf, g \rangle. \end{aligned}$$

Now for $z \in D$ define functions $\alpha(z)$ and $\beta(z)$ by

$$\alpha(z) = \frac{1-z}{p_0} + \frac{z}{p_1}$$

and

$$\beta(z) = \frac{1-z}{q_0} + \frac{z}{q_1}.$$

Set

$$h(z) = \langle Tf_z, g_z \rangle$$

where

$$f_z(x) = |f(x)|^{\alpha(z)/\alpha(t)} e^{i \arg f(x)}$$

and

$$g_z(x) = |g(x)|^{(1-\beta(z))/(1-\beta(t))} e^{i \arg g(x)}.$$

Now $h(z)$ is analytic in D and bounded and continuous on $\overline{D}$. Now we claim that $|h(iy)| \leqslant M_0$ for any $y \in \mathbb{R}$. To see this, observe that by Hölder's Inequality we have $|h(iy)| \leqslant \|g_{iy}\|_{q_0'} \|Tf_{iy}\|_{q_0}$. It is easy to check that $\|g_{iy}\|_{q_0'}$ is equal to $\|g\|_{q_t'}^{q_t'/q_0'}$, which is just 1. Furthermore $\|Tf_{iy}\|_{q_0} \leqslant M_0 \|f_{iy}\|_{p_0}$, a quantity which turns out to equal M_0. A very similar argument establishes that $|h(1+iy)| \leqslant M_1$ for any $y \in \mathbb{R}$. We may now apply Proposition B.1 to get that $|h(t)| \leqslant M_0^{1-t} M_1^t$, and the proof is concluded by observing that $h(t)$ is precisely $\langle Tf, g \rangle$. □

Our first corollary of Riesz-Thorin is a result of Young describing the L^p behaviour of convolutions. We begin with a lemma.

Lemma 14 (Integral Minkowski inequality)**.**

$$\left(\int_X \left|\int_X |F(x,y)|\, d\mu_X(x)\right|^p d\mu_X(y)\right)^{1/p} \leqslant \int_X \left(\int_X |F(x,y)|^p\, d\mu_X(y)\right)^{1/p} d\mu_X(x).$$

Proof. This is really just the ordinary Minkowski inequality in disguise. Set

$$G(y) \;=\; \int_X |F(x,y)| d\mu_X,$$

and for $x \in X$ write

$$G_x(y) \;=\; |F(x,y)|.$$

Then the inequality can be written as

$$\|G\|_p \;\leqslant\; \int_X \|G_x\|_p\, d\mu_X.$$

Remember, though, that the integrals here are simply (weighted) finite sums. □

Theorem 13 (Young's Inequality)**.** *Suppose that $f, g \in B(X)$, and suppose that $p, q, r \in [1, \infty]$ satisfy*

$$1 + r^{-1} \;=\; p^{-1} + q^{-1}. \tag{7.2}$$

Then

$$\|f * g\|_r \;\leqslant\; \|f\|_p \|g\|_q.$$

Proof. Fix $f \in B(X)$. Then the map $T_f : g \mapsto g * f$ is linear. Hölder's inequality immediately yields

$$\|T_f(g)\|_\infty \;\leqslant\; \|f\|_p \|g\|_{p'}. \tag{7.3}$$

Furthermore the integral version of Minkowski's inequality quickly leads to

$$\|T_f(g)\|_p \;\leqslant\; \|f\|_p \|g\|_1. \tag{7.4}$$

Young's inequality follows by using Riesz-Thorin to interpolate the two bounds (7.3) and (7.4). Take $t = p/r$, which obviously lies in the interval $[0, 1]$. □

Our next result is in the spirit of the restriction theorems that are one of the main topics of the course. It relates the L^p behaviour of a function to that of its Fourier transform.

Theorem 14 (The Hausdorff-Young Inequality)**.** *Suppose that $f \in B(\mathbb{F}^n)$ and that $1 \leqslant p \leqslant 2$. Then*

$$\|\hat{f}\|_{p'} \;\leqslant\; \|f\|_p.$$

Proof. The map $T : f \mapsto \hat{f}$ is linear, and we have the bounds $\|f\|_\infty \leqslant \|f\|_1$ and $\|\hat{f}\|_2 = \|f\|_2$. The result follows by using Riesz-Thorin to interpolate between these

two. □

It would not be a bad idea to look at Rudin's red book, where the classical real-variable analogues of all these inequalities are proved (the proofs are the same as the above, modulo some measure theoretic technicalities). One difference between the finite case and the Euclidean case is that Young and Hausdorff-Young are not sharp in the Euclidean case. In the finite case of Hausdorff-Young, for example, equality occurs when f is like a delta function ($f(0) = 1$ and $f(x) = 0$ when $x \neq 0$). In the Euclidean case the extremal functions are gaussians and Hausdorff-Young fails to be sharp by a multiplicative constant.

There is a situation in which Riesz-Thorin allows us to turn a single L^p-L^q bound into a whole family of such bounds. If $T : B(X) \to B(X)$ is an operator then we say that T is *self-adjoint* if $T = T^*$. A natural family of self-adjoint operators is given by the following lemma.

Lemma 15. *Let $K : X \to \mathbb{C}$ be an even function (so that $K(-x) = K(x)$). Then the convolution operator $f \mapsto K * f$ is self-adjoint.*

Proof. Indeed

$$\begin{aligned} \langle f, K * g \rangle &= \int_X f(x) \overline{\int_X K(x-y)\overline{g(y)}\, dy}\, dx \\ &= \int_X g(y) \int_X f(x)\overline{K(x-y)}\, dx\, dy \\ &= \overline{\langle K * f, g \rangle}. \end{aligned}$$

Theorem 15. *Suppose that $T : B(X) \to B(X)$ is self-adjoint and that $p, q \in [1, \infty]$. Let $\theta \in [0, 1]$, and define*

$$\begin{aligned} r &= \frac{pq}{q\theta + (1-\theta)(q-1)p}, \\ s &= \frac{pq}{p\theta + (1-\theta)(p-1)q}. \end{aligned}$$

Then $\|T\|_{r\to s} \leqslant \|T\|_{p\to q}$.

Proof. This follows by using the Riesz-Thorin theorem to interpolate between $\|T\|_{p\to q}$ and $\|T\|_{q'\to p'}$, which are equal by duality and self-adjointness. □

As our final corollary of Riesz-Thorin I want to mention a bilinear interpolation result.

Proposition 7.2 (Bilinear Interpolation)**.** *Let $\psi : B(X) \times B(X) \to \mathbb{C}$ be a hermitian form, and suppose that*

$$|\psi(f, g)| \leqslant C_i \|f\|_{p_i} \|g\|_{q_i}$$

for $i = 0, 1$. Let $t \in [0,1]$ and write $p_t^{-1} = (1-t)p_0^{-1} + tp_1^{-1}$, $q_t^{-1} = (1-t)q_0^{-1} + tq_1^{-1}$. Then

$$|\psi(f,g)| \leqslant C_0^{1-t} C_1^t \|f\|_{p_t} \|g\|_{q_t}.$$

Proof. There exists a linear operator $T : B(X) \to B(X)$ such that $\psi(f,g) = \langle Tf, g\rangle$ for all $f, g \in B(X)$. Furthermore $C_i \geqslant \|T\|_{p_i \to q_i'}$. It follows from Riesz-Thorin that

$$\|T\|_{p_t \to q_t} \leqslant C_0^{1-t} C_1^t,$$

which immediately implies the proposition using Hölder's inequality. □

8. Restriction theory of the discrete paraboloid I

In this set of notes we consider the restriction theory of the discrete paraboloid in $\mathbb{F}_*^3$, defined to be the set of points

$$P = \left\{(\xi_1, \xi_2, \xi_1^2 + \xi_2^2) : \xi_1, \xi_2 \in \mathbb{F}\right\}.$$

We will assume that -1 is not a square in $\mathbb{F}$, so that P does not contain any lines (this is one of the exercises on the second example sheet). The reason for doing this is that we can prove more in this case! Before we begin, here are a few reasons for and against thinking about P instead of spheres like S^2. Reasons for:

- No measure theoretic difficulties; one can compute the L^p norm of any function, and also the Fourier transform.
- We can see the bare bones of some of the Euclidean arguments without the need for smooth bump functions, dyadic decompositions, etc.

Reasons against:

- Some features of the Euclidean case are not really present here. For example, there is no particularly natural notion of curvature in finite fields.
- The relation between restriction and Kakeya is a lot more tenuous in finite fields (though see the paper of Mockenhaupt and Tao), largely because there is no nice notion of tangency or approximation in the neighbourhood of a point in the finite field case.

The Fourier transform of P. Although there is no natural notion of curvature in $\mathbb{F}_*^3$, one can still obtain an analogue of the decay estimates for $d\sigma_{S^{n-1}}$. As $|P| = N^2$, the surface measure σ on P is given by $\sigma(\xi) = N\chi_P(\xi)$.

Lemma 16 (Gauss sums)**.** *Suppose that $a \in \mathbb{F} \setminus \{0\}$, and write $G(a) = \sum_{x \in \mathbb{F}} e(ax^2)$. Then $G(a) = \pm i\sqrt{N}$.*

Proof. We begin by evaluating $|G(a)|$. For a given $t \in \mathbb{F}$, the number of representations $n(t)$ of t as $x^2 - x'^2$ is $N-1$ unless $t=0$, in which case it is $2N-1$. Thus we have

$$\begin{aligned} |G(a)|^2 &= \sum_{x,x'} e(a(x^2-x'^2)) \\ &= \sum_t n(t)e(at) \\ &= N + (N-1)\sum_t e(at) \\ &= N. \end{aligned}$$

Thus certainly $|G(a)| = \sqrt{N}$. Now -1 is not a square in $\mathbb{F}$, and so for any a the sum $G(a) + \overline{G(a)}$, which equals

$$\sum_x e(ax^2) + e(-ax^2),$$

runs over each of the elements of $\mathbb{F}$ exactly twice. It therefore equals zero, and so $G(a)$ is purely imaginary. □

Lemma 17 (Fourier transform of $d\sigma$)**.** *The Fourier transform $\widehat{d\sigma}$ is as follows.*

$$|\widehat{d\sigma}(x)| = \begin{cases} 1 & x = 0 \\ 0 & x_3 = 0, x \neq 0 \\ -N^{-1}e\left((x_1^2+x_2^2)/4x_3\right) & \textit{otherwise} \end{cases}$$

Proof. We have

$$\begin{aligned} \widehat{d\sigma}(x_1,x_2,x_3) &= N^{-2}\sum_{\xi_1,\xi_2} e\left(-\xi_1 x_1 - \xi_2 x_2 - (\xi_1^2+\xi_2^2)x_3\right) \\ &= N^{-2}\sum_{\xi_1} e\left(-\xi_1 x_1 - \xi_1^2 x_3\right)\sum_{\xi_2} e\left(-\xi_2 x_2 - \xi_2^2 x_3\right). \end{aligned} \tag{8.1}$$

It is clear that if $x=0$ then this equals 1. If $x_3 = 0$ but at least one of x_1, x_2 is nonzero then one of the two sums in (8.1) vanishes, and $\widehat{d\sigma}(x) = 0$. If $x_3 \neq 0$ then we may complete the square in both sums, turning them both into Gauss sums. The result then follows from Lemma 16. □

Basic algebra of restriction theory. Let $p, q \in [1,\infty]$. We write $R^*(p \to q)$ for the smallest constant such that we have the restriction estimate

$$\|(fd\sigma)^\vee\|_{L^q(\mathbb{F}^3)} \leqslant R^*(p\to q)\|f\|_{L^p(d\sigma)}.$$

In general, $R^*(p\to q)$ will be a function of the underlying field. If, however, $R^*(p\to q)$ is bounded independently of $|\mathbb{F}|$ then we say that $\mathrm{Res}(p,q)$ holds. We are going to prove several restriction estimates for the discrete parabola. To make this a bit more meaningful, we need to get an idea of which restriction estimates are better than others.

To do this, let is recall the lemma from a previous set of notes which dealt with nesting properties of L^p-norms.

Lemma 18. *Suppose that p_1, p_2 satisfy $1 \leqslant p_1 \leqslant p_2 \leqslant \infty$. Then*
(i) $\|f\|_{L^{p_1}(P)} \leqslant \|f\|_{L^{p_2}(P)}$;
(ii) $\|f\|_{L^{p_2}(\mathbb{F}^3)} \leqslant \|f\|_{L^{p_1}(\mathbb{F}^3)}$. □

The following corollaries are immediate:

Corollary 8.1 (Small p is good)**.** *We have $R^*(p_1 \to q) \leqslant R^*(p_2 \to q)$ whenever $p_1 \geqslant p_2$.* □

Corollary 8.2 (Small q is good)**.** *We have $R^*(p \to q_1) \leqslant R^*(p \to q_2)$ whenever $q_1 \geqslant q_2$.* □

A discrete Tomas-Stein estimate. The Tomas-Stein argument, which we alluded to in a non-examinable handout, concerns the restriction properties of spheres S^{n-1}. In this section we prove $\mathrm{Res}(2,4)$ for the discrete paraboloid using an argument which may be regarded as the discrete analogue of the Tomas-Stein technique (of course, as I am not going to discuss Tomas-Stein, this last claim won't mean a great deal).

Theorem 16. $R^*(2 \to 4) \leqslant 2$.

Let $T : B(P) \to B(\mathbb{F}^3)$ be the extension map $f \mapsto (fd\sigma)^\vee$, and let $T^* : B(\mathbb{F}^3) \to B(P)$ be its dual, the restriction map $g \mapsto \widehat{g}|_P$.

We can get a bound on $\|T\|_{2\to p}$ by using the method of T and T^*. Observe that $TT^*g = (\widehat{g}d\sigma)^\vee = g * \widehat{d\sigma}$ (to check this, take Fourier transforms of both sides), so we are interested in obtaining bounds of the form

$$\|g * \widehat{d\sigma}\|_4 \ \leqslant \ \|g\|_{4/3}. \tag{8.2}$$

To get such bounds we first of all split $\widehat{d\sigma}$ into two pieces

$$\widehat{d\sigma} \ = \ \delta_0 + K,$$

the aim being to use the triangle inequality on the two pieces separately.

It is easy to deal with the δ_0 portion. Indeed $g * \delta_0 = g$, so the required estimate

$$\|g * \delta_0\|_4 \ \leqslant \ \|g\|_{4/3} \tag{8.3}$$

becomes simply $\|g\|_4 \leqslant \|g\|_{4/3}$, which an instance of the nesting-of-norms inequalities under counting measure. (Alternatively – and this has more in common with the Tomas-Stein argument – one could interpolate between the obvious bounds $\|g\|_2 = \|g\|_2$ and $\|g\|_\infty \leqslant \|g\|_1$.)

The key to proving (8.2), then, is the following lemma.

Lemma 19. *We have* $\|g * K\|_4 \leqslant \|g\|_{4/3}$.

Proof. Once again we proceed by interpolation. We have (by a simple case of Young's inequality which is trivial to prove directly) that $\|g * K\|_\infty \leqslant \|K\|_\infty \|g\|_1$. But we know all about the magnitude of K from Lemma 17. Indeed $\|K\|_\infty \leqslant N^{-1}$, leading to

$$\|g * K\|_\infty \ \leqslant \ N^{-1}\|g\|_1. \tag{8.4}$$

It is almost as easy to get an L^2–L^2 bound. One has, using elementary properties of the Fourier transform,

$$\begin{aligned} \|g * K\|_2 &= \|\widehat{g}\overline{K^\vee}\|_2 \\ &\leqslant \|K^\vee\|_\infty \|g\|_2. \end{aligned}$$

However $K^\vee(\xi) = \widehat{d\sigma}^\vee(\xi) - \delta_0^\vee(\xi)$, and this is just $d\sigma(\xi) - 1 = N - 1$. Therefore

$$\|g * K\|_2 \ \leqslant \ N\|g\|_2. \tag{8.5}$$

The lemma follows immediately by interpolating (8.4) and (8.5). □

Combining (8.3) and Lemma 37 using the triangle inequality leads to $\|g * \widehat{d\sigma}\|_4 \leqslant 2\|g\|_{4/3}$. As we have remarked, this gives (by the method of T and T^*) a restriction estimate of the form

$$\|(f d\sigma)^\vee\|_4 \ \leqslant \ 2\|f\|_2,$$

which is exactly the statement of Theorem 16.

**A fairly direct modification of the above proof to the Euclidean case constitutes an argument of Tomas which comes within an ϵ of the Tomas-Stein theorem for spheres that we discussed in the starred handout.. The most important modification is the decomposition of $\widehat{d\sigma}$ into dyadic chunks

$$\widehat{d\sigma_i}(x) \ = \ \widehat{d\sigma}(x)\psi\left(\widehat{d\sigma}(x)/2^i\right),$$

where ψ is a smooth approximation to the unit ball of $\mathbb{R}^n$. One then performs an L^1–L^∞ to L^2–L^2 interpolation, very similar to the above, on each convolution operator $f \mapsto f * \widehat{d\sigma_i}$ separately and adds up the results. In this manner one can prove that $\|(f d\sigma)^\vee\|_{L^p(\mathbb{R}^n)} \ll \|f\|_{L^2(S^{n-1})}$ for any value of p strictly greater than the critical exponent $2(n+1)/(n-1)$. To get the endpoint $p = 2(n+1)/(n-1)$ one needs a technique called complex interpolation, which is very similar to the argument we used to prove Riesz-Thorin. It was introduced in this context by Stein, which explains the reason for the result being called Tomas-Stein. In the finite field case we only need two "dyadic chunks" K and δ_0, so the extra complexity was not a problem**.

What does Theorem 16 actually mean? One way to try and understand this question is to substitute various functions f into the bound $\|(fd\sigma)^\vee\|_4 \ll \|f\|_2$. One example, which really captures the essence of the result, is to substitute $f = \chi_E$ where $E \subseteq P$ is some set. In the next lemma, and in the remainder of this course, we write E for χ_E; that is, we identify sets with their characteristic functions.

Lemma 20. *$N^5\|(Ed\sigma)^\vee\|_4^4$ is equal to the number of* additive quadruples *in E, that is to say the number of quadruples $(\xi_1, \xi_2, \xi_3, \xi_4) \in E^4$ with $\xi_1 + \xi_2 = \xi_3 + \xi_4$.*

Proof. This is a straightforward computation, which illustrates yet again the utility of having L^4, L^6, etc norms around. We have

$$\begin{aligned}\|(Ed\sigma)^\vee\|_4^4 &= \int \left| \int E(\xi) e(x\cdot\xi)\, d\sigma(\xi)\right|^4 dx \\ &= N^{-8} \sum_{\xi_1,\dots,\xi_4} E(\xi_1)\dots E(\xi_4) \sum_x e\left(x\cdot(\xi_1+\xi_2-\xi_3-\xi_4)\right) \\ &= N^{-5} \sum_{\substack{\xi_1+\xi_2=\xi_3+\xi_4 \\ \xi_1,\dots,\xi_4\in E}} 1,\end{aligned}$$

which concludes the proof. □

It is rather easy to see that $\|E\|_2^4 = N^{-4}|E|^2$, and so our $R^*(2\to 4)$ estimate tells us that if $E \subseteq P$ then the number of additive quadruples in E is bounded above by a constant multiple of $N|E|^2$. Now the number of additive quadruples in a set is a particular way of describing how much arithmetic structure that set has. Thus, very qualitatively, our restriction estimate tells us that P has no subspaces with lots of arithmetic structure. This is not a surprise if one imagines P as a sort of curved surface.

In fact, Lemma 20 holds the key to a better restriction estimate, giving a good bound on $R^*(8/5 \to 4)$.

Theorem 17. *$R^*(8/5 \to 4) \leqslant 10 \log N$.*

Proof. The following lemma is the key to the proof.

Lemma 21. *Suppose that $E \subseteq P$. Then the number of additive quadruples in E^4 is at most $5|E|^{5/2}$.*

Proof. Fix $\xi_1 \in E$. We will show that the number of pairs $(\xi_2, \xi_3) \in E^2$ with $\xi_2, \xi_3 \neq \xi_1$ and $\xi_2 - \xi_3 \in P - \xi_1$ is no more than $|E|^{3/2}$. The number of additive quadruples in E^4 which involve ξ_1 is then at most $|E|^{3/2} + 2|E|$, and the proposition follows on summing over ξ_1. Now fix ξ_3, and let us ask what condition on ξ_2 guarantees that $\xi_2 - \xi_3 \in P - \xi_1$. Writing $\xi_i = (\alpha_i, \beta_i, \alpha_i^2 + \beta_i^2)$ it transpires that we must have

$$(\alpha_3 - \alpha_1)(\alpha_3 - \alpha_2) \;=\; -(\beta_3 - \beta_1)(\beta_3 - \beta_2).$$

This implies that (α_2, β_2) lies on a line $l(\xi_3) \subseteq \mathbb{F}_p^2$ defined by ξ_3. After a linear change in coordinates so that (α_1, β_1) becomes the origin of $\mathbb{F}_p^2$, this line may be written in the form

$$\alpha_2\alpha_3 + \beta_2\beta_3 = \alpha_3^2 + \beta_3^2.$$

Consider some other line $l(\xi_3')$ given by

$$\alpha_2'\alpha_3 + \beta_2'\beta_3 = \alpha_3'^2 + \beta_3'^2.$$

In order for this to coincide with $l(\xi_3)$ we must have, say, $\alpha_3' = \lambda\alpha_3$ and $\beta_3' = \lambda\beta_3$. But then we would need to have

$$\lambda^2(\alpha_3^2 + \beta_3^2) = \lambda(\alpha_3^2 + \beta_3^2).$$

Since -1 is not a square in $\mathbb{F}^2$ and neither (α_3, β_3) nor (α_3', β_3') is the zero vector, this forces $\xi_3 = \xi_3'$.

We have, then, $m = |E|$ points p_i and m lines l_i, and wish to place an upper bound on the number of pairs (i, j) with $p_i \in l_j$. To do this, relabel so that each of the points $p_1, \ldots, p_{m'}$ meets *some* line, and suppose in fact that p_i meets $n_i \geqslant 1$ lines for $i = 1, \ldots, m'$. The number of triples (p_i, l_j, l_k) with $j < k$ and $p_i = l_j \cap l_k$ is then $\sum_i \binom{n_i}{2}$, which is at least

$$\frac{1}{2m'}\left(\sum_i n_i - m'\right)^2$$

by Cauchy-Schwarz. On the other hand it is at most the number of pairs of lines (l_j, l_k) with $j < k$, which is at most m^2. Thus

$$\frac{1}{2m'}\left(\sum_i n_i - m'\right)^2 \leqslant m^2,$$

which implies that $\sum n_i \leqslant 3m^{3/2}$. The lemma follows. □

The deduction of Theorem 17 from this is via a dyadic decomposition. Let $f \in B(P)$, and assume that $\|f\|_\infty = 1$. For each non-negative integer j write E_j for the set of x for which $2^{-j-1} < f(x) \leqslant 2^{-j}$. We have

$$\begin{aligned} \|f\|_{8/5}^{8/5} &= N^{-2}\sum |f(x)|^{8/5} \\ &\geqslant N^{-2}\sum_j 2^{-8(j+1)/5}|E_j|, \end{aligned}$$

so that

$$|E_j|^{5/2} \leqslant N^5 2^{4(j+1)}\|f\|_{8/5}^4.$$

Write $f_j = f\chi_{E_j}$. It follows from Lemma 21 and the preceding that

$$\begin{aligned}\|(f_j d\sigma)^\vee\|_4^4 &= N^{-5} \sum_{\substack{\xi_1,\dots,\xi_4 \in E_j \\ \xi_1+\xi_2=\xi_3+\xi_4}} f(\xi_1)f(\xi_2)f(\xi_3)f(\xi_4) \\ &\leqslant N^{-5}2^{-4j} \sum_{\substack{\xi_1,\dots,\xi_4 \in E_j \\ \xi_1+\xi_2=\xi_3+\xi_4}} 1 \\ &\leqslant 5 \cdot N^{-5} \cdot 2^{-4j} \cdot |E_j|^{5/2} \\ &\leqslant 80\|f\|_{8/5}^4.\end{aligned}$$

Now since $\|f\|_\infty = 1$, $\|f\|_{8/5}^4$ is at least N^{-5}. Thus, writing $g = f\chi_{\{x:f(x)<2^{-m}\}}$ we have

$$\begin{aligned}\|(gd\sigma)^\vee\|_4^4 &= N^{-5} \sum_{\substack{\xi_1,\dots,\xi_4 \in P \\ \xi_1+\xi_2=\xi_3+\xi_4}} g(\xi_1)g(\xi_2)g(\xi_3)g(\xi_4) \\ &\leqslant 2^{-4m} \\ &\leqslant \|f\|_{8/5}^4,\end{aligned}$$

provided that $m \geqslant 3\log N$. Finally, then,

$$\begin{aligned}\|(fd\sigma)^\vee\|_4 &\leqslant \|(gd\sigma)^\vee\|_4 + \sum_{j \leqslant 3\log N} \|(f_j d\sigma)^\vee\|_4 \\ &\leqslant 10\log N\|f\|_{8/5},\end{aligned}$$

as claimed. □

9. Restriction Theory of the discrete paraboloid, II

In this set of notes we will show that $R^*(2 \to 18/5)$ (for the discrete paraboloid) grows more slowly than any power of N. It is a very simple matter to adapt the arguments to show that in fact $\mathrm{Res}(2, 18/5+\epsilon)$ holds for any $\epsilon > 0$. The proof of this will use as its main ingredient the bound on $R^*(8/5 \to 4) \leqslant 10\log N$ from the previous set of notes, together with quite a bit of interpolation and some specific properties of the Fourier transform of the paraboloid.

We begin by using TT^*. Once again let $T : B(P) \to B(\mathbb{F}^3)$ be the extension map $g \mapsto (gd\sigma)^\vee$, and let $T^* : B(\mathbb{F}^3) \to B(P)$ be its dual, the restriction map $f \mapsto \hat{f}|_P$. Observe that TT^* is the map from $B(\mathbb{F}^3)$ which sends f to $f * \widehat{d\sigma}$.

Theorem 18. *We have $R^*(2 \to 18/5) \leqslant \log N$.*

By the method of TT^*, it certainly suffices to prove that

$$\|f * \widehat{d\sigma}\|_{18/5} \leqslant 10\log N\|f\|_{18/13}. \tag{9.1}$$

Now we know quite a lot about $\widehat{d\sigma}$ from our work in the previous set of notes. Recall that $\widehat{d\sigma}$ splits as $\delta_0 + K$, where the so-called *Bochner-Riesz* kernel $K(x_1, x_2, x_3)$ equals 0 when $x_3 = 0$ and

$$K(x_1, x_2, x_3) \ = \ -\frac{1}{N} e\left(\frac{x_1^2 + x_2^2}{4x_3}\right).$$

Now $f * \delta_0 = f$, and so the δ_0 part of (9.1) follows immediately from the fact that norms $\|\cdot\|_r$ on $\mathbb{F}^3$ are nested as r decreases. That is, we have

$$\|f * \delta_0\|_{18/5} \ \leqslant \ \|f\|_{18/13}. \tag{9.2}$$

The main content of this set of notes is the following.

Proposition 9.1. *We have the estimate*

$$\|f * K\|_4 \ \ll \ 10N^{1/8} \log N \|f\|_{8/5} \tag{9.3}$$

Proof. Let $f \in B(\mathbb{F}^3)$. For each $u \in \mathbb{F}$ write f_u for the restriction of f to the hyperplane $x_3 = u$. We will prove that for each u

$$\|f_u * K\|_4 \ \leqslant \ 10N^{-1/4} \log N \|f_u\|_{8/5}. \tag{9.4}$$

The left-hand side, raised to the power four, is

$$\sum_y \left| \sum_x f_u(x) K(y - x) \right|^4.$$

Substituting $x = x' + (0, 0, u)$ in the inner sum we see that it may be assumed, with no loss of generality, that $u = 0$. Writing $x = (\mathbf{x}, x_3)$ and $y = (\mathbf{y}, y_3)$, and recalling that

$$K(\mathbf{x}, x_3) \ = \ -N^{-1} e\left(\mathbf{x}^2/4x_3\right)$$

when $x_3 \neq 0$ and $K(\mathbf{x}, 0) = 0$, this equals

$$N^{-4} \sum_{y, y_3 \neq 0} \left| \sum_{\mathbf{x}} f(\mathbf{x}, 0) e\left(\frac{(\mathbf{x} - \mathbf{y})^2}{4y_3}\right) \right|^4. \tag{9.5}$$

Write $\mathbf{z} = -\mathbf{y}/2y_3$ and $t = 1/4y_3$. Then

$$\frac{(\mathbf{y} - \mathbf{x})^2}{4y_3} \ = \ \mathbf{z}^2 y_3 + \mathbf{x}.\mathbf{z} + t\mathbf{x}^2,$$

and so (9.5) becomes

$$N^{-4} \sum_{(\mathbf{z}, t), t \neq 0} \left| \sum_{\mathbf{x}} f(\mathbf{x}, 0) e\left((\mathbf{x}, \mathbf{x}^2).(\mathbf{z}, t)\right) \right|^4.$$

But this is precisely

$$N^{-4} \sum_{(\mathbf{z},t),t\neq 0} \left|\widehat{Gd\sigma}(\mathbf{z},t)\right|^4 \;=\; N^{-4}\|\widehat{Gd\sigma}\|_4^4,$$

where the function $G : P \to \mathbb{C}$ is defined by $G(\mathbf{x},\mathbf{x}^2) = N^2 f(\mathbf{x},0)$. This may be estimated using the inequality $R^*(18/5 \to 4) \leqslant 10\log N$ that we obtained in the last set of notes. Using the fact that $\|G\|_{8/5} = N^{3/4}\|f_u\|_{8/5}$ (n.b. the two norms here are on different spaces) we see that (9.4) does indeed hold.

Now simply observe that

$$\begin{aligned}
\|f * K\|_4 &\leqslant \sum_{u\in\mathbb{F}} \|f_u * K\|_4 \\
&\leqslant 10N^{-1/4}\log N \sum_u \|f_u\|_{8/5} \\
&\leqslant 10N^{1/8}\log N \left(\sum_u \|f_u\|_{8/5}^{8/5}\right)^{5/8} \\
&= 10N^{1/8}\log N\|f\|_{8/5}.
\end{aligned}$$

This concludes the proof of Proposition B.1. □

We are now in a position to prove (9.1) for the Bochner-Riesz part of $\widehat{d\sigma}$.

Proposition 9.2. *We have the estimate* $\|f * K\|_{18/5} \leqslant \log N\|f\|_{18/13}$.

Proof. The operator $f \mapsto f * K$ is self-adjoint, because K is the Fourier transform of a real-valued function. Thus (9.3) automatically gives an estimate

$$\|f * K\|_{8/3} \leqslant 10N^{1/8}\log N\|f\|_{4/3}.$$

Interpolating between this and (9.3) (with $t = 1/2$ in the Riesz-Thorin theorem) gives

$$\|f * K\|_{16/5} \leqslant 10N^{1/8}\log N\|f\|_{16/11}. \tag{9.6}$$

Now Young's inequality togther with the estimate $\|K\|_\infty \leqslant 1/N$ gives

$$\|f * K\|_\infty \leqslant N^{-1}\|f\|_1. \tag{9.7}$$

Interpolating between this and (9.6), with $t = 8/9$ in the Riesz-Thorin theorem, gives the bound claimed. □

This brings to an end our discussion of restriction phenomena for the discrete paraboloid. It might be conjectured that Res(2, 3) holds, at least in the sense that there are no obvious examples of functions f for which $\|(fd\sigma)^\vee\|_3$ is substantially larger than $\|f\|_2$. Needless to say, such a result is not known to be true.

10. Montgomery's conjecture and Kakeya

I am all too aware that this course contains rather a large number of conjectures. The inclusion of this section makes the situation even worse in this regard. On the plus side we will learn some interesting probabilistic lemmas, Khintchine's inequality, and we will also see that the Kakeya problem is related to an old conjecture of Montgomery concerning Dirichlet sums, the truth of which would have implications for the Riemann Hypothesis.

We begin with some preliminary lemmata. These are all of considerable importance in their own right (probably more so than the main theme of this section!).

Lemma 22. *Let $a_1, \dots, a_n$ be real numbers, and let $\epsilon_1, \dots, \epsilon_n$ be independent Bernouilli random variables (that is, ϵ_i takes each value ± 1 with probability $1/2$). Let $t > 0$ be a real number. Then*

$$\mathbb{P}\left(\sum_{i=1}^n \epsilon_i a_i \geqslant t\right) \leqslant \exp\left(-t^2/2\sum_{i=1}^n a_i^2\right).$$

Proof. Let $\mu > 0$ be a real number to be chosen later. Observe that for any real number x one has the inequality $\cosh x \leqslant \exp(x^2/2)$; to prove this, write both functions as power series, so that

$$\cosh x = \sum_{j \geqslant 0} \frac{x^{2j}}{(2j)!}$$

and

$$\exp(x^2/2) = \sum_{j \geqslant 0} \frac{x^{2j}}{2^j j!},$$

and compare term-by-term. Write $X = \sum_{i=1}^n \epsilon_1 a_i$. We have, then,

$$\begin{aligned} \mathbb{E}e^{\mu X} &= \prod_{i=1}^n \mathbb{E}e^{\mu \epsilon_i a_i} \\ &= \prod_{i=1}^n \cosh(\mu a_i) \\ &\leqslant \exp\left(\mu^2 \sum_{i=1}^n a_i^2/2\right). \end{aligned}$$

However by Markov's inequality one has

$$\mathbb{P}(X \geqslant t) = \mathbb{P}\left(e^{\mu X} \geqslant e^{\mu t}\right) \leqslant e^{-\mu t}\mathbb{E}e^{\mu X}.$$

Combining this with the above gives

$$\mathbb{P}(X \geqslant t) \leqslant \exp\left(-\mu t + \frac{1}{2}\mu^2 \sum_{i=1}^n a_i^2\right).$$

Choose $\mu = t/\sum_{i=1}^n a_i^2$, and the lemma follows immediately. □

Corollary 10.1. *Let $a_1, \dots, a_n$ be complex numbers, and let $\epsilon_1, \dots, \epsilon_n$ be independent Bernouilli random variables. Let $t > 0$ be a real number. Then*

$$\mathbb{P}\left(\left|\sum_{i=1}^n \epsilon_i a_i\right| \geqslant t\right) \leqslant 4\exp\left(-t^2/8\sum_{i=1}^n a_i^2\right).$$

Proof. Write $U = \Re\sum_{i=1}^n \epsilon_i a_i$ and $V = \Im\sum_{i=1}^n \epsilon_i a_i$. By the lemma, one has

$$\mathbb{P}(U \geqslant t/2) \leqslant \exp\left(-t^2/8\sum_{i=1}^n |\Re a_i|^2\right) \leqslant \exp\left(-t^2/8\sum_{i=1}^n |a_i|^2\right).$$

Similar inequalities hold for $\mathbb{P}(U \leqslant -t/2)$, $\mathbb{P}(V \geqslant t/2)$ and $\mathbb{P}(V \leqslant -t/2)$. However if $|U+V| \geqslant t$ then at least one of these four events must hold. □

The next result is called Khintchine's inequality. We have already seen it in action in the case $p = 4$, but there a direct expanding-out type of proof was available.

Lemma 23. *Let $p > 1$ be a real number. Then there are positive constants $C_1(p)$ and $C_2(p)$ with the following property. Let $g_i : X \to \mathbb{C}$ be functions on some measure space X. Suppose that $\epsilon_1, \dots, \epsilon_n$ are independent Bernouilli random variables. Then*

$$C_1(p)\int_X \left(\sum_{i=1}^n |g_i(x)|^2\right)^{p/2} dx \leqslant \mathbb{E}\left\|\sum_{i=1}^n \epsilon_i g_i\right\|_p^p \leqslant C_2(p)\int_X \left(\sum_{i=1}^n |g_i(x)|^2\right)^{p/2} dx.$$

Proof. We work with one point of X at a time. One has

$$\mathbb{E}\left|\sum_{i=1}^n \epsilon_i g_i(x)\right|^p = p\int_0^\infty t^{p-1}\mathbb{P}\left(\left|\sum_{i=1}^n \epsilon_i g_i(x)\right| \geqslant t\right) dt.$$

Writing $M = \sum_{i=1}^n |g_i(x)|^2$ and using Corollary 10.1, it becomes apparent that this is bounded by

$$4p\int_0^\infty t^{p-1}e^{-t^2/8M}\, dt.$$

Make the substitution $u = t^2/8M$. Then this integral becomes

$$\frac{1}{2}(8M)^{p/2}\int_0^\infty u^{p/2-1}e^{-u},$$

which is precisely

$$C_2(p)\left(\sum_{i=1}^n |g_i(x)|^2\right)^{p/2},$$

where $C_2(p) = \frac{1}{2}8^{p/2}\Gamma(p/2)$. The exact value of $C_2(p)$ is irrelevant, of course. Integrating over X gives the upper bound of the lemma. To get the lower bound one can simply

use Hölder's inequality. One has, for any $x \in X$,

$$\begin{aligned}\sum_{i=1}^n |g_i(x)|^2 &= \mathbb{E}\left|\sum_{i=1}^n \epsilon_i g_i(x)\right|^2 \\ &\leqslant \left(\mathbb{E}\left|\sum_{i=1}^n \epsilon_i g_i(x)\right|^p\right)^{1/p}\left(\mathbb{E}\left|\sum_{i=1}^n \epsilon_i g_i(x)\right|^{p'}\right)^{1/p'}. \\ &\leqslant C_2(p')^{1/p'}\left(\sum_{i=1}^n |g_i(x)|^2\right)^{1/2}.\end{aligned}$$

Cancelling the common factor, raising to the power p and integrating over X gives the lower bound claimed, with $C_1(p) \geqslant C_2(p')^{-p/p'}$. □

The next lemma, which is often replaced by the words "a simple averaging argument", is of very wide applicability.

Lemma 24. *Let Γ be an abelian group with cardinality N, and let $S \subseteq \Gamma$ have cardinality k. Then there is a set $T \subseteq \Gamma$, $|T| \leqslant \lceil N \log N/k \rceil$, such that the translates $S + t$, $t \in T$, cover Γ.*

Proof. Pick elements $x_1, \dots, x_r \in \Gamma$ uniformly at random. If $i_1 < i_2 < \cdots < i_s$ then the expected size of

$$|(S + x_{i_1}) \cap (S + x_{i_2}) \cap \cdots \cap (S + x_{i_s})|$$

is k^s/N^{s-1}. Indeed for any $y \in \Gamma$ there are exactly k^s choices of the x_{i_j} such that $y \in S + x_{i_j}$ for each $j = 1, \dots, s$. Now we may use inclusion-exclusion to calculate

$$\begin{aligned}\mathbb{E}\,|(S + x_1) \cup \cdots \cup (S + x_r)| &= \mathbb{E}\sum_{s=1}^r (-1)^{s+1} \sum_{1 \leqslant i_1 < \cdots < i_s \leqslant r} |(S + x_{i_1}) \cap \cdots \cap (S + x_{i_s})| \\ &= \sum_{s=1}^r (-1)^{s+1}\binom{r}{s}\frac{k^s}{N^{s-1}} \\ &= N\left(1 - \left(1 - \frac{k}{N}\right)^r\right) \\ &> N\left(1 - e^{-kr/N}\right).\end{aligned}$$

Set $r = \lceil N \log N/k \rceil$, and we see that

$$\mathbb{E}\,|(S + x_1) \cup \cdots \cup (S + x_r)| > N - 1.$$

In particular, there is some specific choice of $x_1, \dots, x_r$ for which the translates $S + x_i$ cover Γ completely. □

Montgomery's conjecture. Let M be a positive integer. A *Dirchlet series* of length

M is a series of the form

$$D(s) \;=\; \sum_{n=1}^{M} a_n n^{is} \;=\; \sum_{n=1}^{M} a_n e^{is \log n}.$$

This is normally thought of as a function of a *real* variable s. Now for a typical value of s one might expect the phase n^{is} to depend in a rather random manner on n. In that case one might hope for square-root cancellation, so that $D(s) \approx \sqrt{M}$. Montgomery, over thirty years ago, made a conjecture asserting that this is indeed the case in a certain sense. Montgomery's original conjecture was actually false, but the following modified form due to Bourgain seems rather plausible.

Conjecture 10.2. *Let $\nu \geqslant 1$ be a real number. Let $\epsilon > 0$ be a real number. Then there is a constant C_ϵ with the following property. Let $\{a_n\}_{n=1}^M$ be any sequence of complex numbers with $|a_n| \leqslant 1$, and suppose that $T \geqslant M^\nu$. Then*

$$\frac{1}{T}\int_0^T |D(s)|^{2\nu}\, ds \;\leqslant\; C_\epsilon M^{\nu+\epsilon}.$$

It is necessary to insist that $T \geqslant M^\nu$, or else the contribution to the integral from small values of s can dominate (for example when $a_n = 1$ for all n). It is not particularly difficult to prove the conjecture when ν is an integer, and in fact the cases $\nu = 1, 2$ are on the third example sheet. When $1 < \nu < 2$, however, the conjecture seems to lie much deeper. Indeed, as we shall show in this section, it implies the Kakeya conjecture. If that is not evidence enough for its difficulty, we remark that the conjecture seems closely related to some well-known open problems concerning the Riemann ζ-function (such as the so-called *density hypothesis*).

Montgomery's conjecture and the Kakeya problem. We begin by combinatorializing the Kakeya problem. We have already done this in various ways (by using slices, for example). Our method here will be rather different. For simplicity of notation we will think about the Kakeya problem in $\mathbb{R}^3$, but everything works in exactly the same way in higher dimensions. Let $B \subseteq \mathbb{R}^3$ be Besicovitch, and suppose without loss of generality that $B \subseteq [0,4]^3$.

Proposition 10.3. *Suppose that $\eta > 0$ has the property that $|N_\delta(B)| \leqslant \delta^{20\eta}$ for some sequence of δs tending to zero. Then there is a sequence of primes $N \to \infty$ and subsets $A \subseteq \mathbb{Z}_N$ with the following properties:*

- $|A| \leqslant N^{1-2\eta}$
- *For any $d \in \mathbb{Z}_N^*$ there is an arithmetic progression $P \subseteq A$ with length N^η and common difference d.*

Proof. Let δ be such that $|N_\delta(B)| \leqslant \delta^{20\eta}$, set $W = \lceil 2/\delta \rceil$ and divide $[0,4]^3$ into $8W^3$ little cubes of sidelength $1/W$. Index these cubes, using their lower left corners, by lattice points (i_1, i_2, i_3) with $0 \leqslant i_1, i_2, i_3 < 4W$, let $\Sigma \subseteq \mathbb{Z}^3$ be the set of cubes which have non-empty intersection with B, and let $\tilde{\Sigma} \subseteq \mathbb{Z}^3$ be the corresponding set of lattice points. Observe that Σ is contained in the neighbourhood $N_\delta(B)$, since the diameter of a cube with sidelength $1/W$ is $\sqrt{3}/W$, which is at most δ. Clearly

$$|\tilde{\Sigma}| \;\leqslant\; W^3|\Sigma| \;\leqslant\; 10\delta^{-3}|N_\delta(B)| \;\leqslant\; 10\delta^{20\eta-3}.$$

Now suppose that $d = (d_1, d_2, d_3)$ is a lattice vector with $|d| \leqslant W^{1-4\eta}$. The Besicovitch set B contains a unit line $l = x + [0,1]d/|d|$ in direction d. Suppose that x lies in the cube (i_1, i_2, i_3). Then it is not hard to see that $\tilde{\Sigma}$ contains the *arithmetic progression*

$$\left\{(i_1, i_2, i_3) + j(d_1, d_2, d_3) \,|\, j = 0, 1, \ldots, \lfloor W^{4\eta} \rfloor\right\}. \tag{10.1}$$

Consider the map

$$\psi : \mathbb{Z}^3 \longrightarrow \mathbb{Z}$$

defined by

$$\psi(i_1, i_2, i_3) \;=\; i_1 + 4Wi_2 + 16W^2 i_3.$$

This map is one-to-one on $\{0, \ldots, 4W-1\}^3$, which it maps to the set $\{0, \ldots, 64W^3\}$. Furthermore the map sends an arithmetic progression such as the one in (10.1) to an arithmetic progression in $\mathbb{Z}$ with common difference $d_1 + 4Wd_2 + 16W^2 d_3$, and thus the set $X = \psi(\tilde{\Sigma})$ contains plenty of arithmetic progressions with distinct common differences. In fact, since $B(0, W^{1-4\eta})$ contains at least $W^{3-12\eta}$ lattice points, X contains at least $W^{3-12\eta}$ arithmetic progressions of length at least $W^{4\eta}$. X has the same cardinality as $\tilde{\Sigma}$, and so $|X| \leqslant 10\delta^{20\eta-3}$.

Pick a prime $N \in [64W^3, 128W^3]$, and regard X as a subset of $\mathbb{Z}_N$. Using Lemma 24 one can construct a new set A which is still quite thin, but which contains arithmetic progressions in all directions. Indeed X contains $W^{3-12\eta}$ arithmetic progressions with length at least $W^{4\eta}$. Apply Lemma 24 to the set S of common differences of these progressions, where $S \subseteq \mathbb{Z}_N^*$. This tells us that there is some union A of $\lceil N \log N / W^{3-12\eta} \rceil$ dilates of X which contains a progression of length $W^{4\eta}$ is *every* direction. The proof of the proposition is concluded by observing that, for N sufficiently large, $N^{1-2\eta} \geqslant 10\delta^{20\eta-3}\lceil N \log N / W^{3-12\eta} \rceil$ and $N^\eta < W^{4\eta}$. □

The point of Proposition 10.3, of course, is that if $\underline{d}(B) < 3$ then there must exist some positive η satisfying the conditions. Thus if we could show that no set A of the form described can exist then we would have a proof of Kakeya in 3 dimensions (and, by an almost identical argument, in any dimension). The exact conclusion of Proposition 10.3 is not completely convenient for applying the Montgomery conjecture. Suppose

we have a set $A \subseteq \mathbb{Z}_N$ with $|A| \leqslant N^{1-2\eta}$ containing a progression of length N^η with every common difference. By subjecting A to the *unwrapping map* $\mathbb{Z}_N \hookrightarrow \{1, \dots, N\}$ one ends up with a set $A \subseteq \{1, \dots, N\}$ of the same size which contains an arithmetic progression of length L and common difference d for every $d \in [N/6L, N/3L]$, where $L = \lfloor N^\eta/2 \rfloor$. In what follows we shall, regarding η as fixed and N as large, consider such a set A. Using the Montgomery conjecture, we shall show that A cannot exist.

The Short Sum Construction. Suppose that $a, d \in \mathbb{Z}$, and consider the arithmetic progression

$$P \ = \ \{a + 2\pi d, \dots, a + 2\pi L d\} \tag{10.2}$$

of real numbers. We associate to P a so-called "short" Dirichlet sum $d_P(s,h)$ defined by

$$d_P(s,h) \ = \sum_{d \leqslant n < d+h} e^{-ia(n-d)/d} n^{is}. \tag{10.3}$$

To understand why we have made this definition, let us prove a technical lemma which will be required later.

Lemma 25. *Suppose that x, y are positive real numbers with $x \geqslant y$. Then*

$$\left| \log x - \log y - \frac{x-y}{y} \right| \ \leqslant \ \left(\frac{x-y}{y} \right)^2.$$

Proof. Indeed, we have

$$\begin{aligned} \left| \log x - \log y - \frac{x-y}{y} \right| \ &= \ \left| \int_y^x \frac{t-x}{t^2} \, dt \right| \\ &\leqslant \ |x-y| \sup_{t \in [y,x]} \left| \frac{t-x}{t^2} \right| \\ &\leqslant \ \left| \frac{x-y}{y} \right|^2. \end{aligned}$$

Let us forget for a moment the exact form of the lemma, and write $\log x \approx \log y + (x - y)/y$. Substituting into the definition (10.3) gives, heuristically at least,

$$d_P(s,h) \ \approx \ d^{is} \sum_{0 \leqslant n' < h} e^{i(s-a)n'/d}. \tag{10.4}$$

Now if s lies in P, or is close to an element of P, then the quantity $(s-a)n'/d$ is roughly 2π times an integer. Thus the exponentials in (10.4) are all roughly unity and the sum $d_P(s,h)$ is approximately h in magnitude.

We wish to extend the definition of short sum to any progression of real numbers, and in particular to progressions of integers which will interest us later. Suppose then that $P = \{a + 2\pi d, \dots, a + 2\pi L d\}$, where d need not be an integer. Associate to P the

progression $P' = \{a + 2\pi\lfloor d\rfloor, \dots, a + 2\pi L\lfloor d\rfloor\}$, which is of the special form (B.1). Set $d_P(s,h) = d_{P'}(s,h)$.

Now let P be a progression of integers of length L, $P = \{a + t, \dots, a + Lt\}$. Let $\tilde{P}$ denote the *fattening* of P, that is the set of all integers of the form $p + x$ where $p \in P$ and $|x| \leqslant N$.

Proposition 10.4. *Suppose that $\tilde{P} \subseteq \{1, \dots, 2N^2\}$, that $t \geqslant N^2/6L$ and that $h \leqslant N/200L$. Then $|d_P(s,h)| \geqslant h/2$ for all $s \in \tilde{P}$.*

Proof. Write $d = \lfloor t/2\pi\rfloor$, so that $d \geqslant N^2/40L$. If $s \in \tilde{P}$ then we can write

$$s \;=\; a + 2\pi kd + x, \tag{10.5}$$

where $k \in \{1, \dots, L\}$ and $|x| \leqslant 2N$. Of course, the short sum $d_P(s,h)$ is defined in terms of the progression $P' = \{a + 2\pi d, \dots, a + 2\pi Ld\}$.

Now using the inequality $|e^{i\theta_1} - e^{i\theta_2}| \leqslant |\theta_1 - \theta_2|$ together with Lemma 27 gives

$$\begin{aligned}\left| d_P(s,h) - d^{is} \sum_{d \leqslant n < d+h} e^{i(s-a)(n-d)/d} \right| &\leqslant \frac{|s|}{d^2} \sum_{0 \leqslant n' < h} n'^2 \\ &\leqslant h/8,\end{aligned}$$

this last fact following because $|s| \leqslant 2N^2$. With s as in (B.2), we have furthermore that

$$\begin{aligned}\left| d^{is} \sum_{d \leqslant n < d+h} e^{i(s-a)(n-d)/d} \right| &= \left| \sum_{0 \leqslant n' < h} e^{ixn'/d} \right| \\ &\geqslant \Re \sum_{0 \leqslant n' < h} e^{ixn'/d} \\ &\geqslant h\sqrt{2}/2\end{aligned}$$

Since $|hx| \leqslant \pi d/4$. The proposition follows immediately. □

Remark. This is rather closely analogous to the so-called *Knapp example* which we discussed when considering restriction phenomena of spheres.

Recall now our set $A \subseteq \{1, \dots, N\}$ containing lots of APs. Let us "fatten up" A to a subset $A' \subseteq \{1, \dots, 2N^2\}$ by decreeing that if $x \in A$ then all of the points $Nx + 1, \dots, Nx + 2NL$ are in A'. Clearly, we have

$$|A'| \;\leqslant\; 2NL|A|. \tag{10.6}$$

Now we claim that if $d \in [N^2/6L, N^2/3L]$ then A' contains $\tilde{P}$, where P is some progression with common difference d and length L. Indeed, write $d = Nd' + r$, where $d' \in [N^2/6L, N^2/3L]$ and $0 \leqslant r < N$, and suppose that the progression $\{a + d', \dots a + Ld'\}$

lies in A. Then $\tilde{P} \subseteq A'$, where $P = \{N(a+d')+r, \dots, N(a+Ld')+Lr\}$. Now select N values $d_1, \dots, d_N \in [N^2/6L, N^2/3L]$ which are $N/7L$-spaced and consider the corresponding progressions $P_1, \dots, P_N$. Associated to each of these is a short sum $d_i(s,h)$ supported on the range $\left[\left\lfloor \frac{d_i}{2\pi} \right\rfloor, \left\lfloor \frac{d_i}{2\pi} \right\rfloor + h\right]$. Take $h = \lfloor N/200L \rfloor$. Because of our choice of parameters, these intervals do not overlap. Thus any sum of the form

$$D(s) \ = \ \sum_{i=1}^{N} \epsilon_i d_i(s,h),$$

where $\epsilon_i = \pm 1$, is a Dirichlet sum of length at most N^2/L, all of whose coefficients have unit modulus. Choose the signs ϵ_i independently at random. By Khintchine's inequality we have, for any ν and any T,

$$\mathbb{E} \int_0^T |D(s)|^{2\nu} \, ds \ \geqslant \ C_1(2\nu) \int_0^T \left(\sum_{i=1}^{N} |d_i(s;h)|^2 \right)^{\nu} \, ds \tag{10.7}$$

Choose $T = 2N^2$. Then if $s \in \tilde{P}_i$, s lies in the range of the integral. Therefore the right-hand side of (10.7) is, by Proposition 10.4, at least

$$(h/2)^{2\nu} C_1(2\nu) \int_0^T \left(\sum_{i=1}^{N} \chi_{\tilde{P}_i} \right)^{\nu} \, ds.$$

Now, using Hölder's inequality, we have

$$\begin{aligned} N^2 L \ &\leqslant \ \int_0^T \left(\sum_{i=1}^{N} \chi_{\tilde{P}_i} \right) \chi_A \, ds \\ &\leqslant \ \left(\int_0^T \left(\sum_{i=1}^{N} \chi_{\tilde{P}_i} \right)^{\nu} \, ds \right)^{1/\nu} |A'|^{1/\nu'} . \end{aligned} \tag{10.8}$$

Now there is a specific choice of the signs ϵ_i for which (10.7) holds without the expectation symbol. Making such a choice, and combining (10.7) with (10.8), one gets that

$$\frac{1}{T} \int_0^T |D(s)|^{2\nu} \, ds \ \gg_{\nu} \ h^{2\nu} N^{2\nu-2} L^{\nu} |A'|^{1-\nu}.$$

The length of $D(s)$ is at most N^2/L, and the value of T is $2N^2$. The hypothesis of Montgomery's conjecture is therefore satisfied provided $\nu < 2/(2-\eta)$. For such a value of ν, one has (assuming Montgomery's conjecture) that

$$(N^2/L)^{\nu+\epsilon} \ \gg_{\nu,\epsilon} \ h^{2\nu} N^{2\nu-2} L^{\nu} |A'|^{1-\nu}.$$

Recalling that $h \gg N/L$ one gets after a slight rearrangement that

$$|A'| \ \gg_{\nu,\epsilon} \ N^{2-\frac{2\epsilon}{\nu-1}} L^{\epsilon}.$$

Since $L \gg N^\eta$ it follows that

$$|A| \gg_{\nu,\epsilon} N^{1-\eta+\epsilon\left(\eta-\frac{2}{\nu-1}\right)}.$$

For any fixed $\nu < 2/(2-\eta)$ we may choose ϵ so that this is at least $N^{1-2\eta}$.

Let us summarize what has been achieved. We have shown that Montgomery's conjecture implies that there does not exist $\eta > 0$, infinitely many primes N and subsets $A \subseteq \mathbb{Z}_N$ with $|A| \leqslant N^{1-2\eta}$ and A containing a progression of length N^η with every common difference. We showed that any Besicovitch set $B \subseteq \mathbb{R}^3$ with dimension less than 3 could be used to find such an η. Hence, assuming Montgomery's conjecture, Besicovitch subsets of $\mathbb{R}^3$ have Minkowski dimension 3.

**I suspect that any subset of $\mathbb{Z}_N$ containing a progression of length N^η with every common difference must in fact have cardinality at least $N/2$. Of course, I have no idea how to prove this since it would imply the Kakeya conjecture. Worse than that, however, I cannot even see how it would follow from Montgomery's conjecture. This question may be stated in the following alternative form (it is easy to see that they are equivalent).

Question 10.5. Let $A \subseteq \mathbb{Z}_N$ have cardinality $\lceil N/2 \rceil$. Is it true that there is some dilate λA whose maximum gap is $\alpha(N)$, where $\alpha(N)$ is a function more slowly growing than any power of N?

There are some interesting sets of size $\lfloor N/2 \rfloor$. One that intrigues me particularly is the set $A \subseteq \mathbb{Z}_N$ of quadratic residues. For every quadratic non-residue a, the dilated set aA equals $A^c \cup \{0\}$ and hence has a gap of length $n(N)$, the least quadratic non-residue modulo N. It is suspected that $n(N) \ll_\epsilon N^\epsilon$: in fact, assuming the Generalised Riemann Hypothesis (GRH), one has $n(N) \ll (\log N)^2$. Nothing of this strength is known however (the best bound, due to Burgess, states that $n(N) \ll N^{1/4\sqrt{e}+\epsilon}$).

Could it be the case, then, that to attack the Kakeya conjecture one firsts needs to despatch the GRH? I suspect that the answer to this question is no. To see why, look again at the quadratic residue example. If $n(N)$ was large then A^c would contain plenty of arithmetic progressions in different directions, but these progressions are all centred on the point 0. The analagous geometric situation, in which one has a large number of lines emanating from a single point, is well understood (the lines intersect only at that point). Thus I believe that the link between the Kakeya problem and number theory is slightly artifical as regards mounting an attack on Kakeya. It might well be the case, however, that by understanding geometrical versions of restriction and Kakeya phenomena one might be able to better understand Montgomery's conjecture.**

11. $\Lambda(p)$-sets

In this set of notes we begin to move away from what might classically be called restriction phenomena. All that follows is, however, very much in the spirit of restriction theory – that is, trying to understand the structural properties of some set S by looking at L^p norms of sums of exponentials supported on S. I hope that the work we did on the discrete parabola will make at least some of what follows seem natural, perhaps even obvious.

Let $p > 2$, and suppose that $S \subseteq \mathbb{Z}$. We say that S is a $\Lambda(p)$ set with constant $K_p = K_p(S)$ if

$$\left\| \sum_{n \in S} a_n e^{in\theta} \right\|_{L^p(S^1)} \leqslant K_p \left(\sum_{n \in S} |a_n|^2 \right)^{1/2} \tag{11.1}$$

for all choices of complex numbers $\{a_n\}_{n \in S}$. We will normally assume that $K_p(S)$ is the least constant for which this inequality is always satisfied. Observe that by Parseval's identity

$$\left\| \sum_{n \in S} a_n e^{in\theta} \right\|_{L^2(S^1)} = (2\pi)^{1/2} \left(\sum_{n \in S} |a_n|^2 \right)^{1/2}.$$

Thus a set has small $\Lambda(p)$-constant if the L^p-norm of any function $f : S^1 \to \mathbb{C}$ which is *S-spectral* (that is, its Fourier transform is supported on S) is comparable to its L^2-norm.

Various natural questions (in addition to the question "why make such a definition?") present themselves. Are there any $\Lambda(p)$-sets with small constant K_p? If so, what is the largest $\Lambda(p)$-subset of $\{1, \dots, N\}$? We address this last question first of all.

Lemma 26. *Suppose that $S \subseteq \{1, \dots, N\}$. Then $|S| \leqslant 4K_p(S)^2 N^{2/p}$.*

Proof. Test (11.1) with the constant sequence $a_n = 1$. When $|\theta| \leqslant 1/2N$ we have $\cos n\theta \geqslant 1/2$ for all $n \in S$, and so for this range of θ

$$\left| \sum_{n \in S} e^{in\theta} \right| \geqslant |S|/2.$$

This implies that

$$\left\| \sum_{n \in S} e^{in\theta} \right\|_p \geqslant \frac{|S|}{2N^{1/p}}.$$

However, the definition of $\Lambda(p)$ implies that the left-hand side is also at most $K_p(S)|S|^{1/2}$. Thus we have the inequality

$$\frac{|S|}{2N^{1/p}} \leqslant K_p(S)|S|^{1/2},$$

which immediately implies the lemma. □

The question of whether, for a fixed p, there exists a sequence of sets $S_N \subseteq \{1,\dots,N\}$ with cardinalities $\gg N^{2/p}$ and with uniformly bounded $\Lambda(p)$ constant was famously unsolved for many years. In 1988 Bourgain showed that in fact a *random* subset of $\{1,\dots,N\}$ of cardinality about $N^{2/p}$ has bounded $\Lambda(p)$ constant. The proof is a tour de force and uses tools such as entropy and decoupling (about which I do not pretend to understand very much). It is, however, not particularly difficult to construct $\Lambda(p)$-sets of essentially optimal size when p is an even integer.

Proposition 11.1. *Let N be a positive integer and let $p = 2h$ be an even integer. Then there is a set $S \subseteq \{1,\dots,N\}$ with $|S| \geqslant \frac{1}{2}N^{2/p}$ and $K_p(S) \leqslant 3\sqrt{h}$.*

Proof. We begin by constructing something called a B_h-set. This is a large subset of $\{1,\dots,N\}$ which has the property that the only solutions to the equation

$$x_1 + \cdots + x_h \;=\; x'_1 + \cdots + x'_h$$

are ones in which the set $\{x_1,\dots,x_h\}$ is just a rearrangement of $\{x'_1,\dots,x'_h\}$. To do this, take a prime q such that $2^{-h}N < q^h \leqslant N$ (this is possible by Bertrand's theorem). Consider the finite field $K = \mathbb{F}_{q^h}$, the subfield $L \subseteq K$ with $L \cong \mathbb{F}_q$ and a generator a of the cyclic group K^*. Let $X \subseteq \mathbb{Z}/(q^h-1)\mathbb{Z}$ be the set of all θ for which

$$a^\theta - a \;\in\; L.$$

We claim that X is a B_h-subset of $\mathbb{Z}/(q^h-1)\mathbb{Z}$. Suppose, indeed, that

$$\theta_1 + \cdots + \theta_h \;=\; \theta'_1 + \cdots + \theta'_h,$$

where $a^{\theta_i} = a + \lambda_i$ and $a^{\theta'_i} = a + \lambda'_i$. Then we have

$$(a+\lambda_1)\times\cdots\times(a+\lambda_h) \;=\; (a+\lambda'_1)\times\cdots\times(a+\lambda'_h).$$

Since the minimal polynomial of a over L has degree h, this identity must collapse so that we have

$$\sum_i \lambda_i \;=\; \sum_i \lambda'_i,$$
$$\sum_{i<j} \lambda_i\lambda_j \;=\; \sum_{i<j} \lambda'_i\lambda'_j$$

and so on for the other symmetric functions. Thus $\{\lambda_1,\dots,\lambda_h\}$ and $\{\lambda'_1,\dots,\lambda'_h\}$ are both the set of h roots of some polynomial of degree h. Hence these sets must be rearrangements of one another, which confirms our claim that X is a B_h-set.

At the moment X is a subset of $\mathbb{Z}/(q^h-1)\mathbb{Z}$, but by picking the least positive residue of each class in X we may construct a B_h-subset $X \subseteq \{1,\dots,N\}$ with cardinality at least

$N^{1/h}/2$. We claim that $K_p(X) \leqslant (2\pi h!)^{1/p} \leqslant 3\sqrt{h}$. Indeed we have, for any sequence $\{a_n\}_{n\in X}$,

$$\begin{aligned}\left\|\sum_{n\in X} a_n e^{in\theta}\right\|_p^p &= 2\pi \sum_{\substack{n_1+\cdots+n_h=\\ n_1'+\cdots+n_h'}} a_{n_1}\dots a_{n_h}\overline{a_{n_1'}}\dots\overline{a_{n_h'}} \\ &\leqslant 2\pi h! \sum_{n_1,\dots,n_h} |a_{n_1}|^2\dots|a_{n_h}|^2 \\ &= 2\pi h!\left(\sum_{n\in X}|a_n|^2\right)^{p/2}.\end{aligned}$$

Checking that $(2\pi h!)^{1/p} \leqslant 3\sqrt{h}$ is a rather simple matter which we leave to the reader. □

Eigenvectors of the Laplacian. The definition of $\Lambda(p)$ constant makes perfect sense in $\mathbb{Z}^d$, except that instead of taking norms on S^1 one must take them on the d-dimensional torus $\mathbb{T}^d$. There are various arithmetically-defined sets for which estimates of the $\Lambda(p)$ constant of that set would have very interesting consequences. One such arises in discussing eigenvectors of the Laplacian on the 2-torus $\mathbb{T}^2$.

Functions on $\mathbb{T}^2$ may be regarded as functions of $x = (x_1, x_2)$ which have any vector $(2\pi n_1, 2\pi n_2)$ as a period. The Laplacian is defined, for functions $f \in C^\infty(\mathbb{T}^2)$, by

$$\Delta f = \frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2}.$$

An eigenvector of the Laplacian is a function f such that $\Delta f + \lambda f = 0$ for some λ. A fact, which we shall use without proof, is that the only eigenvalues are $\lambda = m_1^2 + m_2^2$, where $m_1, m_2 \in \mathbb{Z}$. The eigenspace corresponding to λ is generated by all exponentials $e^{im.x}$, where $|m|^2 = m_1^2 + m_2^2 = \lambda$. You can easily check that these exponentials *are* eigenvectors of the Laplacian.

A general eigenvector of Δ with eigenvalue λ therefore has the form

$$\phi_\lambda(x) = \sum_{|m|^2=\lambda} a_m e^{im.x}. \tag{11.2}$$

This makes it clear that the the L^p norms of eigenvectors of Δ are tied up with the $\Lambda(p)$-properties of sets such as $\{(m_1, m_2) : m_1^2 + m_2^2 = \lambda\}$.

Proposition 11.2. *Suppose that ϕ_λ is an eigenvector of the Laplacian on $\mathbb{T}^2$ with eigenvalue λ. Then we have the estimate $\|\phi_\lambda\|_4 \leqslant 3^{1/4}\|\phi_\lambda\|_2$.*

Proof. Write ϕ_λ in the form (11.2). Then

$$\|\phi_\lambda\|_4^4 \;=\; 4\pi^2 \sum_{\substack{m_1+m_2=m_3+m_4 \\ |m_i|^2=\lambda}} a_{m_1} a_{m_2} \overline{a_{m_3} a_{m_4}}. \tag{11.3}$$

Now there are rather few solutions to the equation $m_1 + m_2 = m_3 + m_4$. Indeed, think geometrically, so that the m_i are points on a circle of radius $\sqrt{\lambda}$. Suppose that $m_1 + m_2 = m_3 + m_4$, and consider the midpoint of $m_1 m_2$. If this is not the origin then we must have $\{m_3, m_4\} = \{m_1, m_2\}$, because given a point inside a circle and not at its centre there is a unique chord bisected by it. Thus if $\{m_3, m_4\} \neq \{m_1, m_2\}$ then we must have $m_2 = -m_1$ and $m_4 = -m_3$. Referring back to (11.3), this implies that

$$\begin{aligned}
(4\pi^2)^{-1}\|\phi_\lambda\|_4^4 &\leqslant 2 \sum_{m_1,m_2} |a_{m_1}|^2 |a_{m_2}|^2 + \sum_{m_1,m_2} |a_{m_1} a_{-m_1} a_{m_2} a_{-m_2}| \\
&\leqslant 3 \left(\sum_m |a_m|^2 \right)^2 \\
&= 3 \cdot (4\pi^2)^{-1} \|\phi_\lambda\|_2^4,
\end{aligned}$$

which is what we wished to prove. □

Amazingly, it is not known whether a similar result holds for $\|\phi_\lambda\|_6$. This seems to involve some tricky number theory. Some other rather nice open problems are as follows:

Problem 11.3. Is there any $p > 2$ such that one has an estimate $\|\phi_\lambda\|_p \ll \|\phi_\lambda\|_2$ for all eigenvectors ϕ_λ of the Laplacian on $\mathbb{T}^3$?

Problem 11.4. Are the squares a $\Lambda(p)$-set for any $p > 2$?

Regarding this last question, it is not particularly hard to see that the squares are *not* a $\Lambda(4)$-set.

12. Beckner's inequality

In this set of notes we work with the finite field $F = \mathbb{F}_2$ and the vector spaces F^n over it. Rather annoyingly our policy of using the counting measure on the spatial side and normalised counting measure on the Fourier side is not particularly natural here. For this reason we adopt the convention that the sum symbol Σ refers to counting measure (as usual) whilst the integral symbol $\int$ refers to normalised counting measure. Norms, $\|.\|_p$, will be taken with respect to the *normalised* counting measure.

If X is a finite measure space then $B(X)$ will refer to the *real-valued* functions $f : X \to \mathbb{R}$. For each $\xi \in F_*^n$ define the character $u_\xi(x) = (-1)^{\xi.x}$, and given $f \in B(F^n)$ define

the Fourier transform

$$\widehat{f}(\xi) \;=\; \int f(x)u_\xi(x)\,dx.$$

Observe that the Fourier transform of a real-valued function is always real-valued. The possibility of dealing with the Fourier transform without worrying about complex-valued functions is a special feature of working in characteristic two. We will celebrate by omitting complex conjugation from any inner products or instances of Parseval's identity that we might encounter. Note also that $\widehat{f}(\xi) = \widehat{f}(-\xi)$, so there is no need to use the symbol $f^\vee$ when working in characteristic two.

Given $\xi \in F_*^n$ we write $|\xi|$ for the number of non-zero components of ξ with respect to the standard basis of F_*^n.

To begin with suppose that $n = 1$. Let $\epsilon \in (0,1)$, and define a map $T : B(F) \to B(F)$ by

$$Tf(x) \;=\; \sum f(y)K(x,y),$$

where

$$K(x,y) \;=\; \int_\xi \epsilon^{|\xi|}u_\xi(x)u_\xi(y)\,d\xi.$$

It is easy to write down a more explicit form for T. Indeed it is a trivial matter to check that $K(0,0) = K(1,1) = \frac{1}{2}(1+\epsilon)$ and that $K(0,1) = K(1,0) = \frac{1}{2}(1-\epsilon)$. Suppose that a function $f \in B(F)$ is given by $f(0) = a - b$ and $f(1) = a + b$ (this way of writing things is very convenient). Then we see that $(Tf)(0) = a - \epsilon b$ and $(Tf)(1) = a + \epsilon b$.

Lemma 27. *One has* $\|Tf\|_2 \leqslant \|f\|_{1+\epsilon^2}$.

Proof. Set $p = 1 + \epsilon^2$. We need to prove the inequality

$$(a^2 + \epsilon^2 b^2)^{p/2} \;\leqslant\; \frac{|a+b|^p + |a-b|^p}{2}.$$

If $a = 0$ this is rather easy, and if $a \neq 0$ we may de-homogenize the inequality by dividing both sides by a^p, reducing it to

$$(1 + \epsilon^2 y^2)^{p/2} \;\leqslant\; \frac{|1+y|^p + |1-y|^p}{2}. \tag{12.1}$$

Suppose first of all that $|y| < 1$. Then the right-hand side of (B.1) may be expanded using the binomial theorem as

$$\sum_{k=0}^{\infty} \binom{p}{2k} y^{2k}. \tag{12.2}$$

Now for fixed $\lambda < 1$ the function $g(x) = (1+x)^\lambda - 1 - \lambda x$ has negative derivative on the interval $(0,\infty)$. It follows that $g(x) \leqslant 0$ for all non-negative x, and so the left-hand side of (B.1) is no more than $1 + p\epsilon^2 y^2/2$. This expression is equal to the first two terms of the binomial expansion (B.2). To confirm (B.1) in the case $|y| < 1$ one only need

observe that, since $1 < p < 2$, all of the binomial coefficients in (B.2) are positive.

When $y = \pm 1$ the inequality (B.1) follows from what we have just done by a limiting argument. Suppose, then, that $|y| > 1$. Set $z = 1/y$, and rewrite (B.1) as

$$(z^2 + \epsilon^2)^{p/2} \leqslant \frac{|1+z|^p + |1-z|^p}{2}.$$

Observe, however, that

$$\begin{aligned}(1 + \epsilon^2 z^2) - (z^2 + \epsilon^2) &= (1 - z^2)(1 - \epsilon^2) \\ &\geqslant 0,\end{aligned}$$

and so this new inequality follows immediately from (B.1) in the case $|y| < 1$ (which we have already proved). □

Now given linear maps $T_i : B(X_i) \to B(X_i)$ written in the form $T_i f(x) = \sum f(y) K_i(x, y)$, $i = 1, 2$, one may define the product map

$$T_1 \otimes T_2 : B(X_1 \times X_2) \to B(X_1 \times X_2)$$

by

$$\left((T_1 \otimes T_2) f\right)(x_1, x_2) = \sum f(y_1, y_2) K_1(x_1, y_1) K_2(x_2, y_2).$$

Lemma 28. *Suppose that T_1, T_2 are two maps with $\|T_i\|_{p\to 2} \leqslant 1$. Then $\|T_1 \otimes T_2\|_{p\to 2} \leqslant 1$.*

Proof. Using the integral Minkowski inequality (see the 7th set of notes) one has, for any $f \in B(X_1 \times X_2)$,

$$\begin{aligned}\|(T_1 \otimes T_2) f\|_2^2 &= \int\int \left|(T_1 \otimes T_2) f(x_1, x_2)\right|^2 dx_1\, dx_2 \\ &\leqslant \int \left(\int |T_2 f(x_1, x_2)|^p \, dx_2 \right)^{2/p} dx_1 \\ &\leqslant \left(\int \left(\int |T_2 f(x_1, x_2)|^2 \, dx_2 \right)^{p/2} dx_1 \right)^{2/p} \\ &\leqslant \left(\int\int |f(x_1, x_2)|^p \, dx_1\, dx_2 \right)^{2/p} \\ &= \|f\|_p^2.\end{aligned}$$

Now using the map $T : B(F) \to B(F)$ we may construct its nth power $T^{\otimes n} : B(F^n) \to B(F^n)$. Combining lemmas 27 and 29, it can be seen that $\|T^{\otimes n}\|_{1+\epsilon^2 \to 2} \leqslant 1$. But what

is this map? We have

$$\begin{aligned}(T^{\otimes n}f)(x_1,\dots,x_n) &= \sum f(y_1,\dots,y_n)\int_{\xi_i}\epsilon^{|\xi_1|+\cdots+|\xi_n|}\prod u_{\xi_i}(x_i)\prod u_{\xi_i}(y_i)\prod d\xi_i\\ &= \sum f(y)\int_\xi \epsilon^{|\xi|}u_\xi(x)u_\xi(y)\,d\xi\\ &= \int_\xi \epsilon^{|\xi|}\widehat{f}(\xi)u_\xi(x)\,d\xi.\end{aligned}$$

Write $T = T^{\otimes n}$ for simplicity of notation. This formula shows that T maps f to a function whose Fourier transform is precisely $\epsilon^{|\xi|}\widehat{f}(\xi)$. Observe how the map gives weight to Fourier coefficients at frequencies which have small support (meaning very few non-zero coefficients when written in the standard basis on F_*^n). Let us record what what we have proved.

Proposition 12.1 (Beckner's inequality)**.** *Let $\epsilon \in (0,1)$. Define a map $T : B(F^n) \to B(F^n)$ by mapping f to the function whose Fourier transform at $\xi \in F_*^n$ is $\epsilon^{|\xi|}\widehat{f}(\xi)$. Then $\|Tf\|_2 \leqslant \|f\|_{1+\epsilon^2}$.*

Now as with any map which acts simply as a multiplier on the Fourier side, T is self-adjoint. Indeed for any functions $f, g \in B(F^n)$ one has, using Parseval's identity,

$$\begin{aligned}\sum f(x)Tg(x) &= \int \widehat{f}(\xi)\widehat{Tg}(\xi)\,d\xi\\ &= \int \epsilon^{|\xi|}\widehat{f}(\xi)\widehat{g}(\xi)\,d\xi\\ &= \int \widehat{Tf}(\xi)\widehat{g}(\xi)\,d\xi\\ &= \sum Tf(x)g(x).\end{aligned}$$

Thus the general theory of operators that we developed earlier in the course allows us to conclude from Proposition B.3 that

$$\|Tf\|_{1+\epsilon^{-2}} \leqslant \|f\|_2. \tag{12.3}$$

In this dual form, Beckner's inequality may be used to deduce a statement (which is pretty much equivalent to the original inequality) concerning the $\Lambda(p)$-constant of the subset of F_*^n consisting of all ξ with $|\xi| = k$.

Proposition 12.2. *Let k be a positive integer, and let Λ be the set of all $\xi \in F_*^n$ for which $|\xi| = k$. Then for any $p > 2$ and any sequence $\{a_\xi\}_{\xi\in\Lambda}$ one has*

$$\left\|\sum_{\xi\in\Lambda} a_\xi(-1)^{\xi.x}\right\|_p \leqslant (p-1)^{k/2}\left(\sum_{\xi\in\Lambda}|a_\xi|^2\right)^{1/2}.$$

Proof. Define a function $f : F^n \to \mathbb{R}$ by setting $\widehat{f}(\xi) = 2^n a_\xi$ if $\xi \in \Lambda$ and $\widehat{f}(\xi) = 0$ otherwise (it is prefectly valid to define a function via its Fourier transform - just use the inversion formula). Observe that

$$\sum_{\xi\in\Lambda} a_\xi (-1)^{\xi.x} \;=\; \int_\xi \widehat{f}(\xi) u_\xi(x)\, d\xi. \tag{12.4}$$

Set $\epsilon = (p-1)^{-1/2}$, and observe that the right-hand side of (12.4) is equal to

$$\epsilon^{-k} \int_\xi \epsilon^{|\xi|} \widehat{f}(\xi) u_\xi(x)\, d\xi \;=\; \epsilon^{-k} Tf.$$

Since $p = 1 + \epsilon^{-2}$ this means, using (12.3), that

$$\begin{aligned} \left\| \sum_{\xi\in\Lambda} a_\xi (-1)^{\xi.x} \right\|_p &= (p-1)^{k/2} \|Tf\|_p \\ &\leqslant (p-1)^{k/2} \|f\|_2 \\ &= (p-1)^{k/2} \left(\sum_{\xi\in\Lambda} |a_\xi|^2 \right)^{1/2}. \end{aligned}$$

This completes the proof of the proposition. □

In keeping with the general philosophy of this course, we can interpret this proposition as a Fourier-analytic manifestation of the fact that the set of all $\xi \in F_*^n$ with $|\xi| = k$ does not have a great deal of additive structure. This is most evident when $k = 1$, when this set has no non-trivial additive relations whatsoever.

13. The influence of boolean functions

In this set of notes write $N = 2^n$.

A *boolean function* in n variables is a map from $\{0,1\}^n$ to $\{0,1\}$. Such an object may equivalently be regarded as a map from $\mathbb{F}_2^n$ to $\{0,1\}$, or as a subset of the power set $\mathcal{P}([n])$. In this latter incarnation, which is also known as a set-system, one identifies a function f with the collection of sets A such that $f(\chi_A(1), \dots, \chi_A(n)) = 1$.

We will typically suppose that the variables of our boolean function are $x_1, \dots, x_n$. For each $k = 1, \dots, n$ write

$$(\sigma_k f)(x_1, \dots, x_n) \;=\; f(x_1, \dots, x_{k-1}, x_k, x_{k+1}, \dots, x_n) - f(x_1, \dots, x_{k-1}, 1-x_k, x_{k+1}, \dots, x_n).$$

Define the kth influence $I_k(f)$ by

$$I_k(f) \;=\; N^{-1} \sum_{x_i} |(\sigma_k f)(x_1, \dots, x_n)|\,.$$

This may be interpreted as the probability that if $x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n$ are chosen at random then $f(x_1, \dots, x_n)$ is still undetermined (that is, depends on the value of x_k).

Example. Define the function f by

$$f(x_1, \dots, x_n) \ = \ x_1 + \dots + x_n$$

(where, of course, the addition is taken modulo 2). Then the value of f is never determined by the values of $x_2, \dots, x_n$ and so $I_1(f) = 1$. Similarly, $I_k(f) = 1$ for all k. □

Write $\mathbb{P}$ for the uniform probability measure on $\mathbb{F}_2^n$ (which is the same as the normalised counting measure). Write $E = \mathbb{E}f$. If $1/4 \leqslant E \leqslant 3/4$ (a slightly arbitrary choice) we say that f is *quite fair*.

Example. Consider the function $f(x_1, \dots, x_n) = x_1 x_2 \dots, x_n$ (a computer scientist might write this $x_1 \wedge \dots \wedge x_n$). Then f is determined by $x_2, \dots, x_n$ unless $x_2 = \dots = x_n = 1$, which occurs with probability just 2^{1-n}. Thus in this case all of the influences of f are tiny. However, f is nowhere near being quite fair. □

Proposition 13.1. *There is a quite fair boolean function f, all of whose influences are at most $2 \log n/n$.*

Proof. Let m be an integer to be chosen later, let $r = \lfloor n/m \rfloor$ and set

$$f(x_1, \dots, x_n) \ = \ x_1 x_2 \dots x_m + x_{m+1} x_{m+2} \dots x_{2m} + \dots + x_{(r-1)m+1} x_{(r-1)m+2} \dots x_{rm}.$$

The influences $I_{rm+1}(f), \dots, I_n(f)$ are clearly all zero, and all of the other influences are exactly 2^{1-m}. To see this, observe that the only choices of $x_2, \dots, x_{rm}$ for which f remains undetermined are those in which $x_2 = \dots = x_m = 1$. The function f is only quite fair if m is relatively small, and the majority of the work in this proposition goes into decided just how small.

Now if the x_i are chosen independently and at random then the r variables

$$Y_i \ = \ x_{(i-1)m+1} \dots x_{im}$$

are independent Bernouilli random variables with $\mathbb{P}(Y_i = 1) = q = 2^{-m}$. $\mathbb{P}(f = 0)$ is then precisely

$$\sum_{0 \leqslant j \leqslant r/2} \binom{r}{2j} q^{2j} (1-q)^{r-2j},$$

which is just

$$\frac{1}{2} \left(1 + (1-2q)^r\right).$$

Set $m = \log n - \log\log n + C$, where the logarithms are to the base two and $C \in [0,1)$ is selected so that m is an integer. Then

$$\begin{aligned}(1-2q)^r &= \left(1 - \frac{\log n}{2^{C-1}n}\right)^{\lfloor n/m \rfloor} \\ &= \left(1 - \frac{\log n}{2^{C-1}n}\right)^{\frac{n}{\log n}(1+o(1))} \\ &= (1+o(1))\exp\left(-2^{1-C}\right) \\ &\leqslant 1/2\end{aligned}$$

for n sufficiently large. Thus, for this value of m, the function f is quite fair.

Now we have

$$\begin{aligned}\max_{1\leqslant k\leqslant n} I_k(f) &= 2^{1-m} \\ &\leqslant \frac{2\log n}{n},\end{aligned}$$

as claimed. □

Our main objective in this set of notes will be to show that this example is basically sharp; *every* boolean function has at least one variable whose influence is $\gg E\log n/n$.

Influences and Fourier analysis. To link the quantities $I_k(f)$ with the Fourier transform, observe that

$$I_k(f) = N^{-1}\sum_x \sigma_k f(x)^2.$$

Now $\sigma_k f$ is equal to $f * J_k$, where J_k is the function defined by $J_k(0,\dots,0) = 1$,

$$J_k(0,\dots,0,1,0,\dots,0) = 1$$

(where the 1 is in the kth position) and $J_k(x) = 0$ otherwise. Thus

$$\begin{aligned}I_k(f) &= N^{-1}\sum_x \sigma_k f(x)^2 \\ &= N^{-1}\sum_x (f * J_k)(x)^2 \\ &= N^{-1}\int_\xi \widehat{f}(\xi)^2\widehat{J}_k(\xi)^2\,d\xi.\end{aligned}$$

The Fourier transform $\widehat{J}_k(\xi)$ is easily computed. It is equal to 2 if $\xi_k = 1$, and 0 if $\xi_k = 0$. Thus we have the important formula

$$I_k(f) = 4\cdot N^{-1}\int_{\xi:\xi_k=1}\hat{f}(\xi)^2 d\xi. \tag{13.1}$$

This immediately gives another important formula,

$$\sum_{k=1}^{n} I_k(f) \;=\; 4 \cdot N^{-1} \int_{\xi} |\xi| \hat{f}(\xi)^2 \, d\xi. \tag{13.2}$$

The Fourier transform of $\sigma_k f$ is also worth recording in its own right:

$$\widehat{\sigma_k f}(\xi) \;=\; \begin{cases} 2\widehat{f}(\xi) & (\xi_k = 1) \\ 0 & (\xi_k = 0). \end{cases} \tag{13.3}$$

Theorem 19. *Let f be any boolean function with $E \leqslant 1/2$. Then there is a value of k such that $I_k(f) \geqslant E \log n/15n$.*

Now the function $\sigma_k f$ takes the values ± 1. One can think of this function being rather like the characteristic function of some set, the size of this set being proportional to $I_k(f)$. Suppose that $g : F^n \to \{-1, 0, 1\}$ is some function for which the set $A_g = \{x : g(x) \neq 0\}$ has size αN. Then

$$\int_{\xi} \widehat{g}(\xi)^2 \, d\xi \;=\; \sum_{x} g(x)^2 \;=\; \alpha N. \tag{13.4}$$

We will use Beckner's inequality to show that if α is small then only a tiny proportion of this L^2-norm can be concentrated at frequencies ξ with $|\xi|$ less than about $\log n$. Thus either there is some large $I_k(f)$ or else the contribution to (13.2) from small $|\xi|$ is negligable. But then the right-hand side of (13.2) must be quite large, and in fact $\sum_{k=1}^{n} I_k(f)$ must be at least $CE \log n$. This will lead to a proof of Theorem 19.

Observe that this slightly complicated strategy is necessary. Indeed it is not the case that $\sum_k I_k(f) \gg \log n$ for all functions f, the function $f(x_1, \ldots, x_n) = x_1$ being an example where such an inequality fails. In this example, however, there is one influence, $I_1(f)$, which is very large indeed.

Proposition 13.2. *Consider a function g as above, and let M be a positive integer. Then*

$$\int_{\xi : |\xi| \leqslant M} \widehat{g}(\xi)^2 \, d\xi \;\leqslant\; 4^M \alpha^{8/5} N.$$

Remark. If M and α are both small, then this is an insignificant fraction of αN.

Proof. Using Beckner's inequality,

$$\begin{aligned}\int_{\xi:|\xi|\leqslant M} \widehat{g}(\xi)^2\,d\xi &\leqslant 4^M \int_{\xi} (1/2)^{2|\xi|}\widehat{g}(\xi)^2\,d\xi \\ &= 4^M \int \widehat{T_{1/2}g}(\xi)^2\,d\xi \\ &= 4^M \sum_x T_{1/2}g(x)^2 \\ &= 4^M N\|T_{1/2}g\|_2^2 \\ &\leqslant 4^M N\|g\|_{5/4}^2 \\ &= 4^M \alpha^{8/5} N.\end{aligned}$$

Remark. In fact, by carefully choosing some ϵ to play the rôle of $1/2$, one can improve this proposition so as to conclude that

$$\int_{\xi:|\xi|\leqslant M} \widehat{g}(\xi)^2\,d\xi \;\leqslant\; \left(\frac{10\log(1/\alpha)}{M}\right)^M \alpha^2 N,$$

provided that $M \leqslant \log(1/\alpha)$. We do not, however, need this strong form and dealing with it turns out to be slightly messier than might be desired.

Another remark. Although we have used Beckner's inequality directly, Proposition 13.2 is really a $\Lambda(p)$-set phenomenon. That is, a ± 1 function with small support cannot have a lot of L^2-Fourier weight on any set with small $\Lambda(p)$-constants. We will see another instance of this phenomenon in the next set of notes, where we will proceed directly from the $\Lambda(p)$ property.

Now if $I_k(f) \geqslant En^{-3/4}$ for any k then Theorem 19 is proved. Suppose, then, that $I_k(f) \leqslant En^{-3/4}$ for all k. The following is an immediate consequence of (13.3) and Proposition 13.2.

Proposition 13.3. *For each k,*

$$\int_{\substack{\xi:\xi_k=1\\ |\xi|\leqslant \log n/20}} \widehat{f}(\xi)^2\,d\xi \;\leqslant\; n^{-11/10}EN. \qquad \square$$

Summing over k gives

$$\int_{\substack{|\xi|\leqslant \log n/20\\ \xi\neq 0}} \widehat{f}(\xi)^2\,d\xi \;\leqslant\; n^{-1/10}EN.$$

However, by Parseval's identity, we have

$$\int_{\xi\neq 0} \widehat{f}(\xi)^2\,d\xi \;=\; (E-E^2)N \;\geqslant\; EN/2.$$

Thus

$$\int_{\substack{|\xi|\geqslant \log n/20\\ \xi\neq 0}} \widehat{f}(\xi)^2\,d\xi \;\geqslant\; EN/3$$

and, by (13.2),

$$\begin{aligned}\sum_k I_k(f) &= 4N^{-1}\int |\xi|\widehat{f}(\xi)^2\,d\xi \\ &\geqslant E\log n/15.\end{aligned}$$

This completes the proof of Theorem 19. □

14. Sumsets in $\mathbb{F}_2^n$

Let C and D be subsets of an abelian group Γ. The *sumset* $C+D$ is defined to be the set of all elements of Γ which have the form $c+d$, with $c\in C$ and $d\in D$. My original intention was to present a result which says that if $\Gamma = \mathbb{Z}/N\mathbb{Z}$ and if $|C|, |D| \gg N$ then $C+D$ contains an arithmetic progression of length about $e^{\sqrt{\log N}}$ (this is to be thought of as quite long). I can supply any interested readers with a paper where this is done, but I thought that I would take the opportunity to prove an analagous result for $\Gamma = \mathbb{F}_2^n$. This has two advantages; firstly, we can use results (such as Beckner's inequality) from the previous two sets of notes and secondly the argument comes out much more cleanly than it does for $\mathbb{Z}/N\mathbb{Z}$. The main theorem of this set of notes, then, is the following.

Theorem 20. *Suppose that γ and δ are real numbers with $\gamma\delta \geqslant 1/\sqrt{n}$. Let $C, D \subseteq \mathbb{F}_2^n$ have cardinalities γN and δN respectively. Then $C+D$ contains a translate of some subspace of $\mathbb{F}_2^n$ with dimension at least $n\gamma\delta/80$.*

**Before proving this result I would like to spend some time discussing a few examples. These should give you some idea of what the result is saying, and why it is (I think) interesting. Thanks to Oleg Pikhurko and Tom Körner for pushing me into discussing these things in lectures.

Let us begin by remarking that we will think of γ and δ in Theorem 20 as being fixed, and of n as being large (though, as stated, the theorem covers some situations where γ and δ tend to zero slowly with n). In all our examples γ and δ will be fixed reals of moderate size.

Example 1. $C = D$ is a subspace of codimension 2. Then $\gamma = \delta = 1/4$ and $C+D$ contains a subspace of dimension $n-2$, rather obviously.

Example 2. C is generated by choosing each element of $\mathbb{F}_2^n$ to lie in C independently and at random with probability $1/4$. D is generated similarly. Then with very high probability both γ and δ will be about $1/4$ (exercise). Now, for any fixed $x\in\mathbb{F}_2^n$ there

are $\lfloor N/2 \rfloor$ completely disjoint pairs (u,v) with $u+v=x$. For each such pair there is a probability $1/16$ that $u \in C$ and $v \in D$, and so the probability that $x \notin C+D$ is bounded above by $(15/16)^{\lfloor N/2 \rfloor}$. It follows that

$$\mathbb{E}|C+D| \geqslant N\left(1-(15/16)^{\lfloor N/2 \rfloor}\right),$$

which is very, very, close to N. In fact one can show that $C+D$ is all of $\mathbb{F}_2^n$ with high probability by bounding $\mathbb{E}|(C+D)^c|^2$ and using Chebyshev's inequality (I have put this on the third example sheet).

Examples 1 and 2 show that if we take either very structured sets *or* very random sets then $C+D$ will contain a translate of a subspace having very large dimension indeed. Quite often, in a combinatorial problem, either very structured examples or random examples are extremal. This problem is different, as the following modification of an example of Ruzsa shows.

Example 3. Let $C=D$ be the set of all vectors $x \in \mathbb{F}_2^n$ with at least $n/2+\sqrt{n}/2$ ones with respect to the standard basis. By the central limit theorem the number of ones in a random vector $(x_1,\dots,x_n)$ is roughly normally distributed with mean $n/2$ and standard deviation $\sqrt{n}/2$, and so for large n both γ and δ are at least $1/4$. Now any vector $x \in C+D$ must have at least $\sqrt{n}$ zeros. Using this fact, we shall prove that $C+D$ meets all translates of all $(n-\lfloor\sqrt{n}\rfloor)$-dimensional subspaces. Indeed, write $d=\lfloor\sqrt{n}\rfloor$ and suppose that U is a translate of some subspace of dimension $n=d$. U can be written as

$$U = \{a_0+\lambda_1 a_1+\dots+\lambda_{n-d}a_{n-d} : \lambda_i \in \mathbb{F}_2\},$$

where the a_i are linearly independent. Write a_i in component form as $(a_i^{(j)})_{j=1}^n$. The column rank of the matrix (a_{ij}) is $n-d$, and hence so is the row rank. Without loss of generality, suppose that the first $n-d$ rows $(a_1^{(j)},\dots,a_{n-d}^{(j)})$, $j=1,\dots,n-d$, are linearly independent. Then we can solve the $n-d$ equations

$$a_0^{(j)}+\lambda_1 a_1^{(j)}+\dots+\lambda_{n-d}a_{n-d}^{(j)} = 1$$

for the λ_i, giving a vector in U with no more than d zeros.**

Hereditary non-uniformity. If $\kappa>0$ is a real number then we say that a set $A \subseteq \mathbb{F}_2^n$ is κ-hereditarily non-uniform (HNU) if for any $S \subseteq A$ one has

$$\sup_{\xi \neq 0} |\widehat{S}(\xi)| \geqslant \kappa |S|.$$

Let us make a couple of remarks on the meaning of this statement. In $\mathbb{F}_2^n$ it is particularly easy to understand what it means for a set to have a large Fourier coefficient. Indeed if

$\hat{A}(\xi)$ is large then A has a *bias* relative to the two hyperplanes $\xi.x = 0$ and $\xi.x = 1$; one of these hyperplanes contains rather more than half the points of A, the other rather less. To say that a set is HNU means that every subset of A is biased with respect to some pair of hyperplanes. The notion of uniformity (having no large Fourier coefficients except at $\xi = 0$) will be explored in a lot more detail in Tim Gowers' course "Additive and Combinatorial Number Theory" in the Lent Term.

Proposition 14.1. *Let $C, D \subseteq \mathbb{F}_2^n$ have cardinalities γN and δN respectively. Then $(C + D)^c$ is $\sqrt{\gamma\delta}$-HNU.*

Proof. Suppose that $S \subseteq (C + D)^c$. Then

$$\sum_x S(x)(C * D)(x) \;=\; 0.$$

Writing this in terms of Fourier coefficients using Parseval's identity and the formula for Fourier transforms of convolutions gives

$$\sum_\xi \widehat{S}(\xi)\widehat{C}(\xi)\widehat{D}(\xi) \;=\; 0.$$

By the triangle inequality and Parseval (again) one then has

$$\begin{aligned}
\alpha\beta N^2|S| &= |\widehat{C}(0)||\widehat{D}(0)||\widehat{S}(0)| \\
&\leqslant \sum_{\xi\neq 0} |\widehat{C}(\xi)||\widehat{D}(\xi)||\widehat{S}(\xi)| \\
&\leqslant \sup_{\xi\neq 0} |\widehat{S}(\xi)| \left(\sum_\xi |\hat{C}(\xi)|^2\right)^{1/2} \left(\sum_\xi |\hat{D}(\xi)|^2\right)^{1/2} \\
&\leqslant \sup_{\xi\neq 0} |\hat{S}(\xi)| \cdot (\gamma\delta)^{1/2} N^2.
\end{aligned}$$

The result follows. □

The following beautiful result will be our way of harnessing Beckner's inequality in this instance.

Lemma 29. *Let $A \subseteq \mathbb{F}_2^n$ have cardinality αN, let $\rho \in (0, 1)$ be a real number and let Λ be the set of all ξ for which $|\hat{A}(\xi)| \geqslant \rho|A|$. Then Λ is contained in a subspace of dimension at most $8\rho^{-2}\log(1/\alpha)$.*

Proof. Suppose not. Then we may select linearly independent vectors $\xi_1, \ldots, \xi_d \in \Lambda$. If $M : \mathbb{F}_2^n \to \mathbb{F}_2^n$ is an invertible linear map and if a set A' is defined by $A'(x) = A(M^{-1}x)$ then one has $\widehat{A'}(\xi) = \widehat{A}(M^T\xi)$. By choosing M suitably and replacing A with A' we may assume that $\xi_1, \ldots, \xi_d$ are the standard basis vectors $e_1, \ldots, e_d$. We may now repeat an

argument from the previous set of notes almost verbatim. For any $\epsilon \in (0,1)$ we have

$$\begin{aligned} d\rho^2\alpha^2N^2 &= \sum_{i=1}^{d} \widehat{A}(e_i)^2 \\ &\leqslant \epsilon^{-2}\sum_{\xi} \epsilon^{2|\xi|}\hat{A}(\xi)^2 \\ &= \epsilon^{-2}N^2\|TA\|_2^2 \\ &\leqslant \epsilon^{-2}N^2\|A\|_{1+\epsilon^2}^2 \\ &= \epsilon^{-2}N^2\alpha^{2/(1+\epsilon^2)}. \end{aligned}$$

The passage between the third and fourth lines is effected via Beckner's inequality. Now cancelling the α^2N^2 from both sides and setting $\epsilon = (\log(1/\alpha))^{-1/2}$ gives $d \leqslant e^2\rho^{-2}\log(1/\alpha)$, from which the lemma follows immediately. □

We'll need the following result, which can be proved in an almost identical manner to the Bernstein/Chernoff bound we saw in Notes 10. With more foresight I could have proved this more general version back then, but I didn't, and I don't want to repeat the argument. A full (non-examinable) proof may be found on the *Expositions* page of my website.

Lemma 30. *Let $X_1, \dots, X_n$ be independent complex-valued random variables with $\mathbb{E}X_i = 0$ and $\mathbb{E}|X_j|^2 = \sigma_j^2$. Write $\sigma^2 = \sigma_1^2 + \cdots + \sigma_n^2$, and suppose that $|X_j| \leqslant 1$ uniformly in j. Suppose that $\sigma^2 \geqslant 6nt$. Then we have the inequality*

$$\mathbb{P}\left(|\overline{X}| \geqslant t\right) \leqslant 4e^{-n^2t^2/8\sigma^2},$$

where $\overline{X} = (X_1 + \cdots + X_n)/n$.

Theorem 21. *Suppose that $\kappa > n^{-1/4}$, and that $A \subseteq \mathbb{F}_2^n$ is κ-HNU. Then A^c contains a translate of some subspace of $\mathbb{F}_2^n$ of dimension at least $n\kappa^2/80$.*

Remark. The condition on κ could be relaxed slightly, but our interest is in constant κ anyhow.

Proof. Let $\beta = 2^{-n\kappa^2/40}$. We consider two cases.

Case 1. $|A| \leqslant \beta N$. Pick any subspace $U \subseteq \mathbb{F}_2^n$ of dimension less than $\log(1/\beta)$. $\mathbb{F}_2^n$ is a disjoint union of more than βN translates of U, so one of these must miss A entirely.

Case 2. $|A| \geqslant \beta N$. Let $B \subseteq A$ be a subset of cardinality βN for which $\sup_\xi |\widehat{B}(\xi)|$ is as small as possible. Since A is κ-HNU this quantity cannot be *too* small; in fact, if it equals $\eta|B|$ then we must have $\eta \geqslant \kappa$. Define Λ to be the set of all ξ for which $|\hat{B}(\xi)| \geqslant \eta|B|/2$. By Lemma 29, Λ is contained in a subspace $V \subseteq \mathbb{F}_2^n$ of dimension at

most $32\kappa^{-2}\log(1/\beta)$. Let $V^\perp$ be the orthogonal complement of V with respect to the standard inner product $x.y = x_1y_1 + \cdots + x_ny_n$. Then $V^\perp$ is a subspace of dimension at least $n - 32\kappa^{-2}\log(1/\beta)$. Pass to a subspace $W \subseteq V^\perp$ with dimension $n\kappa^2/50$. To ensure that this is possible, it must be checked that

$$\frac{n\kappa^2}{50} + 32\kappa^{-2}\log\left(\frac{1}{\beta}\right) \ < \ n,$$

a trivial matter.

Lemma 31. *For at least $(1-\eta/16)\,N$ values of x we have $B \cap (W+x) = \emptyset$.*

Proof. Suppose not. Then $|B \cap (W+x)| \geqslant 1$ for more than $\eta N/16$ values of x, and so

$$\begin{aligned} |W||B| &= \sum_x |(x+W)\cap B| \\ &> \eta N/16. \end{aligned}$$

We claim this is at least $|W||B|$. To see this, note that it suffices to check that $\eta/16 > 2^{-n\kappa^2/200}$, which is immediate from the fact that $\eta \geqslant \kappa \geqslant n^{-1/4}$. This contradiction establishes the lemma. □

Call the set C of such x *good*; the above lemma tells us that $|C| \geqslant (1-\eta/16)N$. As C is very large, it cannot have any really huge Fourier coefficients. Indeed if $\xi \neq 0$ then

$$\begin{aligned} |\widehat{C}(\xi)| &= |\widehat{C^c}(\xi)| \\ &\leqslant |C^c| \\ &\leqslant \eta N/16 \\ &\leqslant \eta|C|/8. \end{aligned} \tag{14.1}$$

We are now going to choose a subset $D \subseteq C$ of size t. We will do this by picking elements of C at random with probability $t/|C|$. It turns out that, provided t is large enough, D inherits from C the property of not having any really large Fourier coefficients.

Lemma 32. *Let $t \geqslant 2^{14}\eta^{-2}\log N$. Then there is a subset $D \subseteq C$ with size t such that $\sup_{\xi\neq 0}|\widehat{D}(\xi)| \leqslant \eta t/4$.*

Proof. As promised, choose a set $E \subseteq C$ by letting each $x \in C$ be in E with probability $p = t/|C|$, these choices being independent. The Fourier coefficient $\widehat{E}(\xi)$ is then a sum of $|C|$ independent random variables $X_j^{(\xi)} = E(x)(-1)^{\xi.x}$ with variances at most p. It follows from Lemma 37 and (14.1) that

$$\begin{aligned} \mathbb{P}\left(|\widehat{E}(\xi)| \ \geqslant \ \eta t/6\right) &\leqslant \mathbb{P}\left(\left|\widehat{E}(\xi) - \mathbb{E}\widehat{E}(\xi)\right| \ \geqslant \ \eta t/24\right) \\ &< 4e^{-\eta^2 t/5000}. \end{aligned}$$

By the same token

$$\mathbb{P}\left(||E| - t| \geqslant \eta t/24\right) < 4e^{-\eta^2 t/5000}.$$

If $t \geqslant 2^{14}\eta^{-2}\log N$, then, there is a positive probability that none of the above events happen. By adding or deleting at most $\eta t/12$ elements from E we get a set D satisfying the conclusion of the lemma. □

An almost identical argument proves the following.

Lemma 33. *Let $N^{1/2} \geqslant t \geqslant 2^{14}\eta^{-2}\log N$. Then there is a subset $X \subseteq B$ with $|X| = t$ and*

$$\left|\widehat{X}(\xi) - \frac{t\widehat{B}(\xi)}{|B|}\right| \leqslant \eta t/12$$

for all $\xi \neq 0$. □

Lemma 34. *Let S be the (multi)set $(B \setminus X) \cup D$. Then*

$$\sup_{\xi\in\Lambda} |\widehat{S}(\xi)| \leqslant \eta|S| - \eta t/6,$$

whilst

$$|\widehat{S}(\xi)| \leqslant \frac{\eta|S|}{2} + \frac{\eta t}{3}$$

for all other $\xi \neq 0$.

Proof. We have

$$\begin{aligned}\widehat{S}(\xi) &= \widehat{B}(\xi) - \widehat{X}(\xi) + \widehat{D}(\xi)\\ &= \left(1 - \frac{t}{|B|}\right)\widehat{B}(\xi) + Q,\end{aligned}$$

where $|Q| \leqslant \eta t/3$ by the previous two lemmas. If $\xi \in \Lambda$ then $|\widehat{B}(\xi)|/|B| \geqslant \eta/2$ by definition, and the first part of the result follows easily. For the second part of the result observe that if $\xi \notin \Lambda$ then

$$\begin{aligned}|\widehat{S}(\xi)| &\leqslant \left|1 - \frac{t}{|B|}\right||\widehat{B}(\xi)| + |Q|\\ &\leqslant \frac{\eta|S|}{2} + \frac{\eta t}{3}.\end{aligned}$$

This proves the lemma. □

Now let $D = \{d_1, \ldots, d_t\}$. Let D' be any set obtained by replacing d_j $(j = 1, \ldots, t)$ with $d_j + x_j$, where $x_j \in W$ (now might be a good opportunity to recall the definition of W).

Lemma 35. *Suppose that $t \leqslant \eta\beta N/10$. Let S' be the (multi)set $(B \setminus X) \cup D'$. Then*

$$\sup_{\xi \neq 0} |\widehat{S'}(\xi)| \ < \ \eta |S'|.$$

Proof. We deal first with the easy case $\xi \notin \Lambda$. However we change the elements of D the contribution to $\widehat{S}(\xi)$ cannot vary by more than $2t$. It follows from Lemma 34 that, for $\xi \notin \Lambda$,

$$|\widehat{S'}(\xi)| \ \leqslant \ \frac{\eta|S'|}{2} + 5t \ < \ \eta|S'|.$$

However if $\xi \in \Lambda$ then

$$\begin{aligned} \widehat{S'}(\xi) &= \widehat{S}(\xi) + \sum_{j=1}^{t} \left((-1)^{\xi(d_j+x_j)} - (-1)^{\xi d_j}\right) \\ &= \widehat{S}(\xi), \end{aligned}$$

Since $x_j \in W = V^{\perp}$ and $\Lambda \subseteq V$. The result follows immediately from Lemma 34. □

If we could choose $x_1, \ldots, x_t$ so that S' was actually a set (as opposed to a multiset) and also so that $S' \subseteq A$ then we would have a contradiction of our earlier assumption about the minimality of B. It follows that there is no such choice of $x_1, \ldots, x_t$.

Lemma 36. *There is some j such that $d_j + W$ is contained in A^c, except for at most t elements.*

Proof. Suppose not, and recall that none of the $d_j + W$ intersects B. Thus we may choose $x_1 \in W$ so that $d_1 + x_1 \in A \setminus B$, and then $x_2 \in W$ so that $d_2 + x_2 \in A \setminus (B \cup \{d_1 + x_1\})$. Continue in this way; at the last stage we will still be able to choose $x_t \in W$ so that

$$d_t + x_t \in A \setminus \left(B \cup \bigcup_{j=1}^{t-1} \{d_j + x_j\} \right).$$

This gives us an S' of the type that we argued couldn't exist. The lemma follows. □

Pick a $j \in [t]$ be such that the conclusion of this lemma holds. There is clearly a subspace $U \subseteq W$ of dimension at least $\dim W - \log t$ such that some translate $x + U$ lies *entirely* in A^c. Certainly $\dim U \geqslant n\kappa^2/80$, and the proof of Theorem 21 is complete.

Appendix A. A brief discussion of spheres

There is a large and highly-regarded literature concerning generalisations of results like the L^∞–L^4 local restriction theorem for circles to higher dimensions. I thought long and hard about whether to include any of this in the course, but in the end I decided that I could not achieve a satisfactory level of rigour without assuming more measure theory than I would like to. This set of notes, which is entirely non-examinable, is

intended to give you a brief overview of what is known and what we would like to know. We will return to the restriction phenomena in the next set of notes, where we will investigate the paraboloid in $\mathbb{F}^3$ in some detail. Although this "toy model" is in some ways analagous to the Euclidean 2-sphere, there are a number of features of the Euclidean situation which cannot be adequately appreciated in the finite field case.

The reader who feels bothered by all this, and who would like to understand the Euclidean case better, should consult Terry Tao's lecture notes for "Math 254B", particularly the first two chapters. These are available on his website.

The first thing to understand is the Fourier transform of $d\sigma$, where σ is the natural measure on the sphere $S^{n-1} \subseteq \mathbb{R}^n$. It turns out not to be particularly difficult to generalise the results we proved in the plane. Using stationary phase, one can get an asymptotic which implies the estimate

$$|\widehat{d\sigma}(\lambda)| \ll \min(1, |\lambda|^{-(n-1)/2}).$$

To prove this, it once again suffices to consider the case $\lambda = (0, 0, \ldots, \lambda)$. In estimating

$$\int_{S^{n-1}} e^{-2\pi i \lambda x_n} \, d\sigma(x)$$

there are just two stationary phase points, the north and south poles of the sphere. At these points one can write the surface of the sphere in local coordinates as $z = y_1^2 + \cdots + y_{n-1}^2$, much as we did for the circle, and then use the stationary phase lemma to estimate the resulting integral. When $n = 3$ one can actually evaluate $\widehat{d\sigma}(\lambda)$ in closed form (this is on the example sheet).

By analogy with the results we proved for the circle, we might ask the following.

Problem A.1 (Local restriction problem)**.** Let $f : S^{n-1} \to \mathbb{C}$ be measurable. For which p do we have the estimate

$$\|(f d\sigma)^\vee\|_{L^p(B(0,R))} \ll_\epsilon R^\epsilon \|f\|_{L^\infty(S^{n-1})}? \tag{A.1}$$

If such an estimate holds we say that $\mathrm{LR}(n, p)$ is true. It is easy to see that $\mathrm{LR}(n, p) \Rightarrow \mathrm{LR}(n, p')$ if $p' \geqslant p$, so it is natural to ask for the smallest value of p for which $\mathrm{LR}(n, p)$ holds. We know that $\mathrm{LR}(2, 4)$ holds, and that this is sharp (simply test (A.1) with $f = 1$). By analogy one might conjecture that $\mathrm{LR}(n, 2n/(n-1))$ holds. This is called the *local restriction conjecture* and is not known to be true, even in three dimensions. The trick that made our proof work in two dimensions was the observation that an L^4 norm can be worked out simply, as $|z|^4 = z^2\overline{z}^2$. There is no equivalent of this for, say, an L^3 norm, and this is why a naïve modification of our arguments does not work.

By modifying the argument of the previous set of notes one can show that $\mathrm{LR}(n,p)$ implies that the Minkowski dimension of Kakeya sets in $\mathbb{R}^n$ is at least $\frac{2p}{p-2} - n$, and in particular that the local restriction conjecture implies the Minkowski Kakeya conjecture. The so-called "Knapp example" generalises to say that the Fourier transform of a δ-cap of S^{n-1} is large on a tube with dimensions roughly $\delta^{-2} \times \delta^{-1} \times \cdots \times \delta^{-1} \times \delta^{-1}$. The main difficulty in generalising the argument is that one has to prove an estimate of the form

$$\mathbb{E}\left\|\sum_{i=1}^k \epsilon_i g_i\right\|_p^p \gg \left\|\left(\sum_{i=1}^k |g_i|^2\right)^{1/2}\right\|_p^p,$$

where the ϵ_i are random phases as they were before, and the g_i are measurable functions on Euclidean space. In the previous set of notes we needed this for $p = 4$, and could prove it by simply multiplying out. For p not an even integer such a trick is not available and a different proof is necessary. One is given in Chapter 1 of Tao's notes; I may also discuss the result in an examples class. It is called Khintchine's inequality.

Are there any results of the form $\mathrm{LR}(n,p)$ that we *can* prove when $n > 2$? One strategy is to aim for a weaker type of result of the form

$$\|(fd\sigma)^\vee\|_{L^p(B(0,R))} \ll_\epsilon R^\epsilon \|f\|_{L^2(S^{n-1})}. \tag{A.2}$$

This certainly implies $\mathrm{LR}(n,p)$, and it is well-known to analysts that L^2 norms are nice things to have around. It turns out that we can prove an essentially optimal result of the form (A.2), but sadly this does not have the exponent $p = 2n/(n-1)$. Indeed by testing (A.2) with a Knapp example (that is, f is the charactistic function of a δ-ball) one can see that such a result can only hold for $p \geqslant 2(n+1)/(n-1)$. The *Stein-Tomas theorem* states that this is exactly the correct range, and that there is actually no need to localise to a ball:

Theorem 22 (Stein-Tomas)**.** *We have the bound*

$$\|(fd\sigma)^\vee\|_{L^{2(n+1)/(n-1)}(\mathbb{R}^n)} \ll \|f\|_{L^2(S^{n-1})}.$$

Stein-Tomas is proved by the method of T and T^*, which will be discussed in the discrete setting in the next set of notes. This is very specific to the L^2 setting.

Thus $\mathrm{LR}(n, 2(n+1)/(n-1))$ is true. Sadly, however, this leads only to the rather disappointing bound $d(n) \geqslant 1$ for Kakeya sets! This is far worse than the bounds we know for Kakeya, particularly bounds like $d(3) \geqslant 5/2$ given by the Wolff Hairbrush argument[6].

In the early 1990's an argument was introduced by Bourgain which allows one to get a

[6]admittedly, in this course we only proved this in the "toy" setting of finite fields

(rather partial) reverse implication of the form

$$\text{Kakeya estimates} \Rightarrow \text{Restriction estimates.}$$

I believe that Bourgain's result and subsequent modifications of it allow one to prove something at least as strong as $\mathrm{LR}(3, 4-\frac{2}{7})$, which is an improvement on the Tomas-Stein result $\mathrm{LR}(3,4)$. There are also improvements in higher dimensions. Sadly Bourgain's argument is outside the scope of this course, but you can read about it in the fifth chapter of Terry Tao's notes for Math 254B.

I should conclude this section by explaining something about where the term *restriction* comes from. Tomas-Stein asserts that the *extension* operator $T : f \mapsto (f d\sigma)^\vee$ is bounded as a map from $L^2(S^{n-1})$ to $L^p(\mathbb{R}^n)$, $p = 2(n+1)/(n-1)$. As in the discrete case T has an adjoint T^*, which is just the restriction map $f \mapsto \hat{f}|_{S^{n-1}}$. Thus the Tomas-Stein theorem says, in dual form, that there is an inequality

$$\|\hat{f}|_{S^{n-1}}\|_{L^2(S^{n-1})} \ll \|f\|_{L^{2(n+1)/(n+3)}(\mathbb{R}^n)}. \tag{A.3}$$

When $n = 3$, the exponent on the right equals $4/3$. Tomas-Stein is often phrased in the form "the Fourier transform of an $L^{4/3}$ function on $\mathbb{R}^3$ can be meaningfully restricted to the 2-sphere". This is perhaps rather nonsensical, but at least (A.3) makes it clear where the word "restriction" comes from.

Appendix B. Stationary phase

1. Introduction. The Principle of Stationary Phase (PSP) is a means of estimating *oscillatory integrals* of the form

$$I(\lambda) \;=\; \int_{\mathbb{R}^n} e^{i\pi\lambda f(x)} a(x)\, dx \tag{B.1}$$

asymptotically as $\lambda \to \infty$. It is particularly simple to state and prove results under the assumptions that $f \in C^\infty$ and $a \in C_0^\infty$, and furthermore many interesting examples are covered by this case. For example we will prove later on the following result.

Theorem 23. *Let μ be the measure on S^{n-1} induced from Lebsegue measure on $\mathbb{R}^n$. Then*

$$\hat{\mu}(\lambda) \;=\; 2|\lambda|^{-(n-1)/2} \cos\, \pi\left(|\lambda| - \frac{n-1}{4}\right) + O\left(|\lambda|^{-(n+1)/2}\right).$$

If f is rapidly varying then we might expect there to be lots of cancellation in (B.1), so that $I(\lambda)$ decays very rapidly. This indeed proves to be the case, and we will formulate this more exactly as the Principle of Non-Stationary Phase (PNSP). If $\nabla f = 0$, however, there is not nearly such rapid decay. The analysis of this case consititutes the Principle of Stationary Phase as we will study it. We will only study functions f

with non-degenerate singularities (that is to say singularities at which the Hessian is non-singular). A fully general analysis would be immensely complicated.

These notes draw heavily on lecture notes of T. Wolff and on lectures of E. Stein.

2. Two 1-Dimensional Estimates. In this section we state and prove two 1-dimensional estimates which may be regarded as the heart of the PNSP and the PSP respectively. The first is rather standard.

Proposition B.1. *Let $a \in C_0^\infty(\mathbb{R})$ and let*

$$J(\lambda) \;=\; \int_{\mathbb{R}} e^{i\pi\lambda x} a(x)\, dx.$$

Then $J(\lambda) = O\left(\lambda^{-N}\right)$ as $\lambda \to \infty$ for any positive integer N.

Proof. This is nothing more than the statement that the Fourier transform of a C_0^∞ function decays superpolynomially. It may be proved by repeated integration by parts. □

The second result is a little less standard.

Proposition B.2. *Let $a \in C_0^\infty(\mathbb{R})$ and let*

$$K(\lambda) \;=\; \int_{\mathbb{R}} e^{i\pi\lambda x^2} a(x)\, dx.$$

Let $\lambda > 0$. Then

$$K(\lambda) \;=\; \lambda^{-1/2} e^{i\pi/4} a(0) + O\left(\lambda^{-3/2}\right).$$

Proof. Applying Parseval's formula gives

$$\int_{\mathbb{R}} e^{-zx^2} a(x)\, dx \;=\; (4\pi z)^{-1/2} \int_{\mathbb{R}} \hat{a}(\xi) e^{-\xi^2/4z}\, d\xi \tag{B.2}$$

for $z > 0$. Now the left hand side here can easily be extended to an analytic function on all of $\mathbb{C}$. The right hand side cannot, but applying the dominated convergence theorem it is possible to see that it can be extended to a continuous function on the set

$$S \;=\; \{z : \Re z \geqslant 0, z \neq 0\}$$

which is analytic on S^o. It follows by the identity principle that (B.2) holds for all $z \in S$, and hence in particular for $z = -i\pi\lambda$. When $\lambda > 0$ this gives the identity

$$\int_{\mathbb{R}} e^{i\pi\lambda x^2} a(x)\, dx \;=\; \lambda^{-1/2} e^{i\pi/4} \int_{\mathbb{R}} \hat{a}(\xi) e^{i\xi^2/4\pi\lambda}\, d\xi.$$

The proposition follows immediately by observing that

$$\begin{aligned}\left|\int_{\mathbb{R}} \hat{a}(\xi)e^{i\xi^2/4\pi\lambda}\,d\xi - \int_{\mathbb{R}} \hat{a}(\xi)\,d\xi\right| &\leqslant \int_{\mathbb{R}} |\hat{a}(\xi)|\left|1-e^{i\xi^2/4\pi\lambda}\right|\,d\xi \\ &\leqslant \frac{1}{4\pi\lambda}\int_{\mathbb{R}} |\hat{a}(\xi)|\,|\xi^2|\,d\xi \\ &\ll \lambda^{-1}\end{aligned}$$

and that

$$\int_{\mathbb{R}} \hat{a}(\xi)\,d\xi \;=\; 2\pi a(0).$$

There is no reason to be mysterious about why Propositions B.1 and B.2 encapsulate the PNSP and PSP respectively. The reason is that (at least in $\mathbb{R}$) the functions $x \mapsto x$ and $x \mapsto x^2$ are in some sense the "generic" functions with non-vanishing derivative and with non-degenerate singular point respectively. We can formalise this in the following standard lemma from differential geometry, the second part of which is due to Morse.

Lemma 37. (i) *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a smooth function and that $\nabla f(p) \neq 0$ at some point $p \in \mathbb{R}^n$. Then there is a neighbourhood U containing p, a neighbourhood V containing 0 and a diffeomorphism $\phi : V \to U$ with $\phi(0) = p$ such that $f(\phi(x)) = f(p) + x_1$.*

(ii) *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a smooth function such that $\nabla f(p) = 0$. Suppose that the Hessian $H_p(f) = (\partial^2 f/\partial x_j \partial x_k)_{j,k}$ is non-singular. Then there is an $r \leqslant n$, neighbourhoods U, V with $p \in U$,$0 \in V$ and a diffeomorphism $\phi : V \to U$ with $\phi(0) = p$ such that*

$$f(\phi(x)) \;=\; f(p) + x_1^2 + \cdots + x_r^2 - x_{r+1}^2 - \cdots - x_n^2.$$

We call the quantity $\sigma = 2r - n$ the *signature* of f at the critical point p. It turns out to be an invariant of f (in other words it is independent of the diffeomorphism ϕ).

In view of this lemma it will come as no surprise that we are interested in the following two results, which are simple generalisations of Propositions B.1 and B.2 proved in exactly the same way.

Proposition B.3. *Let $a \in C_0^\infty(\mathbb{R}^n)$ and let*

$$J(\lambda) \;=\; \int_{\mathbb{R}^n} e^{i\pi\lambda x_1} a(x)\,dx.$$

Then $J(\lambda) = O\left(\lambda^{-N}\right)$ as $|\lambda| \to \infty$ for any positive integer N.

Proposition B.4. *Let $a \in C_0^\infty(\mathbb{R}^n)$ and let*

$$K(\lambda) \;=\; \int_{\mathbb{R}^n} e^{i\pi\lambda\left(x_1^2+\cdots+x_r^2-x_{r+1}^2-\cdots-x_n^2\right)} a(x)\,dx.$$

Let $\lambda > 0$. Then

$$K(\lambda) \;=\; \lambda^{-n/2} e^{i\pi\sigma/4} a(0) + O\left(\lambda^{-(n+2)/2}\right).$$

3. The Principle of Non-Stationary Phase.

Theorem 24. *Let $a \in C_0^\infty(\mathbb{R}^n)$ and $f \in C^\infty(\mathbb{R}^n)$ be such that $\nabla f \neq 0$ in $\mathrm{Supp}(a)$. Write*

$$I(\lambda) \;=\; \int_{\mathbb{R}^n} e^{i\pi\lambda f(x)} a(x)\, dx.$$

Then $I(\lambda) = O(\lambda^{-N})$ for any N as $\lambda \to \infty$.

Proof. Let us first of all work locally. Let $p \in \mathrm{Supp}(a)$ and let U_p, V_p be the neighbourhoods featuring in Lemma 37. Let $\phi_p : U_p \to V_p$ be the diffeomorphism described in that Lemma. Let $b_p : \mathbb{R}^n \to \mathbb{R}$ be smooth and supported on U_p. Then by change of variables

$$\begin{aligned}\int_{U_p} e^{i\pi\lambda f(x)} b_p(x)\, dx &= e^{i\pi\lambda f(p)} \int_{V_p} e^{i\pi\lambda x_1} b_p(\phi(x))\, |\mathcal{J}_\phi(x)|\, dx \\ &\ll \lambda^{-N}\end{aligned}$$

for any N by Proposition B.1. Now pick a finite set of points p_j such that the corresponding neighbourhoods $U_j = U_{p_j}$ cover $\mathrm{Supp}(a)$. Take a partition of unity g_j relative to the U_j, so that

- The g_j are C^∞;
- $\mathrm{Supp}(g_j) \subseteq U_j$;
- $\sum_j g_j(x)$ is identically equal to 1 on $\bigcup_j U_j$.

Set $b_j(x) = a(x)g_j(x)$. Then the theorem follows from the observation that

$$I(\lambda) \;=\; \sum_j \int_{U_p} e^{i\pi\lambda f(x)} b_j(x)\, dx.$$

4. The Principle of Stationary Phase.

Theorem 25. *Let $a \in C_0^\infty(\mathbb{R}^n)$ and suppose that $f \in C^\infty(\mathbb{R}^n)$ has only non-degenerate critical points. Let these be $p_1, \ldots, p_m$ and suppose that the signature of f at p_j is σ_j. Let Δ_j denote the absolute value of $\det \mathcal{H}_f(p_j)$, the determinant of the Hessian at p_j. Write*

$$I(\lambda) \;=\; \int_{\mathbb{R}^n} e^{i\pi\lambda f(x)} a(x)\, dx.$$

Then

$$I(\lambda) \;=\; \lambda^{-n/2} \left(\sum_{j=1}^{m} e^{i\pi\lambda f(p_j)} e^{i\pi\sigma_j/4} \Delta_j^{-1/2} a(p_j) \right) + O\left(\lambda^{-(n+2)/2}\right).$$

Proof. Once again we work locally in the first instance. Let p be a critical point of f and suppose we are in the situation described by (ii) of Lemma 37. To spell it out, we have neighbourhoods U, V with $p \in U$, $0 \in V$ and a diffeomorphism $\phi : V \to U$ with $\phi(0) = p$ such that

$$f(\phi(x)) \;=\; f(p) = x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_n^2.$$

Let b be any C^∞ function supported in U. By change of variables and Proposition B.2 we have

$$\begin{aligned} \int_U e^{i\pi\lambda f(x)} b(x)\, dx &= e^{i\pi\lambda f(p)} \int_V e^{i\pi\lambda(x_1^2+\cdots+x_k^2-x_{k+1}^2-\cdots-x_n^2)} b(\phi(x)) \left|\mathcal{J}_\phi(x)\right| \, dx \\ &= e^{i\pi\lambda f(p)} e^{i\pi\sigma/4} \lambda^{-n/2} b(p) \left|\mathcal{J}_\phi(0)\right| + O\left(\lambda^{-(n+2)/2}\right). \end{aligned} \tag{B.3}$$

It turns out that we can express $|\mathcal{J}_\phi(0)|$ in terms of intrinsic properties of f. To do this one uses the chain rule and the fact that $\nabla f = 0$ to check that

$$\mathcal{H}_{(f\circ\phi)}(0) \;=\; \phi'(0)^t \mathcal{H}_f(p) \phi'(0).$$

Noting that $|\det \mathcal{H}_{(f\circ\phi)}(0)| = 1$ this implies that

$$|\mathcal{J}_\phi(0)| \;=\; |\det \phi'(0)| \;=\; \Delta^{-1/2},$$

and so (B.3) constitutes a sort of local version of the theorem. To deduce a global version, choose the neighbourhoods $U_j = U_{p_j}$ to be disjoint, and let b_j be a C^∞ function which is supported in U_j and which equals $a(x)$ in a neighbouhood of p_j. Let $c(x) = a(x) - \sum_j b_j(x)$, so that $\mathrm{Supp}(c)$ is contained in an open set on which $\nabla f \neq 0$. We have then that

$$I(\lambda) \;=\; \sum_j \int_{U_j} e^{i\pi\lambda f(x)} b_j(x)\, dx + \int_{\mathbb{R}^n} e^{i\pi\lambda f(x)} c(x)\, dx,$$

and the full strength of the theorem follows from the local version together with Theorem 24. □

The form of Theorem 25 looks complicated. We would like to emphasise again that

there is really nothing at all difficult going on here. One looks at the behaviour in the simplest possible case of a non-degenerate stationary point (Proposition B.2) and then reduces the general case to this by showing that the behaviour of these oscillatory integrals is in some sense invariant under diffeomorphisms. The complicated look of our calculations is not helped by the need to invoke partitions of unity, but these are merely a convenient way of allowing one to think locally. Local then becomes global because the amplitude function a has compact support, allowing us to add up a finite number of local estimates.

Before moving on to an application we remark that these asymptotic expansions can be continued to greater accuracy by expanding the exponential $e^{i\xi^2/4\pi\lambda}$ appearing in the proof of Proposition B.2. One can also say something about the dependence of the implied constants on f, a and their derivatives. We will not, however, discuss this matter here.

5. Application: Fourier Transforms of Spherical Measures. An important use of PSP is in the estimation of the Fourier Transforms of measures μ supported on the sphere $S^{n-1} \subseteq \mathbb{R}^n$.

Theorem 26. *Let μ be the measure on S^{n-1} induced from Lebsegue measure on $\mathbb{R}^n$. Then*

$$\hat{\mu}(\lambda) \;=\; 2|\lambda|^{-(n-1)/2} \cos\, \pi\left(|\lambda| - \frac{n-1}{4}\right) + O\left(|\lambda|^{-(n+1)/2}\right).$$

Proof. It is clear that $\hat{\mu}$ is a radial function (that is, one which depends only on distance from the origin) because S^{n-1} is radially symmetric. It suffices, then, to evaluate $\hat{\mu}$ at λe_n for $\lambda \in \mathbb{R}^+$.

Lemma 38. *Let*

$$\phi_j : (x_1, \ldots, x_n) \longrightarrow (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)$$

denote the jth projection map. Then we can express S^{n-1} as a (non-disjoint) union $V \cup W \cup \bigcup_{k=1}^m U_k$ of open sets, where V and W are small caps about the south and north poles $P_\pm = (0, 0, \ldots, \pm 1)$ of S^{n-1}, and for each k there is $j \leqslant n-1$ such that $\phi_j : U_k \to \phi_j(U_k)$ is a diffeomorphism.

Proof. $\phi_j|_{S^{n-1}}$ is a local diffeomorphism everywhere except on $x_j = 0$. Hence if p is not one of the poles then there is a neighbourood U containing p on which some ϕ_j ($j \leqslant n-1$) is a diffeomorphism. The lemma follows by a compactness argument. □

Let $\{v, w, u_k\}$ be a partition of unity for the cover $V \cup W \cup \bigcup_{k=1}^m U_k$ so that $v(P_-) =$

$w(P^+) = 1$. Then

$$\hat{\mu}(\lambda e_n) \;=\; \int_V v(x)e^{i\pi\lambda x_n}\,d\mu + \int_W w(x)e^{i\pi\lambda x_n}\,d\mu + \sum_{k=1}^{m}\int_{U_k} u_k(x)e^{i\pi\lambda x_n}\,d\mu. \qquad \text{(B.4)}$$

Working in local coordinates $(x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)$ it is easy to see that each integral

$$\int_{U_k} u_k(x)e^{i\pi\lambda x_n}\,d\mu$$

is $O(\lambda^{-N})$ for any N by the PNSP. The interesting integrals are those involving v and w, and we will treat the former. The local coordinates at P_- are $(x_1,\ldots,x_{n-1})$ and in terms of these one has

$$x_n \;=\; -\left(1-|x|^2\right)^{1/2}$$

and

$$d\mu \;=\; (1-|x|^2)^{-1/2}dx_1\ldots dx_{n-1}.$$

It is easy to check that x_n, considered as a function of $x_1,\ldots,x_{n-1}$, is singular only at $(0,\ldots,0)$ and that the Hessian at this point is just the identity matrix. It follows from PSP that

$$\int_V v(x)e^{i\pi\lambda x_n}\,d\mu \;=\; \lambda^{-(n-1)/2}e^{-i\pi\lambda}e^{i\pi(n-1)/4} + O\left(\lambda^{-(n+1)/2}\right).$$

Similarly

$$\int_W w(x)e^{i\pi\lambda x_n}\,d\mu \;=\; \lambda^{-(n-1)/2}e^{i\pi\lambda}e^{-i\pi(n-1)/4} + O\left(\lambda^{-(n+1)/2}\right).$$

Adding these together and recalling (B.4) gives the result. □

It can be shown that if $S \in \mathbb{R}^n$ is a smooth $(n-1)$-dimensional hypersurface with non-vanishing Gauss curvature at every point, and if μ is a smooth measure supported on S, then $\hat{\mu}(\lambda) \ll |\lambda|^{-(n-1)/2}$ for large λ. The reader will appreciate that this is essentially a stationary phase argument if we say what it means for a point $p \in S$ to be a point of non-zero Gauss curvature. Take a tangent plane Π to S at p, and use the orthogonal projection of S onto Π as a method of defining local coordinates. If the equation of S in these local coordinates is $x_n = F(x_1,\ldots,x_{n-1})$ then we say that p is a point of non-zero curvature if the Hessian $(\partial^2 F/\partial x_j \partial x_k)_{1\leqslant j,k\leqslant n-1}$ has non-zero determinant.

6. Van der Corput's Estimates. This section covers a topic related to stationary phase, but with a slightly different flavour.

Theorem 27 (Van der Corput's Estimates)**.** *Let $[a,b]$ be an interval on the real line, let k be a positive integer, and let $f : [a,b] \to \mathbb{R}$ be continuous. Write*

$$I \;=\; \int_a^b e^{if(x)}\,dx.$$

(i) *Suppose that $k = 1$ and that $f \in C^2(a,b)$. Suppose furthermore that f' is monotonic and that $|f'(x)| \geqslant \lambda \geqslant 0$ uniformly on (a,b). Then $|I| \leqslant 4/\lambda$.*

(ii) *Suppose that $k \geqslant 2$ and that $f \in C^k(a,b)$. Suppose that $|f^{(k)}(x)| \geqslant \lambda \geqslant 0$ uniformly on (a,b). Then $|I| \leqslant 3 \cdot 2^k \cdot \lambda^{-1/k}$.*

Proof. (i) Observe that

$$e^{if(x)} \;=\; \frac{1}{if'(x)}\frac{d}{dx}\left(e^{if(x)}\right),$$

this expression being valid everywhere because of our assumption about f. Thus

$$\begin{aligned}
|I| &= \left|\int_a^b \frac{d}{dx}\left(e^{if(x)}\right)\frac{dx}{if'(x)}\right| \\
&= \left|\left[\frac{e^{if(x)}}{if'(x)}\right]_a^b - \int_a^b e^{if(x)}\frac{d}{dx}\left(\frac{1}{if'(x)}\right)\,dx\right| \\
&\leqslant \frac{2}{\lambda} + \int_a^b \left|\frac{d}{dx}\left(\frac{1}{if'(x)}\right)\right|\,dx \\
&= \frac{2}{\lambda} + \left|\int_a^b \frac{d}{dx}\left(\frac{1}{if'(x)}\right)\,dx\right| \\
&\leqslant \frac{4}{\lambda}.
\end{aligned}$$

The crucial step here of taking the absolute value signs back outside the integral was permissible, of course, because f' was assumed to be monotonic. The integrations by parts were all valid because $f \in C^2(a,b)$.

(ii) We will prove the stronger statement that $|I| \leqslant \left(3 \cdot 2^k - 2\right) \cdot \lambda^{-1/k}$ by induction on k. The inductive step will also serve to deduce the case $k = 2$ from part (i), as the reader will verify. Suppose then that $f \in C^k(a,b)$ and that $\left|f^{(k)}(x)\right| \geqslant \lambda$ on (a,b). Since $f^{(k)}$ is continuous it is clear that either $f^{(k)}(x) \geqslant \lambda$ or $f^{(k)}(x) \leqslant -\lambda$ for all x. Without loss of generality we suppose the former holds. Now observe that for any $\delta > 0$ we may write

$$I \;=\; I_1 \cup I_2 \cup I_3$$

where $f^{(k-1)}(x) \leqslant -\delta\lambda$ on I_1, $f^{(k-1)}(x) \geqslant \delta\lambda$ on I_3 and $|I_2| \leqslant \delta\lambda$. Furthermore on each of I_1 and I_3 the function $f^{(k-1)}$ will be monotonic (this remark is only relevant for

$k = 2$). It follows from our inductive hypothesis that

$$\begin{aligned} |I| &\leqslant \left|\int_{I_1} e^{if(x)}\,dx\right| + \left|\int_{I_2} e^{if(x)}\,dx\right| + \left|\int_{I_3} e^{if(x)}\,dx\right| \\ &\leqslant 2\left(3\cdot 2^{k-1} - 2\right)(\delta\lambda)^{-1/(k-1)} + 2\delta. \end{aligned}$$

Choosing $\delta = \lambda^{-1/k}$ gives the result. □

A very useful feature of this result is that the estimates are independent of the interval $[a, b]$. The Van der Corput estimates imply a very general stationary phase result in 1 dimension, in which the phase function f is allowed to have finitely many stationary points each with order at most k. We sketch the argument briefly. For an estimate on

$$I(\lambda) \;=\; \int_{\infty}^{\infty} e^{i\lambda f(x)} a(x)\,dx$$

where $a \in C_0^\infty(\mathbb{R})$ one can begin by splitting the range of integration into finitely many intervals on which one of the derivatives $f^{(j)}$, $j \leqslant k$, is bounded below in absolute value by some positive constant. One can then apply the Van der Corput estimate to each interval in turn to get an estimate of form $I(\lambda) \ll \lambda^{-1/k}$. In doing this one writes $a(x)$ in integral form as

$$a(x) \;=\; \int_c^x b(t)\,dt;$$

substituting into the integral for I and swapping the order of summation brings us into the realm of Theorem 27. The details we leave to the reader.

Appendix C. Exercises from the course

1. Show that for every $\epsilon > 0$ there is a subset of $\mathbb{R}^2$ of area at most ϵ in which one can continuously rotate a unit needle through 180 degrees.

2. Prove that $d_F(n) \geqslant (n+1)/2$ in a different way by adapting the "slicing" argument we used to prove $d(n) \geqslant (n+1)/2$ (you should find that the argument is much simpler, and quite similar to part of the argument we used to show that $d_F(n) \geqslant (4n+3)/7$).

3. Let m, n be positive integers. Show that $d_F(m+n) \leqslant d_F(m) + d_F(n)$, and that $d_F(n+1) \leqslant d_F(n)$.

4. Prove in detail that the arithmetic Kakeya conjecture implies that $d_F(n) = n$.

5. Construct a subset $E \subseteq [0,1]$ with $\underline{\dim}(E) \leqslant 1/10$ and $\overline{\dim}(E) \geqslant 9/10$.

6. Show that any subset of $\mathbb{R}^3$ which contains a unit plane in each direction (that is, a unit square normal to every direction) has positive measure.

7. Construct a set $G \subseteq \mathbb{Z} \times \mathbb{Z}$ with

$$|\pi_{-1}(G)| \geqslant \max\left(|\pi_0(G)|, |\pi_1(G)|, |\pi_\infty(G)|\right)^{11/10}.$$

(Compare with the arithmetic Kakeya conjecture.)

8.** The base 4 Cantor set C is the set of all real numbers between 0 and 1 whose base 4 expansion contains only zeros and ones. Define a subset of $\mathbb{R}^2$ as follows: take the sets $A = C \times \{0\}$ and $B = 2C \times \{1\}$ and join everything in A to everything in B by a line segment. This gives a set E with line segments in many different directions. Show that E has measure zero.

9.** Show that any subset of $\mathbb{R}^4$ which contains a unit plane in each direction has positive measure.

Things to read. To get some idea of what the course is about, try having a look at Terry Tao's article *From rotating needles to stability of waves: emerging connections between combinatorics, analysis and PDE*, Available at
`http://www.ams.org/notices/200103/fea-tao.pdf`.

Alex Iosevich's article *Curvature, combinatorics and the Fourier transform*, available at
`http://www.ams.org/notices/200106/fea-iosevich.pdf`,
will give you some idea of the material we will cover in lectures 6 and 7.

1. Evaluate the Fourier transform $\widehat{d\sigma}(\xi)$ explicitly when σ is the surface measure on the sphere $S^2 \subseteq \mathbb{R}^3$.

2. Show that the discrete paraboloid $P \subseteq \mathbb{F}_*^3$ contains a line if and only if -1 is a square in $\mathbb{F}$.

3. Suppose that -1 is a square in $\mathbb{F}$, and let $2 < p < 4$ be a real number. Supppose that $\|\widehat{fd\sigma}\|_p \leqslant C\|f\|_2$ for all functions $f \in B(P)$. Prove that $C \gg N^{2/p-1/2}$ (so, in particular, $\mathrm{Res}(2,p)$ does not hold when $p < 4$).

4. Use stationary phase to get an asymptotic for $\widehat{\chi}(\xi)$, where χ is the characteristic function of the unit ball in $\mathbb{R}^2$.

5. Show that the set $A \subseteq (\mathbb{Z}/p\mathbb{Z})^3$ consisting of all triples (x, x^2, x^3) is a B_3-set. Using such sets construct (for any large N) a set $X \subseteq \{1, \ldots, N\}$ with $|X| \geqslant \frac{1}{10}N^{1/3}$ and $K_6(X) \leqslant 20$.

6*. (The Phragmén–Lindelöf Theorem) Suppose that f is analytic, and that f satisfies

an estimate

$$|f(z)| \leqslant Ce^{|\Im z|^\alpha}$$

for all z with $0 \leqslant \Re z \leqslant 1$, where $\alpha > 1$ and C are fixed constants. Suppose also that f is bounded on the two lines $\Re z = 0$ and $\Re z = 1$. Prove that f is in fact bounded in the whole strip $0 \leqslant \Re z \leqslant 1$.

7*. Let $\psi \in C_0^\infty(\mathbb{R})$. Show that

$$\int e^{i\lambda x^3}\psi(x)\,dx \ll |\lambda|^{-1/3}.$$

8*.** Try and prove that there is an absolute constant C such that $\|\phi\|_6 \leqslant C\|\phi\|_2$ for any eigenfunction ϕ of the Laplacian on $\mathbb{T}^2$.

In this set of exercises p is always a prime and $\mathbb{Z}_p$ is shorthand for $\mathbb{Z}/p\mathbb{Z}$.

1. Let $n = 2m$ be an even integer, and consider the *majority* function $f : \mathbb{F}_2^n \to \{0,1\}$ defined by $f(x_1,\dots,x_n) = 0$ if less than m of the x_i equal 1 and $f(x_1,\dots,x_n) = 1$ otherwise. Show that

$$\widehat{f}((1,1,0,0,\dots,0)) = -2^{2m}\pi^{-1/2}m^{-3/2}(1+o(1)),$$

where the $o(1)$ denotes a function which tends to 0 as $m \to \infty$.

2. Find a set $A \subseteq \mathbb{Z}_p$ with $|A| > p/20$, but such that $A + A$ does not contain an arithmetic progression of length more than $10\sqrt{p}$. Can you find such a set with $|A| > p/2$? *What about with $|A| > p/3$?

3. Suppose that $A \subseteq \mathbb{Z}_p$ is a set with cardinality αp with the property that the only three-term arithmetic progressions in A are the trivial ones of the form (a,a,a). Show that A is α-HNU.

4. Let $f : \{0,1\}^n \to \{0,1\}$ be a boolean function. Show that there is a monotone increasing function $g : \{0,1\}^n \to \{0,1\}$ such that $I_k(f) \geqslant I_k(g)$ for all $k \in \{1,\dots,n\}$, and such that $\mathbb{E}f = \mathbb{E}g$. (Recall that a *monotone increasing* function is a function $h : \{0,1\}^n \to \{0,1\}$ such that if $h(x_1,\dots,x_n) = 1$, and if we change some x_i from zero to one, then h still equals 1).

5. Let p be very large. Show that there is a set $A \subseteq \mathbb{Z}_p$ with $|A| > p/2$, but such that there do not exist 2003 translates $A + x_j$ with

$$\bigcup_{j=1}^{2003}(A + x_j) = \mathbb{Z}_p.$$

6. Let A be a subset of $\mathbb{Z}_p$ containing, for every common difference $d \in \mathbb{Z}_p^*$, an arithmetic progression of length at least $10\sqrt{p}$. Show that $|A| \geqslant p/2$.

Appendix D. Errata to the notes

Thanks to Tom Sanders, Julia Wolf, Graham Lee.

1. Notes2, p3 line 10: $k^2|R_i|^2$ is comparable to 1, not δ^{-1} as previously advertised. Also, in (6) I have changed $\delta^{-\epsilon}$ to δ^{ϵ}.

2. Notes 8, p5. The two occurrences of $R^*(18/5 \to 4)$ have been replaced by $R^*(8/5 \to 4)$. Also an errant factor of $\|f\|_{8/5}$ has been removed from the statement of Theorem 9.

3. Notes 14, p1. "Cardinalities C and D" changed to "Cardinalities γN and δN". p5, line -6: "all of the above" changed to "none of the above".

4. Notes 2, p2. $\underline{d}$ changed to $\overline{d}$ in statement of Problem 2 (Kakeya problem). The definition of $d_F(n)$ has also been changed slightly, by changing the phrase "infinitely many p" to "all p". This means that $d_F(n+m) \leqslant d_F(n) + d_F(m)$, a desirable property, is easy to demonstrate.

5. Notes 5, p3 "length $|BX|$" changed to "length $|AX|$".

6. Notes 9, p1. "For each $u \in \mathbb{F}^3$" changed to "For each $u \in \mathbb{F}$".

7. Notes 11, p4. In equation (3), it should be (and now is) $|m_i|^2 = \lambda$.

8. Notes 12, p3. $\|T^{\otimes n}\|_{1+\epsilon^2 \to 2} \leqslant 1$ replaces $\|T^{\otimes n}\|_{2 \to 1+\epsilon^2} \leqslant 1$