GEOMETRIC MEASURE THEORY

TUOMAS ORPONEN

ABSTRACT. These are the lecture notes for the course *Geometric measure theory*, given at the University of Helsinki in fall semester 2018. The presentation is largely based on the books of Falconer [6] and Mattila [16, 18].

CONTENTS

1. Some discrete problems

This course is about studying the geometric properties of fractals using tools from measure theory (in addition to combinatorial and geometric arguments). What is a fractal? There is no rigorous definition, but usually people have something like Figure 1 in mind. A fractal is typically uncountable, has zero Lebesgue measure, and is very "non-

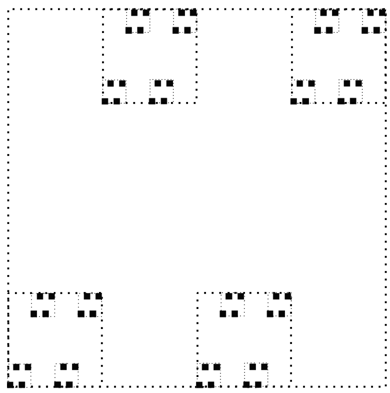

FIGURE 1. A fractal.

smooth". Given a fractal, typical questions we will study are: how big is it, and how do various transformations affect its size? Since fractals can't be (sensibly) measured by either counting or Lebesgue measure, we will need *Hausdorff measures* and *Hausdorff dimension* to quantify their size. These concepts are introduced in Section 2.

The questions we ask about fractals often have their origins in a field of combinatorics known as *incidence geometry*. So, before heading to measures, dimension and so on, we discuss three problems in incidence geometry, whose "fractal versions" will later occupy us during most of the lectures. Let $P \subset \mathbb{R}^2$ be a set with cardinality $|P| = n \in \mathbb{N}$.

Question 1 (Projections). *Let $e \in S^1$ (the unit circle), and let π_e be the projection to the line $\ell_e = \mathrm{span}(e)$. How many directions $e \in S^1$ can there be such that $|\pi_e(P)| \leq n/8$?*

Question 2 (Distance sets). *How many distinct distances does P span? In other words, find a lower bound for the size of the "distance set"*

$$\Delta(P) := \{|p - q| : p, q \in P\}.$$

Question 3 (The Kakeya problem). *Let $m \in \mathbb{N}$, and let $\mathcal{L}$ is a family of m distinct lines in $\mathbb{R}^2$. Assume that*

$$|P \cap \ell| \geq m, \qquad \ell \in \mathcal{L}.$$

Find a lower bound for $n = |P|$.

Questions 1 and 3 are quite easy in $\mathbb{R}^2$ (but not in higher dimensions). To solve them, consider the following notion:

Definition 1.1 (Incidences)**.** Let $P \subset \mathbb{R}^d$ be a set of points, and let $\mathcal{L}$ be a family of lines in $\mathbb{R}^d$. The *incidences* between P and $\mathcal{L}$ are the pairs

$$\mathcal{I}(P, \mathcal{L}) := \{(p, \ell) \in P \times \mathcal{L} : p \in \ell\}.$$

The fundamental question of incidence geometry is: how large can $\mathcal{I}(P, \mathcal{L})$ be (also when $\mathcal{L}$ is replaced by a collection of more complicated sets than lines)? The basic result in $\mathbb{R}^2$ is the following sharp bound of Szemerédi and Trotter [20] from the 80's:

Theorem 1.2. *For any $P \subset \mathbb{R}^2$ and $\mathcal{L}$ a finite set of lines in $\mathbb{R}^2$, we have*

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

We will prove something weaker, but still useful:

Proposition 1.3. *Let $P \subset \mathbb{R}^2$ be a finite set of points, and let $\mathcal{L}$ be a finite set of lines in $\mathbb{R}^2$. Then*

$$|\mathcal{I}(P, \mathcal{L})| \leq 4 \min\{|P||\mathcal{L}|^{1/2} + |\mathcal{L}|, |P|^{1/2}|\mathcal{L}| + |P|\}.$$

Note that the Szemerédi-Trotter theorem seems to "interpolate" between the two terms appearing in the proposition: it's a much better bound when $|P| \approx |\mathcal{L}|$.

Proof of Proposition 1.3. We only prove that

$$|\mathcal{I}(P, \mathcal{L})| \leq 4(|P||\mathcal{L}|^{1/2} + |\mathcal{L}|),$$

since the other bound can be obtained with a similar argument, interchanging the roles of points and lines. We first estimate as follows, using the definition of $\mathcal{I}(P, \mathcal{L})$, and Cauchy-Schwarz:

$$|\mathcal{I}(P, \mathcal{L})| = \sum_{\ell \in \mathcal{L}} |P \cap \ell| \leq |\mathcal{L}|^{1/2} \left(\sum_{\ell \in \mathcal{L}} |P \cap \ell|^2 \right)^{1/2}.$$

Then, we write $|P \cap \ell|^2 = |\{(p, q) \in P \times P : p, q \in \ell\}|$, and exchange the order of summation.

$$\sum_{\ell \in \mathcal{L}} |P \cap \ell|^2 = \sum_{p,q \in P} |\{\ell \in \mathcal{L} : p, q \in \ell\}|.$$

Finally, we separate the "diagonal terms" where $p = q$:

$$\sum_{p,q \in P} |\{\ell \in \mathcal{L} : p, q \in \ell\}| = \sum_{p \in P} |\{\ell \in \mathcal{L} : p \in \ell\}| + \sum_{\substack{p,q \in P \\ p \neq q}} |\{\ell \in \mathcal{L} : p, q \in \ell\}|.$$

The first sum is simply $|\mathcal{I}(P, \mathcal{L})|$ again! For the second sum, note that

$$|\{\ell \in \mathcal{L} : p, q \in \ell\}| \leq 1,$$

whenever $p, q \in \mathbb{R}^2$ are distinct. Putting everything together, and using the basic inequality $(a + b)^{1/2} \leq a^{1/2} + b^{1/2}$ yields

$$|\mathcal{I}(P, \mathcal{L})| \leq |\mathcal{L}|^{1/2} \left(|\mathcal{I}(P, \mathcal{L})| + |P|^2 \right)^{1/2} \leq |\mathcal{L}|^{1/2}|\mathcal{I}(P, \mathcal{L})|^{1/2} + |\mathcal{L}|^{1/2}|P|.$$

If the first term on the right is larger, then $|\mathcal{I}(P, \mathcal{L})| \leq 2|\mathcal{L}|^{1/2}|\mathcal{I}(P, \mathcal{L})|^{1/2}$, which gives

$$|\mathcal{I}(P, \mathcal{L})| \leq 4|\mathcal{L}| \leq 4(|\mathcal{L}|^{1/2}|P| + |\mathcal{L}|)$$

In the opposite case also $|\mathcal{I}(P, \mathcal{L})| \leq 2|\mathcal{L}|^{1/2}|P| \leq 4(|\mathcal{L}|^{1/2}|P| + |\mathcal{L}|)$. □

1.1. **Applying the elementary incidence bound.** We will now apply Proposition 1.3 to Questions 1 and 3 in $\mathbb{R}^2$.

1.1.1. *Solution to Question 1.* Let's first look at Question 1: consider the set of vectors

$$B := \{e \in S^1 : |\pi_e(P)| \leq \tfrac{n}{8}\},$$

where $n = |P|$. By deleting at most half of the vectors in B, we may assume that no distinct vectors in B are parallel. For fun, let's note that a "trivial" bound for $|B|$ is

$$|B| \leq \binom{|P|}{2} \leq |P|^2, \tag{1.4}$$

because for every $e \in B$ there exist distinct points $p, q \in P$ such that $\pi_e(p) = \pi_e(q)$, and there are only $\binom{|P|}{2}$ possible choices of such $\{p, q\}$. Moreover, a fixed couple $\{p, q\}$ can work for at most one vector in $e \in B$, namely the one with $e \perp (p - q)/|p - q|$.

Proposition 1.3 will give something much better than (1.4). For each $e \in B$, consider the family of lines $\mathcal{L}_e := \{\pi_e^{-1}\{t\} : t \in \pi_e(P)\}$. Then

$$|\mathcal{I}(P, \mathcal{L}_e)| = |P| = n, \tag{1.5}$$

because every point in P lies on a unique line in $\mathcal{L}_e$. Then, consider further

$$\mathcal{L} := \bigcup_{e \in B} \mathcal{L}_e.$$

Note that the families $\mathcal{L}_e$ are disjoint for different choices of $e \in B$, because we assumed that B contains no pairs of parallel vectors. It now follows from (1.5) that $|\mathcal{I}(P, \mathcal{L})| = |B|n$. Since $|\mathcal{L}_e| \leq n/8$ for all $e \in B$, Proposition 1.3 gives

$$|B|n = |\mathcal{I}(P, \mathcal{L})| \leq 4(|P||\mathcal{L}|^{1/2} + |\mathcal{L}|) \leq 4n \cdot \left(\frac{|B|n}{8}\right)^{1/2} + \frac{|B|n}{2}.$$

Subtracting $|B|n/2$ from both sides and rearranging gives

$$|\{e \in S^1 : |\pi_e(P)| \leq \tfrac{|P|}{8}\} \leq 64|P|.$$

It's not hard to see that this estimate is sharp in the sense that $|P|$ (on the right hand side) can't be replaced by $|P|^s$ for any $s < 1$. Probably 8 and 64 are not the best constants.

1.1.2. *Solution to Question 3.* Next, we consider Question 3. This is a straightforward application of Proposition 1.3 to the sets $P, \mathcal{L}$. Since $|P \cap \ell| \geq m$ for all $\ell \in \mathcal{L}$, we have

$$m^2 \leq |\mathcal{I}(P, \mathcal{L})| \leq 4(|P|^{1/2}m + |P|),$$

now using the second inequality in Proposition 1.3. It follows that $n = |P| \gtrsim m^2$, which is clearly optimal (up to multiplicative constants).

Question 3 is much harder in higher dimensions. A naive formulation is the following: assume that $\mathcal{L}$ is a collection of m^{d-1} lines in $\mathbb{R}^d$, each containing m points of P. How big is P? One might first guess that

$$|P| \gtrsim m^{d-1} \cdot m = m^d,$$

in analogy with the planar result. This already fails in $\mathbb{R}^3$. In fact, if all the m^2 lines in $\mathcal{L}$ are all allowed to lie on a common plane $V \subset \mathbb{R}^3$, then it's possible to arrange them so, and find a corresponding set $P \subset V \subset \mathbb{R}^3$, such that $|P \cap \ell| \geq m$ for all $\ell \in \mathcal{L}$ and

$$|P| \approx m^{5/2} \ll m^3.$$

However, this is the only obstacle in $\mathbb{R}^3$: Guth and Katz [9] have shown that if no plane in $\mathbb{R}^3$ contains more than m lines, then $|P| \gtrsim m^3$. In dimensions $d \geq 4$, the question is open: one needs to assume that the lines are not concentrated on algebraic varieties of low degree, but it's not known if this is sufficient to guarantee $|P| \gtrsim n^d$.

Regarding the distance set problem, Question 2, Proposition 1.3 says nothing: the problem concerns incidences between points and circles, and that's much more difficult. We leave the following bounds as exercises:

Exercise 1.6. For given $n \in \mathbb{N}$, find a set $P \subset \mathbb{R}^2$ such that $|\Delta(P)| \lesssim n$. Then, prove that $|\Delta(P)| \gtrsim n^{1/2}$. *Hint*: consider two distinct points $p_1, p_2 \in P$ and show that either

$$|\{|p_1 - q| : q \in P\}| \gtrsim n^{1/2} \quad \text{or} \quad |\{|p_2 - q| : q \in P\}| \gtrsim n^{1/2}.$$

Remark 1.7. Arranging P in a $(\sqrt{n} \times \sqrt{n})$-grid gives $|\Delta(P)| \sim n/\sqrt{\log n}$, and Erdős conjectured that this would be the sharp lower bound. It has been recently shown by Guth and Katz [10] that $|\Delta(P)| \gtrsim n/\log n$, which is almost sharp!

2. Dimension

We now leave the discrete world behind, and start looking for "fractal versions" of Questions 1, 2, and 3, concerning infinite sets P. It no longer makes sense to measure the sizes of $P, \pi_e(P)$ or $\Delta(P)$ with cardinality. There are various options we could employ as a replacement: the most widely used is, no doubt, *Hausdorff dimension*.

Most people have an intuitive idea of what "dimension" means: a point should have dimension zero, a curve should have dimension one, a surface should have dimension two and so on. Before heading for a definition, let's think for a moment: where does this intuition actually come from?

Definition 2.1 (Covering number)**.** Given a bounded set $E \subset \mathbb{R}^d$, and a radius $\delta > 0$, let $N(E, \delta)$ be the minimal number of closed balls of diameter δ needed to cover E.

Typically, when $\delta \to 0$, the covering number $N(E, \delta) \to \infty$, and the **rate of increase of** $N(E, \delta)$ is closely related with the dimension of E.

Exercise 2.2. Recall that a map $f\colon (X, d) \to (Y, d')$ is biLipschitz, if there exists a constant $C \geq 1$ such that

$$\frac{d'(f(x), f(y))}{C} \leq d(x, y) \leq Cd'(f(x), f(y)), \qquad x, y \in X.$$

Let $f\colon [0,1]^d \to S \subset \mathbb{R}^D$ be biLipschitz, $D \geq d$, where $S := f([0,1]^d)$. Show that

$$N(S, \delta) \sim \delta^{-d}, \qquad \delta > 0.$$

So, it seems that if $N(E, \delta) \sim \delta^{-s}$ for some $s \geq 0$, then the exponent s reflects the "dimension" of E. We can turn this into a definition:

Definition 2.3 (Box dimensions)**.** Given a bounded set $E \subset \mathbb{R}^d$, we write

$$\overline{\dim}_{\mathrm{B}} E := \limsup_{\delta \to 0} \frac{\log N(E, \delta)}{-\log \delta} \quad \text{and} \quad \underline{\dim}_{\mathrm{B}} E := \liminf_{\delta \to 0} \frac{\log N(E, \delta)}{-\log \delta}.$$

If $\overline{\dim}_{\mathrm{B}} E = \underline{\dim}_{\mathrm{B}} E$, we denote the common value by $\dim_{\mathrm{B}} E$. These three concepts are called the upper box dimension, the lower box dimension, and the box dimension, of E. Note that the latter need not always exist.

While the box dimension is perhaps the most "intuitive" notion of dimension, and it gives the "correct" value in simple cases like the ones in Exercise 2.2, it has problems. Here are two obvious ones (there will be more later):

- It doesn't always exist, and it's not too pretty to work with lower and upper box dimensions separately. Also, it doesn't work for unbounded sets E (because it can happen that $N(E,\delta) \equiv \infty$ for $\delta > 0$).
- $\dim_{\mathrm{B}} \mathbb{Q} \cap [0,1] = 1$, even though $\mathbb{Q} \cap [0,1]$ is countable. This is bit weird: a measure theorist always likes when her concepts are "countably stable". So, she would prefer a dimension "dim" which satisfies

$$\dim \bigcup_{i \in \mathbb{N}} E_i = \sup_{i \in \mathbb{N}} \dim E_i. \tag{2.4}$$

  Then, because every point should have dimension zero, also countable sets (such as $\mathbb{Q} \cap [0,1]$) should have dimension zero.

2.1. **Hausdorff measures and Hausdorff dimension.** The concept of Hausdorff dimension fixes these problems, and leads to a very pleasant theory. Before talking about Hausdorff dimension, we introduce *Hausdorff measures*.

Definition 2.5 (Hausdorff measures)**.** Let $s \geq 0$ and $\delta \in (0, \infty]$. Given a set $E \subset \mathbb{R}^d$, a *δ-cover of E* is any countable family of sets $\{U_j\}_{j \in \mathbb{N}}$ such that

$$E \subset \bigcup_{j \in \mathbb{N}} U_j \quad \text{and} \quad \mathrm{diam}(U_j) \leq \delta \text{ for all } j \in \mathbb{N}. \tag{2.6}$$

Note that when $\delta = \infty$, the second requirement becomes vacuous. The *s-dimensional Hausdorff δ-measure* is

$$\mathcal{H}^s_\delta(E) := \inf \left\{ \sum_{j \in \mathbb{N}} \mathrm{diam}(U_j)^s : \{U_j\}_{j \in \mathbb{N}} \text{ is a } \delta\text{-cover of } E \right\}.$$

Note that when $\delta \searrow 0$, the condition of being a δ-cover becomes more restrictive, so there are fewer candidates in the "inf" above, and hence $\mathcal{H}^s_\delta(E)$ increases:

$$\mathcal{H}^s_{\delta_1}(E) \geq \mathcal{H}^s_{\delta_2}(E), \qquad 0 < \delta_1 \leq \delta_2.$$

Consequently the limit

$$\mathcal{H}^s(E) := \lim_{\delta \to 0} \mathcal{H}^s_\delta(E) = \sup_{\delta > 0} \mathcal{H}^s_\delta(E) \in [0, \infty]$$

exists, and $\mathcal{H}^s(E)$ is called the *s-dimensional Hausdorff measure of E*.

Exercise 2.7. This exercise collects some basic facts about Hausdorff measures:

(i) Show that, for $d \in \mathbb{N}$, there is a constant $C_d > 0$ such that

$$\frac{\mathcal{H}^d(E)}{C_d} \leq \mathcal{L}^d(E) \leq C_d \mathcal{H}^d(E), \qquad E \subset \mathbb{R}^d.$$

Here $\mathcal{L}^d$ is Lebesgue measure on $\mathbb{R}^d$. In fact, one can even show (but you don't need to) that $\mathcal{H}^d(E) = C_d \mathcal{L}^d(E)$ for all $E \subset \mathbb{R}^d$.

(ii) Show that if $0 \leq s_1 \leq s_2$, then

$$\mathcal{H}^{s_1}(E) \geq \mathcal{H}^{s_2}(E), \qquad E \subset \mathbb{R}^d.$$

(iii) Show that if $0 \le s_1 < s_2 < \infty$ and $E \subset \mathbb{R}^d$, then
$$\mathcal{H}^{s_1}(E) = \infty \quad \text{or} \quad \mathcal{H}^{s_2}(E) = 0.$$
(iv) If $s > d$, show that $\mathcal{H}^s(\mathbb{R}^d) = 0$.
(v) Prove that if $E \subset \mathbb{R}^d$ and $s \ge 0$, then
$$\mathcal{H}^s(E) = 0 \quad \Longleftrightarrow \quad \mathcal{H}^s_\infty(E) = 0.$$
(vi) Prove that if $d_1, d_2 \in \mathbb{N}$, $A \subset \mathbb{R}^{d_1}$ is a set, and $f \colon A \to \mathbb{R}^{d_2}$ is L-Lipschitz, then
$$\mathcal{H}^s(f(A)) \lesssim_L \mathcal{H}^s(A).$$

Exercises 2.7(ii)-(iii) show that the function
$$s \mapsto g(s) := \mathcal{H}^s(E)$$
is non-increasing, and if $g(s_1) < \infty$ for any $s_1 \ge 0$, then $g(s) = 0$ for all $s > s_1$. So, informally, first $g(s) = \infty$ for a while, and then g abruptly drops to zero. The value of s at which g drops to zero is called the *Hausdorff dimension* of E:

Definition 2.8 (Hausdorff dimension)**.** The Hausdorff dimension of a set $E \subset \mathbb{R}^d$ is the number
$$\dim_{\mathrm{H}} E := \sup\{s \ge 0 : \mathcal{H}^s(E) = \infty\} = \inf\{s \ge 0 : \mathcal{H}^s(E) = 0\}.$$

Some remarks are in order:

- By definition, we have
$$\mathcal{H}^s(E) \in \{0, \infty\}, \qquad s \in [0, \infty) \setminus \{\dim_{\mathrm{H}} E\}.$$
For $s = \dim_{\mathrm{H}} E$, all three possibilities can occur: $\mathcal{H}^s(E) = 0$, or $0 < \mathcal{H}^s(E) < \infty$, or $\mathcal{H}^s(E) = \infty$.
- It is clear from Exercise 2.7(i)&(iv) that $\dim_{\mathrm{H}} \mathbb{R}^d = d$, and hence $\dim_{\mathrm{H}} E \in [0, d]$ for all $E \subset \mathbb{R}^d$.
- By Exercise 2.7(v), another definition of Hausdorff dimension is
$$\dim_{\mathrm{H}} E = \inf\{s \ge 0 : \mathcal{H}^s_\infty(E) = 0\} = \sup\{s \ge 0 : \mathcal{H}^s_\infty(E) > 0\}. \tag{2.9}$$
We could **not** define Hausdorff dimension by $\sup\{s \ge 0 : \mathcal{H}^s_\infty(E) = \infty\}$, because $\mathcal{H}^s_\infty(E) \le \operatorname{diam}(E)^s$, so $\mathcal{H}^s_\infty(E)$ is actually finite for all bounded sets.

Exercise 2.10. Verify that Hausdorff dimension is monotone,
$$E_1 \subset E_2 \quad \Longrightarrow \quad \dim_{\mathrm{H}} E_1 \le \dim_{\mathrm{H}} E_2$$
and satisfies the "countable stability" (2.4).

Remark 2.11. Let's compare Hausdorff measures and dimension with the δ-covering numbers and box dimensions mentioned in the previous section. Let $E \subset \mathbb{R}^d$ be a bounded set. Any (finite) cover $\{B(x_j, \frac{\delta}{2})\}_{j=1}^N$ of E by balls of diameter δ is clearly a δ-cover, in the sense of (2.6). Hence
$$\mathcal{H}^s_\delta(E) \le N \cdot \delta^s,$$
and taking the inf over all such finite ball-covers gives
$$\mathcal{H}^s_\delta(E) \le N(E, \delta) \cdot \delta^s,$$
by definition of $N(E, \delta)$. We can now easily prove the following inequality:
$$\dim_{\mathrm{H}} E \le \underline{\dim}_{\mathrm{B}} E. \tag{2.12}$$

Indeed, fix

$$s > \underline{\dim}_{\mathrm{B}} E = \liminf_{\delta \to 0} \frac{\log N(E,\delta)}{-\log \delta}.$$

Then there exist arbitrarily small scales δ_j such that

$$N(E,\delta_j) \leq \delta_j^{-s},$$

and consequently

$$\mathcal{H}^t_{\delta_j}(E) \leq N(E,\delta_j) \cdot \delta_j^s \leq 1.$$

This gives $\mathcal{H}^s(E) \leq 1 < \infty$, and hence $\dim_{\mathrm{H}} E \leq s$. This proves (2.12).

2.2. **Some general measure theory.** The objects defined in the previous section were called Hausdorff **measures**. In this section, we clarify that this really makes sense, and collect some generalities about measures in $\mathbb{R}^d$.

Definition 2.13. In these lecture notes, a *measure* is a function $\mu \colon \mathcal{P}(\mathbb{R}^d) \to [0,\infty]$ with the following three properties:

(M1) $\mu(\emptyset) = 0$,
(M2) if $E_1 \subset E_2$, then $\mu(E_1) \leq \mu(E_2)$, and
(M3) if $E_1, E_2, \ldots \subset \mathbb{R}^d$ is a countable sequence of sets with $E := \bigcup_j E_j$, then

$$\mu(E) \leq \sum_{j=1}^{\infty} \mu(E_j).$$

Sometimes, our measures are also called *outer measures*.

Lemma 2.14. *For any $s \geq 0$ and $\delta \in (0,\infty]$ the set functions $\mathcal{H}^s_\delta$ and $\mathcal{H}^s$ are measures.*

Proof. Voluntary exercise: if you've seen it for Lebesgue measure, you won't be too surprised. □

Why do we bother considering $\mathcal{H}^s$ at all? Why not just work with $\mathcal{H}^s_\delta$? The main problem is that $\mathcal{H}^s_\delta$ does not have many measurable sets. Recall the definition:

Definition 2.15 (Measurable sets)**.** Let μ be a measure. A set $A \subset \mathbb{R}^d$ is called μ measurable if

$$\mu(E) = \mu(E \cap A) + \mu(E \setminus A), \qquad E \subset \mathbb{R}^d.$$

Exercise 2.16. If $0 \leq s \leq 1$ and $d \geq 2$ show that

$$\mathcal{H}^s_2(U(0,1)) = \mathcal{H}^s_2(B(0,1)) = \mathcal{H}^s_2(\partial B(0,1)),$$

where $U(0,1)$ is the open unit ball in $\mathbb{R}^d$, $B(0,1) = \overline{U(0,1)}$. Conclude that $U(0,1)$ is **not** $\mathcal{H}^s_2$ measurable.

Definition 2.17 (Borel and Borel regular measures)**.** A measure μ on $\mathbb{R}^d$ (or any topological space) is called a Borel measure, if all Borel sets are μ measurable. A measure μ is called Borel regular, if it is a Borel measure, and for an arbitrary set $A \subset \mathbb{R}^d$ there exists a Borel set $B \supset A$ with $\mu(A) = \mu(B)$.

Since $U(0,1)$ is open, in particular Borel, Exercise 2.16 shows that $\mathcal{H}^s_\delta$ is typically **not** a Borel measure. This is one of the main reasons why $\mathcal{H}^s$ is so much nicer than $\mathcal{H}^s_\delta$:

Lemma 2.18. *$\mathcal{H}^s$ is a Borel regular measure for any $s \geq 0$. In particular, if $B_1, B_2, \ldots$ is a sequence of disjoint Borel sets, and $B := \bigcup_j B_j$, then*

$$\mathcal{H}^s(B) = \sum_{j \in \mathbb{N}} \mathcal{H}^s(B_j).$$

Proof. We omit the proof that $\mathcal{H}^s$ is a Borel measure, see for example Theorem 1.5 and Section 1.2 in Falconer's book [6]. The idea is to show (easy exercise) that Hausdorff measures are *metric outer measures* in the sense that $\mathcal{H}^s(A \cup B) = \mathcal{H}^s(A) + \mathcal{H}^s(B)$ for all sets $A, B \subset \mathbb{R}^d$ with $\operatorname{dist}(A, B) > 0$. Then, a general result (Theorem 1.5 in Falconer's book) states that metric outer measures are Borel measures. The Borel regularity of Hausdorff measures is an **exercise**. The final claim follows from basic measure theory: any measure μ has this property for disjoint μ measurable sets. □

We conclude this section with a very useful "approximation" result for Borel measures:

Lemma 2.19. *Let μ be a Borel measure on $\mathbb{R}^d$, $\varepsilon > 0$, and let $A \subset \mathbb{R}^d$ be Borel.*

(i) *If $\mu(A) < \infty$, there exists a compact set $K \subset A$ such that $\mu(A \setminus K) < \varepsilon$.*
(ii) *If A can be covered by countably many open sets with finite μ measure, then there exists an open set $U \supset A$ with $\mu(U \setminus A) < \varepsilon$.*

Proof. We start by proving (i) It clearly suffices to prove this part of the theorem for the restriction $\mu|_A$, so we may assume without loss of generality that $\mu(\mathbb{R}^d) < \infty$ (but in (i) only; of course the extra assumption in (ii) is vacuous for finite measures, so it's not something we can assume w.l.o.g. there). We then consider the collection $\mathcal{F}$ of sets $B \subset \mathbb{R}^d$ with the following property: for any $\varepsilon > 0$, there exist a closed set $C \subset B$ and an open set $U \supset B$ such that $\mu(U \setminus C) < \varepsilon$. We will prove that $\mathcal{F}$ contains all the Borel sets: this will conclude the proof of (i), because then $A \in \mathcal{F}$, hence A can be approximated by closed sets from inside, and finally also by compact sets, because closed sets are σ-compact.

First note that $\mathcal{F}$ contains all the closed sets, because every closed set is the countable intersection of a decreasing sequence open sets (the $(1/j)$-neighbourhoods for example). So, it remains to check that $\mathcal{F}$ is a σ-algebra. It's clear (check as an exercise if it isn't) that $\mathcal{F}$ contains the empty set, and is closed under countable unions. Also, $\mathcal{F}$ is closed under taking complements: indeed, fix $B \in \mathcal{F}$ and $\varepsilon > 0$, and start by finding a closed set $C \subset B$ and an open set $U \supset B$ such that $\mu(U \setminus C) < \varepsilon$. Then note that

$$C' := U^c \subset B^c \quad \text{and} \quad U' := C^c \supset B^c,$$

and $\mu(U' \setminus C') = \mu(U \setminus C) < \varepsilon$. This shows that $B^c \in \mathcal{F}$, and the proof of (i) is done.

To prove (ii), let $V_1, V_2, \ldots$ be a sequence of open sets with $A \subset \bigcup V_i$ and $\mu(V_i) < \infty$. Since $V_i \setminus A$ is Borel and has finite measure, we may nearly exhaust it with a closed set according to (i): there exists $C_i \subset V_i \setminus A$ closed such that $\mu([V_i \setminus A] \setminus C_i) < \epsilon/2^i$. Then the sets $U_i := V_i \setminus C_i$ are open, and

$$A \subset U := \bigcup U_i \quad \text{and} \quad \mu(U \setminus A) \leq \sum \mu(U_i \setminus A) < \sum \frac{\epsilon}{2^i} = \epsilon.$$

This concludes the proof of (ii). □

2.3. **Continuous analogues of Questions 1, 2 and 3.** Now we have introduced the central notions of the course – Hausdorff measures and dimension – so we may formulate "fractal versions" of the three problems mentioned in the first section.

Question 4 (Projections). *Assume that $E \subset \mathbb{R}^2$ is a compact set. Is it true that*

$$\dim_{\mathrm{H}} \pi_e(E) = \min\{\dim_{\mathrm{H}} E, 1\}$$

for all $e \in S^1$? Or at least some $e \in S^1$?

It's clear that $\dim_{\mathrm{H}} \pi_e(E) \leq 1$, since $\pi_e(E)$ is a subset of ℓ_e, and $\dim_{\mathrm{H}} \ell_e = 1$. This is why there's the "min" above, and also below:

Question 5 (Distance sets). *Assume that $E \subset \mathbb{R}^2$ is a compact set. Is it true that*

$$\dim_{\mathrm{H}} \Delta(E) = \min\{\dim_{\mathrm{H}} E, 1\}?$$

Question 6 (The Kakeya problem). *Assume that $E \subset \mathbb{R}^d$ is a set containing a unit line segment in every direction. Is it true that $\mathcal{H}^d(E) > 0$. Or at least $\dim_{\mathrm{H}} E = d$?*

In the upcoming sections, we will develop some tools to answer these problems. The answers will be only partial: in particular the two latter questions are still open!

3. The mass distribution principle and Frostman's lemma

In all of the questions in Section 2.3, we need to estimate the Hausdorff dimension of certain sets from below (the estimates from above are often much easier, and they are essentially trivial in the problems above). The next simple lemma gives a very useful criterion:

Lemma 3.1 (Mass distribution principle). *Assume that $E \subset \mathbb{R}^d$, and there a exists measure μ on $\mathbb{R}^d$ such that*

$$\mu(B(x,r)) \leq Cr^s, \qquad x \in E, \ 0 < r \leq r_0. \tag{3.2}$$

Then $\mathcal{H}^s_{r_0}(E) \geq \mu(E)/C$. In particular, if $\mu(E) > 0$, then $\mathcal{H}^s(E) > 0$, and hence $\dim_{\mathrm{H}} E \geq s$.

Proof. Let $\{U_j\}_{j\in\mathbb{N}}$ be an r_0-cover of E. It suffices to show that

$$\sum_{j\in\mathbb{N}} \operatorname{diam}(U_j)^s \geq \frac{\mu(E)}{C}. \tag{3.3}$$

Start by discarding all the sets U_j which do not meet E; the remaining sets still cover E, and we will show that (3.3) holds for these remaining sets U_j. Now, for each U_j, pick $x_j \in E \cap U_j$, and note that

$$U_j \subset B(x_j, \operatorname{diam}(U_j)).$$

In particular, the balls $B(x_j, \operatorname{diam}(U_j))$ cover E, and have radius at most r_0. Consequently

$$\sum_{j\in\mathbb{N}} \operatorname{diam}(U_j)^s \overset{(3.2)}{\geq} \frac{1}{C} \sum_{j\in\mathbb{N}} \mu(B(x_j, \operatorname{diam}(U_j))) \geq \frac{\mu(E)}{C},$$

using the subadditivity of μ in the second inequality. ∎

We will eventually apply this criterion to Questions 4 and 5. To do so, we need to find measures supported on $\pi_e(E)$ and $\Delta(E)$, satisfying (3.2). The general idea will be similar in both cases: for example, we note that $\pi_e(E)$ is the image of E under the map π_e. So, if we first find a measure supported on E, then we can *push-forward* it to a measure supported on $\pi_e(E)$. We now review the relevant concepts to make the discussion more rigorous.

Definition 3.4 (Support of a measure)**.** Let μ be a measure on a separable metric space (X, d). The *support* of μ is the set

$$\operatorname{spt} \mu := \{x \in X : \mu(B(x, r)) > 0 \text{ for all } r > 0\}.$$

The support of μ is evidently closed: if $\{x_i\}_{i \in \mathbb{N}} \subset \operatorname{spt} \mu$ is a sequence converging to some point $x \in X$, and $r > 0$, then $B(x, r)$ contains $B(x_i, r/2)$ for some $i \in \mathbb{N}$ large enough, and hence $\mu(B(x, r)) > 0$. This means that $x \in \operatorname{spt} \mu$. Another common definition of $\operatorname{spt} \mu$ is the following: $\operatorname{spt} \mu$ is the smallest closest set F such that $\mu(X \setminus F) = 0$. These definitions agree on all separable metric spaces.

Definition 3.5 (Push-forward)**.** Let μ be a measure on a space X, and let $f \colon X \to Y$ be a map, where Y is another arbitrary space. We define the *push-forward of μ under f* as the measure $f\mu$ defined by

$$f\mu(A) := \mu(f^{-1}(A)), \qquad A \subset Y.$$

It is easy to check that μ is a measure. The following estimate gives subadditivity:

$$f\mu\left(\bigcup E_i\right) = \mu\left(f^{-1}\left(\bigcup E_i\right)\right) = \mu\left(\bigcup f^{-1}(E_i)\right) \le \sum \mu(f^{-1}(E_i)) = \sum f\mu(E_i),$$

while it is trivial that $f\mu$ is monotone, and $f\mu(\emptyset) = 0$. We will only apply the concept of push-forward in the case where f is a Borel map (between two topological spaces X, Y) and μ is a Borel measure. Then $f\mu$ is also a Borel measure: if $A \subset Y$ is Borel, and $E \subset Y$ is arbitrary, then $f^{-1}(A) \subset X$ is Borel, and hence

$$\begin{aligned} f\mu(E) = \mu(f^{-1}(E)) &= \mu(f^{-1}(E) \cap f^{-1}(A)) + \mu(f^{-1}(E) \setminus f^{-1}(A)) \\ &= \mu(f^{-1}(E \cap A)) + \mu(f^{-1}(E \setminus A)) = f\mu(E \cap A) + f\mu(E \setminus A). \end{aligned}$$

In particular, if $A_1, A_2, \ldots$ is a sequence of disjoint Borel sets with $A := \bigcup A_i$, then $f\mu(A) = \sum f\mu(A_i)$. Finally, we record the following lemma:

Lemma 3.6. *Assume that $(X, d), (Y, d')$ are separable, $f \colon X \to Y$ is continuous, and μ is a measure on X with compact support. Then*

$$\operatorname{spt} f\mu = f(\operatorname{spt} \mu),$$

and in particular $\operatorname{spt} f\mu$ is compact. Moreover, if $g \colon Y \to [0, \infty]$ is a non-negative Borel function, and μ is a Borel measure, then

$$\int_Y g \, df\mu = \int_X (g \circ f) \, d\mu. \tag{3.7}$$

Proof. Exercise. □

3.1. **Frostman's lemma.** We return to discussing the Questions. Our first goal will be to tackle Question 4 about projections. As already explained above, the idea will be to start with a compact set $E \subset \mathbb{R}^2$, find a suitable non-zero measure μ with $\operatorname{spt} \mu \subset E$, and then consider the push-forward measures $\pi_e \mu$ for various $e \in S^1$. By Lemma 3.6, we know that

$$\operatorname{spt} \pi_e \mu \subset \pi_e(\operatorname{spt} \mu) \subset \pi_e(E),$$

So, if we manage to show that $\pi_e \mu$ satisfies an estimate of the form (3.2) from the mass distribution principle, we will find a lower bound for $\dim_{\mathrm{H}} \pi_e(E)$!

One crucial missing piece from this story is: how to find a "suitable non-zero measure μ with $\operatorname{spt} \mu \subset E$?" The next result, due to Frostman [7], provides a very useful answer:

Lemma 3.8 (Frostman's lemma)**.** *Assume that $E \subset \mathbb{R}^d$ is a compact set with $\mathcal{H}^s(E) > 0$. Then, there exists a compactly supported Borel measure μ with $\operatorname{spt} \mu \subset E$ and $\mu(E) \gtrsim_d \mathcal{H}^s_\infty(E)$, such that*

$$\mu(B(x,r)) \leq r^s, \qquad x \in \mathbb{R}^d, \ r > 0. \tag{3.9}$$

Remark 3.10. Note that this lemma is a converse to the mass distribution principle, Lemma 3.1. The assumption that E is compact can be relaxed: Frostman's lemma holds for all Borel sets E, and even for analytic sets E, but the proof is substantially trickier.

Since the decay condition $\mu(B(x,r)) \lesssim r^s$ will play such a prominent role in the sequel, we give it a name:

Definition 3.11 (Frostman measures)**.** A Borel measure on $\mathbb{R}^d$ is called an s-Frostman measure, if there is a constant $C \geq 1$ such that $\mu(B(x,r)) \leq Cr^s$ for all $x \in \mathbb{R}^d$ and $r > 0$.

Thus, Frostman's lemma states that a Borel set with positive s-dimensional Hausdorff measure supports an s-Frostman measure. After the lemma has been proven, we will often use it to find finite but non-trivial Borel measures μ which have compact support contained in a Borel set E. It gets a bit long to write all of that repeatedly, so we introduce the following notation:

Notation 3.12. Given a set $E \subset \mathbb{R}^d$, we write $\mathcal{M}(E)$ for the family of finite Borel measures μ with compact support satisfying

$$\operatorname{spt} \mu \subset E \quad \text{and} \quad \mu(E) > 0.$$

The proof of Frostman's lemma has two parts: first, a "combinatorial" argument, where one finds a sequence of measures μ_i supported on smaller and smaller neighbourhoods of E, and each μ_i satisfying (3.9) with a uniform constant C. In the second part, one needs to get convinced that this sequence (or rather its subsequence) actually converges somewhere: the limit object will be the measure $\mu \in \mathcal{M}(E)$ claimed in the lemma.

3.2. **Weak convergence and compactness of measures.** The second part of the proof requires some technology, which we discuss first.

Definition 3.13 (Weak convergence of measures)**.** Let $\{\mu_j\}_{j \in \mathbb{N}}$ and μ be locally finite[1] Borel measures on a metric space (X,d). We say that the measures μ_j converge weakly to μ if

$$\int \varphi \, d\mu_j \to \int \varphi \, d\mu$$

[1]Recall that a *locally finite mesure* gives finite measure to all compact sets.

for all $\varphi \in \mathcal{C}_0(X)$, i.e. all continuous functions $\varphi\colon X \to \mathbb{C}$ with compact support. In this case, we write $\mu_j \rightharpoonup \mu$.

Example 3.14. *The measures $\delta_{\{i\}}$, $i \in \mathbb{N}$, converge weakly to zero as $i \to \infty$. Indeed, if $\varphi \in \mathcal{C}_0(\mathbb{R})$, then $i \notin \operatorname{spt} \varphi$ for $i \geq i_0$, large enough, and then*

$$\int \varphi \, d\delta_{\{i\}} = \varphi(i) = 0, \qquad i \geq i_0.$$

As another example, the measures $\mu_k := \frac{1}{k}\sum_{j=1}^{k} \delta_{j/k}$ converge weakly to Lebesgue measure on $[0,1]$, because

$$\int \varphi \, d\mu_k = \frac{1}{k}\sum_{j=1}^{k} \varphi(\tfrac{j}{k}) \to \int_0^1 \varphi(x)\, dx, \qquad \varphi \in \mathcal{C}_0(\mathbb{R}).$$

Warning: The first example above shows that weak convergence can sometimes be a little unintuitive: there we saw a sequence of measures $\mu_j = \delta_{\{j\}}$ with $\mu_j(\mathbb{R}) \equiv 1$, converging weakly to a measure $\mu \equiv 0$ with $\mu(\mathbb{R}) = 0$. The following is still true:

Lemma 3.15. *Assume that $\{\mu_j\}_{j\in\mathbb{N}}$ is a sequence of locally finite Borel measures in a locally compact metric space (X, d) converging weakly to a locally finite Borel measure μ. If $K \subset X$ is compact, and $U \subset X$ is open and σ-compact, then the following inequalities hold:*

$$\mu(K) \geq \limsup_{j\to\infty} \mu_j(K) \quad \text{and} \quad \mu(U) \leq \liminf_{j\to\infty} \mu_j(U).$$

Proof. We start with the first inequality. Fix $\varepsilon > 0$. Since K is compact and X is locally compact, there exists an open set $U' \supset K$ with compact closure, and hence $\mu(U') < \infty$. Then, it follows that $\mu(\{x : \operatorname{dist}(x, K) < 1/j\}) \to \mu(K)$ as $j \to \infty$ (the $(1/j)$-neighbourhoods are contained in U' for $j \in \mathbb{N}$ large enough), and consequently we may find an open set $U \supset K$ with $\mu(K) \geq \mu(U) - \varepsilon$. Then, pick a function $\varphi \in \mathcal{C}_0(X)$ satisfying

$$\chi_K \leq \varphi \leq \chi_U, \tag{3.16}$$

for example

$$\varphi(x) = \max\left\{0, 1 - \frac{\operatorname{dist}(x, K)}{\operatorname{dist}(K, U^c)}\right\}.$$

By definition of weak convergence, we then have

$$\mu(K) \geq \mu(U) - \varepsilon \geq \int \varphi \, d\mu - \varepsilon = \lim_{j\to\infty} \int \varphi \, d\mu_j \geq \limsup_{j\to\infty} \mu_j(K) - \varepsilon.$$

Let $\varepsilon \to 0$ to complete the proof of the first part. The proof of the second inequality is similar: pick any compact set $K \subset U$, and find, again, a function $\varphi \in \mathcal{C}_0(X)$ satisfying (3.16). Then

$$\mu(K) \leq \int \varphi \, d\mu = \lim_{j\to\infty} \int \varphi \, d\mu_j \leq \liminf_{j\to\infty} \mu_j(U).$$

Since U is σ-compact, this implies that $\mu(U) \leq \liminf_{j\to\infty} \mu_j(U)$. □

The next result is the most crucial one: it says (cheating a little) that the set of locally finite Borel measures is compact in the topology of weak convergence:

Lemma 3.17. *Let $\{\mu_j\}_{j\in\mathbb{N}}$ be a sequence of Borel measures on $\mathbb{R}^d$ satisfying*

$$\sup_{j\in\mathbb{N}} \mu_j(K) < \infty \tag{3.18}$$

for all compact sets $K \subset \mathbb{R}^d$. Then, there exists a locally finite Borel measure μ, and a subsequence $\{\mu_{j_i}\}_{i\in\mathbb{N}}$, such that

$$\mu_{j_i} \rightharpoonup \mu.$$

Proof. Let $\|\cdot\|$ be the sup-norm in the space $\mathcal{C}_0(\mathbb{R}^d)$. We will use the easy fact that $(\mathcal{C}_0(\mathbb{R}^d), \|\cdot\|)$ is separable, that is, there exists a countable dense subset $\{\varphi_k\}_{k\in\mathbb{N}} \subset \mathcal{C}_0(\mathbb{R}^d)$. The sequence can also be chosen so that any function $\varphi \in \mathcal{C}_0(\mathbb{R}^d)$ supported in $B(0,M)$ can be approximated by functions supported in $B(0,2M)$, for $M \in \mathbb{N}$.

Then, for each $k \in \mathbb{N}$, choose a subsequence $\{j_i^k\}_{i\in\mathbb{N}}$ such that

$$\exists \alpha_k := \lim_{i\to\infty} \int \varphi_k \, d\mu_{j_i^k}.$$

Such subsequences exist, because the set of real numbers

$$\left\{ \int \varphi_k \, d\mu_j : j \in \mathbb{N} \right\}$$

is bounded by (3.18), and recalling that φ_k has compact support. Moreover, we may always take $\{j_i^{k+1}\}_{i\in\mathbb{N}}$ to be a subsequence of $\{j_i^k\}_{i\in\mathbb{N}}$, for any $k \in \mathbb{N}$, just by picking the sequences one at a time. Then, the sequence $\{j_m^m\}_{m\in\mathbb{N}}$ is an "eventual" subsequence of every sequence $\{j_i^k\}$, $k \in \mathbb{N}$: more precisely $j_m^m \in \{j_i^k\}_{i\in\mathbb{N}}$ for all $m \geq k$, because then $j_m^m \in \{j_i^m\}_{i\in\mathbb{N}} \subset \{j_i^k\}_{i\in\mathbb{N}}$. It follows that

$$\alpha_k = \lim_{m\to\infty} \int \varphi_k \, d\mu_{j_m^m}, \qquad k \in \mathbb{N}. \tag{3.19}$$

Now, it follows from the density of the sequence $\{\varphi_k\}_{k\in\mathbb{N}}$ that actually the limit

$$\lim_{m\to\infty} \int \varphi \, d\mu_{j_m^m} =: \Lambda(\varphi) \tag{3.20}$$

exists for every $\varphi \in \mathcal{C}_0(\mathbb{R}^d)$. To see this, pick $\varphi \in \mathcal{C}_0(\mathbb{R}^d)$ with support in $B(0,M)$, and let $\varphi_{k_1}, \varphi_{k_2}, \ldots$ be a subsequence of $\{\varphi_k\}_{k\in\mathbb{N}}$ with $\operatorname{spt} \varphi_{k_l} \subset B(0,2M)$ and $\|\varphi - \varphi_{k_l}\| \to 0$ as $l \to \infty$. Then for any $l, L \in \mathbb{N}$,

$$\begin{aligned} |\alpha_{k_l} - \alpha_{k_L}| &\leq \limsup_{m\to\infty} \left(\left| \alpha_{k_l} - \int \varphi \, d\mu_{j_m^m} \right| + \left| \alpha_{k_L} - \int \varphi \, d\mu_{j_m^m} \right| \right) \\ &= \limsup_{m\to\infty} \left(\left| \int \varphi_{k_l} \, d\mu_{j_m^m} - \int \varphi \, d\mu_{j_m^m} \right| + \left| \int \varphi_{k_L} \, d\mu_{j_m^m} - \int \varphi \, d\mu_{j_m^m} \right| \right) \\ &\leq \sup_{m\geq 0} \mu_{j_m^m}(B(0,2M)) \left(\|\varphi - \varphi_{k_l}\| + \|\varphi - \varphi_{k_L}\| \right), \end{aligned}$$

which shows by (3.18) that $\{\alpha_{k_l}\}_{l\in\mathbb{N}}$ is a Cauchy sequence with a limit $\alpha \in \mathbb{R}$. Finally, fix $\epsilon > 0$ and pick $k_l \in \mathbb{N}$ so large that $|\alpha - \alpha_{k_l}| < \epsilon$. Then, using the triangle inequality, (3.19), and (3.18), we obtain

$$\limsup_{m\to\infty}\left|\alpha - \int \varphi\, d\mu_{j_m^m}\right| = \limsup_{m\to\infty}\left|\int \varphi_{k_l}\, d\mu_{j_m^m} - \int \varphi\, d\mu_{j_m^m}\right| + \epsilon$$
$$\leq \limsup_{m\to\infty} \mu_{j_m^m}(B(0,2M))\|\varphi_{k_m} - \varphi_k\| + \epsilon = \epsilon,$$

Letting $\epsilon \to 0$ proves that the left hand side of (3.20) exists, and $\Lambda(\varphi) = \alpha$.

The operator $\varphi \mapsto \Lambda(\varphi)$ is clearly positive and linear: $\Lambda(\varphi) \geq 0$ if $\varphi \geq 0$, and $\Lambda(a\varphi_1 + b\varphi_2) = a\Lambda(\varphi_1) + b\Lambda(\varphi_2)$. The Riesz representation theorem (see Theorem 2.14 in Rudin's book [19]) now states that the functional Λ is given by a positive Borel measure μ:

$$\Lambda(\varphi) = \int \varphi\, d\mu, \qquad \varphi \in \mathcal{C}_0(\mathbb{R}^d).$$

The weak convergence $\mu_{j_m^m} \rightharpoonup \mu$ follows immediately from (3.20). □

3.3. **Proof of Frostman's lemma.** We are now in a position to prove Frostman's lemma. We will use dyadic cubes, so here is first a quick reminder of their basic properties:

3.3.1. *Dyadic set families.* Let X be a space. A set family $\mathcal{D} \subset \mathcal{P}(X)$ is called *dyadic* if it has the following property:

(D) If $Q, Q' \in \mathcal{D}$ intersect, then either $Q' \subset Q$ or $Q \subset Q'$.

Note that any subcollection of a dyadic family of sets is still dyadic. For $X = \mathbb{R}^d$, one often talks about the *dyadic cubes*

$$\mathcal{D}_{2^{-n}} := \left\{2^{-n}[k_1, k_1+1) \times \cdots \times [k_d, k_d+1) : k_1, \ldots, k_d \in \mathbb{Z}\right\},$$

and

$$\mathcal{D} := \mathcal{D}_{\mathbb{R}^d} := \bigcup_{n\in\mathbb{Z}} \mathcal{D}_{2^{-n}}.$$

This is a basic, and very useful, example of a dyadic set family in $\mathbb{R}^d$. One of the most useful properties of dyadic set families is the following lemma:

Lemma 3.21. *Let $\mathcal{D} \subset \mathcal{P}(X)$ be a dyadic family, and let $\mathcal{D}^{\max} \subset \mathcal{D}$ be the maximal sets in $\mathcal{D}$ (w.r.t. set inclusion). Then, the family $\mathcal{D}^{\max}$ consists of disjoint sets.*

Warning: There may be no maximal sets in $\mathcal{D}$: this is the case for the family $\mathcal{D}_{\mathbb{R}^d}$. Instead, one typically applies the lemma to subfamilies $\mathcal{D} \subset \mathcal{D}_{\mathbb{R}^d}$, where the existence of maximal sets (or cubes) is clear.

Proof of Lemma 3.21. Let $Q, Q' \in \mathcal{D}^{\max}$. If $Q \cap Q' \neq \emptyset$, then either $Q \subset Q'$ or $Q' \subset Q$ by axiom (D), and then maximality forces $Q = Q'$. □

We now prove Frostman's lemma:

Proof of Lemma 3.8. Recall that $E \subset \mathbb{R}^d$ is a compact set with

$$\mathcal{H}^s_\infty(E) =: \varepsilon > 0 \tag{3.22}$$

for some $0 \leq s \leq d$. The goal is to find $\mu \in \mathcal{M}(E)$ with with $\mu(\mathbb{R}^d) \gtrsim_d \varepsilon$, and satisfying

$$\mu(B(x,r)) \leq r^s, \qquad x \in \mathbb{R}^d,\ r > 0. \tag{3.23}$$

We will first argue that $E \subset [0,1)^d$ without loss of generality. Indeed, assume that the lemma has already been proven for compact all $E \subset [0,1)^d$. Then, let $E \subset \mathbb{R}^d$ be an arbitrary compact set satisfying (3.22). In this case we may still find some (half-open) cube $Q \subset \mathbb{R}^d$ with finite side-length $M \geq 1$ such that $E \subset Q$. Let $T \colon Q \to [0,1)^d$ be a similarity map with $|T(x) - T(y)| = |x-y|/M$. Then, it is easy to check that $\mathcal{H}^s_\infty(T(E)) = \mathcal{H}^s(E)/M^s$, and we may find (by assumption) a measure $\mu_0 \in \mathcal{M}(T(E))$ satisfying (3.22) and (3.23) with $\varepsilon = \mathcal{H}^s_\infty(E)/M^s$. Finally, consider the measure $\mu := M^s \cdot T^{-1}\mu_0 \in \mathcal{M}(E)$. Then

$$\mu(E) = M^s\mu_0(E) \gtrsim_d \mathcal{H}^s_\infty(E),$$

and

$$\mu(B(x,r)) = M^s\mu_0(T(B(x,r)) \leq M^s(r/M)^s = r^s.$$

So, μ is the desired measure in $\mathcal{M}(E)$.

Now, we prove the lemma under the assumption $E \subset [0,1)^d$. Let $\delta = 2^{-n}$ for some $n \in \mathbb{N}$, and let $\mathcal{D}_\delta$ be the collection of dyadic cubes of side-length $\ell(Q) = \delta$, which are contained in $[0,1)^d$. Also, let $\mathcal{D}_\delta(E) := \{Q \in \mathcal{D}_\delta : Q \cap E \neq \emptyset\}$, and write

$$E_\delta := \bigcup_{Q \in \mathcal{D}_\delta(E)} Q \subset [0,1)^d.$$

We will first construct a measure $\mu_\delta \in \mathcal{M}(\overline{E_\delta})$, and satisfying (3.23) for all $x \in \mathbb{R}^d$, and all $\delta \leq r < \infty$. For $Q \in \mathcal{D}_\delta$, start by finding a measure $\mu^0_\delta \in \mathcal{M}(\overline{E_\delta})$ such that

$$\mu^0_\delta(Q) := \begin{cases} \ell(Q)^s, & \text{if } Q \in \mathcal{D}_\delta(E), \\ 0, & \text{if } Q \in \mathcal{D}_\delta \setminus \mathcal{D}_\delta(E). \end{cases}$$

It does not matter very much how the measure μ^0_δ behaves inside the individual cubes $Q \in \mathcal{D}_\delta$: to make the definition rigorous, let $\mu^0_\delta|_Q$ be a suitably weighted copy of Lebesgue measure on Q. It is clear that μ^0_δ satisfies (3.23) for all $x \in \mathbb{R}^d$ and $\delta \leq r < 2\delta$, say, but we have no control (yet) for $\mu^0_\delta(B(x,r))$ when $r \gg \delta$. This is why we (perhaps) need to modify μ^0_δ on scales larger than δ, and the rigorous way to do this is by induction.

Assume that μ^k_δ has already been defined for some $k \geq 0$, and consider a cube $Q \in \mathcal{D}_{2^{k+1}\delta}$. If

$$\mu^k_\delta(Q) \leq \ell(Q)^s = (2^{k+1}\delta)^s, \tag{3.24}$$

do nothing at all. In other words, set

$$\mu^{k+1}_\delta|_Q := \mu^k_\delta|_Q.$$

If, however,

$$\mu^k_\delta(Q) > \ell(Q)^s, \tag{3.25}$$

set

$$\mu^{k+1}_\delta|_Q := \frac{\ell(Q)^s}{\mu^k_\delta(Q)} \cdot \mu^k_\delta|_Q, \tag{3.26}$$

so that now $\mu^{k+1}_\delta(Q) = \ell(Q)^s$. This completes the definition of μ^{k+1}_δ. It's worth pointing out that

$$\mu^{k+1}_\delta(A) \leq \mu^k_\delta(A), \qquad A \subset \mathbb{R}^d, \ k \geq 0, \tag{3.27}$$

since $\ell(Q)^s/\mu^k_\delta(Q) < 1$ in (3.26).

Let $k_0 \geq 0$ be the index such that $2^{k_0}\delta = 1$, and set $\mu_\delta := \mu^{k_0}_\delta$. Then $\mu_\delta([0,1)^d) \leq 1$ by construction, and since $\mu_\delta(\mathbb{R}^d \setminus [0,1)^d) = 0$, we also have

$$\mu_\delta(Q) \leq \ell(Q)^s \quad \text{for all } Q \text{ dyadic with } \ell(Q) \geq 1. \tag{3.28}$$

There are now two things to check: first, that μ_δ satisfies the bound (3.23) and, second, that μ_δ still has reasonably large total mass, despite all the modifications. We start by verifying a version of (3.23) for dyadic cubes: fix a cube $Q \in \mathcal{D}_{2^k\delta}$ for some $k \geq 0$. If $k \geq k_0$, just recall (3.28). If $k < k_0$, the algorithm ensured that $\mu_\delta^k(Q) \leq \ell(Q)^s$, and then (3.27) implies that $\mu_\delta(Q) \leq \mu_\delta^k(Q) \leq \ell(Q)^s$. So, we conclude that

$$\mu_\delta(Q) \leq \ell(Q)^s$$

for all dyadic cubes of side-length $\geq \delta$. Since every ball $B(x,r)$, with $x \in \mathbb{R}^d$ and $r \geq \delta$, can be covered by $N \lesssim_d 1$ such dyadic cubes $Q_1, \ldots, Q_N$ of side-lengths $\ell(Q_j) \in [r, 2r)$, we infer that (3.23) holds for all $x \in \mathbb{R}^d$ and $r \geq \delta$.

Next, we would like to show that $\mu_\delta(E) \gtrsim_d \mathcal{H}^s_\infty(E) = \varepsilon$. To see this, note that every point $x \in E_\delta$ is contained in some dyadic cube $Q_x \in \mathcal{D}_\delta$ such that $\mu_\delta(Q_x) = \ell(Q_x)^s$. Indeed, this is the biggest cube $Q \ni x$ for which alternative (3.25) occurred, because then $\mu_\delta(Q) = \ell(Q)^s$; and if alternative (3.25) never occurred for cubes containing x, then the initial cube $Q \in \mathcal{D}_\delta$ containing x satisfies $\mu_\delta(Q) = \ell(Q)^s$. Since the cubes Q_x are all contained in $[0,1)^d$, every Q_x is certainly contained in some maximal element of $\{Q_x : x \in E_\delta\}$. Denoting these by $\mathcal{M}$, we recall from Lemma 3.21 that $\mathcal{M}$ consists of disjoint cubes. These cubes also cover $\cup\{Q_x : x \in E_\delta\} \supset E_\delta$, so

$$\mu_\delta(E_\delta) = \sum_{Q \in \mathcal{M}} \mu_\delta(Q) = \sum_{Q \in \mathcal{M}} \ell(Q)^s \gtrsim_d \sum_{Q \in \mathcal{M}} \operatorname{diam}(Q)^s \geq \mathcal{H}^s_\infty(E). \tag{3.29}$$

The remainder of the proof is an abstract application of the weak compactness result, Lemma 3.17. It is clear from (3.23) that the sequence of measures $\{\mu_{2^{-n}}\}_{n\in\mathbb{N}}$ satisfies the hypothesis (3.18). Hence, passing to a subsequence if necessary, there exists a Borel measure μ such that

$$\mu_{2^{-n}} \rightharpoonup \mu.$$

The measure $\mu_{2^{-n}}$ is supported on the $\overline{E_{2^{-n}}}$, so it follows easily from Lemma 3.15 that $\mu(\mathbb{R}^d \setminus E) = 0$. This gives $\operatorname{spt}\mu \subset E$. Another application of Lemma 3.15 shows that if $E \subset B(0,M)$, then

$$\mu(\mathbb{R}^d) \geq \mu(B(0,2M)) \geq \limsup_{n\to\infty} \mu_{2^{-n}}(B(0,2M)) \geq \limsup_{n\to\infty} \mu_{2^{-n}}(E_{2^{-n}}) \overset{(3.29)}{\gtrsim_d} \mathcal{H}^s_\infty(E).$$

This completes the proof of the lemma. □

3.4. **Hausdorff dimension of product sets.** If $A \subset \mathbb{R}^{d_1}$ and $B \subset \mathbb{R}^{d_2}$, what is the relation between the dimensions of A, B and $A \times B$? The next result gives some inequalities:

Theorem 3.30. *Assume that $A \subset \mathbb{R}^{d_1}$ and $B \subset \mathbb{R}^{d_2}$ are Borel sets, and $s,t \geq 0$. Then*

$$\mathcal{H}^{s+t}_\infty(A \times B) \gtrsim_{d_1,d_2} \mathcal{H}^s_\infty(A)\mathcal{H}^t_\infty(B). \tag{3.31}$$

In particular,

$$\dim_{\mathrm{H}}(A \times B) \geq \dim_{\mathrm{H}} A + \dim_{\mathrm{H}} B. \tag{3.32}$$

For arbitrary sets $A \subset \mathbb{R}^{d_1}$ and $B \subset \mathbb{R}^{d_2}$, with B bounded, we also have

$$\dim_{\mathrm{H}}(A \times B) \leq \dim_{\mathrm{H}} A + \overline{\dim}_{\mathrm{B}} B. \tag{3.33}$$

Proof. We start with (3.31). We may assume that $\mathcal{H}^s_\infty(A) > 0$ and $\mathcal{H}^t_\infty(B) > 0$, since otherwise there is nothing to prove. Using Frostman's lemma (for Borel sets), pick Borel measures $\mu_A \in \mathcal{M}(A)$ and $\mu_B \in \mathcal{M}(B)$ such that

$$\mu_A(\mathbb{R}^{d_1}) \gtrsim_{d_1} \mathcal{H}^s_\infty(A) \quad \text{and} \quad \mu_B(\mathbb{R}^{d_2}) \gtrsim_{d_2} \mathcal{H}^t_\infty(B),$$

and

$$\mu_A(B_{d_1}(x_1, r)) \lesssim_{d_1} r^s \quad \text{and} \quad \mu_B(B_{d_2}(x_2, r)) \lesssim_{d_2} r^t$$

for all balls $B_{d_j}(x_j, r)$ (in the correct spaces). Now, the product measure $\mu_A \times \mu_B$ is evidently supported on $A \times B$, and satisfies

$$(\mu_A \times \mu_B)(B((x_1, x_2), r)) \lesssim_{d_1,d_2} r^{s+t}, \qquad (x_1, x_2) \in \mathbb{R}^{d_1+d_2},\ r > 0,$$

because $B((x_1, x_2), r) \subset B_{d_1}(x_1, r) \times B_{d_2}(x_2, r)$. It follows from the mass distribution principle, Lemma 3.1, that

$$\mathcal{H}^{s+t}_\infty(A \times B) \gtrsim_{d_1,d_2} (\mu_A \times \mu_B)(A \times B) \gtrsim \mathcal{H}^s_\infty(A)\mathcal{H}^t_\infty(B),$$

as claimed. The inequality (3.32) for Hausdorff dimensions follows from the characterisation of Hausdorff dimension using the $\mathcal{H}^s_\infty$-measure, recall (2.9). Indeed, if $s < \dim_{\mathrm{H}} A$ and $t < \dim_{\mathrm{H}} B$, then $\mathcal{H}^s_\infty(A) > 0$ and $\mathcal{H}^t_\infty(B) > 0$, which implies using (3.31) that $\mathcal{H}^{s+t}_\infty(A \times B) > 0$, and so

$$\dim_{\mathrm{H}}(A \times B) \geq s + t.$$

We leave the upper bound (3.33) as an exercise. □

Remark 3.34. The dimensional inequality $\dim_{\mathrm{H}}(A \times B) \geq \dim_{\mathrm{H}} A + \dim_{\mathrm{H}} B$ holds for **arbitrary** sets $A \subset \mathbb{R}^{d_1}$, $B \subset \mathbb{R}^{d_2}$, as shown by Marstrand.

4. Riesz energies

Back to Question 4. We already have most of the necessary ingredients to to show that $\dim_{\mathrm{H}} \pi_e(E) = \min\{\dim_{\mathrm{H}} E, 1\}$ for almost all $e \in S^1$. Here is the general idea. Assuming that

$$\min\{\dim_{\mathrm{H}} E, 1\} > s \geq 0,$$

we note that $\mathcal{H}^s_\infty(E) > 0$, and hence Frostman's lemma gives us an s-Frostman meausre $\mu \in \mathcal{M}(E)$. We would, roughly, like to show that $\pi_e\mu$ is also an s-Frostman measure for almost all $e \in S^1$, and then conclude from the mass distribution principle that $\dim_{\mathrm{H}} \pi_e(E) \geq \dim_{\mathrm{H}} \operatorname{spt} \pi_e\mu \geq s$ for these e. A proof along these lines could be completed: the biggest unknown is, of course, the implication

$$\mu(B(x, r)) \lesssim r^s \quad \Longrightarrow \quad \pi_e\mu(B(x, r)) \lesssim r^s \text{ for "many" } e \in S^1.$$

In fact, this implication is probably not quite accurate, and should be understood as a heuristic here. To make things more rigorous, we introduce two more fundamental tools, the *Riesz potential* and the *Riesz energy*.

Definition 4.1 (Riesz potential and energy). Let $0 \leq s \leq d$, and let μ be a Borel measure on $\mathbb{R}^d$. The *s-dimensional Riesz potential of μ* is the convolution

$$V_s(\mu)(x) = (\mu * k_s)(x) = \int \frac{d\mu(y)}{|x - y|^s},$$

where k_s is the s-dimensional Riesz kernel

$$k_s(x) = \frac{1}{|x|^s}.$$

The *s-dimensional Riesz energy of μ* is the μ integral of $V_s(\mu)$:

$$I_s(\mu) = \int V_s(\mu)\, d\mu = \iint \frac{d\mu(x)d\mu(y)}{|x-y|^s}.$$

The finiteness of $I_s(\mu)$ is closely related with μ satisfying the s-Frostman decay condition. The next result clarifies the connection:

Proposition 4.2. *Let $\mu \in \mathcal{M}(\mathbb{R}^d)$ an s-Frostman measure for some $s > 0$. Then $\|V_t(\mu)\|_\infty < \infty$ for all $0 \leq t < s$, and in particular $I_t(\mu) < \infty$ (because μ is a finite measure). Conversely, if $I_s(\mu) < \infty$, then there exists a Borel set $B \subset \mathbb{R}^d$ with $\mu(B) > 0$ such that the restriction $\mu|_B$ is s-Frostman.*

Remark 4.3. Before proving the lemma, we recall a useful formula. Let μ be a Borel measure on $\mathbb{R}^d$, and let $f \colon \mathbb{R}^d \to [0,\infty]$ be a Borel function. Then

$$\int f\, d\mu = \int_0^\infty \mu(\{x \in \mathbb{R}^d : f(x) \geq \lambda\})\, d\lambda. \tag{4.4}$$

The proof is a quick application of Fubini's theorem. Starting from the right hand side, we write

$$\begin{aligned}\int_0^\infty \mu(\{x \in \mathbb{R}^d : f(x) \geq \lambda\})\, d\lambda &= \int_0^\infty \int \chi_{\{(x,\lambda) : f(x) \geq \lambda\}}(x,\lambda)\, d\mu(x)\, d\lambda \\ &= \int \int_0^\infty \chi_{\{(x,\lambda) : f(x) \geq \lambda\}}(x,\lambda)\, d\lambda\, d\mu(x) \\ &= \int f(x)\, d\mu(x).\end{aligned}$$

Proof of Lemma 4.2. Fix $0 \leq t < s$ and $x \in \mathbb{R}^d$. We will show that $V_t(\mu)(x) \lesssim_{\mu,s,t} 1$. We first apply (4.4) to the function $y \mapsto k_t(x-y) = |x-y|^{-t}$:

$$\begin{aligned}V_t(\mu)(x) &= \int_0^\infty \mu(\{y \in \mathbb{R}^d : |x-y|^{-t} \geq \lambda\})\, d\lambda \\ &= \int_0^\infty \mu(\{y \in \mathbb{R}^d : |x-y| \leq \lambda^{-1/t}\})\, d\lambda \\ &= \int_0^\infty \mu(B(x,\lambda^{-1/t}))\, d\lambda.\end{aligned}$$

For $\lambda \in (0,1]$, we simply estimate the integrand by $\mu(\mathbb{R}^d) < \infty$. For $\lambda > 1$, we use the s-Frostman assumption on μ:

$$\int_1^\infty \mu(B(x,\lambda^{-1/t}))\, d\lambda \leq C \int_1^\infty \lambda^{-s/t}\, d\lambda \lesssim_{s,t} 1,$$

recalling that $t < s$. This completes the proof of $V_t(\mu) \lesssim_{\mu,s,t} 1$, and hence $I_t(\mu) < \infty$.

Next, we assume that

$$\int V_s(\mu)\, d\mu = I_s(\mu) < \infty.$$

Using Chebyshev's inequality with $M = 2I_s(\mu)/\mu(\mathbb{R}^d) < \infty$ gives

$$\mu(\{x : V_s(\mu)(x) > M\}) \leq \frac{I_s(\mu)}{M} = \frac{\mu(\mathbb{R}^d)}{2},$$

and hence the set

$$B := \{x \in \mathbb{R}^d : V_s(\mu) \leq M\}$$

has $\mu(B) \geq \mu(\mathbb{R}^d)/2 > 0$. We leave it to the reader to check that B is a Borel set. Now, for any $x \in B$ and $r > 0$, we observe that

$$\frac{\mu(B(x,r))}{r^s} = \frac{1}{r^s} \int_{B(x,r)} d\mu(y) \leq \int_{B(x,r)} \frac{d\mu(y)}{|x-y|^s} \leq V_s(\mu)(x) \leq M,$$

so we have proven that

$$\mu|_B(B(x,r)) \leq \mu(B(x,r)) \leq Mr^s, \qquad x \in B,\ r > 0. \tag{4.5}$$

It remains to verify a similar estimate for points $x \in \mathbb{R}^d \setminus B$, so fix one of these, and a radius $r > 0$. If $B(x,r) \cap B = \emptyset$, then trivially $\mu|_B(B(x,r)) = 0 \leq Mr^s$. On the other hand, if $B(x,r) \cap B \neq \emptyset$, then $B(x,r) \subset B(x',2r)$ for some $x' \in B$, and consequently

$$\mu|_B(B(x,r)) \leq \mu(B(x',2r)) \leq (2^s M)r^s$$

by (4.5). All in all, the restriction $\mu|_B$ satisfies the decay estimate (3.23) with constant $C = 2^s M$. □

Combining the previous proposition with Frostman's lemma and the mass distribution principle gives the following important corollary:

Corollary 4.6. *Let $E \subset \mathbb{R}^d$ be a Borel set such that $\dim_{\mathrm{H}} E > s \geq 0$. Then, there exists a measure $\mu \in \mathcal{M}(E)$ with $I_s(\mu) < \infty$. Conversely, if E is an arbitrary set and there exists $\mu \in \mathcal{M}(E)$ with $I_s(\mu) < \infty$, then $\dim_{\mathrm{H}} E \geq s$.*

Proof. Pick $\dim_{\mathrm{H}} E > t > s$, so that $\mathcal{H}^t_\infty(E) > 0$. By Frostman's lemma, there exists a t-Frostman measure $\mu \in \mathcal{M}(E)$. It follows from the previous proposition that $I_s(\mu) < \infty$, as claimed.

Conversely, if $\mu \in \mathcal{M}(E)$ satisfies $I_s(\mu) < \infty$, then the previous proposition implies that $\mu|_B$ is an s-Frostman measure for some Borel set $B \subset \mathbb{R}^d$ with $\mu(B) > 0$. Since $\mu|_B(E) > 0$, the mass distribution principle gives $\mathcal{H}^s_\infty(E) > 0$, and hence $\dim_{\mathrm{H}} E \geq s$. □

4.1. **Marstrand's projection theorem: part I.**

The following theorem answers Question 4 quite well. It is due to Marstrand [14] from 1954:

Theorem 4.7 (Marstrand)**.** *Let $E \subset \mathbb{R}^2$ be a Borel set. Then*

$$\dim_{\mathrm{H}} \pi_e(E) = \min\{\dim_{\mathrm{H}} E, 1\}$$

for $\mathcal{H}^1$ almost every $e \in S^1$. Moreover, if $\dim_{\mathrm{H}} E > 1$, then $\mathcal{H}^1(\pi_e(E)) > 0$ for $\mathcal{H}^1$ almost every $e \in S^1$.

Notation 4.8. We will, from now on, write

$$\pi_e(x) = x \cdot e, \qquad e \in S^1.$$

In other words, we interpret the orthogonal projection $\mathbb{R}^2 \to \ell_e$ as a map $\mathbb{R}^2 \to \mathbb{R}$. This will be technically convenient in the proofs, but of course quantities like $\dim_{\mathrm{H}} \pi_e(E)$ and $\mathcal{H}^1(\pi_e(E))$ are unaffected, because $t \mapsto te$ is an isometry $\mathbb{R} \to \ell_e$.

The second statement of Theorem 4.7 will require some further tools, which we develop in Section 6, so we only prove the first part of the theorem in this section. The upper bound $\dim_{\mathrm{H}} \pi_e(E) \leq \min\{\dim_{\mathrm{H}} E, 1\}$ is true for **every** $e \in S^1$: this follows from Exercise 2.7(v) which in particular implies that Lipschitz maps (such as π_e) do not increase Hausdorff dimension.

Proof of the first part of Marstrand's projection theorem. It suffices to prove the following statement: if

$$\min\{\dim_{\mathrm{H}} E, 1\} > s > 0,$$

then $\dim_{\mathrm{H}} \pi_e(E) \geq s$ for $\mathcal{H}^1$ almost every $e \in S^1$. By considering a sequence of values $s_i \nearrow \min\{\dim_{\mathrm{H}} E, 1\}$, this will imply that $\dim_{\mathrm{H}} \pi_e(E) \geq \min\{\dim_{\mathrm{H}} E, 1\}$ for $\mathcal{H}^1$ almost every $e \in S^1$.

So, fix $0 < s < \min\{\dim_{\mathrm{H}} E, 1\}$, and use the first part of Corollary 4.6 to find a measure $\mu \in \mathcal{M}(E)$ with $I_s(\mu) < \infty$. The aim will be to show that

$$\int_{S^1} I_s(\pi_e \mu)\, d\mathcal{H}^1(e) < \infty. \tag{4.9}$$

This will, in particular, show that $I_s(\pi_e\mu) < \infty$ for $\mathcal{H}^1$ almost every $e \in S^1$. Since $\pi_e\mu \in \mathcal{M}(\pi_e(E))$ for all $e \in S^1$, the second part of Corollary 4.6 will then imply that $\dim_{\mathrm{H}} \pi_e(E) \geq s$ for $\mathcal{H}^1$ almost every $e \in S^1$, completing the proof.

We start the proof of (4.9) by spelling out the definition:

$$\int I_s(\pi_e\mu)\, d\mathcal{H}^1(e) = \int_{S^1} \iint \frac{d(\pi_e\mu)(x) d(\pi_e\mu)(y)}{|x-y|^s}\, d\mathcal{H}^1(e).$$

To proceed, we recall from (3.7) the general formula

$$\int_Y g df\mu = \int_X (g \circ f)\, d\mu,$$

valid whenever f is a continuous, and g is a non-negative Borel function. We apply this twice with $f = \pi_e$:

$$\iint \frac{d(\pi_e\mu)(x) d(\pi_e\mu)(y)}{|x-y|^s}\, d\mathcal{H}^1(e) = \iint \frac{d\mu(x) d\mu(y)}{|\pi_e(x) - \pi_e(y)|^s}. \tag{4.10}$$

Next, since π_e is linear, we may write

$$\frac{1}{|\pi_e(x) - \pi_e(y)|^s} = \frac{1}{|x-y|^s} \frac{1}{\left|\pi_e\left(\frac{x-y}{|x-y|}\right)\right|^s}, \qquad x \neq y. \tag{4.11}$$

Hence, collecting the previous computations, and using Fubini's theorem, we find that

$$\begin{aligned} \int I_s(\pi_e\mu)\, d\mathcal{H}^1(e) &= \int_{S^1} \left(\iint \frac{d\mu(x) d\mu(y)}{|\pi_e(x) - \pi_e(y)|^s} \right) d\mathcal{H}^1(e) \\ &= \iint \frac{1}{|x-y|^s} \left(\int_{S^1} \frac{\mathcal{H}^1(e)}{\left|\pi_e\left(\frac{x-y}{|x-y|}\right)\right|^s} \right) d\mu(x)\, d\mu(y). \end{aligned} \tag{4.12}$$

We have cheated here a little, because (4.11) only holds when $x \neq y$. This is not a problem, however, because clearly

$$I_s(\mu) < \infty \quad \Longrightarrow \quad (\mu \times \mu)\{(x,y) : x = y\} = 0,$$

and hence the set $\{(x,y) : x = y\}$ can be omitted from the right hand side of (4.10).

Now, examining the right hand side of (4.12), we write $(x-y)/|x-y| =: z \in S^1$. If we manage to show that

$$\int_{S^1} \frac{\mathcal{H}^1(e)}{|\pi_e(z)|^s} \lesssim_s 1, \qquad z \in S^1, \tag{4.13}$$

we can simply plug in this estimate to (4.12), and conclude the whole proof as follows:

$$\int I_s(\pi_e \mu)\, d\mathcal{H}^1(e) \lesssim_s \iint \frac{d\mu(x)\, d\mu(y)}{|x-y|^s} = I_s(\mu) < \infty.$$

So, we turn to the proof of (4.13): note that, curiously, this is the only part of the proof where some geometry appears! We start by fixing $z \in S^1$, and applying the useful formula (4.4) to write

$$\begin{aligned}\int_{S^1} \frac{\mathcal{H}^1(e)}{|\pi_e(z)|^s} &= \int_0^\infty \mathcal{H}^1(\{e \in S^1 : |\pi_e(z)|^s \geq \lambda\})\, d\lambda \\ &= \int_0^\infty \mathcal{H}^1(\{e \in S^1 : |\pi_e(z)| \leq \lambda^{-1/s}\})\, d\lambda \\ &\leq \mathcal{H}^1(S^1) + \int_1^\infty \mathcal{H}^1(\{e \in S^1 : |\pi_e(z)| \leq \lambda^{-1/s}\})\, d\lambda\end{aligned}$$

Given $r \in [0,1]$, what does the set

$$B_z(r) := \{e \in S^1 : |\pi_e(z)| \leq r\}$$

look like? Clearly $B_z(0) = S^1 \cap z^\perp$, and we claim that $B_z(r)$ consists of two arcs of length $\sim r$, centred at the two antipodal points in $S^1 \cap z^\perp$. Let's consider the special case $z = (1,0)$, and note that $\pi_e(z) = e_1$ for $e = (e_1, e_2) \in S^1$. So, $|\pi_e(z)| \leq r$ implies $|e_1| \leq r$, which forces e to lie in a $\sim r$ neighbourhood of $\{(0,1), (0,-1)\}$. The general case can be reduced to this by applying a rotation.

So, we have argued that $|B_z(r)| \lesssim r$ for $r \in [0,1]$, and hence

$$\int_1^\infty \mathcal{H}^1(\{e \in S^1 : |\pi_e(z)| \leq \lambda^{-1/s}\})\, d\lambda \lesssim \int_1^\infty \lambda^{-1/s}\, d\lambda \lesssim_s 1,$$

recalling that $s < 1$ in the last estimate. This completes the proof of (4.13), and the (first half of) Marstrand's projection theorem. □

Exercise 4.14. This exercise asks you to prove a sharper version of the previous result, which is due to R. Kaufman [13]. Namely, prove that if $E \subset \mathbb{R}^2$ is a Borel set with $\dim_{\mathrm{H}} E =: s \in [0,1]$, then

$$\dim_{\mathrm{H}}\{e \in S^1 : \dim_{\mathrm{H}} \pi_e(E) < s\} \leq s.$$

Hint: Pick $t < s$ and a measure $\mu \in \mathcal{M}(E)$ with $I_t(\mu) < \infty$. Make a counter assumption that $\dim_{\mathrm{H}}\{e \in S^1 : \dim_{\mathrm{H}} \pi_e(E) < t\} > t$. Why is this actually a counter assumption? Find an t-Frostman measure $\sigma \in \mathcal{M}(\{e \in S^1 : \dim_{\mathrm{H}} \pi_e(E) < t\})$. You may take for granted that $\{e \in S^1 : \dim_{\mathrm{H}} \pi_e(E) < t\}$ is Borel. Then, integrate $I_t(\pi_e \mu)$ with respect to σ and see what happens. Why do you reach a contradiction?

5. Rectifiable sets and the Besicovitch projection theorem

In this section, we study the projection properties of compact sets $E \subset \mathbb{R}^2$ with $0 < \mathcal{H}^1(E) < \infty$. We call such sets simply 1*-sets* in this section. By Marstrand's projection theorem, we already know that $\dim_{\mathrm{H}} \pi_e(E) = 1$ for $\mathcal{H}^1$ almost every $e \in S^1$, but the following is open:

Question 7. *If $E \subset \mathbb{R}^2$ is a compact 1-set, then is it true that $\mathcal{H}^1(\pi_e(E)) > 0$ for $\mathcal{H}^1$ almost every $e \in S^1$?*

Note that even the second half of Marstrand's projection theorem does not answer the question, as it assumes $\dim_{\mathrm{H}} E > 1$. Answering Question 7 leads to a deep dichotomy for planar 1-sets, discovered by Besicovitch in the early 20th century. It turns out that there are essentially two kinds of planar 1-sets: *rectifiable* and *unrectifiable* ones, and the answer to Question 7 is opposite for the two kinds of sets.

5.1. A density lemma.

Before proceeding to rectifiable and unrectifiable sets, we record the following useful lemma:

Lemma 5.1. *Assume that $s \geq 0$, and $A \subset \mathbb{R}^d$ with $\mathcal{H}^s(A) < \infty$. Write*

$$\Theta^{s,*}(A,x) := \limsup_{r \to 0} \frac{\mathcal{H}^s(A \cap B(x,r))}{r^s} \quad \text{and} \quad \Theta^{s}_{*}(A,x) := \liminf_{r \to 0} \frac{\mathcal{H}^s(A \cap B(x,r))}{r^s}.$$

These quantities are called the upper and lower s-dimensional densities of A at x. *Then*

(1) $1 \leq \Theta^{s,*}(A,x) \leq 2^s$ *for $\mathcal{H}^s$ almost all $x \in A$.*
(2) *If A is $\mathcal{H}^s$ measurable, then $\Theta^{s,*}(A,x) = 0$ for $\mathcal{H}^s$ almost all $x \in \mathbb{R}^d \setminus A$.*

Proof. The only statement we will need is the lower bound in (1), so we omit the other proofs, see Theorem 6.1 in [16] for the remaining arguments. For the lower bound, we don't even need the assumption $\mathcal{H}^s(A) < \infty$.

To prove that $\Theta^{s,*}(A,x) \geq 1$ for $\mathcal{H}^s$ almost every $x \in A$, we consider the sets

$$A_{R,\delta,\tau} := \{x \in A \cap B(0,R) : \mathcal{H}^s(A \cap B(x,r)) \leq \tau r^s \text{ for all } 0 < r < \delta\}, \qquad \delta, \tau > 0.$$

It suffices to show that $\mathcal{H}^s(A_{R,\delta,\tau}) = 0$ for all $R, \delta > 0$ and $0 < \tau < 1$, because

$$\{x \in A : \Theta^{s,*}(A,x) < 1\} \subset \bigcup_{R>0} \bigcup_{0<\tau<1} \bigcup_{\delta>0} A_{R,\delta,\tau}.$$

Let $U_1, U_2, \ldots$ be a δ-cover for $A_{R,\delta,\tau}$ with the property that $U_k \cap A_{R,\delta,\tau} \neq \emptyset$ for all $k \in \mathbb{N}$. Then, picking $x_k \in U_k \cap A_{R,\delta,\tau}$, we have

$$U_k \subset B(x_k, \mathrm{diam}(U_k)) =: B_k.$$

It follows (using the subadditivity $\mathcal{H}^s_\delta$ and the inequality $\mathcal{H}^s_\delta \leq \mathcal{H}^s$) that

$$\mathcal{H}^s_\delta(A_{R,\delta,\tau}) \leq \sum_{k \in \mathbb{N}} \mathcal{H}^s(A_{R,\delta,\tau} \cap B_k) \leq \tau \sum_{k \in \mathbb{N}} \mathrm{diam}(U_k)^s.$$

Taking an inf over all δ-covers of $A_{R,\delta,\tau}$ gives $\mathcal{H}^s_\delta(A_{R,\delta,\tau}) \leq \tau \mathcal{H}^s_\delta(A_{R,\delta,\tau})$. Because

$$\mathcal{H}^s_\delta(A_{R,\delta,\tau}) \lesssim \left(\frac{R}{\delta}\right)^d < \infty$$

and $\tau < 1$, this implies $\mathcal{H}^s_\delta(A_{R,\delta,\tau}) = 0$, and hence $\mathcal{H}^s(A_{R,\delta,\tau}) = 0$ by Exercise 2.7(v). □

5.2. **Rectifiable and unrectifiable sets.** We will only consider the projections of rectifiable and unrectifiable sets in $\mathbb{R}^2$, but at least we can give the definitions in general dimensions.

Definition 5.2. Let $0 < n < d$ be integers, and let $E \subset \mathbb{R}^d$. We say that E is *n-rectifiable*, if $\mathcal{H}^n$ almost all of E can be covered by Lipschitz images of $\mathbb{R}^n$. In other words, there exist Lipschitz maps $f_j \colon \mathbb{R}^n \to \mathbb{R}^d$, $j \in \mathbb{N}$, such that

$$\mathcal{H}^n\left(E \setminus \bigcup_{j\in\mathbb{N}} f_j(\mathbb{R}^n)\right) = 0.$$

The set E is called *purely n-unrectifiable*, if

$$\mathcal{H}^n(E \cap f(\mathbb{R}^d)) = 0$$

for all Lipschitz $f \colon \mathbb{R}^n \to \mathbb{R}^d$. Equivalently, E is purely n-unrectifiable if and only if $\mathcal{H}^n(E \cap R) = 0$ for all n-rectifiable sets $R \subset \mathbb{R}^d$.

The definition and all the results in this section are due to Besicovitch. The following decomposition result is easy but quite fundamental:

Theorem 5.3. *Let $0 < n < d$, and let $E \subset \mathbb{R}^d$ with $\mathcal{H}^n(E) < \infty$. Then there exists a Borel set $R \subset \mathbb{R}^d$ such that $U = E \setminus R$ is purely n-unrectifiable. It follows that E can be written as a disjoint union*

$$E = [E \cap R] \cup U,$$

where $E \cap R$ is n-rectifiable, and U is purely n-unrectifiable. The decomposition is unique up to null sets.

Remark 5.4. Note that if E was assumed Borel (or $\mathcal{H}^n$ measurable), the fact that $R \subset \mathbb{R}^d$ is Borel implies that also the sets $E \cap R$ and U in the decomposition are Borel (or $\mathcal{H}^n$ measurable).

Proof of Theorem 5.3. Write

$$M := \sup\{\mathcal{H}^n(E \cap R) : R \text{ is Borel and } n\text{-rectifiable}\} < \infty,$$

and choose a sequence of numbers $0 < M_j < M$ with $M_j \nearrow M$. For each M_j, choose a Borel n-rectifiable set R_j with $\mathcal{H}^n(E \cap R_j) \geq M_j$. Writing

$$R := \bigcup_{j\in\mathbb{N}} R_j,$$

we see that $E \cap R$ is Borel, n-rectifiable, and $\mathcal{H}^n(E \cap R) = M$. It remains to verify that U is purely n-unrectifiable. To this end, fix $f \colon \mathbb{R}^n \to \mathbb{R}^d$ Lipschitz, and note that $R^+ := R \cup f(\mathbb{R}^n)$ is Borel (even σ-compact). So, using the $\mathcal{H}^n$ measurability of R, we may write

$$\mathcal{H}^n(E \cap R^+) = \mathcal{H}^n([E \cap R^+] \cap R) + \mathcal{H}^n([E \cap R^+] \setminus R) = \mathcal{H}^n(E \cap R) + \mathcal{H}^n(U \cap f(\mathbb{R}^n)).$$

Noting that $\mathcal{H}^n(E \cap R^+) = M = \mathcal{H}^n(E \cap R)$, we infer that $\mathcal{H}^n(U \cap f(\mathbb{R}^n)) = 0$. The proof is now complete, except for the uniqueness of the decomposition. If $E = R' \cup U'$ is another decomposition, where R' is n-rectifiable and U' is purely n-unrectifiable, then

$$\mathcal{H}^n(R \cap U') = 0 = \mathcal{H}^n(R' \cap U).$$

This is exactly what was claimed. □

5.3. **Projections of rectifiable sets.** By the decomposition result in Theorem 5.3, we can split Question 7 into two sub-questions: what are the projections of E like if E is (a) 1-rectifiable, (b) purely 1-unrectifiable? Question (a) is a lot easier, and we will sketch an answer in this section. Question (b) will be answered in the next sections.

Proposition 5.5. *Assume that $E \subset \mathbb{R}^2$ is 1-rectifiable with $0 < \mathcal{H}^1(E) < \infty$. If $e_1, e_2 \in S^1$ are distinct vectors with $\mathcal{H}^1(\pi_{e_1}(E)) = 0 = \mathcal{H}^1(\pi_{e_2}(E))$, then $e_1 = -e_2$.*

Remark 5.6. The example of a line segment shows that even a rectifiable set can have null projection in one direction. As a second remark, note the following corollary: if $0 < \mathcal{H}^1(E) < \infty$, and E has two projections of zero length on non-parallel lines, then E is purely 1-unrectifiable.

Proof of Proposition 5.5. Since E is 1-rectifiable, there exist a bounded open interval $J \subset \mathbb{R}$ and an L-Lipschitz map $f\colon J \to \mathbb{R}^2$, for some $L > 0$, such that

$$\mathcal{H}^1(E \cap f(J)) = \epsilon > 0.$$

We employ an approximation result of Lusin, which states that for any $\delta > 0$, there exists a $\mathcal{C}^1$-map $g\colon J \to \mathbb{R}^2$ such that $\mathcal{H}^1(B) < \delta$, where

$$B = \{t \in J : f(t) \neq g(t)\}.$$

Recalling that f is L-Lipschitz, we find that $\mathcal{H}^1(f(B)) < L\delta$, and choosing $\delta < \epsilon/L$ we get

$$\mathcal{H}^1(E \cap g(J)) \geq \mathcal{H}^1(E \cap f(J \setminus B)) \geq \mathcal{H}^1(E \cap f(J)) - \mathcal{H}^1(f(B)) > 0.$$

We would now like to find a point $t_0 \in I$ such that

$$g'(t_0) \neq 0 \quad \text{and} \quad \liminf_{r \to 0} \frac{\mathcal{H}^1(E \cap g(B(t_0, r)))}{r} > 0. \tag{5.7}$$

Let $F_1, F_2 \subset J$ be the set of points t_0 failing the first and second parts of (5.7), respectively. It suffices to show that $\mathcal{H}^1(g(F_1)) = 0 = \mathcal{H}^1(E \cap g(F_2))$. We leave it as an exercise to check that $\mathcal{H}^1(g(F_1)) = 0$, and we now argue that $\mathcal{H}^1(E \cap g(F_2)) = 0$.

Fix $\delta > 0$, and for every $t \in F_2$, find $r > 0$ such that $B(t, r) \subset J$ and $\mathcal{H}^1(E \cap g(B(t, r))) < \delta r$. Then, use the $5r$-covering theorem to find a countable sub-family of intervals $\{B(t_j, r_j)\}_{j \in \mathbb{N}}$ such that the intervals $B(t_j, r_j/5)$ are disjoint and

$$F_2 \subset \bigcup_{j \in \mathbb{N}} B(t_j, r_j).$$

It follows that

$$\mathcal{H}^1(E \cap g(F_2)) \leq \sum_{j \in \mathbb{N}} \mathcal{H}^1(E \cap g(B(t_j, r_j))) < 5\delta \sum_{j \in \mathbb{N}} \frac{r_j}{5} \leq 5\delta \mathcal{H}^1(J),$$

and hence $\mathcal{H}^1(E \cap g(F_2)) = 0$.

Now, fix $t_0 \in J \setminus [F_1 \cup F_2]$ as in (5.7), and let $e_0 := g'(t_0)/|g'(t_0)| \in S^1$. Since $g \in \mathcal{C}^1$, the set $g(B(t_0, r))$ starts to resemble a segment on $g(t_0) + \operatorname{span}(e_0)$ as $r \to 0$. More precisely, if

$$e \in S^1 \setminus e_0^{\perp},$$

then there exists $r > 0$ such that the projection π_e restricted to $g(B(t_0, r))$ is biLipschitz (of course with constants getting worse as e approaches $e_0^\perp$). Since $t_0 \notin F_2$, it follows that for such $r > 0$

$$\mathcal{H}^1(\pi_e(E)) \geq \mathcal{H}^1(\pi_e[E \cap g(B(t_0, r))]) \gtrsim \mathcal{H}^1(E \cap g(B(t_0, r))) \gtrsim r > 0.$$

This proves that $\{e \in S^1 : \mathcal{H}^1(\pi_e(E)) = 0\} \subset e_0^\perp$. □

5.4. **Conical densities of purely 1-unrectifiable sets.** Now, we start studying purely 1-unrectifiable sets. Recall from Lemma 5.1 that if $s \geq 0$ is arbitrary, and $A \subset \mathbb{R}^d$ with $0 < \mathcal{H}^s(A) < \infty$, then the s-dimensional upper density

$$\Theta^{s,*}(A, x) = \limsup_{r \to 0} \frac{\mathcal{H}^s(A \cap B(x, r))}{r^s}$$

is positive for many points $x \in A$. Could we replace $B(x, r)$ by some smaller set and still have the same conclusion? To avoid geometric complications, we restrict attention to the plane; for analogous results in higher dimensions, see Section 15 in [16]. For $S \subset S^1$ (typically an arc), consider the "cone"

$$C(x, S) := \bigcup_{e \in S} \ell_e(x),$$

where $\ell_e(x)$ is the line parallel to e which contains x. Now, if $J \subset S^1$ is a fixed arc, we can ask if

$$\Theta_J^{s,*}(A, x) := \limsup_{r \to 0} \frac{\mathcal{H}^s(A \cap B(x, r) \cap C(x, J))}{r^s} > 0 \tag{5.8}$$

for some – or most – points $x \in A$. By Lemma 5.1, this is evidently true if $J = S^1$. In general, however, (5.8) can fail rather badly: just consider $A = \mathrm{span}(1, 0) \subset \mathbb{R}^2$, and let

$$J := S^1 \setminus A = \{(e_1, e_2) \in S^1 : e_2 \neq 0\}.$$

Then $A \cap B(x, r) \cap C(x, J) = \{x\}$ for all $x \in A$, so (5.8) fails for any $x \in A$. With this example in mind, it's quite remarkable that (5.8) holds for all arcs $J \subset S^1$, if A is purely 1-unrectifiable:

Proposition 5.9. *There is an absolute constant $c > 0$ such that the following holds. Let $E \subset \mathbb{R}^2$ be purely 1-unrectifiable with $0 < \mathcal{H}^1(E) < \infty$. Then, for $\mathcal{H}^1$ almost all $x \in E$,*

$$\Theta_J^{1,*}(E, x) \geq c\mathcal{H}^1(J)$$

for all arcs $J \subset S^1$.

We begin the proof with a simple test of rectifiability:

Lemma 5.10. *Let $E \subset \mathbb{R}^2$ be a set, and let $J \subset S^1$ be a non-degenerate arc. If*

$$E \cap C(x, J) = \{x\}, \qquad x \in E,$$

then E is 1-rectifiable.

Proof. Write $J = B(e, \rho) \cap S^1$, with $\rho > 0$, and let $\xi \in S^1 \cap e^\perp$. Assume without loss of generality that $e = (0, 1)$ and $\xi = (1, 0)$. We claim that the orthogonal projection $\pi_\xi \colon E \to \ell_\xi$ is $\sim \rho$-biLipschitz on E. It will follow that $\pi_\xi^{-1} \colon \pi_\xi(E) \to E$ is well-defined and $\sim (1/\rho)$-Lipschitz, and hence $E = \pi^{-1}(\pi_\xi(E))$ is 1-rectifiable.

Figure 2 explains the whole proof. To show that π_ξ is biLipschitz on E, we need to fix

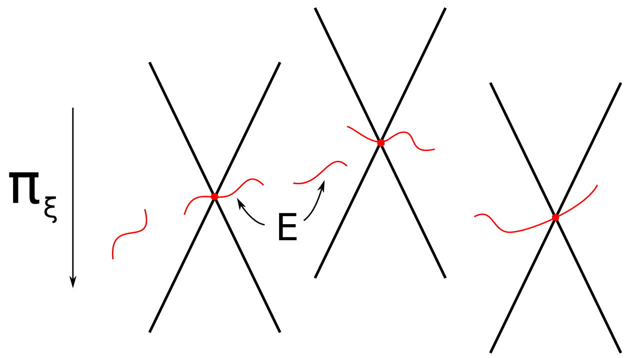

FIGURE 2. The set E mostly avoids the cones $C(c, J)$ with $x \in E$.

$x, y \in E$ with $x \neq y$ and show that $|\pi_\xi(x) - \pi_\xi(y)| \gtrsim \rho|x - y|$, which is equivalent to

$$\left|\pi_\xi\left(\frac{x-y}{|x-y|}\right)\right| \gtrsim \rho.$$

Note that $e' := (x - y)/|x - y| \in S^1 \setminus [J \cup (-J)]$ by assumption, and it follows easily that the first coordinate of e' is $|e'_1| \gtrsim \rho$. Hence

$$|\pi_\xi(e')| = |e'_1| \gtrsim \rho,$$

as claimed. □

We are now prepared to prove Proposition 5.9:

Proof of Proposition 5.9. Let $E \subset \mathbb{R}^2$ be purely 1-unrectifiable, and recall that $\mu = \mathcal{H}^1|_E$. We also write $\sigma := \mathcal{H}^1|_{S^1}$. It suffices to show that if $J \subset S^1$ is a **fixed** arc, then

$$\Theta_J^{1,*}(E, x) \geq c\sigma(J) \tag{5.11}$$

for μ almost every $x \in E$, with constants independent of J. This reduction follows by choosing a countable family of arcs $\{J_i\}_{\mathbb{N}}$ with the property that for any arc $J \subset S^1$ there exists $J_i \subset J$ with $\sigma(J_i) \geq \sigma(J)/2$. Then (5.11) holds μ a.e. for all the arcs J_i simultaneously, and it follows that (5.11) holds μ a.e. for **all** arcs $J \subset S^1$ simultaneously.

So, fix an arc $J \subset S^1$. Without loss of generality, we assume that J is a "north cap"

$$J = B((0, 1), \rho) \cap S^1$$

for some $\rho > 0$. Set

$$B_{\delta,\tau} := \{x \in E : \mu(B(x, r) \cap C(x, J)) < \tau r \text{ for all } 0 < r < \delta\}.$$

Since

$$\{x \in E : \Theta_J^{1,*}(E, x) < \tau\} \subset \bigcup_{\delta > 0} B_{\delta,\tau},$$

it suffices to show that $\mu(B_{\delta,\tau}) = 0$ for any $\delta > 0$, and

$$\tau := \frac{\sigma(J)}{C},$$

where $C \geq 1$ is a suitable absolute constant. Fix $\delta > 0$. Instead of showing directly that $\mu(B_{\delta,\tau}) = 0$, we prove that B has low density everywhere in the following sense:

$$\mu(B_{\delta,\tau} \cap B(x_0, r_0/4)) \lesssim \frac{r_0}{C}, \qquad x_0 \in B_{\delta,\tau}, \ 0 < r_0 \leq \delta. \tag{5.12}$$

For $C \geq 1$ sufficiently large (depending on the implicit constants in (5.12)), this gives $\Theta^{1,*}(B_{\delta,\tau}, x_0) < 1$ for all $x_0 \in B_{\delta,\tau}$, and then finally $\mu(B_{\delta,\tau}) = 0$ by Lemma 5.1(1).

To prove (5.12), fix $x_0 \in B_{\delta,\tau}$ and $0 < r_0 \leq \delta$. Let

$$B \subset B_{\delta,\tau} \cap B(x_0, r_0/4)$$

be an arbitrary compact set: if we manage to prove that $\mu(B) \lesssim r_0/C$, then Lemma 2.19 gives (5.12). All we know is that

$$\mu(B(x,r) \cap C(x,J)) < \frac{\sigma(J)r}{C}, \qquad 0 < r \leq r_0, \ x \in B \subset B_{\delta,\tau}. \tag{5.13}$$

The information (5.13) is a little awkward to use: we would be in much better shape if we knew that the density of μ in **vertical tubes** is small, instead of similar information about vertical cones (recall that $J = B((0,1), \rho) \cap S^1$, so $C(x,J)$ is indeed a cone around the vertical line $\ell_{(0,1)}(x)$). In fact, we will prove the following claim:

Claim 5.14. *For μ almost every $x \in B$, there exists a vertical tube $T_x \ni x$ around the line $\ell_{(0,1)}(x)$ of width $0 < w(T_x) \leq r_0$ such that*

$$\mu(B \cap T_x) \lesssim \frac{w(T_x)}{C}. \tag{5.15}$$

Let's see how to complete the proof with this claim in hand. Let $G \subset B$ be the set of points x such that T_x exists. For every $x \in G$, we pick T_x as in (5.15), and use the $5r$-covering theorem to find a disjoint countable subcollection $T_1, T_2, \ldots$ such that the tubes $T_j/5$ are disjoint, and

$$G \subset \bigcup_{j \in \mathbb{N}} T_j.$$

Why can we apply the $5r$-covering theorem here? Note that the projection of $T_j/5$ to the x-axis is an interval of length $w(T_j)/5$, and apply the $5r$-covering theorem on the x-axis to these intervals. Now, we have

$$\mu(B) = \mu(G) \leq \sum_{j \in \mathbb{N}} \mu(B \cap T_j) \lesssim \frac{1}{C} \sum_{j \in \mathbb{N}} w(T_j) = \frac{5}{C} \sum_{j \in \mathbb{N}} w(T_j/5) \lesssim \frac{r_0}{C},$$

where the last inequality follows from the disjointness of the tubes $T_j/5$, and the fact that they all intersect $B \subset B(x_0, r_0)$. This concludes the proof of (5.12).

It remains to verify Claim 5.14. The main challenge is that information of the form "cones have low μ mass", as in (5.13), must be converted into information of the form "tubes have low μ mass". Note that this is certainly not true for rectifiable sets!

A basic observation, depicted in Figure 3, is the following. For $x \in B$ and $0 < r < r_0$, take a vertical tube T of width $w(T) = \sigma(J)r$ around x. Then, by elementary geometry,

$$T \setminus B(x, 2r) \subset C(x,J) \tag{5.16}$$

(this corresponds to the the green part of the tube in Figure 3). So, we might have some estimates for $\mu(T \setminus B(x,2r))$, but an additional idea is needed to handle the part of T inside $B(x, 2r)$.

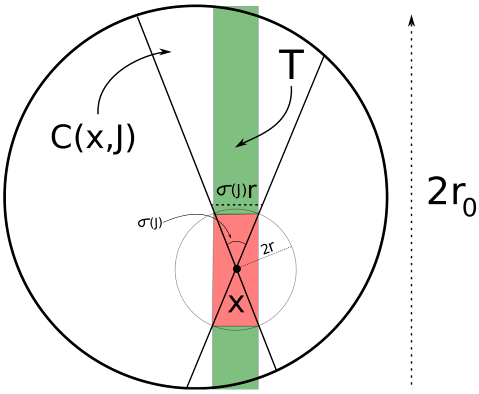

FIGURE 3. The green part of the tube is under control, but who knows about the red part.

The clever solution of Besicovitch is to use two cones instead of one, see Figure 4. Start

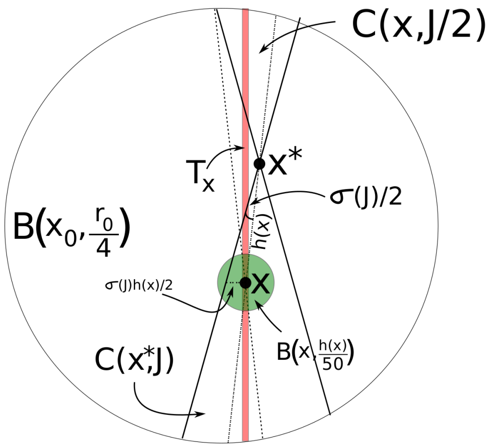

FIGURE 4. Covering a tube around x with two cones.

by recalling that B is compact, so also $B \cap C(x, J/2)$ is compact, where $J/2$ stands for the closed arc which is concentric with J and has $\sigma(J/2) = \sigma(J)/2$. Consequently, there exists a point $x^* \in B \cap C(x, J/2)$ with

$$|x^* - x| = h(x) := \max\{|y - x| : y \in B \cap C(x, J/2)\} \leq \operatorname{diam}(B) \leq r_0/2.$$

Assume for a moment that $h(x) > 0$, so that $B(x, h(x))$ is a non-degenerate disc with the property

$$B \cap C(x, J/2) \subset B(x, h(x)). \tag{5.17}$$

We discuss the assumption $h(x) > 0$ in a moment. Let T_x be a tube around the line $\ell_{(0,1)}(x)$ of width $w(T_x) = h(x)\sigma(J)/100$. We now make two remarks about the geometry of the situation.

(i) First, we certainly have $T_x \setminus B(x, h(x)) \subset C(x, J/2)$ by (5.16) (this inclusion isn't particularly sharp, as the tube T_x is very narrow compared to $h(x)\sigma(J/2)$). So,

$$\begin{aligned} B \cap T_x &\subset B(x, h(x)) \cup [(B \cap T_x) \setminus B(x, h(x))] \\ &= B(x, h(x)) \cup [B \cap (T_x \setminus B(x, h(x)))] \\ &\subset B(x, h(x)) \cup [B \cap C(x, J/2)] \overset{(5.17)}{\subset} B(x, h(x)). \end{aligned}$$

(ii) Second, we have $T_x \subset C(x, J) \cup C(x^*, J)$. The proof consists of two observations: first, that $T_x \setminus B(x, h(x)/50) \subset C(x, J)$ by (5.13), so it suffices to argue that

$$T_x \cap B(x, h(x)/50) \subset C(x^*, J).$$

But the ball $B(x, h(x)/50)$ lies at distance $\geq 49h(x)$ from x^*, and then one needs a little elementary geometry (see Figure 4), to conclude that $T_x \cap B(x, h(x)/50)$ is well inside $C(x^*, J)$. *Note*: we are **not** claiming that $C(x, h(x)/50)$ would be contained in $B(x^*, J)$!

Combining (i) and (ii), and noting that $B(x, h(x)) \subset B(x^*, 2h(x))$, we arrive at the inclusion

$$B \cap T_x \subset [C(x, J) \cap B(x, h(x))] \cup [C(x^*, J) \cap B(x^*, 2h(x))],$$

so consequently

$$\begin{aligned} \mu(B \cap T_x) &\leq \mu([C(x, J) \cap B(x, h(x))] \cup [C(x^*, J) \cap B(x^*, 2h(x))]) \\ &\overset{(5.13)}{\lesssim} \frac{h(x)\sigma(J)}{C} \sim \frac{w(T_x)}{C}, \end{aligned}$$

This proves Claim 5.14 under the assumption that $h(x) > 0$.

Finally, we use the 1-unrectifiability of E to guarantee that $h(x) > 0$ for μ almost every $x \in B$. Indeed, consider the set

$$R = \{x \in B : h(x) = 0\}.$$

Recalling the definition of $h(x)$, we have $x \in R$ if and only if $B \cap C(x, J/2) = \{x\}$. In particular, $R \cap C(x, J/2) = \{x\}$ for all $x \in R$, which implies by Lemma 5.10 that R is rectifiable. But $R \subset B \subset E$, and E is purely 1-unrectifiable, so $\mu(R) = 0$. This completes the proof of Claim 5.14, and of the proposition. □

5.5. **Projections of unrectifiable sets.** Now we will answer Question 7 negatively for purely 1-unrectifiable sets. The following theorem is due to Besicovitch [2] from 1939:

Theorem 5.18 (Besicovitch)**.** *Assume that $E \subset \mathbb{R}^2$ is Borel purely 1-unrectifiable 1-set. Then $\mathcal{H}^1(\pi_e(E)) = 0$ for $\mathcal{H}^1$ almost every $e \in S^1$.*

Since Borel 1-sets E can be approximated in $\mathcal{H}^1$ measure by compact 1-sets $K \subset E$ according Lemma 2.19, it suffices to prove Theorem 5.18 for compact 1-unrectifiable 1-sets. Unless otherwise specified, in this section E will always stand for a set with these properties. We consistently write

$$\mu := \frac{1}{\mathcal{H}^1(E)} \mathcal{H}^1|_E \quad \text{and} \quad \sigma := \frac{1}{\mathcal{H}^1(S^1)} \mathcal{H}^1|_{S^1}.$$

For $e \in S^1$, we also write $\ell_e(x)$ for the line which is parallel to e and contains x.

Definition 5.19 (Directions of high multiplicity and high density)**.** Let $x \in E$ and $e \in S^1$.

- We say that e is a *direction of high multiplicity at* x, denoted $e \in H_x$, if
$$|E \cap \ell_e(x) \cap B(x,r)| \geq 2 \qquad \text{for all } r > 0.$$
The directions $e \in S^1$ such that the above holds for some fixed $r > 0$ are denoted by $H_x(r)$. Clearly $H_x(r_1) \subset H_x(r_2)$ if $r_1 \leq r_2$, and
$$H_x = \bigcap_{r>0} H_x(r), \qquad x \in E. \tag{5.20}$$
Since $E \cap \ell_e(x) \cap B(x,r)$ always contains x itself, $e \in H_x$ means that $E \cap \ell_e(x)$ contains **other points of** E arbitrarily close to x.
- We say that e is a *direction of high density at* x, denoted $e \in D_x$, if for all $r_0, M, \varepsilon > 0$ there exists a radius $0 < r < r_0$ and an arc $J \subset S^1$ with $e \in J$ and $0 < \sigma(J) < \varepsilon$ such that
$$\frac{\mu(C(x,J) \cap B(x,r))}{r} \geq M\sigma(J). \tag{5.21}$$
For $r_0, \epsilon, M > 0$ fixed, we also write $D_x(r_0, \varepsilon, M)$ for those directions $e \in S^1$ such that there exist $0 < r < r_0$ and $J \subset S^1$ with $e \in J$ and $0 < \sigma(J) < \epsilon$ satisfying (5.21). Thus
$$D_x = \bigcap_{r_0,\varepsilon,M>0} D_x(r_0,\varepsilon,M), \qquad x \in E. \tag{5.22}$$

The directions of high multiplicity and high density are two alternative ways of quantifying that close to x, there is "plenty" of E concentrated in arbitrarily narrow cones around $\ell_e(x)$: you can imagine that such information might be useful in showing that E has small projections. Indeed, the proof of Theorem 5.18 now has two main components: the first, more difficult, step is to argue that for μ almost every $x \in E$, σ almost all directions $e \in S^1$ are either directions of high multiplicity or high density. The second, easier, step will be to show that the points with the property from the first step have small projections in almost all directions.

Lemma 5.23. *For μ almost all $x \in E$, σ almost every $e \in S^1$ is either a direction of high multiplicity or a direction of high density at x (or possibly both).*

Proof. Fix $r_0, \varepsilon, M > 0$. By (5.20)-(5.22), it suffices to show that
$$\sigma(S^1 \setminus [H_x(r_0) \cup D_x(r_0,\varepsilon,M)]) = 0 \tag{5.24}$$
for μ almost every $x \in E$. In fact, we can specify these "good" points $x \in E$ immediately. They are the points $x \in E$ satisfying conclusion of Proposition 5.9, namely that for any arc $J \subset S^1$,
$$\limsup_{r \to 0} \frac{\mu(B(x,r)) \cap C(x,J))}{r} \geq c\sigma(J). \tag{5.25}$$
Note that the condition (5.25) is closely related to (5.21): the key difference is that in (5.25) there is no *a priori* guarantee that $c > 0$ can be taken large.

Now, let $x \in E$ be any point satisfying (5.25). We will show that for **every** $e \in S^1$ one of the following alternatives hold:
$$\Theta^{1,*}(H_x(r_0), e) > 0 \quad \text{or} \quad \Theta^1_*(D_x(r_0,\varepsilon,M), e) > 0. \tag{5.26}$$
Since σ almost all $e \in S^1 \setminus [H_x(r_0) \cup D_x(r_0,\varepsilon,M)]$ satisfies
$$\Theta^1(H_x(r_0), e) = 0 = \Theta^1(D_x(r_0,\varepsilon,M), e) \tag{5.27}$$

by the Lebesgue density theorem, this will prove (5.24).

To prove (5.26), fix $e \in S^1$, and assume that the first alternative fails:

$$\Theta^1(H_x(r_0), e) = 0. \tag{5.28}$$

We then need to demonstrate that $\Theta^1_*(D_x(r_0, \varepsilon, M), e) > 0$. By (5.28), for all sufficiently short arcs $J \subset S^1$ with $e \in J$ we have

$$\sigma(H_x(r_0) \cap J) < \frac{c\sigma(J)}{4M}, \tag{5.29}$$

where $c > 0$ is the constant from (5.25). Fix any such arc $J \subset S^1$, furthermore with $0 < \sigma(J) < \epsilon$. For later technical convenience, assume that J is "half-open", i.e. contains the left endpoint but not the right one. We will now prove that

$$\sigma(D_x(r_0, \epsilon, M) \cap J) \geq \frac{\sigma(J)}{4M}, \tag{5.30}$$

which will establish $\Theta^1_*(D_x(r_0, \epsilon, M), e) \gtrsim 1/M > 0$ and hence complete the proof.

Recalling (5.25), we may find $0 < r \leq r_0$ such that $\mu(B(x,r) \cap C(x,J)) \geq cr\sigma(J)$. Since $r \leq r_0$, we infer from (5.29) that also

$$\sigma(H_x(r) \cap J) < \frac{c\sigma(J)}{4M}. \tag{5.31}$$

Then, to prove (5.30), we will find many directions $\xi \in J$ such that for some arc $I \subset S^1$ with $x \in I$ and $0 < \sigma(I) < \epsilon$ we have

$$\frac{\mu(C(x,I) \cap B(x,r))}{r} \geq M\sigma(I). \tag{5.32}$$

Before showing the details, we discuss the idea. Recall the definition of $H_x(r)$. If $\xi \in J \setminus H_x(r)$, then

$$E \cap \ell_\xi(x) \cap B(x,r) = \{x\}.$$

So $E \cap C(x,J) \cap B(x,r)$ is "packed" to the lines ℓ_ξ with $\xi \in H_x(r)$, see Figure 5. But

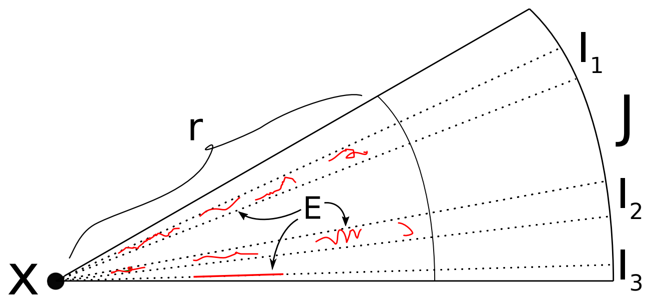

FIGURE 5. The set $E \cap C(x,J) \cap B(x,r)$, in red, is contained in a few narrow cones.

there are very few such lines by (5.31)! Since, however, there is a reasonable amount of μ mass on $C(x,J) \cap B(x,r)$ according to (5.25), and $\operatorname{spt} \mu \subset E$, these observations lead to "crowding": all the μ mass must be packed to a few narrow cones $C(x,I)$, which will then satisfy (5.32).

Now we fill in the details. Let $\mathcal{D}$ be a system of "dyadic arcs" inside J: thus $J \in \mathcal{D}$, and whenever $I \in \mathcal{D}$, then the left and right (half-open) halves of I are in $\mathcal{D}$. We cover $H_x(r) \cap J$ by arcs $I_1, I_2, \ldots \in \mathcal{D}$ such that

$$\sum_{j=1}^{\infty} \sigma(I_j) < \frac{c\sigma(J)}{2M}, \tag{5.33}$$

which can be done by (5.31). We can assume that the arcs I_j are disjoint (this is one benefit of using dyadic arcs). By definition of $H_x(r)$, we note that

$$E \cap C(x, J) \cap B(x, r) \subset \bigcup_{j=1}^{\infty} C(x, I_j) \cap B(x, r). \tag{5.34}$$

We are only interested in the "heavy" arcs I_j with the property

$$\frac{\mu(C(x, I_j) \cap B(x, r))}{r} \geq M\sigma(I_j).$$

We denote these arcs by $\mathcal{G}$, and note that

$$\sum_{I_j \notin \mathcal{G}} \frac{\mu(C(x, I_j) \cap B(x, r))}{r} < M \sum_{I_j \notin \mathcal{G}} \sigma(I_j) \overset{(5.33)}{<} \frac{c\sigma(J)}{2}.$$

It follows from (5.25) and (5.34) that the cones associated to the arcs in $\mathcal{G}$ still cover a substantial amount of $E \cap C(x, J) \cap B(x, r)$:

$$\frac{\mu(C(x, G) \cap B(x, r))}{r} = \sum_{I_j \in \mathcal{G}} \frac{\mu(C(x, I_j) \cap B(x, r))}{r} \geq c\sigma(J) - \frac{c\sigma(J)}{2} = \frac{c\sigma(J)}{2}, \tag{5.35}$$

where

$$G := \bigcup_{I_j \in \mathcal{G}} I_j.$$

Note that $G \subset D_x(r_0, \epsilon, M) \cap J$ because if $\xi \in I_j \in \mathcal{G}$, then I_j is an arc with $\xi \in I_j$ and $0 < \sigma(I_j) \leq \sigma(J) < \epsilon$ such that $\mu(C(x, I_j) \cap B(x, r)) \geq Mr\sigma(I_j)$ (recall also that we chose $0 < r < r_0$). So, to verify (5.30) and conclude the proof of the lemma, it would suffice to argue that $\sigma(G) \gtrsim \sigma(J)/M$.

The bad news is that isn't necessarily true. It could happen that the heavy arcs are extremely heavy compared to their length, such as I_3 in Figure 5, and $\sigma(G)$ can be arbitrarily small. The solution is to replace all "overweight" arcs I_j by dyadic ancestors $\hat{I}_j \subset J$ satisfying $\mu(C(x, \hat{I}_j) \cap B(x, r)) \sim M\sigma(I_j)$, and then their union will satisfy (5.30).

Indeed, for every $I_j \in \mathcal{G}$, choose the maximal dyadic arc $\hat{I}_j \subset J$ such that $I_j \subset \hat{I}_j$ and

$$\frac{\mu(C(x, \hat{I}_j) \cap B(x, r))}{r} \geq M\sigma(\hat{I}_j). \tag{5.36}$$

If $J = \hat{I}_j$ for some $I_j \in \mathcal{G}$, then (5.36) implies that $J \subset D_x(r_0, \epsilon, M)$, and (5.30) is clear. Otherwise the dyadic parent of every $\hat{I}_j$ satisfies the inequality opposite to (5.36), which implies

$$\frac{\mu(C(x, \hat{I}_j) \cap B(x, r))}{r} \leq 2M\sigma(\hat{I}_j).$$

Now, the arcs $\hat{I}_j$ are disjoint, contained $D_x(r_0, \epsilon, M) \cap J$ by (5.36), and their union clearly covers G. Consequently

$$\begin{aligned}\sigma(D_x(r_0, \epsilon, M) \cap J) &\geq \sum \sigma(\hat{I}_j) \\ &\geq \sum \frac{\mu(C(x, \hat{I}_j) \cap B(x,r))}{2Mr} \\ &\geq \frac{\mu(C(x,G) \cap B(x,r))}{2Mr} \overset{(5.35)}{\geq} \frac{c\sigma(J)}{4M}.\end{aligned}$$

This concludes the proof of (5.30), and the lemma. □

We are now almost equipped to prove the Besicovitch projection theorem, but we still record a few easier lemmas separately:

Lemma 5.37. *Let $E \subset \mathbb{R}^2$ be compact. Then, the sets*

$$H := \{(x,e) \in E \times S^1 : e \text{ is a high multiplicity direction at } x\} = \{(x,e) : e \in H_x\}$$

and

$$D := \{(x,e) \in E \times S^1 : e \text{ if a high density direction at } x\} = \{(x,e) : e \in D_x\}$$

are Borel, and in particular $\mu \times \sigma$ measurable.

Proof. Exercise. □

Lemma 5.38. *Let $E \subset \mathbb{R}^2$ be an arbitrary set. Then, for any Lipschitz map $f \colon E \to \mathbb{R}$, and for any $s \geq 1$, we have*

$$\int_{\mathbb{R}}^{*} \mathcal{H}^{s-1}(E \cap f^{-1}\{t\})\,dt \leq \mathrm{Lip}(f)\mathcal{H}^s(E).$$

Remark 5.39. There is no reason to believe that the integrand $t \mapsto \mathcal{H}^{s-1}(E \cap f^{-1}\{t\})$ is Lebesgue integrable under these assumptions, so we use the *upper integral*

$$\int^{*} h\,dt := \inf_{\psi} \int \psi\,dt,$$

where the inf runs over all Lebesgue measurable functions $\psi \geq h$. If h happens to be measurable, then clearly $\int h\,dt \leq \int^* h\,dt$, so proving upper bounds for the upper integral is harder – and the results better – than proving them for the "usual" integral.

Proof of Lemma 5.38. For every $\delta > 0$, let $\{U_j^\delta\}_{j\in\mathbb{N}}$ be a δ-cover for E by open sets with

$$\sum_j \mathrm{diam}(U_j^\delta)^s \leq \mathcal{H}^s_\delta(E) + \delta.$$

Then, for $t \in \mathbb{R}$ fixed, $\{U_j^\delta \cap f^{-1}\{t\}\}_{j\in\mathbb{N}}$ is a δ-cover of $E \cap f^{-1}\{t\}$ so

$$\mathcal{H}^{s-1}(E \cap f^{-1}\{t\}) \leq \liminf_{\delta\to 0} \sum_j \mathrm{diam}(U_j^\delta \cap f^{-1}\{t\})^{s-1} =: \psi(t).$$

Since ψ is evidently Lebesgue measurable, even Borel, we have

$$\int^{*} \mathcal{H}^{s-1}(E \cap f^{-1}\{t\})\,dt \leq \int \psi(t)\,dt \leq \liminf_{\delta\to 0} \int \sum_j \mathrm{diam}(U_j^\delta \cap f^{-1}\{t\})^{s-1}\,dt. \tag{5.40}$$

To estimate the right hand side further, consider the sets

$$F_j^\delta := \{t \in \mathbb{R} : U_j^\delta \cap f^{-1}\{t\} \neq \emptyset\}, \qquad j \in \mathbb{N}.$$

Note that if $t_1, t_2 \in F_j^\delta$, then there exist $x_1, x_2 \in U_j^\delta$ such that $f(x_i) = t_i$, which implies that

$$|t_1 - t_2| \leq \mathrm{Lip}(f)|x_1 - x_2|,$$

and consequently

$$\mathcal{L}^1(F_j^\delta) \leq \mathrm{diam}(F_j^\delta) \leq \mathrm{Lip}(f)\,\mathrm{diam}(U_j^\delta).$$

It follows that

$$\int \sum_j \mathrm{diam}(U_j^\delta \cap f^{-1}\{t\})^{s-1}\,dt = \sum_j \int_{F_j^\delta} \mathrm{diam}(U_j^\delta)^{s-1}\,dt \leq \mathrm{Lip}(f)[\mathcal{H}^s_\delta(E) + \delta].$$

The lemma follows by combining this estimate with (5.40). □

Lemma 5.41. *Let $\nu \in \mathcal{M}(\mathbb{R}^d)$, and let $M\nu$ be the Hardy-Littlewood maximal function*

$$M\nu(x) := \sup_{r>0} \frac{\nu(B(x,r))}{r^d}.$$

Then

$$\mathcal{L}^d(\{x \in \mathbb{R}^d : M\nu(x) > \Lambda\}) \lesssim_d \frac{\nu(\mathbb{R}^d)}{\Lambda}, \qquad \Lambda > 0.$$

Proof. This follows from the usual proof of the fact that $M \colon L^1 \to L^{1,\infty}$ boundedly, familiar from previous courses. Fix $\Lambda > 0$ and write $E_\Lambda := \{x : M\nu(x) > \Lambda\}$. For every $x \in E_\Lambda$ find some radius $r > 0$ with $\nu(B(x,r)) > \Lambda r^d$. Note that $r < (\nu(\mathbb{R}^d)/\Lambda)^{1/d} < \infty$, so we may use the $5r$ covering theorem to extract a countable disjoint sub-family $B(x_1, r_1), B(x_2, r_2), \ldots$ such that

$$E_\Lambda \subset \bigcup_{j \in \mathbb{N}} B(x_j, 5r_j).$$

Then

$$\mathcal{L}^d(E_\Lambda) \lesssim_d \sum_{j\in\mathbb{N}} r_j^d \leq \sum_{j \in \mathbb{N}} \frac{\nu(B(x_j, r_j))}{\Lambda} \leq \frac{\nu(\mathbb{R}^d)}{\Lambda},$$

as claimed. □

Then, we prove the Besicovitch projection theorem:

Proof of Theorem 5.18. Fix $E \subset \mathbb{R}^2$ compact with $0 < \mathcal{H}^1(E) < \infty$, and let $H, D \subset E \times S^1$ be the sets from Lemma 5.37. From Lemma 5.23, we know that

$$\sigma(\{e \in S^1 : e \in H \cup D\}) = 1 \qquad \text{for } \mu \text{ a.e. } x \in E.$$

Hence $\mu \times \sigma(\{(x,e) \in E \times S^1 : e \in H \cup D\}) = 1$, and by Fubini's theorem

$$\mu(\{x \in E : e \in H \cup D\}) = 1 \qquad \text{for } \sigma \text{ a.e. } e \in S^1.$$

Fix $e \in S^1$ as above. We claim that $\mathcal{H}^1(\pi_\xi(E)) = 0$, where $\xi \in S^1 \cap e^\perp$. To see this, write $E = N \cup E_H \cup E_D$, where

$$\mu(N) = 0 \quad \text{and} \quad E_H = \{x \in E : e \in H\} \quad \text{and} \quad E_D = \{x \in E : e \in D\}.$$

To complete the proof, it suffices to show that

(a) $\mathcal{H}^1(\pi_e(N)) = 0$,
(b) $\mathcal{H}^1(\pi_e(E_H)) = 0$,
(c) $\mathcal{H}^1(\pi_e(E_D)) = 0$.

Here (a) is clear, recalling that $\mu = c\mathcal{H}^1|_E$, and the 1-Lipschitz map π_ξ does not increase any Hausdorff measure. The claim (b) follows from Lemma 5.38 applied with $s = 1$ and $f = \pi_e$, and the observation that for all $x \in E_H$, the line $\ell_e(x)$ satisfies $|E \cap \ell_e(x)| = \infty$. This is equivalent to saying that for all $t \in \pi_\xi(E_H)$, the line $\pi_\xi^{-1}\{t\}$ satisfies $|E \cap \pi_\xi^{-1}\{t\}| = \infty$. Hence

$$\mathcal{H}^1(\pi_\xi(E_H)) \leq \frac{1}{A} \int_{\pi_\xi(E_H)}^{*} |E \cap \pi_\xi^{-1}(t)| \, dt \lesssim \frac{\mu(E)}{A}, \qquad A > 0,$$

which gives (b). Finally, (c) follows from Lemma 5.41 by considering the measure $\nu = \pi_\xi \mu \in \mathcal{M}(\mathbb{R})$. Indeed, we claim that

$$M\nu(t) = \infty, \qquad t \in \pi_\xi(E_D). \tag{5.42}$$

To see this, consider $t \in \pi_\xi(E_D)$, and pick $x \in E_D$ with $t = \pi_\xi(x)$. By definition of E_D, we have $e \in D_x$, and hence for any $A \geq 1$ there are $r > 0$ and $e \in J \subset S^1$ such that

$$\frac{\mu(C(x,J) \cap B(x,r))}{r} \geq A\mathcal{H}^1(J)$$

By elementary geometry, there is an interval $I \subset \mathbb{R}$, centred at t and of length $\mathcal{H}^1(I) \sim \mathcal{H}^1(J)r$ such that

$$C(x,J) \cap B(x,r) \subset \pi_\xi^{-1}(I).$$

Hence

$$\nu(I) = \pi_\xi\mu(I) = \mu(\pi_\xi^{-1}(I)) \geq \mu(C(x,J) \cap B(x,r)) \geq A\mathcal{H}^1(J)r \sim A\mathcal{H}^1(I).$$

This implies (5.42) by letting $A \to \infty$, and proves the theorem. □

We end the section with a corollary of Proposition 5.5 and Theorem 5.18:

Corollary 5.43. *Let $E \subset \mathbb{R}^2$ be a Borel 1-set. Then E is purely 1-unrectifiable if and only if E projects to zero length in two distinct directions.*

Proof. If E is purely 1-unrectifiable, the conclusion certainly follows from Theorem 5.18. Conversely, if E is not purely 1-unrectifiable, then E contains a rectifiable piece of positive measure, and then, by Lemma 5.5, E can have at most one zero-length projection. □

6. Fourier transforms of measures

Recall that we only proved the first half of Theorem 4.7 in Section 4.1 To prove the second half, we introduce a useful tool: the Fourier transform of measures. In fact, there is also a proof of Marstrand's theorem without the Fourier transform, but the technique will also have other – indispensable – applications later.

Definition 6.1. Let $\mu \in \mathcal{M}(\mathbb{R}^d)$. The Fourier transform of μ is the function

$$\hat{\mu}(\xi) = \int e^{-i\xi \cdot x} \, d\mu(x).$$

Since $\mu \in \mathcal{M}(\mathbb{R}^d)$ is a finite measure, and $|e^{-i\xi\cdot x}| \equiv 1$, the expression for $\hat{\mu}(\xi)$ is well-defined. We will need a couple of basic facts about the Fourier transform, which we gather in the following lemma:

Lemma 6.2. *Let $\mu \in \mathcal{M}(\mathbb{R}^d)$, and let $f \in C_0(\mathbb{R}^d)$.*

(i) *The following Parseval and Plancherel formulae hold:*

$$\int f\, d\mu = (2\pi)^{-d} \int \hat{f}\bar{\hat{\mu}}\, d\mathcal{L}^d \quad \textit{and} \quad \|f\|_2^2 = (2\pi)^{-d}\|\hat{f}\|_2^2.$$

(ii) *Fourier transform turns convolution into multiplication:*

$$\widehat{\mu * f}(\xi) = \hat{\mu}(\xi)\hat{f}(\xi),$$

(iii) *If $\hat{\mu} \in L^2(\mathbb{R}^d)$, then $\mu \ll \mathcal{L}^d$ and $\mu \in L^2(\mathbb{R}^d)$.*

(iv) *If $T\colon \mathbb{R}^d \to \mathbb{R}^d$ is an invertible linear map, then*

$$\widehat{f \circ T} = |\det(T)|^{-1}\hat{f} \circ (T^{-1})^{\perp},$$

where $(T^{-1})^{\perp}$ is the transpose of the inverse T^{-1}.

Proof. We omit the proof of (i), which can be found on any basic text on Fourier analysis. The proof of (ii) is a short computation:

$$\begin{aligned}
\widehat{\mu * f}(\xi) &= \int e^{-ix\cdot\xi}\left[\int f(x-y)\, d\mu(y)\right] d\mathcal{L}^d(x)\\
&= \int \left[\int e^{-ix\cdot\xi} f(x-y)\, d\mathcal{L}^d(x)\right] d\mu(y)\\
&\overset{x\mapsto u+y}{=} \int \left[\int e^{-i(u+y)\cdot\xi} f(u)\, d\mathcal{L}^d(u)\right] d\mu(y)\\
&= \int e^{-iy\cdot\xi}\, d\mu(y) \int e^{-iu\cdot\xi} f(u)\, d\mathcal{L}^d(u) = \hat{\mu}(\xi)\hat{f}(\xi).
\end{aligned}$$

The proof of (iii) is based on duality and (i): we note that

$$\left|\int f\, d\mu\right| = (2\pi)^{-d}\left|\int \hat{f}\bar{\hat{\mu}}\, d\mathcal{L}^d\right| \leq \|\hat{f}\|_2\|\hat{\mu}\|_2 \sim \|f\|_2\|\hat{\mu}\|_2$$

for all smooth and compactly supported $f\colon \mathbb{R}^d \to \mathbb{C}$. So, if $\hat{\mu} \in L^2$, then $f \mapsto \int f\, d\mu$ defines a bounded linear functional on the dense subspace of L^2 consisting of smooth compactly supported functions f. Such a functional has (by the Hahn-Banach theorem) a unique extension to a continuous linear functional Λ on L^2. Further, continuous functionals on L^2 are always represented by L^2 functions, i.e. there is $g \in L^2$ such that $\Lambda(f) = \int fg\, d\mathcal{L}^d$ for all $f \in L^2(\mathbb{R}^d)$. In particular,

$$\int f\, d\mu = \Lambda(f) = \int fg\, d\mathcal{L}^d$$

for all smooth and compactly supported $f\colon \mathbb{R}^d \to \mathbb{C}$. This easily implies that $d\mu = g\,d\mathcal{L}^d$, as claimed. The proof of (iv) is again a short computation:

$$\widehat{f \circ T}(\xi) = \int e^{-ix\cdot\xi} f(Tx)\,d\mathcal{L}^d(x)$$
$$\overset{x \mapsto T^{-1}(y)}{=} \frac{1}{|\det(T)|} \int e^{iT^{-1}(y)\cdot\xi} f(y)\,d\mathcal{L}^d(y)$$
$$= \frac{1}{|\det(T)|} \int e^{-iy\cdot(T^{-1})^{\perp}\xi} f(y)\,d\mathcal{L}^d(y) = \frac{1}{|\det(T)|} \hat{f}((T^{-1})^{\perp}\xi).$$

This completes the proof of the lemma. □

A key reason why the Fourier transform is so useful for us is that the s-dimensional Riesz energy of μ, for $0 < s < d$, can be expressed in terms of the Fourier transform:

Lemma 6.3. *Let $\mu \in \mathcal{M}(\mathbb{R}^d)$, and let $0 < s < d$. There exists a constant $c = c(s,d) > 0$ such that*

$$I_s(\mu) = c \int |\hat{\mu}(\xi)|^2 |\xi|^{s-d}\,d\mathcal{L}^d(\xi). \tag{6.4}$$

Proof. Recall that

$$I_s(\mu) = \int V_s(\mu)\,d\mu = \int (\mu * k_s)(x)\,d\mu(x),$$

where $k_s(x) = |x|^{-s}$ is the s-dimensional Riesz kernel. The least trivial part of the proof is to find out the Fourier transform of k_s: it is given by

$$\widehat{k_s}(\xi) = c(s,d)k_{d-s}(\xi) = c(s,d)|\xi|^{s-d}, \tag{6.5}$$

whenever $s \in (0,d)$. With (6.5) in hand, the proof of the lemma is very short – at least if we brush all integrability questions under the carpet: applying the Parseval formula from Lemma 6.2(i) to the function $f = \mu * k_s$, and then the convolution formula from Lemma 6.2(ii), we find that

$$I_s(\mu) = (2\pi)^{-d} \int \widehat{\mu * k_s}\bar{\hat{\mu}}\,d\mathcal{L}^d$$
$$= (2\pi)^{-d} \int |\hat{\mu}(\xi)|^2 \widehat{k_s}(\xi)\,d\mathcal{L}^d(\xi)$$
$$\overset{(6.5)}{=} (2\pi)^{-d} c(s,d) \int |\hat{\mu}(\xi)|^2 |\xi|^{s-d}\,d\mathcal{L}^d(\xi).$$

We will omit the precise rigorous justification of the computation above (see Chapter 12 in Mattila's book [16] for more details), but we will discuss (6.5) a little.

First, recall that Fourier transform maps radially symmetric functions to radially symmetric functions: this follows easily from Lemma 6.2(v) applied to orthogonal transformations $O\colon \mathbb{R}^d \to \mathbb{R}^d$. Namely, if $f \in L^1 + L^2$ is radially symmetric, then

$$\hat{f}(O\xi) = |\det((O^{\perp})^{-1}|(\widehat{f \circ (O^{\perp})^{-1}})(\xi) = \hat{f}(\xi),$$

because $(O^{\perp})^{-1}$ is still an orthogonal transformation, so $f \circ (O^{\perp})^{-1} = f$.

Observing that $k_s \in L^1 + L^2$ is radially symmetric, we conclude that $\widehat{k_s}$ is also a radially symmetric function. Next, let T be the self-adjoint linear map $x \mapsto x/R$, where $r > 0$ is a fixed constant. Again using Lemma 6.2(v), we find that

$$\widehat{k_s}(R\xi) = \widehat{k_s} \circ T^{-1}(\xi) = |\det(T)|\widehat{k_s \circ T}(\xi) = R^{s-d}\widehat{k_s}(\xi),$$

using that $(k_s \circ T)(x) = |Tx|^{-s} = |x/R|^{-s} = R^s|x|^{-s} = R^s k_s(x)$. So, $\widehat{k_s}$ is a radially symmetric function satisfying

$$\widehat{k_s}(R\xi) = R^{s-d}\widehat{k_s}(\xi), \qquad \xi \in \mathbb{R}^d, \ R > 0.$$

This implies that $\widehat{k_s}(\xi) = c|\xi|^{s-d}$, where c is the constant

$$c = \widehat{k_s}(x) \text{ for any } x \in S^{d-1}. \tag{6.6}$$

The precise value of c is not that important for us, so won't even attempt to calculate it; it is however clear from the formula (6.4) that $c > 0$, because the left hand side is positive, and the constant c in (6.4) is $(2\pi)^{-d}$ times the constant in (6.6). □

6.1. **Marstrand's projection theorem: part II.** The final technical tool before tackling the second part of Theorem 4.7 is the following lemma, which relates the Fourier transform of $\pi_e\mu$ to the Fourier transform of μ:

Lemma 6.7. *Let $\mu \in \mathcal{M}(\mathbb{R}^d)$, $\xi \in \mathbb{R}$, and $e \in S^{d-1}$. Then*

$$\widehat{\pi_e\mu}(\xi) = \hat{\mu}(\xi e).$$

Proof. This is a simple computation, using formula (3.7):

$$\widehat{\pi_e\mu}(\xi) = \int e^{-ix\xi} d\pi_e\mu(x) = \int e^{-i\pi_e(x)\xi}\, d\mu(x) = \int e^{-ix\cdot\xi e}\, d\mu(x) = \hat{\mu}(\xi e).$$

□

Now, we can finish the proof of Marstrand's projection theorem. We will use integration in polar coordinates, which should be familiar from previous analysis courses:

$$\int f(x)\, d\mathcal{L}^d(x) = \int_{\mathbb{R}} r^{d-1} \int_{S^{d-1}} f(re)\, d\sigma(e)\, dr = \int_{\mathbb{R}} \int_{S(0,r)} f(e)\, d\sigma_r(e)\, dr. \tag{6.8}$$

Here $f \in L^1(\mathbb{R}^d)$ and $\sigma = c\mathcal{H}^{d-1}|_{S^{d-1}}$ and $c\sigma_r = \mathcal{H}^{d-1}|_{S(0,r)}$, where the constant $c = c_d > 0$ is a suitable normalisation so that (6.8) holds.

Proof of the second part of Theorem 4.7. Let $E \subset \mathbb{R}^d$ be a Borel set with $\dim_{\mathrm{H}} E > 1$. By Corollary 4.6, we can pick a measure $\mu \in \mathcal{M}(E)$ with $I_1(\mu) < \infty$. Using first the previous lemma, then integration in polar coordinates, and finally Lemma 6.3, we find that

$$\begin{aligned}\int_{S^{d-1}} \int_{\mathbb{R}} |\widehat{\pi_e\mu}(\xi)|^2\, d\xi\, d\sigma(e) &= \int_{S^{d-1}} \int_{\mathbb{R}} |\hat{\mu}(\xi e)|^2\, d\xi\, d\sigma(e)\\ &\sim \int_{\mathbb{R}^d} \frac{|\hat{\mu}(\xi)|^2}{|\xi|^{d-1}}\, d\mathcal{L}^d(\xi) \sim I_1(\mu) < \infty.\end{aligned}$$

This shows that $\widehat{\pi_e\mu} \in L^2$ for $\mathcal{H}^1$ almost every $e \in S^1$, which by Lemma 6.2(iv) implies that $\pi_e\mu \in L^2$ for these $e \in S^1$. Since $\operatorname{spt} \pi_e\mu \subset \pi_e(E)$, and the support of a non-trivial L^2 function certainly has positive Lebesgue measure (as quantified in the next lemma), we have shown that $\mathcal{L}^1(\pi_e(E)) > 0$ for $\mathcal{H}^1$ almost every $e \in S^1$. □

We claimed above that the support of a non-trivial L^2 function has positive Lebesgue measure. More precisely, we have the following estimate:

Lemma 6.9. *Let $f \in L^2$. Then*

$$\mathcal{L}^d(\operatorname{spt} f) \geq \frac{\|f\|_1^2}{\|f\|_2^2}$$

Proof. The proof is a single application of Cauchy-Schwarz:

$$\|f\|_1 = \int_{\operatorname{spt} f} |f| \, d\mathcal{L}^d \leq \mathcal{L}^d(\operatorname{spt} f)^{1/2} \|f\|_2.$$

The claim follows by rearranging the terms. □

We also record a useful corollary of the proof of Marstrand's projection theorem:

Corollary 6.10. *Assume that $\mu \in \mathcal{M}(\mathbb{R}^d)$. Then $\pi_e \mu \in L^2$ for $\mathcal{H}^1$ almost every $e \in S^1$.*

The proof of (the second part of) Marstrand's projection theorem exhibits one of the most frequently used tricks in the field: to show that a certain set $E \subset \mathbb{R}^d$ has positive measure, try to find a natural measure $\mu \in \mathcal{M}(E)$, show that actually $\mu \in L^2$, and apply Lemma 6.9. Sometimes one can do even better, and show that $\mu \in \mathcal{C}_0(\mathbb{R}^d)$: in this case E contains the support of a non-trivial continuous function, which implies that E has non-empty interior. We will see an example of such an argument in the next section.

6.2. **Distance sets.** As a second application of the Fourier transform, we now turn our attention to Question 5, which is actually an open problem. Here is a collection of the best partial results:

Theorem 6.11. *Let $E \subset \mathbb{R}^2$ be a Borel set with $\dim_H E = s \in [0, 2]$.*

- *If $s > 0$, then $\dim_H \Delta(E) \geq s/2 + \varepsilon$ for some small positive constant $\varepsilon(s) > 0$. This is a result of Bourgain, Katz and Tao from 2000-2003.*
- *If $s > 4/3$, then $\mathcal{L}^1(\Delta(E)) > 0$. This is a result of Wolff from 1999.*
- *If $s > 1$, then $\dim_H \Delta(E) > 0.685$. This is a result of Keleti and Shmerkin from 2018.*

Apart from Question 5, which asks about the Hausdorff dimension of $\Delta(E)$, the following conjecture is open in all dimensions $d \geq 2$:

Conjecture 6.12 (Falconer's distance set conjecture)**.** *Let $E \subset \mathbb{R}^d$, $d \geq 2$, be a Borel set with $\dim_H E > d/2$. Then $\mathcal{L}^1(\Delta(E)) > 0$.*

The proofs of the results in Theorem 6.11 are too complicated for this course, and we will only prove the following partial result, due to Mattila and Sjölin [17] from 1999:

Theorem 6.13. *Let $d \geq 2$, and let $E \subset \mathbb{R}^d$ be a Borel set with $\dim_H E > (d+1)/2$. Then $\Delta(E)$ has non-empty interior, and in particular $\mathcal{L}^1(\Delta(E)) > 0$.*

6.2.1. *Approximating measures by smooth functions.* Before starting the proof of Theorem 6.13, we discuss the useful technique of approximating measures $\mu \in \mathcal{M}(\mathbb{R}^d)$ by smooth functions. Let $\psi \colon \mathbb{R}^d \to [0, \infty)$ be any fixed non-negative $\mathcal{C}^\infty$-smooth function with

$$\int \psi \, d\mathcal{L}^d = 1 \quad \text{and} \quad \operatorname{spt} \psi \subset B(0, 1).$$

Then, define the functions $\{\psi_\varepsilon\}_{\varepsilon>0}$ by $\psi_\varepsilon(x) := \varepsilon^{-d}\psi(x/\varepsilon)$, so that

$$\int \psi_\varepsilon \, d\mathcal{L}^d = 1 \quad \text{and} \quad \operatorname{spt} \psi_\varepsilon \subset B(0,\varepsilon), \qquad \varepsilon > 0.$$

Now, if $\mu \in \mathcal{M}(\mathbb{R}^d)$, we write $\mu_\varepsilon := \mu * \psi_\varepsilon$. The next lemma gathers some useful properties of the measures μ_ε:

Lemma 6.14. *The measures μ_ε have C^∞-densities, they are supported on the ε-neighbourhood of μ, and they approximate μ weakly: $\mu_\varepsilon \rightharpoonup \mu$ as $\varepsilon \to 0$. Further, the Fourier transform of μ_ε is given by*

$$\widehat{\mu_\varepsilon}(\xi) = \widehat{\psi}(\varepsilon\xi)\hat{\mu}(\xi),$$

and the Riesz energies $I_s(\mu_\varepsilon)$, $0 < s < d$, satisfy $I_s(\mu_\varepsilon) \leq I_s(\mu)$ for all $\varepsilon > 0$.

Proof. Exercise. □

Remark 6.15. So, if $I_s(\mu) < \infty$ and $0 < s < d$, then μ can be approximated by smooth functions with uniformly bounded s-energy!

6.2.2. *Back to distance sets.* Now we can prove Theorem 6.13:

Proof of Theorem 6.13. The plan is to find a measure $\nu \in \mathcal{M}(\Delta(E))$, which is absolutely continuous with $\nu \in C_0(\mathbb{R})$. We begin by applying Corollary 4.6 to the set E: since $\dim_{\mathrm{H}} E > (d+1)/2$, there exists $\mu \in \mathcal{M}(E)$ with $I_{(d+1)/2}(\mu) < \infty$. Then, since the map $\delta \colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$,

$$\delta(x,y) = |x-y|,$$

is continuous and $\delta(E \times E) \subset \Delta(E)$, we have $\nu := \delta(\mu \times \mu) \in \mathcal{M}(\Delta(E))$. The main goal of the proof will be to establish that $\nu \ll \mathcal{L}^1$, and

$$\nu(r) = r^{d-1} \int \widehat{\sigma}(r\xi)|\hat{\mu}(\xi)|^2 \, d\mathcal{L}^d(\xi) \quad \text{for a.e. } r > 0, \tag{6.16}$$

where σ is the the normalised $(d-1)$-dimensional Hausdorff measure on S^{d-1}, introduced after (6.8). Before proving (6.16), we argue that the right hand side of (6.16) is continuous, so then (6.16) implies that there is a continuous representative in the L^1 equivalence class of ν. This representative is the measure in $\mathcal{M}(\Delta(E))$ we were after. The continuity of (6.16) is based on the known decay estimate for $\widehat{\sigma}$:

$$|\widehat{\sigma}(\xi)| \lesssim_d |\xi|^{(1-d)/2}, \tag{6.17}$$

which we unfortunately don't have time to prove in these lectures. Hence the integral on the right hand side of (6.16) is bounded in absolute value by

$$\lesssim \int |r\xi|^{(1-d)/2}|\hat{\mu}(\xi)|^2 \, d\mathcal{L}^d(\xi) \sim r^{(1-d)/2} I_{(d+1)/2}(\mu) < \infty, \tag{6.18}$$

using Lemma 6.3. Now the continuity easily of (6.16) for $r > 0$ follows from the dominated convergence theorem.

It remains to prove the formula (6.16). The plan is to prove the equation first for the smooth measures μ_ε from the previous subsection, and then let $\varepsilon \to 0$. To this end, let

$$\nu_\varepsilon := \delta(\mu_\varepsilon \times \mu_\varepsilon), \qquad \varepsilon > 0.$$

For $r > 0$, let σ_r be the normalised $(d-1)$-dimensional Hausdorff measure on $S(0,r)$, introduced in connection with the integration in polar coordinate formula (6.8). We claim that $\nu_\varepsilon \ll \mathcal{L}^1$, and the density of ν_ε is given by

$$\nu_\varepsilon(r) = \int (\sigma_r * \mu_\varepsilon)(x)\mu_\varepsilon(x)\, d\mathcal{L}^d(x) \quad \text{for a.e. } r > 0. \tag{6.19}$$

To see this, fix $g \in C_0(\mathbb{R})$, and compute as follows, using integration in polar coordinates:

$$\begin{aligned}\int_{\mathbb{R}} g(r)\left[\int_{\mathbb{R}^d} (\sigma_r * \mu_\varepsilon)(x)\mu_\varepsilon(x)\, d\mathcal{L}^d(x)\right] dr &= \int_{\mathbb{R}} g(r) \int_{\mathbb{R}^d}\int_{S(0,r)} \mu_\varepsilon(x-e)\, d\sigma_r(e)\mu_\varepsilon(x)\, d\mathcal{L}^d(x)\, dr\\ &= \int_{\mathbb{R}^d}\left[\int_{\mathbb{R}}\int_{S(0,r)} g(|e|)\mu_\varepsilon(x-e)\, d\sigma_r(e)\, dr\right]\mu_\varepsilon(x)\, d\mathcal{L}^d(x)\\ &= \int_{\mathbb{R}^d}\left[\int_{\mathbb{R}^{d-1}} g(|y|)\mu_\varepsilon(x-y)\, d\mathcal{L}^d(y)\right]\mu_\varepsilon(x)\, d\mathcal{L}^d(x)\\ &= \iint g(|x-y|)\mu_\varepsilon(x)\mu_\varepsilon(y)\, d\mathcal{L}^d(x)\, d\mathcal{L}^d(y)\\ &= \int g(r)\, d\delta(\mu_\varepsilon \times \mu_\varepsilon)(r) = \int g(r)\, d\nu_\varepsilon(r).\end{aligned}$$

This proves (6.19). Next, using (6.19), then Lemma 6.2(i)-(ii), and finally Lemma 6.14, we deduce the formula

$$\nu_\varepsilon(r) = \int \widehat{\sigma_r * \mu_\varepsilon}(\xi)\overline{\widehat{\mu_\varepsilon}}(\xi)\, d\mathcal{L}^d(\xi) = \int \widehat{\sigma_r}(\xi)|\widehat{\mu_\varepsilon}(\xi)|^2\, d\mathcal{L}^d(\xi) = \int \widehat{\sigma_r}(\xi)|\hat{\psi}(\varepsilon\xi)|^2|\hat{\mu}(\xi)|^2\, d\mathcal{L}^d(\xi),$$

valid for all $r > 0$ satisfying (6.19). The Fourier transform of σ_r is easy to compute, because if $T_r(x) = rx$, then clearly

$$\sigma_r = r^{d-1}T_r\sigma,$$

and hence

$$\widehat{\sigma_r}(\xi) = r^{d-1}\widehat{T_r\sigma}(\xi) = r^{d-1}\int e^{-ix\cdot\xi}\, dT_r\sigma(x) = r^{d-1}\int e^{-i(rx)\cdot\xi}\, d\sigma(x) = r^{d-1}\widehat{\sigma}(r\xi). \tag{6.20}$$

So, we have now shown that

$$\nu_\varepsilon(r) = r^{d-1}\int \widehat{\sigma}(r\xi)|\hat{\psi}(\varepsilon\xi)|^2|\hat{\mu}(\xi)|^2\, d\mathcal{L}^d(\xi) \quad \text{for a.e. } r > 0. \tag{6.21}$$

Note that

$$|\hat{\psi}(\xi)| \le \int \psi\, d\mathcal{L}^d = 1, \qquad \xi \in \mathbb{R}^d,$$

so the integrand on the right hand side of (6.21) is uniformly (for all $\varepsilon > 0$) bounded by $|\widehat{\sigma}(r\xi)||\hat{\mu}(\xi)|^2$. We already saw in (6.18) that this function is integrable for a fixed $r > 0$. So, since $\widehat{\psi}(0) = \int \psi = 1$, the dominated convergence theorem gives

$$r^{d-1}\int \widehat{\sigma}(r\xi)|\hat{\psi}(\varepsilon\xi)|^2|\hat{\mu}(\xi)|^2\, d\mathcal{L}^d(\xi) \to r^{d-1}\int \widehat{\sigma}(r\xi)|\hat{\mu}(\xi)|^2\, d\mathcal{L}^d(\xi) \tag{6.22}$$

as $\varepsilon \to 0$. On the other hand, it is clear that $\nu_\varepsilon \rightharpoonup \nu$ as $\varepsilon > 0$; more generally, it is true that if $\mu_i \rightharpoonup \mu$, then $f\mu_i \rightharpoonup f\mu$ for any continuous function f, which follows immediately from the definitions. To wrap up, the L^1-functions $r \mapsto \nu_\varepsilon(r)$ converge almost everywhere to the right hand side of (6.22), and on the other hand the measures ν_ε converge weakly to ν. Since almost everywhere convergence implies weak convergence (easy exercise), and weak limits are unique, these facts establish (6.16).

We have now shown that $\nu \in \mathcal{M}(\Delta(E))$ has a continuous representative $\tilde{\nu}$ in its L^1 equivalence class. Clearly $\operatorname{spt} \tilde{\nu} = \operatorname{spt} \nu \in \mathcal{M}(\Delta(E))$. The support of a non-zero continuous function contains an interval, so we have shown that $\Delta(E)$ has non-empty interior. □

6.3. **Fourier dimension and spherical averages.** Recall from Lemma 6.3 that the s-energy of a measure $\mu \in \mathcal{M}(\mathbb{R}^d)$ can be expressed in terms of the Fourier transform as follows:

$$I_s(\mu) = c \int |\hat{\mu}(\xi)|^2 |\xi|^{s-d} \, d\mathcal{L}^d(\xi), \qquad 0 < s < d.$$

This means that if $I_s(\mu) < \infty$, then the Fourier transform of μ has some decay "on average". For example, we have the obvious inequalities

$$\int_{B(0,R)} |\hat{\mu}(\xi)|^2 \, d\xi \leq R^{d-s} \int |\hat{\mu}(\xi)|^2 |\xi|^{s-d} \, d\xi \lesssim R^{d-s} I_s(\mu). \tag{6.23}$$

Since the Lebesgue measure of $\{z : R \leq |z| \leq 2R\}$ is roughly R^d, the above implies that

$$\fint_{\{\xi : R \leq |\xi| \leq 2R\}} |\hat{\mu}(\xi)|^2 \, d\xi \lesssim R^{-s}, \qquad R > 0,$$

whenever $I_s(\mu) < \infty$. One might interpret this by saying that $|\hat{\mu}(\xi)| \lesssim |\xi|^{-s/2}$ **on average**. Now, the natural question arises: are the words "on average" necessary? Perhaps the Fourier transform $\hat{\mu}(\xi)$ decays pointwise as $|\xi| \to \infty$? This is generally not true, but studying the question leads to many interesting concepts.

Definition 6.24 (Fourier dimension)**.** Let $E \subset \mathbb{R}^d$. The *Fourier dimension* of E is

$$\dim_{\mathrm{F}} E := \sup\{s \in [0,d] : \exists \, \mu \in \mathcal{M}(E) \text{ and } C \geq 1 \text{ with } |\hat{\mu}(\xi)| \leq C|\xi|^{-s/2} \text{ for all } \xi \in \mathbb{R}^d \setminus \{0\}\}.$$

Fourier dimension is related to Hausdorff dimension as follows:

Proposition 6.25. $\dim_{\mathrm{F}} E \leq \dim_{\mathrm{H}} E$ *for all* $E \subset \mathbb{R}^d$.

Proof. Exercise. □

Sometimes the inequality is an equality:

Example 6.26. *The Fourier dimension of* S^{d-1} *is* $d-1$*. Indeed, as mentioned during the previous proof, the measure* $\sigma = \mathcal{H}^{d-1}|_{S^{d-1}}$ *satisfies* $|\hat{\sigma}(\xi)| \lesssim |\xi|^{(1-d)/2}$ *for all* $\xi \in \mathbb{R}^d \setminus \{0\}$.

But the inequality can also be very far from an equality:

Example 6.27. *The Fourier dimension of* $E = [0,1] \times \{0\} \subset \mathbb{R}^2$ *is zero. We leave the proof as an exercise: consider how the Fourier transform of any measure* $\mu \in \mathcal{M}(E)$ *behaves on the* y*-axis.*

In particular, the previous example shows that the Fourier transform or $\mathcal{H}^1|_E$ does not decay pointwise, although $I_s(\mathcal{H}^1|_E) < \infty$ for any $0 \leq s < 1$. So, to get decay for $\hat{\mu}(\xi)$ from the finiteness of $I_s(\mu)$, one generally needs to do some averaging. These observations lead to the next natural question: how about *spherical averages* of the form

$$\sigma(\mu)(r) := \fint_{S(0,r)} |\hat{\mu}|^2 \, d\sigma_r?$$

Recall the measure σ_r from (6.8). This turns out to be a highly non-trivial question, and the sharp bounds are only known in the plane. They are the following:

Theorem 6.28. *Let $0 < s < 2$, and let $\mu \in \mathcal{M}(\mathbb{R}^2)$ be an s-Frostman measure. Then:*

(i) *If $s \in (0, \frac{1}{2}]$, then $\sigma(\mu)(r) \lesssim_\varepsilon r^{-s+\varepsilon}$ for any $\varepsilon > 0$.*
(ii) *If $s \in (\frac{1}{2}, 1]$, then $\sigma(\mu)(r) \lesssim r^{-1/2}$.*
(iii) *If $s \in (1, 2)$, then $\sigma(\mu)(r) \lesssim_\varepsilon r^{-s/2+\varepsilon}$ for any $\varepsilon > 0$.*

The first two results are due to Mattila [18], and the last one – by far the hardest – is due to Wolff [22]. We now prove the first two bounds.

Proof of (i)-(ii) in Theorem 6.28. We only prove (i); it will be clear from the proof that also (i) holds. Moreover, we prove the result with the additional (qualitative) assumption that μ is smooth and compactly supported. The general case can be reduced to this by considering the measures $\mu_\epsilon = \mu * \psi_\epsilon$, as in the proof of Theorem 6.13. Then, we may legitimately write

$$\sigma(\mu)(r) \sim \frac{1}{r}\int_{S^1(r)} |\hat{\mu}|^2 \, d\sigma_r = \frac{1}{r}\int (\mu * \widehat{\sigma_r})(x)\mu(x)\, dx \overset{(6.20)}{\leq} \iint |\hat{\sigma}(r(x-y))|\mu(x)\mu(y)\, dx\, dy.$$

Recalling the decay of $\hat{\sigma}$ from (6.17), we note that

$$|\hat{\sigma}(r(x-y))| \lesssim \min\{1, (r|x-y|)^{-1/2}\} \leq (r|x-y|)^{-s}$$

for all $r > 0$, $x, y \in \mathbb{R}^2$, and $0 \leq s \leq 1/2$. Hence,

$$|\sigma(\mu)(r)| \lesssim r^{-s} \iint \frac{\mu(x)\mu(y)}{|x-y|^s}\, dx\, dy = r^{-s} I_s(\mu).$$

Finally, if μ is s-Frostman, then recall that $I_{s-\epsilon}(\mu) < \infty$ for all $\epsilon > 0$. Part (i) follows from this, and the estimate above. □

How about Fourier dimension on the real line? There the idea of "averaging over directions" evidently doesn't make much sense. In $\mathbb{R}$, the Fourier dimension of a set is strongly linked with how "arithmetic" the set is. Heuristically, if the set lies in a small neighbourhood of an arithmetic progression at infinitely many scales, then its Fourier dimension is zero. To see this, note that

$$|\hat{\mu}(\xi)| \geq \left|\int \cos(x\xi)\, d\mu x\right|.$$

So if $\epsilon > 0$, and

$$E \subset \{x \in \mathbb{R} : \cos(x\xi) \geq |\xi|^{-\epsilon}\}, \tag{6.29}$$

then $|\hat{\mu}(\xi)| \geq |\xi|^{-\epsilon}$ for **any** probability measure $\mu \in \mathcal{M}(E)$. For any $0 < s < 1$ and $\epsilon > 0$, it is not too difficult to construct compact sets $E \subset [0,1]$ with $\dim_{\mathrm{H}} E > s$ such that (6.29) holds for all $\xi \in \{\xi_i\}_{i\in\mathbb{N}}$, where $\{\xi_i\}_{i\in\mathbb{N}}$ is a rapidly increasing sequence. So, even on the real line, large Hausdorff dimension does not guarantee large Fourier dimension.

We now prove the following result, which has a nice application to Borel subrings (see Corollary 6.32). The proof is essentially from Bourgain's paper [3], but he calls the theorem "classical and elementary".

Theorem 6.30. *Assume that $A, B \subset \mathbb{R}$ are Borel sets. Then,*

$$\dim_{\mathrm{F}} AB \geq \min\{\dim_{\mathrm{H}} A + \dim_{\mathrm{H}} B - 1, 1\},$$

where $AB = \{xy : x \in A$ and $y \in B\}$.

Proof. We can easily reduce to the situation where $A, B \subset [0,1]$, and $\dim_{\mathrm{H}} A > s > 0$ and $\dim_{\mathrm{H}} B > t > 0$. Then, pick $\mu \in \mathcal{M}(A)$ and $\nu \in \mathcal{M}(B)$ with

$$I_s(\mu) < \infty \quad \text{and} \quad I_t(\mu) < \infty.$$

Choose a smooth compactly supported function $\varphi \colon \mathbb{R} \to [0,\infty)$ with positive Fourier transform, and with $\min\{\varphi(x) : x \in [0,1]\}) > 0$. Such functions are not hard to find: for example, you can start with a non-negative compactly supported function $\eta \colon \mathbb{R} \to [0,\infty)$, and then consider the convolution $\varphi = \eta * \tilde{\eta}$, where $\tilde{\eta}(x) = \eta(-x)$. Then φ is non-negative, smooth, compactly supported, and $\hat{\varphi} = |\hat{\eta}|^2 \geq 0$.

Then, consider the measure $(\varphi\mu)\nu \in \mathcal{M}(AB)$, defined as the push-forward of $\varphi\mu \times \nu$ under the map $(x,y) \mapsto xy$. We claim that

$$|\widehat{(\varphi\mu)\nu}(\xi)| \lesssim |\xi|^{(1-s-t)/2}, \qquad \xi \in \mathbb{R}. \tag{6.31}$$

This will show that $\dim_{\mathrm{F}} AB \geq s + t - 1$ and complete the proof.

Fix $\xi \in \mathbb{R}$, and assume with no loss of generality that $\xi \geq 1$. We start by noting that

$$\widehat{(\varphi\mu)\nu}(\xi) = \int e^{-it\xi}\, d[(\varphi\mu)\nu](t) = \iint e^{-i(xy)\xi}\, d(\varphi\mu)(x)\, d\nu(y) = \int \widehat{\varphi\mu}(y\xi)\, d\nu y,$$

so

$$|\widehat{(\varphi\mu)\nu}(\xi)| \leq \int |\widehat{\varphi\mu}(y\xi)|\, d\nu(y) = \int |\hat{\varphi} * \hat{\mu}(y\xi)|\, d\nu(y) \leq \iint \hat{\varphi}(x - y\xi)|\hat{\mu}(x)|\, dx\, d\nu(y).$$

Here we used the non-negativity of $\hat{\varphi}$. Since $\hat{\varphi}$ is rapidly decreasing, and $\operatorname{spt} \nu \subset [0,1]$, and $\|\hat{\mu}\|_\infty \leq \mu(A) < \infty$, we have

$$\int_{\{x : |x| \geq 2\xi\}} \hat{\varphi}(x - y\xi)|\hat{\mu}(x)|\, dx \lesssim \int_{\{x : |x| \geq \xi\}} \frac{1}{x^2}\, dx \lesssim \frac{1}{\xi}, \qquad y \in \operatorname{spt} \nu.$$

It follows that

$$|\widehat{(\varphi\mu)\nu}(\xi)| \lesssim \int_{B(0,2\xi)} |\hat{\mu}(x)| \int \hat{\varphi}(x - y\xi)\, d\nu(y)\, dx + \xi^{-1} =: \mathrm{I} + \xi^{-1}.$$

To estimate I, we first apply Cauchy-Schwarz and the estimate (6.23):

$$\begin{aligned} \mathrm{I} &\leq \left(\int_{B(0,2\xi)} |\hat{\mu}(x)|^2 \right)^{1/2} \left(\int \left[\int \hat{\varphi}(x - y\xi)\, d\nu y \right]^2 dx \right)^{1/2} \\ &\lesssim \xi^{(1-s)/2} \left(\int \left[\int \hat{\varphi}(x - y\xi)\, d\nu(y) \right]^2 dx \right)^{1/2} =: \xi^{(1-s)/2} (\mathrm{II})^{1/2}. \end{aligned}$$

To estimate the factor II, we note that

$$\hat{\varphi}(x - y\xi) = \hat{\varphi}(\xi(x/\xi - y)) = \widehat{\varphi_\xi}(x/\xi - y),$$

where $\varphi_\xi(x) = \xi^{-1}\varphi(x/\xi)$ as usual. Hence

$$\int \hat{\varphi}(x - y\xi)\, d\nu(y) = \widehat{\varphi_\xi} * \nu(x/\xi).$$

Finally, using Plancherel, the compact support of φ (say $\operatorname{spt}\varphi \subset B(0,C)$), and (6.23), we end up with

$$\begin{aligned} \mathrm{II} = \int (\widehat{\varphi_\xi} * \nu(x/\xi))^2 \, dx = \xi \int (\widehat{\varphi_\xi} * \nu(z))^2 \, dz \leq \xi \int \varphi_\xi(u)^2 |\hat{\nu}(u)|^2 \, du \\ = \xi^{-1} \int \varphi(u/\xi)^2 |\hat{\nu}(u)|^2 \, du \\ \lesssim \xi^{-1} \int_{B(0,C\xi)} |\hat{\nu}(u)|^2 \, du \overset{(6.23)}{\lesssim} \xi^{-t}. \end{aligned}$$

Combining all the estimates, and noting that $\xi^{-1} \leq \xi^{(1-s-t)/2}$ for $\xi \geq 1$, we get

$$|\widehat{(\varphi\mu)\nu}(\xi)| \lesssim \xi^{(1-s)/2}(\mathrm{II})^{1/2} + \xi^{-1} \lesssim \xi^{(1-s-t)/2}.$$

This completes the proof. □

Corollary 6.32. *If $R \subset \mathbb{R}$ is a Borel subring, then either $\dim_{\mathrm{H}} R \leq 1/2$, or $R = \mathbb{R}$.*

Proof. Exercise: start by inferring from the previous theorem that if $\dim_{\mathrm{H}} R > 1/2$, then $\dim_{\mathrm{F}} RR > 0$. □

7. Kakeya sets

We now study Question 6.

Definition 7.1. A set $E \subset \mathbb{R}^d$ is called a *Kakeya set*, if for every $e \in S^{d-1}$ there exists a unit line segment I_e parallel to e such that $I_e \subset K$.

We are **not** assuming that the line segments pass through the origin. The space $\mathbb{R}^d$ is a Kakeya set, and the ball $B(0, \frac{1}{2})$ is also a Kakeya set. Significantly smaller examples are hard to come by, so Kakeya asked, in the early 20th century, whether all such sets have positive Lebesgue measure. This was disproved by Besicovitch [1] in 1919:

Theorem 7.2 (Besicovitch)**.** *For any $d \geq 2$, there exists a compact Kakeya set of zero measure. In fact, there exists a set $B \subset \mathbb{R}^d$ containing a full line (not just a line segment) in every direction.*

Kakeya sets with zero measure are often referred to as *Besicovitch sets*. Besicovitch's construction is very neat but slightly complicated: to see a picture, check out the Wikipedia page for Kakeya sets. Note that it is enough to construct Besicovitch sets in $\mathbb{R}^2$, because if $E \subset \mathbb{R}^2$ is a Besicovitch set, then $E \times \mathbb{R}^{d-2}$ is a Besicovitch set in $\mathbb{R}^d$.

Proof of Theorem 7.2. Let $E \subset [0,1]^2 \subset \mathbb{R}^2$ be a purely 1-unrectifiable set with $0 < \mathcal{H}^1(E) < \infty$ and $\pi_\infty(E) = [0,1]$, where π_∞ is the projection $\pi_\infty(a,b) = a$. The standard example is shown in Figure 6 (more precisely, the set E is the self-similar set obtained by iterating the rule shown in Figure 6 infinitely many times). It is an easy exercise to show that the set E in Figure 6 has $0 < \mathcal{H}^1(E) < \infty$ and $\pi_\infty(E) = [0,1]$, and that E has two projections of zero length. Hence E is purely 1-unrectifiable by Corollary 5.43.

For $t \in \mathbb{R}$, write $\pi_t(a,b) = at + b$, and note that π_t is just a constant times the orthogonal projection to the line $\operatorname{span}(t,1)$. As t varies in $\mathbb{R}$, the slopes of the lines $\operatorname{span}(t,1)$ vary smoothly through all possible slopes, except the horizontal one (the horizontal projection can be interpreted as π_∞). So, by the Besicovitch projection theorem,

$$\mathcal{H}^1(\pi_t(E)) = 0 \quad \text{for a.e. } t \in \mathbb{R}. \tag{7.3}$$

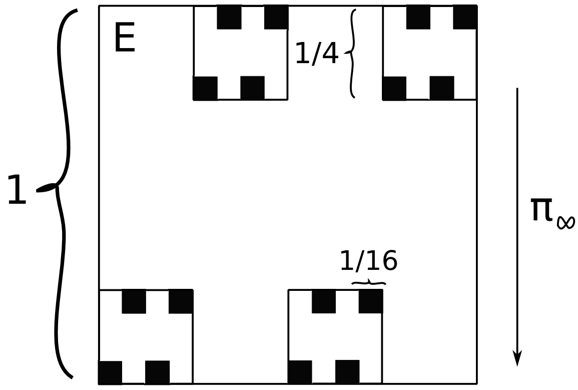

FIGURE 6. A purely 1-unrectifiable set E with one big projection.

Now, for a given point $(a, b) \in E$, consider the line

$$L_{a,b} = \{(x, y) : y = ax + b\},$$

and write

$$B := \bigcup_{(a,b)\in E} L_{a,b}.$$

Since $\pi_\infty(E) = [0, 1]$, for every $a \in [0, 1]$ there corresponds a point $b \in [0, 1]$ such that $(a, b) \in E$, and hence $L_{a,b} \subset B$. So, B contains lines segments of all possible slopes $a \in [0, 1]$, which corresponds to slopes between horizontal and 45 degrees. If you want line segments instead of lines, intersect B with some big closed box; then you also get a compact set. Finally, a union of a few (four) rotated copies of this compact set is a compact Kakeya set. It remains to show that $\mathcal{L}^2(B) = 0$.

To see this, consider the intersection of B with a fixed vertical line $L_t := \{(x, y) : x = t\}$, $t \in \mathbb{R}$. If $(x, y) \in B \cap L_t$, then there exists $(a, b) \in E$ such that $(x, y) \in L_{a,b}$, hence

$$y = ax + b = at + b = \pi_t(a, b) \in \pi_t(E).$$

In other words $B \cap L_t = \{t\} \times \pi_t(E)$, and consequently $\mathcal{H}^1(B \cap L_t) = \mathcal{H}^1(\pi_t(E)) = 0$ for a.e. $t \in \mathbb{R}$ by (7.3). Finally, Fubini's theorem implies that $\mathcal{L}^2(B) = 0$. □

So, Kakeya sets can have measure zero, but how about their dimension?

Question 8. *Do Kakeya sets in $\mathbb{R}^d$ have Hausdorff dimension d?*

A positive answer to this question is known as the *Kakeya conjecture*, which is one of the most famous open problems in geometric measure theory. The answer is only known in the plane. The following theorem is due to Davies [5] from 1971:

Theorem 7.4 (Davies). *Every Kakeya set in $\mathbb{R}^2$ has Hausdorff dimension 2.*

7.1. **The box dimension of planar Kakeya sets is 2.** We will start by proving Davies' theorem for box dimension instead of Hausdorff dimension. Recall from (2.12) that $\dim_{\mathrm{H}} E \leq \underline{\dim}_{\mathrm{B}} E$ for all sets $E \subset \mathbb{R}^d$, so this is an easier task. We will prove the following claim, following the approach of Córdoba [4] from 1977:

Proposition 7.5. *Let $S = \{e_1, \ldots, e_N\} \subset S^1$ be a set of δ-separated vectors, and let T_j, $1 \leq j \leq |S|$, be a $(\delta \times 1)$ tube parallel to e_j. Then, writing $\mathcal{T} := \bigcup T_j$, we have*

$$\mathcal{L}^2\left(\bigcup_{j=1}^{N} T_j\right) \geq \frac{cN\delta}{\log(1/\delta)}$$

for some absolute constant $c > 0$.

First, let's see how this implies the desired box dimension estimate for Kakeya sets. Let $E \subset \mathbb{R}^2$ be a Kakeya set. Pick a **maximal** δ-separated set $S_\delta := \{e_1, \ldots, e_N\} \subset S^1$, so that $N \sim \delta^{-1}$. For every $1 \leq j \leq N$, fix a line segment $I_j \subset E$ parallel to e_j, and let T_j be a $(1 \times \delta)$-tube around I_j. Then, if $B(x_1, \delta/2), \ldots, B(x_M, \delta/2)$ is an arbitrary cover of E by balls of diameter δ, we certainly have

$$\mathcal{T} := \bigcup_{j=1}^{N} T_j \subset \bigcup_{i=1}^{M} B(x_i, 10\delta). \tag{7.6}$$

Now, it follows from Proposition 7.5 that $\mathcal{L}^2(\mathcal{T}) \geq c \log^{-1}(1/\delta)$ for some absolute constant $c > 0$. By (7.6), this implies that $M \geq c\delta^{-2}/\log(1/\delta)$, and consequently

$$\liminf_{\delta \to 0} \frac{\log N(E, \delta)}{-\log \delta} \geq \liminf_{\delta \to 0} \frac{\log \delta^{-2} + \log c - \log\log 1/\delta}{-\log \delta} = 2,$$

as desired.

We then prove Proposition 7.5.

Proof of Proposition 7.5. We study the function

$$f = \sum_{j=1}^{N} \chi_{T_j}.$$

Recalling Lemma 6.9, we have

$$\mathcal{L}^2(\mathcal{T}) = \mathcal{L}^2(\operatorname{spt} f) \geq \frac{\|f\|_1^2}{\|f\|_2^2} = \frac{(N\delta)^2}{\|f\|_2^2}. \tag{7.7}$$

So, we are reduced to proving that $\|f\|_2 \lesssim N\delta \log(1/\delta)$. For slight technical convenience, we assume that all the directions e_j, $1 \leq j \leq N$, lie on one hemisphere of S^1. This allows us to enumerate the vectors e_j so that

$$|e_i - e_j| \gtrsim \delta|j - i|, \qquad 1 \leq i, j \leq N. \tag{7.8}$$

To get rid of the extra assumption, just write $f = f_1 + f_2$ (where the vectors corresponding to f_1 lie on the upper hemisphere, for instance), and then estimate $\|f\|_2^2 \lesssim \|f_1\|_2^2 + \|f_2\|_2^2$.

To bound $\|f\|_2^2$, start by writing

$$\|f\|_2^2 = \int \left(\sum_{j=1}^{N} \chi_{T_j}\right)^2 \, d\mathcal{L}^d = \int \sum_{i,j=1}^{N} \chi_{T_i}\chi_{T_j} \, d\mathcal{L}^d = \sum_{i,j=1}^{N} \mathcal{L}^d(T_i \cap T_j).$$

We make the elementary geometric observation that if the angle between the directions of the segments I_i and I_j is $|e_i - e_j| \gtrsim \delta$, then

$$\mathcal{L}^d(T_i \cap T_j) \lesssim \frac{\delta^2}{|e_i - e_j|}.$$

Recalling (7.8), this allows us to conclude that

$$\sum_{j=2}^{N} \mathcal{L}^2(T_1 \cap T_j) \lesssim \sum_{j=2}^{N} \frac{\delta^2}{(j-1)\delta} \sim \delta \log N \lesssim \delta \log(\tfrac{1}{\delta}).$$

Of course one gets the same estimate with T_1 replaced by any of the other tubes T_i. So, we conclude that

$$\sum_{i,j=1}^{N} \mathcal{L}^d(T_i \cap T_j) = N\delta + \sum_{j \neq i} \mathcal{L}^d(T_i \cap T_j) \lesssim N\delta \log(\tfrac{1}{\delta}), \qquad 1 \leq i \leq N,$$

and consequently

$$\|f\|_2^2 \lesssim N\delta \log(\tfrac{1}{\delta}). \tag{7.9}$$

Recalling (7.7), the proof is complete. □

7.2. **The pigeonhole principle, and the Hausdorff dimension of Kakeya sets.** What is the difference between Hausdorff and box dimension? In the definition of Hausdorff dimension (via Hausdorff measures), coverings by sets of different sizes are allowed, while box dimension only considers coverings (with balls) of a fixed size. As the example of $\mathbb{Q} \cap [0,1]$ demonstrates, sometimes this difference can lead to differences in the values of the two dimensions.

Quite often in practice, however, a proof giving an estimate for the **lower** box dimension can be tinkered to give the same estimate for Hausdorff dimension. The proof strategy is *pigeonholing*. Recall the pigeonhole principle: if you throw $(n+1)$ rocks in n buckets, at least one bucket will end up holding at least two rocks: if this weren't true, then you would clearly have $\leq n$ rocks altogether! Another version – more useful for us – is the following.

Lemma 7.10 (Pigeonhole principle). *Assume that ν is a finite measure on a space X, and $A_1, A_2, \ldots$ is a sequence of arbitrary sets with*

$$\sum_{i \in \mathbb{N}} \nu(A_i) \geq \nu(X). \tag{7.11}$$

Then, there exists an index $i \geq 1$ such that $\nu(A_{i_0}) \geq c\nu(X)/i_0^2$, where $c = 6/\pi^2$.

Proof. If this were not true, then

$$\nu(X) \leq \sum_{i=1}^{\infty} \nu(A_i) < \frac{6\nu(X)}{\pi^2} \sum_{i=1}^{\infty} \frac{1}{i^2} = \frac{6\nu(X)}{\pi^2} \cdot \frac{\pi^2}{6} = \nu(X),$$

a contradiction. □

What does any of this have to do with Hausdorff dimension? Assume that you have a set $E \subset \mathbb{R}^d$, which comes with a measure $\nu \in \mathcal{M}(E)$ (don't think about Frostman measures here, because we don't yet have any lower bounds for $\dim_{\mathrm{H}} E$). Then, assume that you know how to bound from below $N(E', \delta)$ for all $E' \subset E$ with "substantial" ν

measure. Then, it turns out that you can prove the same estimate for $\dim_{\mathrm{H}} E$! Here is a more rigorous statement:

Lemma 7.12. *Let $E \subset \mathbb{R}^d$ be a set, and let $\nu \in \mathcal{M}(E)$. Assume that the following holds for some $s \geq 0$, and for all $i \in \mathbb{N}$: whenever $E' \subset E$ satisfies $\nu(E') \gtrsim \nu(E)/i^2$, then*

$$N(E', 2^{-i}) \gtrsim 2^{is}.$$

Then $\mathcal{H}^s_1(E) \gtrsim 1$, and in particular $\dim_{\mathrm{H}} E \geq s$.

Proof. Let $U_1, U_2, \ldots$ be an arbitrary cover of E with sets of diameter at most 1. For $i \in \mathbb{N}$, write

$$\mathcal{I}_i := \{j \in \mathbb{N} : 2^{-i-1} < \mathrm{diam}(U_j) \leq 2^{-i}\},$$

and let

$$A_i := \bigcup_{j \in \mathcal{I}_i} U_j.$$

Clearly the sets A_i, $i \in \mathbb{N}$, cover E, and hence they satisfy (7.11). It follows that there exists an index $i_0 \in \mathbb{N}$ such that

$$\nu(A_{i_0}) \gtrsim \frac{\nu(E)}{i_0^2}.$$

We now apply the hypothesis with $E' = E \cap A_{i_0}$. The sets $\{U_j\}_{j \in \mathcal{I}_{i_0}}$ form a cover of E' by sets of diameter $\mathrm{diam}(U_j) \sim 2^{-i}$, so

$$|\mathcal{I}_{i_0}| \gtrsim N(E', 2^{-i_0}) \gtrsim 2^{i_0 s}.$$

Consequently,

$$\sum_{j \in \mathbb{N}} \mathrm{diam}(U_j)^s \gtrsim |\mathcal{I}_{i_0}| \cdot 2^{-i_0 s} \gtrsim 1,$$

which shows that $\mathcal{H}^s_1(E) \gtrsim 1$. □

We won't actually need the lemma above as such, but rather the proof idea behind it.

Proof of Theorem 7.4. Let $E \subset \mathbb{R}^2$ be a Kakeya set, and let $U_1, U_2, \ldots$ be an arbitrary cover of E with sets of diameter at most 1. Write

$$\mathcal{I}_i := \{j \in \mathbb{N} : 2^{-i-1} < \mathrm{diam}(U_j) \leq 2^{-i}\}$$

and

$$A_i := \bigcup_{j \in \mathcal{I}_i} U_j,$$

as before. For $e \in S^1$, let $I_e \subset E$ be a unit line segment parallel to e. For $e \in S^1$ fixed, the sets $A_1, A_2, \ldots$ clearly cover I_e, so there exists by Lemma 7.10 an index $i_e \in \mathbb{N}$ such that

$$\mathcal{H}^1(A_{i_e} \cap I_e) \gtrsim \frac{\mathcal{H}^1(I_e)}{i_e^2} = \frac{1}{i_e^2}.$$

If all the indices i_e were the same, we would be in pretty good shape, but now they evidently may depend on e. However, we can run a second pigeonholing argument to fix this, as follows. For $i \in \mathbb{N}$, let

$$S_i := \{e \in S^1 : i_e = i\}.$$

Then, by a second application of Lemma 7.10, we find an index $i_0 \in \mathbb{N}$ such that

$$\sigma(S_{i_0}) \gtrsim \frac{1}{i_0^2},$$

where we recall that $\sigma = c\mathcal{H}^1|_{S^1}$; in this proof, we choose c so that $\sigma(S^1) = 1$. Now, write

$$\delta := 2^{-i_0}.$$

We have shown that there exists a set $S := S_{i_0} \subset S^1$ of measure

$$\sigma(S) \gtrsim \frac{1}{i_0^2} = \log^{-2}(1/\delta) \tag{7.13}$$

such that for every $e \in S$ we have $i_e = i_0$, which by definition means that

$$\mathcal{H}^1(A_{i_0} \cap I_e) \gtrsim \frac{1}{i_0^2} = \log^{-2}(1/\delta). \tag{7.14}$$

All of this may look a little complicated at first, but two observations will clarify things: first, if we had $\log^{-2}(1/\delta) = 1$, then (7.13)-(7.14) would say that $S = S^1$, (up to a null set) and for all $e \in S^1$ the whole segment I_e (up to a null set) is covered by the sets U_j, $j \in \mathcal{I}_0$, which have diameter $\sim \delta$. So, we would be **precisely** in the situation of the proof from the previous section! The second observation is that the number $\log^{-2}(1/\delta)$ is very large compared to quantities like δ^{ε}, so it won't affect any exponents in the following estimates.

Now, for the rest of the details. This will hopefully feel a little repetitive after the previous section! Start by finding a maximal δ-separated subset $\{e_1, \ldots, e_N\} \subset S$, this time with

$$N \gtrsim \frac{\sigma(S)}{\delta} \gtrsim \delta^{-1} \log^{-2}(1/\delta). \tag{7.15}$$

Next, for every $1 \leq k \leq N$, let T_k be a $(1 \times \delta)$-tube around $I_k := I_{e_k}$. Finally, for $j \in \mathcal{I}_{i_0}$, pick some point $x_j \in U_j$, and consider the balls

$$B(x_j, 10\delta) \supset U_j, \qquad j \in \mathcal{I}_{i_0}.$$

For $1 \leq k \leq N$, the union of the sets U_j, $j \in \mathcal{I}_{i_0}$, cover at least a fraction of $\log^{-2}(1/\delta)$ of I_k. It is then easy to see that the balls $B(x_j, 10\delta)$ cover at least a similar fraction of the tube T_k: writing

$$T_k' := T_k \cap \bigcup_{j \in \mathcal{I}_{i_0}} B(x_j, 10\delta),$$

we have

$$\mathcal{L}^2(T_k') \gtrsim \frac{\delta}{\log^2(1/\delta)}. \tag{7.16}$$

Now, we consider the function

$$f := \sum_{k=1}^{N} \chi_{T_k'} \leq \sum_{k=1}^{N} \chi_{T_k}.$$

Recalling (7.9), we have $\|f\|_2^2 \lesssim \log(1/\delta)$, so

$$\mathcal{L}^2\left(\bigcup_{k=1}^{N} T_k'\right) = \mathcal{L}^2(\operatorname{spt} f) \gtrsim \frac{\|f\|_1^2}{\log(1/\delta)} \gtrsim \frac{(N \cdot \mathcal{L}^2(T_k'))^2}{\log(1/\delta)} \gtrsim \frac{1}{\log^9(1/\delta)},$$

combining (7.15)-(7.16) in the last inequality. Since the union $\bigcup T_k'$ is covered by the balls $B(x_j, 10\delta)$, $j \in \mathcal{I}_{i_0}$, each with area $\sim \delta^2$, we infer that

$$|\mathcal{I}_{i_0}| \gtrsim \frac{\delta^{-2}}{\log^9(1/\delta)},$$

and finally, for any $s < 2$, and recalling that $\delta = 2^{-i_0}$, this leads to

$$\sum_{j=1}^{\infty} \operatorname{diam}(U_j)^s \gtrsim |\mathcal{I}_{i_0}| \cdot 2^{-i_0 s} \gtrsim \frac{2^{i_0(2-s)}}{i_0^9} \gtrsim 1.$$

So, we have proven that $\mathcal{H}^s_1(E) \gtrsim_s 1$ for any $s < 2$, and this gives $\dim_{\mathrm{H}} E = 2$. □

7.3. **Kakeya sets in higher dimensions: the hairbrush argument.** As we mentioned in the introduction to the section, Question 8 is open in all dimensions $d \geq 3$. For $d = 3$, the best result at the moment is that Kakeya sets have Hausdorff dimension at least $\frac{5}{2} + \varepsilon$ for some small absolute constant $\varepsilon > 0$. This is a theorem of Katz and Zahl [12] from 2017, and it improves on a previous theorem of Katz, Łaba and Tao [11] from 2000, who showed the same lower bound (perhaps for a different ε) for upper box dimension. The fact that Kakeya sets in $\mathbb{R}^3$ have Hausdorff dimension at least $\frac{5}{2}$ was proven in 1995 by Wolff [21], and we will give a light version of his argument. Namely, we show the idea how to prove the following, but, to avoid too many technicalities, we will unfortunately have to make one very unrealistic simplification during the argument.

Theorem 7.17. *Every Kakeya set in $\mathbb{R}^3$ has lower box dimension at least $\frac{5}{2}$.*

Remark 7.18. Treating $\underline{\dim}_{\mathrm{B}}$ instead of $\dim_{\mathrm{H}}$ is just a matter of convenience: it spares us from the pigeonholing arguments we saw in the previous section. The argument would also work in $\mathbb{R}^d$, for any $d \geq 3$ (modulo the unrealistic simplification), and show that $\dim_{\mathrm{H}} E \geq (d+2)/2$. We present the proof in $\mathbb{R}^3$ for simplicity of notation.

The argument to prove Theorem 7.17 is widely known as Wolff's "hairbrush" argument. If $E \subset \mathbb{R}^3$ were a Kakeya set with low dimension, then one shows that E needs to contain certain configurations, known as "hairbrushes", which are actually quite large. So, assuming (by contradiction) that E is very small gives a fairly large subset $E' \subset E$. Once "fairly large" > "very small", one reaches a contradiction. The way our proof is organised is based on Ben Green's lecture notes [8].

Proof of Theorem 7.17. Assume that $E \subset \mathbb{R}^d$ is a Kakeya set with $\underline{\dim}_{\mathrm{B}} E < \frac{5}{2}$. Then, there exists $\epsilon > 0$ and arbitrarily small $\delta > 0$ such that

$$M := N(E, \delta) \leq \delta^{-5/2+\varepsilon}. \tag{7.19}$$

Fix $\delta > 0$ so that that (7.19) holds, and pick a maximal δ-separated subset $\{e_1, \ldots, e_N\} \subset S^2$. Now $N \sim \delta^{-2}$. For every $1 \leq j \leq N$, pick a unit line segment $I_j \subset E$ parallel to e, and let T_j be a $(\delta \times \delta \times 1)$-tube around I_j. Note that $\mathcal{L}^3(T_j) = \delta^2$. If $B(x_1, \delta/2), \ldots, B(x_M, \delta/2)$ is a collection of balls covering E, then

$$\mathcal{T} := \bigcup_{j=1}^{N} T_j \subset \bigcup_{i=1}^{M} B(x_i, 10\delta).$$

Consequently

$$\mathcal{L}^3(\mathcal{T}) \lesssim M \cdot \delta^3 \leq \delta^{1/2+\varepsilon}, \tag{7.20}$$

according to (7.19). Now, we start looking for the "hairbrush": it will be a union of some (actually quite many) of the tubes T_j with the property that they all intersect a fixed, common, tube T_{j_0}. To find this configuration, we actually show that the "average" tube T_j can be taken as T_{j_0}. We estimate as follows, using (7.20) and the Cauchy-Schwarz inequality:

$$\begin{aligned}\frac{1}{N}\sum_{i=1}^{N}\left[\sum_{j=1}^{N}\mathcal{L}^3(T_i \cap T_j)\right] &= \frac{1}{N}\int_{\mathcal{T}}\left(\sum_{j=1}^{N}\chi_{T_j}\right)^2 d\mathcal{L}^3\\ &\geq \frac{1}{\mathcal{L}^3(\mathcal{T})N}\left(\int_{\mathcal{T}}\sum_{j=1}^{N}\chi_{T_j}\,d\mathcal{L}^3\right)^2\\ &\gtrsim \frac{1}{\mathcal{L}^3(\mathcal{T})N}\left(N\cdot\delta^2\right)^2 \sim \delta^{3/2-\varepsilon},\end{aligned}$$

recalling (7.19) and that $N \sim \delta^{-2}$. Thus, there exists an index $j_0 \in \{1, \ldots, N\}$ such that

$$\sum_{j \neq j_0}\mathcal{L}^3(T_j \cap T_{j_0}) = \sum_{j=1}^{N}\mathcal{L}^3(T_j \cap T_{j_0}) - \mathcal{L}^3(T_{j_0}) \gtrsim \delta^{3/2-\varepsilon} - \delta^2 \sim \delta^{3/2-\varepsilon}. \tag{7.21}$$

Now, we make the unrealistic simplification: all the tubes T_j contributing to the sum above intersect T_{j_0} transversally, that is, the angle between I_j and I_{j_0} is always ~ 1.

With this assumption, we denote by $\mathcal{H}$ the collection of all the tubes T_j intersecting T_{j_0}, with $j \neq j_0$, and we write

$$H := \bigcup_{T_j \in \mathcal{H}} T_j.$$

Now H is the "hairbrush" we were looking for. To reach a contradiction, we want to find a good lower bound for the Lebesgue measure of H, see (7.24) below. To achieve this, we let Z be the line containing the segment I_{j_0}. Translating and rotating, we may assume that Z is the z-axis. So, all the tubes in $\mathcal{H}$ intersect a δ-tube around Z, and consequently every tube in $\mathcal{H}$ is contained in the $C\delta$-neighbourhood of some "vertical" plane containing Z. We now partition the tubes in $\mathcal{H}$ according to these inclusions.

Let $\{\xi_1, \ldots, \xi_m\} \subset S^1 \times \{0\}$ be a maximal δ-separated subset of $S^1 \times \{0\}$, with $m \sim \delta^{-1}$, and let $P_k := \operatorname{span}(\xi_k) \times Z$ be the vertical plane containing ξ_j. Then, as we discussed, every tube in $\mathcal{H}$ is contained in $P_k(C\delta)$ for some $1 \leq k \leq m$. We may hence write $\mathcal{H} := \bigcup_{k=1}^{m} \mathcal{H}_k$, where

$$\mathcal{H}_k := \{T \in \mathcal{H} : T \subset P_k(C\delta)\}.$$

We can assume that the collections $\mathcal{H}_k$ are disjoint, replacing them if necessary by the collections $\mathcal{H}_1' := \mathcal{H}_1$ and $\mathcal{H}_k' := \mathcal{H}_k \setminus [\mathcal{H}_1 \cup \ldots \cup \mathcal{H}_{k-1}]$ for $k \geq 2$. We also write $H_k := \cup\{T : T \in \mathcal{H}_k\}$. Fixing $1 \leq k \leq m$ for the moment, we now draw benefit from the unrealistic assumption: since the tubes in $T \in \mathcal{H}_k$ intersect T transversally, we have the uniform upper bound $\mathcal{L}^3(T \cap T_{j_0}) \lesssim \delta^3$. Consequently, by (7.21), and recalling that the collections $\mathcal{H}_k$ are disjoint, we infer that

$$\delta^3 \sum_{k=1}^{m} |\mathcal{H}_k| \gtrsim \sum_{k=1}^{m} \sum_{T \in \mathcal{H}_k} \mathcal{L}^3(T \cap T_{j_0}) = \sum_{T \in \mathcal{H}} \mathcal{L}^3(T \cap T_{j_0}) \gtrsim \delta^{3/2-\epsilon},$$

and hence

$$\sum_{k=1}^{m} |\mathcal{H}_k| \gtrsim \delta^{-3/2-\varepsilon}. \tag{7.22}$$

To finish the proof, we claim that

$$\mathcal{L}^3\left(H_k \setminus Z(c)\right) \gtrsim |\mathcal{H}_k| \cdot \frac{\delta^2}{\log(1/\delta)}, \qquad 1 \leq k \leq m, \tag{7.23}$$

where $Z(c)$ stands for the c-neighbourhood of the line Z, and $c > 0$ is a small absolute constant, to be specified a little later. The reason we want to exclude the set $Z(c)$ is that the sets $H_k \setminus Z(c)$ have bounded overlap:

$$\sum_{k=1}^{m} \chi_{H_k \setminus Z(c)} \lesssim_c 1.$$

Hence, we will get the following lower bound for $\mathcal{L}^3(H)$ by combining (7.22)-(7.23):

$$\mathcal{L}^3(H) \geq \mathcal{L}^3(H \setminus Z(c)) \gtrsim_c \sum_{k=1}^{m} \mathcal{L}^3(H_k \setminus Z(c)) \gtrsim \frac{\delta^2}{\log(1/\delta)} \sum_{k=1}^{m} |\mathcal{H}_k| \gtrsim \frac{\delta^{1/2-\varepsilon}}{\log(1/\delta)}. \tag{7.24}$$

Recalling that $H \subset \mathcal{T}$, this will immediately contradict (7.20).

So, it remains to prove (7.23), which is essentially Proposition 7.5. That proposition said that if $T_1, \ldots, T_N \subset \mathbb{R}^2$ are $(\delta \times 1)$-tubes with δ-separated directions, then their union has area $\gtrsim N\delta/\log(1/\delta)$. Now, instead, we (after a rotation) have $(\delta \times \delta \times 1)$-tubes $T_1, \ldots, T_N \subset \mathbb{R}^2 \times [-C\delta, C\delta]$ with δ-separated directions. The proof of Proposition 7.5 can be used to show that the union has volume $\gtrsim N\delta^2/\log(1/\delta)$.

The only technical annoyance is the exclusion of the set $Z(c)$: however, $Z(c)$ a tube of width $c > 0$, and the tubes $T \in \mathcal{H}_k$ intersect it transversally. So, if $c > 0$ is small enough, at least half of the volume of every tube $T \in \mathcal{H}_k$ lies in $P_k(C\delta) \setminus Z(c)$, and then one can prove (7.23) using these halves. □

References

[1] A. S. Besicovitch: *On Kakeya's problem and a similar one*, Math. Z. **27** (1928), 312-320

[2] A. S. Besicovitch: *On the fundamental geometrical properties of linearly measurable plane sets of points III*, Math. Ann. **116** (1939), 296-329

[3] J. Bourgain: *The discretized sum-product and projections theorems*, J. Anal. Math. **112** (2010), 193-236

[4] A. Córdoba: *The Kakeya maximal function and the spherical summation of multipliers*, Amer. J. Math. **99** (1977), 1-22

[5] R. O. Davies: *Some remarks on the Kakeya problem*, Proc. Cambridge Philos. Soc. **69** (1971), 417-421

[6] K. Falconer: *The geometry of fractal sets*, Cambridge Tracts in Mathematics **85**, Cambridge University Press 1985

[7] O. Frostman: *Potential d'equilibre et capacité des ensembles avec quelques applications à la théorie des fonctions*, Meddel. Lunds Univ. Mat. Sen. **3** (1935), 1-118

[8] B. Green: *Restriction and Kakeya Phenomena*, notes available at the author's website

[9] L. Guth and N. Katz: *Algebraic methods in discrete analogs of the Kakeya problem*, Adv. Math. **225**(5) (2010), 2828-2839

[10] L. GUTH AND N. KATZ: *On the Erdős distinct distances problem in the plane*, Ann. of Math. **181**(1) (2015), 155-190

[11] N. KATZ, I. ŁABA AND T. TAO: *An improved bound on the Minkowski dimension of Besicovitch sets in* $\mathbb{R}^3$, Ann. of Math. **152** (2000), 383-446

[12] N. KATZ AND J. ZAHL: *An improved bound on the Hausdorff dimension of Besicovitch sets in* $\mathbb{R}^3$, to appear in J. Amer. Math. Soc.

[13] R. KAUFMAN: *On Hausdorff dimension of projections*, Mathematika **15** (1968), 153-155

[14] J. M. MARSTRAND: *Some fundamental geometrical properties of planes sets of fractional dimensions*, Proc. London Math. Soc. (3) **4** (1954), 247-302

[15] P. MATTILA: *Spherical averages of Fourier transforms of measures with finite energy: dimension of intersections and distance sets*, Mathematika **34** (1987), 207-228

[16] P. MATTILA: *Geometry of Sets and Measures on Euclidean Spaces: Fractals and Rectifiability*, Cambridge University Press (1995)

[17] P. MATTILA AND P. SJÖLIN: *Regularity of distance measures and sets*, Math. Nachr. **204** (1999), 157-162

[18] P. MATTILA: *Fourier Analysis and Hausdorff Dimension*, Cambridge Studies in Advanced Mathematics **150**, Cambridge University Press (2015)

[19] W. RUDIN: *Real and complex analysis*, MgGraw-Hill 1986

[20] E. SZEMERÉDI AND W. TROTTER: *Extremal problems in discrete geometry*, Combinatorica **3** (1983), 381-392

[21] T. WOLFF: *An improved bound for Kakeya type maximal functions*, Revista Mat. Iberoamericana **11**(3) (1995), 651-674

[22] T. WOLFF: *Decay of circular means of Fourier transforms of measures*, Int. Math. Res. Notices **1999** (10) (1999), 547-567

UNIVERSITY OF HELSINKI, DEPT. OF MATHEMATICS AND STATISTICS
E-mail address: tuomas.orponen@helsinki.fi