

**THE KAKEYA CONJECTURE IN  $\mathbb{R}^3$**   
**[after Hong Wang and Joshua Zahl]**

by **Larry Guth**

**Contents**

|  |    |
|--|----|
| 1. Statement of main results.....                              | 2  |
| 2. The hero: multiscale analysis.....                          | 4  |
| 3. A key obstacle: the Heisenberg group.....                   | 5  |
| 4. Key structures and outline of the proof.....                | 6  |
| 5. Avoiding technicalities.....                                | 9  |
| 6. Stickiness leads to grain structure.....                    | 10 |
| 7. Grain structure leads to complex conjugation structure..... | 16 |
| 8. Complex conjugation structure leads to a contradiction..... | 20 |
| 9. Leveraging the sum-product theorem at many scales.....      | 23 |
| 10. Sticky vs. not sticky.....                                 | 23 |
| 11. The $L^2$ method.....                                      | 25 |
| 12. The worst case is sticky.....                              | 26 |
| 13. Other possible shapes of $K$ .....                         | 31 |
| 14. The high density lemma.....                                | 34 |
| 15. Final recap.....   | 39 |
| References.....  | 41 |

This is a survey article about the proof of the Kakeya conjecture in three dimensions. The Kakeya conjecture is a problem about the intersection patterns of thin tubes in Euclidean space.

A Kakeya set in  $\mathbb{R}^n$  is a set that contains a unit line segment in every direction. Around 1920, Besicovitch gave an example of a Kakeya set in  $\mathbb{R}^2$  with arbitrarily small Lebesgue measure. Around 1970, Fefferman gave a counterexample to a well-known problem about Fourier multipliers which crucially used Besicovitch’s example. The same idea shows that a number of open problems in Fourier analysis are connected to Kakeya sets. These problems in Fourier analysis connect to quantitative questions about Kakeya sets, such as, “What is the infimal Hausdorff dimension of a Kakeya set in  $\mathbb{R}^n$ ?” For example, the Stein restriction conjecture in Fourier analysis implies that

every Kakeya set in  $\mathbb{R}^n$  has Hausdorff dimension  $n$ . The connection between Fourier analysis and Kakeya problems is described in the survey article Tao (2001).

The Kakeya conjecture for Hausdorff dimension says that every Kakeya set in  $\mathbb{R}^n$  has Hausdorff dimension  $n$ . In the 1980s, Davies proved that every Kakeya set in  $\mathbb{R}^2$  has Hausdorff dimension 2, and the proof is only a couple pages. But proving the conjecture for any  $n \geq 3$  is much more difficult. Wang and Zahl (2025) proved the Kakeya conjecture in dimension 3. Their proof builds on important ideas by many people, including Bourgain, Wolff, Katz, Łaba, Tao, Orponen, and Shmerkin. In dimension  $n \geq 4$ , the conjecture is currently open.

In this survey, we will not use the language of Hausdorff dimension. For understanding the proof, and also for applications in Fourier analysis, the most useful language is in terms of sets of thin tubes. We set up this language and state the main results in the next section.

## 1. Statement of main results

Suppose that  $\mathbb{T}$  is a set of  $\delta$ -tubes in  $\mathbb{R}^n$  with length 1. We write  $U(\mathbb{T})$  for  $\cup_{T \in \mathbb{T}} T$ . One version of the Kakeya conjecture in dimension  $n$  says

**CONJECTURE 1.1.** — *For every  $\epsilon > 0$ , there is a constant  $c(n, \epsilon)$  so that if  $\mathbb{T}$  is a set of  $\sim \delta^{-(n-1)}$   $\delta$ -tubes in  $\mathbb{R}^n$  in  $\delta$ -separated directions, then*

$$|U(\mathbb{T})| \geq c(n, \epsilon) \delta^\epsilon.$$

Wang and Zahl proved this conjecture in dimension  $n = 3$ .

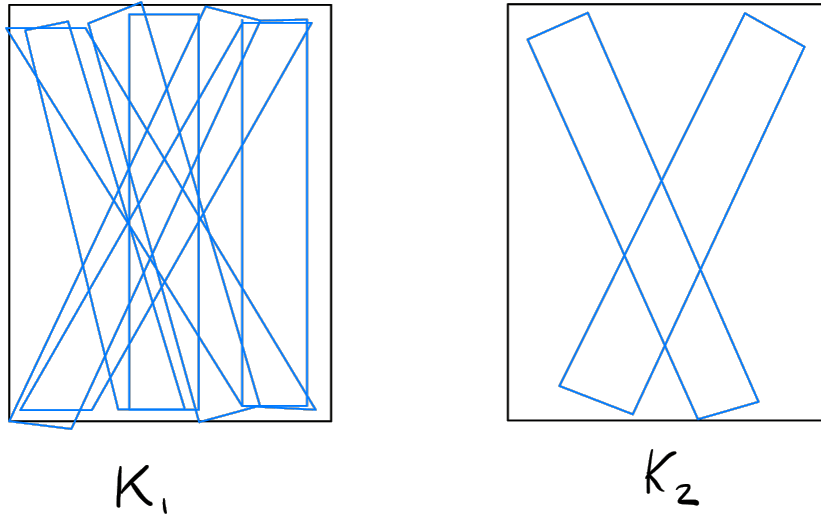
In fact, they proved a more general estimate called the Kakeya conjecture with convex Wolff axioms. This theorem roughly says that the only way a set of tubes in  $\mathbb{R}^3$  can overlap a lot is by clustering into convex sets. If  $K \subset \mathbb{R}^3$  is a convex set, we define

$$\mathbb{T}[K] := \{T \in \mathbb{T} : T \subset K\}.$$

We define the density of  $\mathbb{T}$  in  $K$  as

$$\Delta(\mathbb{T}, K) = \frac{\sum_{T \in \mathbb{T}[K]} |T|}{|K|}.$$

The density of  $\mathbb{T}$  in  $K$  measures how much the tubes of  $\mathbb{T}$  pack into  $K$ . Here is a picture to help illustrate the definition.



In this picture,  $\Delta(\mathbb{T}, K_1) > 1$  and  $\Delta(\mathbb{T}, K_2) < 1$ . Next we consider the maximum density over all convex sets  $K$ .

$$\Delta_{max}(\mathbb{T}) := \max_{K \text{ convex}} \Delta(\mathbb{T}, K).$$

Let us define the typical multiplicity of  $\mathbb{T}$  as

$$\mu(\mathbb{T}) = \frac{\sum_{T \in \mathbb{T}} |T|}{|U(\mathbb{T})|}.$$

On average, a point  $x \in U(\mathbb{T})$  lies in  $\mu(\mathbb{T})$  tubes  $T \in \mathbb{T}$ . Notice that  $\mu(\mathbb{T}[K]) \geq \Delta(\mathbb{T}, K)$ , and so there must be a point  $x$  that lies in at least  $\Delta_{max}(\mathbb{T})$  tubes of  $\mathbb{T}$ .

The main theorem of Wang and Zahl (2025) says that  $\mu(\mathbb{T})$  can only be large when  $\Delta_{max}(\mathbb{T})$  is large.

**THEOREM 1.2** (Wang and Zahl, 2025). — *If  $\mathbb{T}$  is a set of  $\delta$ -tubes in  $\mathbb{R}^3$ , and  $\Delta_{max}(\mathbb{T}) \lesssim 1$ , then*

$$\mu(\mathbb{T}) \lesssim 1.$$

This theorem is called the convex Wolff axioms version of the Kakeya conjecture. It directly implies Conjecture 1.1 for  $n = 3$ . (Wang and Zahl also proved a somewhat more general theorem which implies that a Kakeya set in  $\mathbb{R}^3$  has Hausdorff dimension 3.)

The proof of Theorem 1.2 was completed in Wang and Zahl (2025), but the full proof includes many papers with important contributions by Wolff, Bourgain, Katz, Łaba, Tao, Orponen, and Shmerkin. The goal of this survey is to describe the main ideas of the whole proof.

### 1.1. Notation

Informally, we write  $A \approx B$  to mean that  $A$  and  $B$  are approximately the same size. We write  $A \lesssim B$  to mean that either  $A < B$  or  $A \approx B$ . And we write  $A \ll B$  to mean that  $A$  is much smaller than  $B$ . For more detailed statements and outlines, you can look at Guth (2025) and Guth, Wang, and Zahl (2026).

## 2. The hero: multiscale analysis

The hero of our story is looking at the problem at many different scales. In this section, we explain what this means and give a hint about why it will be important.

Define  $\beta$  to be the infimal exponent so that for every set  $\mathbb{T}$  of  $\delta$ -tubes in  $\mathbb{R}^3$  with  $\Delta_{max}(\mathbb{T}) \lesssim 1$ ,

$$(1) \quad \mu(\mathbb{T}) \lesssim |\mathbb{T}|^\beta.$$

The Kakeya theorem, Theorem 1.2, says that  $\beta = 0$ . We say that  $\mathbb{T}$  is a worst-case Kakeya set if  $\Delta_{max}(\mathbb{T}) \lesssim 1$  and  $\mu(\mathbb{T}) \approx |\mathbb{T}|^\beta$ . The proof will be by contradiction. We suppose  $\beta > 0$ . We let  $\mathbb{T}$  be a worst-case Kakeya set. We will prove that  $\mathbb{T}$  would have to have a lot of geometric and algebraic structure. Using this structure we will eventually get a contradiction.

To exploit the fact that  $\mathbb{T}$  is a worst-case Kakeya set, we will compare  $\mathbb{T}$  with other sets of tubes  $\mathbb{T}'$ . We will find other sets of tubes  $\mathbb{T}'$  which are related to  $\mathbb{T}$  and obey  $\Delta_{max}(\mathbb{T}') \lesssim 1$ . By the definition of  $\beta$  in (1), we know that  $\mu(\mathbb{T}') \lesssim |\mathbb{T}'|^\beta$ . Since  $\mathbb{T}'$  is related to  $\mathbb{T}$ , this bound leads to information about  $\mathbb{T}$ . We will find these sets of tubes  $\mathbb{T}'$  by looking at the original set of tubes at multiple scales.

Suppose that  $\mathbb{T}$  is a set of  $\delta$ -tubes. Given a scale  $\rho \in [\delta, 1]$ , we let  $\mathbb{T}_\rho$  denote the set of  $\rho$ -tubes formed by thickening the  $\delta$ -tubes of  $\mathbb{T}$ . When we thicken two distinct  $\delta$ -tubes, we may get nearly identical  $\rho$ -tubes. When this happens we identify the  $\rho$ -tubes. So we typically have  $|\mathbb{T}_\rho| \ll |\mathbb{T}|$ .

For each  $T_\rho \in \mathbb{T}_\rho$ , we define

$$\mathbb{T}[T_\rho] = \{T \in \mathbb{T} : T \subset T_\rho\}.$$

Figure 1 illustrates these different sets of tubes. The tubes of  $\mathbb{T}$  are the thin tubes, and the tubes of  $\mathbb{T}_\rho$  are the thick tubes. In the picture, each set  $\mathbb{T}[T_\rho]$  consists of 3  $\delta$ -tubes.

Our original set of tubes  $\mathbb{T}$  is the disjoint union of  $\mathbb{T}[T_\rho]$ :

$$(2) \quad \mathbb{T} = \bigsqcup_{T_\rho \in \mathbb{T}_\rho} \mathbb{T}[T_\rho].$$

After some pigeonholing arguments, we can assume that  $|\mathbb{T}[T_\rho]|$  is roughly constant for all  $T_\rho \in \mathbb{T}_\rho$ . So for any  $T_\rho \in \mathbb{T}_\rho$ , we get

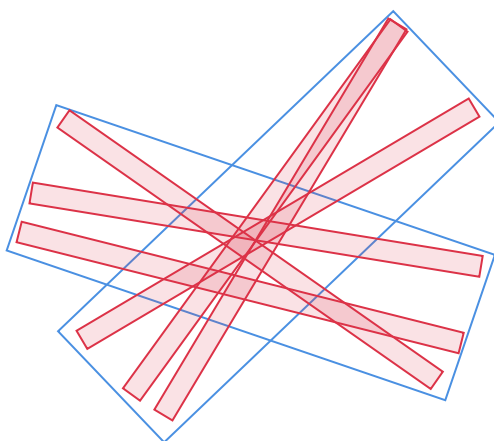


FIGURE 1. Intersecting tubes

$$(3) \quad |\mathbb{T}| \approx |\mathbb{T}[T_\rho]| |\mathbb{T}_\rho|.$$

We will get a lot of information about  $\mathbb{T}$  by applying (1) to  $\mathbb{T}_\rho$  and  $\mathbb{T}[T_\rho]$  for many different scales  $\rho$ . Over the course of the proof, we will also find other sets of tubes  $\mathbb{T}'$  related to  $\mathbb{T}$  and apply (1) to them.

I believe that Tom Wolff was the first person to think about this multiscale structure in the Kakeya problem, in unpublished work shortly before his death. He shared his ideas with Katz, Łaba, and Tao, who developed them further in the remarkable paper (Katz, Łaba, and Tao, 2000), and then the ideas were developed further by many people.

### 3. A key obstacle: the Heisenberg group

Next we introduce one of the key obstacles to proving the main theorem. There are cousin problems that sound quite similar to the Kakeya conjecture but behave differently. For example, the natural analogue of Theorem 1.2 in  $\mathbb{C}^3$  is false. The counterexample is called the Heisenberg group example. It was first published by Katz, Łaba, and Tao (2000). The idea that problems of this type may behave differently over different fields was first noted by Tom Wolff (1999).

We define a metric on  $\mathbb{C}^n$  by identifying it with  $\mathbb{R}^{2n}$ . We define a complex tube in  $\mathbb{C}^n$  with radius  $r$  and length  $L$  by taking a complex line in  $\mathbb{C}^n$ , intersecting it with a ball of radius  $L$ , and then taking its  $r$ -neighborhood. We define a complex  $\delta$ -tube to be a tube of radius  $\delta$  and length 1. The quantities  $\Delta(\mathbb{T}, K)$  and  $\Delta_{max}(\mathbb{T})$  can be defined roughly as above.

In  $\mathbb{C}^3$ , there is a set of complex  $\delta$ -tubes  $\mathbb{T}$  with  $\Delta_{max}(\mathbb{T}) \lesssim 1$  and  $\mu(\mathbb{T}) \approx |\mathbb{T}|^{1/4}$ . This example shows that the complex analogue of Theorem 1.2 is false. This example is called

the Heisenberg group example. It is based on a quadratic real algebraic hypersurface in  $\mathbb{C}^3$ . There are a couple choices for this hypersurface. One is the surface  $H$  defined by

$$H = \{(z_1, z_2, z_3 : |z_1|^2 + |z_2|^2 - |z_3|^2 = 1\}$$

This hypersurface contains many complex lines. For example, if  $\alpha$  is a unit complex number, then the line defined by  $z_1 = 1$ ,  $z_3 = \alpha z_2$  lies in  $H$ . All these lines pass through the point  $(1, 0, 0) \in H$ . The surface  $H$  is very symmetric: it is symmetric under the action of the group  $U(2, 1)$ , which acts transitively on  $H$ . By symmetry, there are infinitely many lines through every point of  $H$ . Taking  $\delta$ -neighborhoods of these lines gives a set of  $\delta$ -tubes  $\mathbb{T}$  where  $\Delta_{max}(\mathbb{T}) \lesssim 1$  and yet  $\mu(\mathbb{T}) \approx \delta^{-1} \approx |\mathbb{T}|^{1/4}$ .

Wolff (1995) proved that if  $\mathbb{T}$  is a set of  $\delta$ -tubes in  $\mathbb{R}^3$  with  $\Delta_{max}(\mathbb{T}) \lesssim 1$ , then  $\mu(\mathbb{T}) \lesssim |\mathbb{T}|^{1/4}$ . It has been very difficult to improve on the exponent  $1/4$ . Katz, Łaba, and Tao (2000) improved  $1/4$  to  $1/4 - \epsilon$  for some tiny  $\epsilon > 0$ , under an additional technical assumption. This technical assumption was removed by Katz and Zahl (2019). The exponent  $1/4 - \epsilon$  was the best known exponent before the recent work of Wang and Zahl, giving the sharp exponent.

Understanding the structure of the Heisenberg group example is essential to the proof of the Kakeya conjecture. Roughly speaking, the proof shows that if the Kakeya conjecture was false, a worst-case Kakeya set would need to have structural properties similar to those of the Heisenberg group. Finally, these strong structural properties lead to a contradiction. In the next section, we discuss these key structures.

## 4. Key structures and outline of the proof

In this section we introduce the key structures of the Heisenberg group example which guide the proof of the Kakeya conjecture.

### 4.1. Grain structure

Write  $N_w(X)$  for the  $w$ -neighborhood of  $X$ :

$$N_w(X) = \{x \text{ so that } \text{dist}(x, X) < w\}.$$

Recall that  $H$  is a smooth real 5-manifold in  $\mathbb{R}^6$ . Therefore, if  $p \in H$ ,  $N_\delta H \cap B_{\sqrt{\delta}}(p)$  is essentially the  $\delta$ -neighborhood of a 5-dimensional disk. We describe the situation as follows:

**Grain structure.** For each  $p \in H$ , we can choose unitary coordinates  $w_1, w_2, w_3$  on  $B_{\sqrt{\delta}}(p)$  so that  $N_\delta H \cap B_{\sqrt{\delta}}(p) = B^2(\sqrt{\delta}) \times A$ , where

- $B^2(\sqrt{\delta})$  is the ball of radius  $\sqrt{\delta}$  in  $\mathbb{C}^2$
- $A$  is the  $\delta$ -neighborhood of  $\mathbb{R}$  in  $B^1(\sqrt{\delta}) \subset \mathbb{C}$ .

We will see that if  $\beta > 0$ , then a worst-case Kakeya set in  $\mathbb{R}^3$  would have a similar grain structure: for a typical ball  $B = B_{\sqrt{\delta}} \subset U(\mathbb{T}_{\sqrt{\delta}}$ , we can choose coordinates so that  $U(\mathbb{T}) \cap B = [0, \sqrt{\delta}]^2 \times A$ , where  $A \subset [0, \sqrt{\delta}]$  is a union of  $\delta$  intervals.

Figure 2 illustrates the situation.

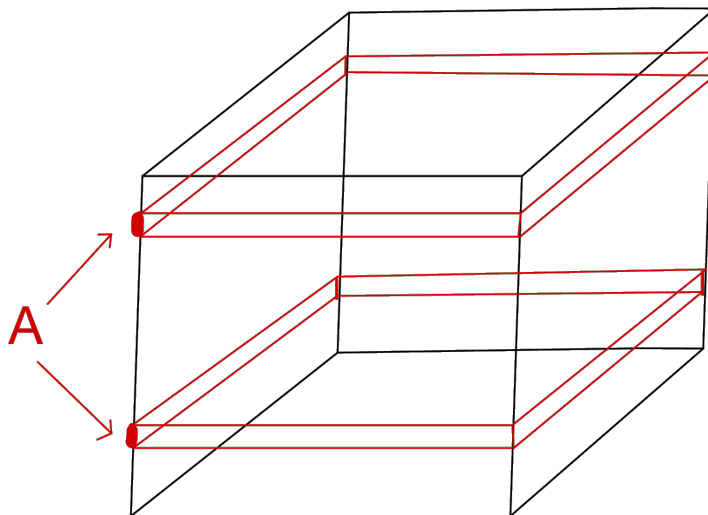


FIGURE 2. Grain structure

The parallel slabs in the picture are parallel to the  $(x_1, x_2)$ -plane and have thickness  $\delta$ . The heights of these slabs correspond to the set  $A$ .

The word grain is supposed to evoke the grains in a piece of wood. Each slab in the picture is called a grain. We will call  $B_{\sqrt{\delta}}(p)$  a grain box. The grains in a grain box are all parallel to each other.

Each point  $z \in H$  lies in a grain, which we call the grain through  $z$ .

#### 4.2. Complex conjugation structure

The definition of  $H$  involves complex conjugation, and the structure of  $H$  is closely related to complex conjugation. One connection between complex conjugation and the geometry of  $H$  comes from considering the slopes of grains along a line.

Fix a line  $\ell \subset H$ . We choose coordinates so that the line  $\ell$  is the  $z_1$  axis. Each point  $z \in \ell$  lies in a grain, which is a complex 2-plane containing  $\ell$ . Such a 2-plane is given by an equation  $z_3 = sz_2$ , where  $s \in \mathbb{C}$ . Suppose the grain of  $H$  through  $z = (z_1, 0, 0) \in \ell$  is given by  $z_3 = s(z_1)z_2$ . So  $s(z_1)$  describes the slope of the grain through  $(z_1, 0, 0) \in \ell$ . For an appropriate choice of coordinates  $z_1, z_2, z_3$ , the slope function is given by:

$$(4) \quad s(z_1) = \bar{z}_1$$

We refer to this equation as the complex conjugation structure of  $H$ .

We can also define a slope function  $s(x)$  related to a worst-case Kakeya set in  $\mathbb{R}^3$ . First we will prove that a worst-case Kakeya set in  $\mathbb{R}^3$  has a grain structure as above.

Suppose that  $\ell$  is the core line of a tube  $T \in \mathbb{T}$ . We choose coordinates so that the line  $\ell$  is the  $x_1$ -axis. Each point  $(x_1, 0, 0) \in \ell$  lies in a grain, which is a plane given by  $x_3 = s(x_1)x_2$ . So  $s(x_1)$  is the slope of the grain through  $(x_1, 0, 0)$ . We will prove that in some sense the function  $s(x_1)$  “is similar to complex conjugation”.

To get a feeling for what this might mean, let’s return to the Heisenberg group example. In the Heisenberg group example,  $A$  is essentially  $\mathbb{R} \subset \mathbb{C}$  and the slope function is  $s(z_1) = \bar{z}_1$ . The set  $A$  and the slope function  $s(z_1)$  interact in a nice way. For instance, for any  $z_1 \in \mathbb{C}$ ,

$$(5) \quad A + s(z_1)z_1 = \mathbb{R} + |z_1|^2 = \mathbb{R} = A$$

For a worst-case Kakeya set, we will show that the set  $A$  and the slope function  $s$  also interact in a nice way in a similar spirit to (5). We postpone the precise statement to Section 7. It is a little more complicated than (5), involving two different slope functions along two different lines.

### 4.3. Stickiness

If  $\mathbb{T}$  is the Heisenberg group example at scale  $\delta$ , then the thicker tubes  $\mathbb{T}_\rho$  are the Heisenberg example at scale  $\rho$ . This leads to some nice numerology about  $|\mathbb{T}_\rho|$ . Let  $|T_\rho^{\mathbb{C}}|$  denote the volume of a complex  $\rho$ -tube in  $\mathbb{C}^3$ . Then for the set  $\mathbb{T}$  of complex tubes from the Heisenberg group, we have  $|\mathbb{T}| \sim |T_\delta^{\mathbb{C}}|^{-1}$  and for each  $\rho \in [\delta, 1]$ ,  $|\mathbb{T}_\rho| \sim |T_\rho^{\mathbb{C}}|^{-1}$ . This numerology is called the ‘sticky’ case, for reasons that we explain below.

**Stickiness.** If  $\mathbb{T}$  is a set of  $\delta$ -tubes in  $\mathbb{R}^n$ , then  $\mathbb{T}$  is sticky if  $|\mathbb{T}_\rho| \approx |T_\rho|^{-1}$  for each  $\rho \in [\delta, 1]$ .

In  $\mathbb{R}^n$ , a  $\rho$ -tube has volume  $|T_\rho| \sim \rho^{n-1}$ , and so  $\mathbb{T}$  is sticky if

$$(6) \quad |\mathbb{T}_\rho| \approx \rho^{-(n-1)} \text{ for all } \rho \in [\delta, 1]$$

Recall from (3) that  $|\mathbb{T}| \approx |\mathbb{T}_\rho| |\mathbb{T}[T_\rho]|$ , and so if  $\mathbb{T}$  is sticky, then

$$(7) \quad |\mathbb{T}[T_\rho]| \approx (\delta/\rho)^{-(n-1)} \text{ for all } \rho \in [\delta, 1], T_\rho \in \mathbb{T}_\rho$$

For comparison, if  $\mathbb{T}$  obeys  $\Delta_{max}(\mathbb{T}) \lesssim 1$ , then we have  $|\mathbb{T}[T_\rho]| \lesssim (\delta/\rho)^{-(n-1)}$ . So a set of tubes  $\mathbb{T}$  with  $\Delta_{max}(\mathbb{T}) \lesssim 1$  is sticky if  $|\mathbb{T}[T_\rho]|$  is as large as possible. The name sticky comes from the following image. If two tubes  $T_1, T_2$  lie in the same fatter tube  $T_\rho$ , then they are “stuck together”. The tubes in a sticky Kakeya set stick together as much as possible, given the condition  $\Delta_{max}(\mathbb{T}) \lesssim 1$ .

To summarize, the Heisenberg group example has three important structures: stickiness, grain structure, and complex conjugation structure.

#### 4.4. Outline of the proof

Now we can outline the proof of the main theorem. Recall that  $\beta$  is the infimal exponent so that for every set  $\mathbb{T}$  of  $\delta$ -tubes in  $\mathbb{R}^3$  with  $\Delta_{max}(\mathbb{T}) \lesssim 1$ ,

$$(8) \quad \mu(\mathbb{T}) \lesssim |\mathbb{T}|^\beta.$$

Our goal is to prove that  $\beta = 0$ . We say that  $\mathbb{T}$  is a worst-case Kakeya set if  $\Delta_{max}(\mathbb{T}) \lesssim 1$  and  $\mu(\mathbb{T}) \approx |\mathbb{T}|^\beta$ . The proof is by contradiction. We suppose  $\beta > 0$ . We let  $\mathbb{T}$  be a worst-case Kakeya set.

- Step 1. Since  $\mathbb{T}$  is worst-case it must be sticky.
- Step 2. Then  $\mathbb{T}$  must have grain structure.
- Step 3. Then  $\mathbb{T}$  must have something like complex conjugation structure.
- Step 4. No such structure exists. (There is no operation on  $\mathbb{R}$  which has properties similar to complex conjugation.)

This is the logical order of the proof but it is not the chronological order. In particular, Step 1 was the last step to be understood.

We will explain the proof in chronological order. Here is an outline.

In the first big part of the paper, we describe the proof of the sticky case of the main theorem. This proof was outlined by Katz and Tao in unpublished work in the 2000s, and shared in a blog post (Tao, 2014). It has several steps.

- In Section 6, we show that stickiness leads to grain structure. (This step is due to Katz, Łaba, and Tao, 2000.)
- In Section 7, we show that grain structure leads to complex conjugation structure. (This step is based on unpublished work of Katz–Tao. The details were carried out in Wang and Zahl, 2026).
- In Section 8, we show that complex conjugation structure leads to a contradiction. (This step is based on work of many people, including Bourgain, Katz, Tao, Orponen, Shmerkin, Wang, Zahl)

After we describe the proof of the sticky case we pause to digest and reflect.

After that, in Sections 12 and 13 and 14, we describe the proof that the worst-case is sticky. (This step is based on work of Wang and Zahl, 2025).

#### 5. Avoiding technicalities

To avoid technical details, we will assume that various quantities are uniform.

For instance, for each  $T_\rho \in \mathbb{T}_\rho$ , we will study the set of tubes  $\mathbb{T}[T_\rho]$ . In general, for different tubes  $T_\rho \in \mathbb{T}_\rho$ ,  $|\mathbb{T}[T_\rho]|$  could be very different. But we will assume that all the sets  $\mathbb{T}[T_\rho]$  have roughly the same cardinality. Similarly we will assume that  $\mu(\mathbb{T}[T_\rho])$  is roughly the same for all  $T_\rho \in \mathbb{T}_\rho$ .

For each point  $x \in U(\mathbb{T})$ , we let  $\mathbb{T}_x := \{T \in \mathbb{T} : x \in T\}$ . In general, for different points  $x \in U(\mathbb{T})$ ,  $|\mathbb{T}_x|$  could be very different. But we will assume that  $|\mathbb{T}_x|$  is roughly the same for all  $x \in U(\mathbb{T})$ . Recall that we defined  $\mu(\mathbb{T}) = \frac{\sum_{T \in \mathbb{T}} |T|}{|U(\mathbb{T})|}$ , which we can interpret as the average size of  $|\mathbb{T}_x|$  over  $x \in U(\mathbb{T})$ . Since we assume that  $|\mathbb{T}_x|$  is roughly constant, we have

$$(9) \quad |\mathbb{T}_x| \approx \mu(\mathbb{T}) \text{ for every } x \in U(\mathbb{T})$$

In this survey, we sketch the proof of the *Keakeya theorem* assuming some uniformity of this kind. The full proof has the same main ideas but there are extra technical details to deal with non-uniformity. For instance, if  $|\mathbb{T}_x|$  is very different at different  $x \in U(\mathbb{T})$ , then we subdivide  $U(\mathbb{T})$  into subsets where  $|\mathbb{T}_x|$  has different sizes. Then we have to keep track of all these subsets. This process is called *pigeonholing*. There is non-trivial technical work involved, but we will not discuss it in this survey of the high level ideas.

When we state lemmas in this survey, the precise statements require some *pigeonholing* and/or *uniformity* hypotheses. To keep the statements simple, we leave out these details. Unfortunately, it means that the statements of the lemmas here are not completely precise. But this choice does keep the statements of lemmas simpler. We hope that it conveys the general strategy to a broader audience.

## 6. Stickiness leads to grain structure

In this section, we will consider a worst-case sticky *Keakeya set* and show that it must have *grain structure*.

We say that  $\mathbb{T}$  is a *sticky Keakeya set* of  $\delta$ -tubes in  $\mathbb{R}^3$  if:

- $\Delta_{max}(\mathbb{T}) \lesssim 1$ .
- For each  $\rho \in [\delta, 1]$ ,  $|\mathbb{T}_\rho| \approx \rho^{-2}$  and for each  $T_\rho \in \mathbb{T}_\rho$ ,  $|\mathbb{T}[T_\rho]| \approx (\rho/\delta)^2$ .

We define  $\beta_{sticky}$  to be the infimal exponent  $\beta$  so that for every sticky *Keakeya set* of tubes

$$(10) \quad \mu(\mathbb{T}) \lesssim |\mathbb{T}|^\beta.$$

**THEOREM 6.1** (*Sticky Keakeya theorem*, Wang and Zahl, 2026)

$\beta_{sticky} = 0$ . In other words, for every sticky *Keakeya set*  $\mathbb{T}$ ,  $\mu(\mathbb{T}) \lesssim 1$ .

The *sticky Keakeya theorem* is the first big part of the proof of the *Keakeya theorem*.

We say that  $\mathbb{T}$  is a *worst-case sticky Keakeya set* if  $\mathbb{T}$  is a *sticky Keakeya set* and  $\mu(\mathbb{T}) \approx |\mathbb{T}|^{\beta_{sticky}}$ . The proof of the *sticky Keakeya theorem* goes by contradiction. We suppose that  $\beta_{sticky} > 0$  and we let  $\mathbb{T}$  be a *worst-case sticky Keakeya set*. By examining  $\mathbb{T}$  at different scales, we will see that it must have a great deal of structure.

### 6.1. Perfect overlap

If  $\mathbb{T}$  is a sticky Kakeya set, then it follows that for each  $\rho \in [\delta, 1]$ ,  $\mathbb{T}_\rho$  and  $\mathbb{T}[T_\rho]$  are also sticky Kakeya sets. This fact makes sticky Kakeya sets well suited for multiscale induction arguments.

By the definition of  $\beta_{\text{sticky}}$  (or by induction on scales), we know that

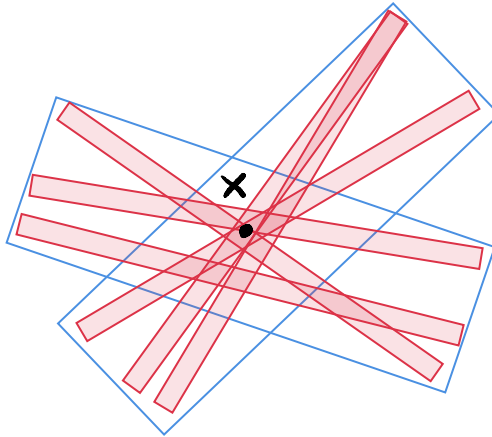
$$(11) \quad \mu(\mathbb{T}_\rho) \lesssim |\mathbb{T}_\rho|^{\beta_{\text{sticky}}}.$$

$$(12) \quad \mu(\mathbb{T}[T_\rho]) \lesssim |\mathbb{T}[T_\rho]|^{\beta_{\text{sticky}}}.$$

Now we state a fundamental lemma relating  $\mu(\mathbb{T})$  with  $\mu(\mathbb{T}_\rho)$  and  $\mu(\mathbb{T}[T_\rho])$ .

LEMMA 6.2. —  $\mu(\mathbb{T}) \lesssim \mu(\mathbb{T}_\rho)\mu(\mathbb{T}[T_\rho])$ .

The following picture illustrates the proof:



*Proof of Lemma 6.2.* — Consider a point  $x \in U(\mathbb{T})$ . The point  $x$  belongs to  $T_\rho$  for  $\approx \mu(\mathbb{T}_\rho)$  fat tubes  $T_\rho \in \mathbb{T}_\rho$ . For each of these  $T_\rho$ , the point belongs to at most  $\mu(\mathbb{T}[T_\rho])$  thin tubes  $T \in \mathbb{T}[T_\rho]$ . So all together,  $x$  belongs to  $\lesssim \mu(\mathbb{T}_\rho)\mu(\mathbb{T}[T_\rho])$  tubes  $T \in \mathbb{T}$ .  $\square$

If  $\mathbb{T}$  is a worst-case sticky Kakeya set, then  $|\mathbb{T}|^{\beta_{\text{sticky}}} \approx \mu(\mathbb{T})$ . Now combining Lemma 6.2 with our bounds for  $\mu(\mathbb{T}_\rho)$  and  $\mu(\mathbb{T}[T_\rho])$  we see that

$$(13) \quad |\mathbb{T}|^{\beta_{\text{sticky}}} \approx \mu(\mathbb{T}) \lesssim \mu(\mathbb{T}_\rho)\mu(\mathbb{T}[T_\rho]) \lesssim |\mathbb{T}_\rho|^{\beta_{\text{sticky}}} |\mathbb{T}[T_\rho]|^{\beta_{\text{sticky}}} \approx |\mathbb{T}|^{\beta_{\text{sticky}}}.$$

Therefore, all the inequalities in the above string must be roughly equalities. This has two important consequences. Whenever  $\mathbb{T}$  is a worst-case sticky Kakeya set, we see that

1.  $\mu(\mathbb{T}_\rho) \approx |\mathbb{T}_\rho|^{\beta_{\text{sticky}}}$  and  $\mu(\mathbb{T}[T_\rho]) \approx |\mathbb{T}[T_\rho]|^{\beta_{\text{sticky}}}$ . Therefore, both  $\mathbb{T}_\rho$  and  $\mathbb{T}[T_\rho]$  are worst-case sticky Kakeya sets.
2.  $\mu(\mathbb{T}) \approx \mu(\mathbb{T}_\rho)\mu(\mathbb{T}[T_\rho])$ , so we have equality in Lemma 6.2.

To digest this second fact, we return to the picture illustrating Lemma 6.2 and add a couple more points.

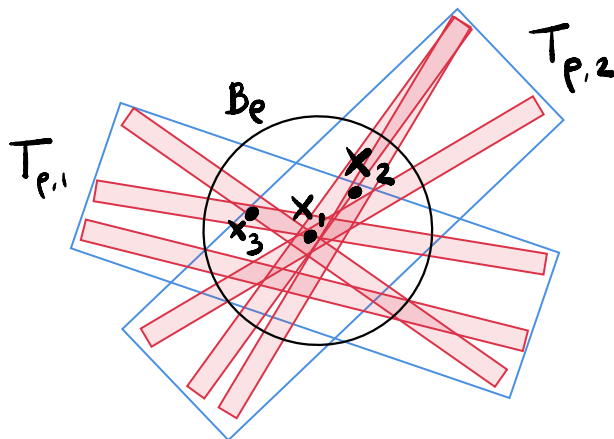


FIGURE 3. When is Lemma 6.2 sharp?

Recall that  $\mathbb{T}_x = \{T \in \mathbb{T} : x \in T\}$ . We also define

$$\mathbb{T}_{\rho, B_\rho} = \{T_\rho \in \mathbb{T}_\rho : B_\rho \cap T_\rho \text{ is non-empty}\}.$$

The picture shows three points  $x_1, x_2, x_3$  all lying in a common  $B_\rho$ . In our picture,  $\mathbb{T}_{\rho, B_\rho}$  is the set of thick tubes. The point  $x_1$  lies in  $U(\mathbb{T}[T_\rho])$  for every  $T_\rho \in \mathbb{T}_{\rho, B_\rho}$ , and so

$$|\mathbb{T}_{x_1}| \approx |\mathbb{T}_{\rho, B_\rho}| |\mathbb{T}[T_\rho]| \approx \mu(\mathbb{T}_\rho) \mathbb{T}[T_\rho].$$

For the point  $x_1$ , Lemma 6.2 would be an equality.

But  $x_2$  and  $x_3$  behave differently. The point  $x_2$  lies in  $U(\mathbb{T}[T_{\rho,2}])$  but not in  $U(\mathbb{T}[T_{\rho,1}])$ . Let us imagine that  $x_2$  lies in  $U(\mathbb{T}[T_\rho])$  for only a small fraction of  $T_\rho \in \mathbb{T}_{\rho, B_\rho}$ . Then we would have

$$|\mathbb{T}_{x_2}| \ll \mu(\mathbb{T}_\rho) \mu(\mathbb{T}[T_\rho]).$$

If Lemma 6.2 is roughly an equality, then most points  $x \in U(\mathbb{T})$  must resemble  $x_1$ . So for each  $T_\rho \in \mathbb{T}_{\rho, B_\rho}$ ,  $U(\mathbb{T}_\rho) \cap B_\rho$  must be essentially the same as  $U(\mathbb{T}) \cap B_\rho$ . We call this property the perfect overlap property.

LEMMA 6.3 (Perfect overlap property). — *If  $\mathbb{T}$  is a worst-case sticky Kakeya set, and  $B_\rho \subset U(\mathbb{T}_\rho)$ , then for each  $T_\rho \in \mathbb{T}_{\rho, B_\rho}$ ,*

$$|U(\mathbb{T}[T_\rho]) \cap B_\rho| \approx |U(\mathbb{T}) \cap B_\rho|.$$

*Morally, the sets  $U(\mathbb{T}[T_\rho]) \cap B_\rho$  are all the same as  $T_\rho$  varies in  $\mathbb{T}_{\rho, B_\rho}$ .*

The perfect overlap property is a very strong condition. If you look back at Figure 3, the tubes in the picture fail the perfect overlap property: most points in  $U(\mathbb{T}) \cap B_\rho$  are like  $x_2$  or  $x_3$  and only a few are like  $x_1$ . Recall that we supposed that  $\beta_{\text{sticky}} > 0$ , and so  $|U(\mathbb{T})| \ll 1$ . It's not hard to see that since  $\mathbb{T}$  is a worst-case sticky Kakeya set,  $|U(\mathbb{T}) \cap B_\rho| \ll |B_\rho|$  (and we will prove this in Section 6.3). So  $U(\mathbb{T}[T_\rho]) \cap B_\rho$  are all small subsets of  $B_\rho$ . There are many different  $T_\rho \in \mathbb{T}_{\rho, B_\rho}$ , so we have many small subsets  $U(\mathbb{T}[T_\rho]) \cap B_\rho$ . According to the perfect overlap property, all of these small subsets coincide almost exactly. This is a strong condition and it gives a lot of information about the Kakeya set.

## 6.2. Grain structure

The perfect overlap property is easier to analyze when  $\rho = \sqrt{\delta}$  because of the following property.

LEMMA 6.4. — *If  $T_1, T_2 \in \mathbb{T}[T_{\sqrt{\delta}}]$  and  $T_1, T_2$  both intersect  $B_{\sqrt{\delta}}$ , then  $T_1 \cap B_{\sqrt{\delta}}$  and  $T_2 \cap B_{\sqrt{\delta}}$  are essentially parallel tubes of length  $\sqrt{\delta}$  and radius  $\delta$ . More precisely, there are parallel tubes  $S_1, S_2$  with length  $\sqrt{\delta}$  and radius  $2\delta$  so that  $T_j \cap B_{\sqrt{\delta}} \subset S_j$ .*

We illustrate the situation in Figure 4

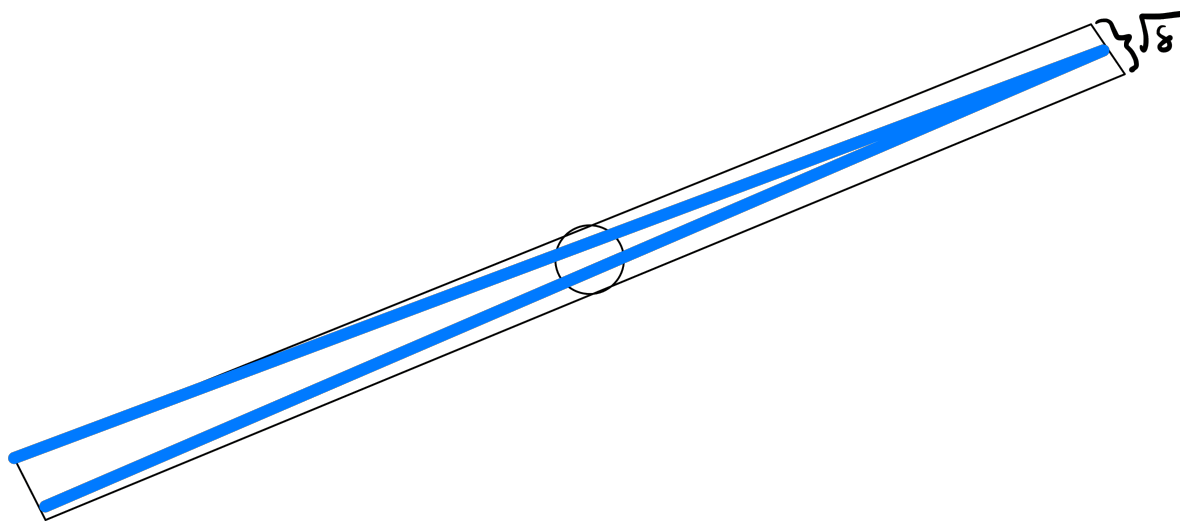
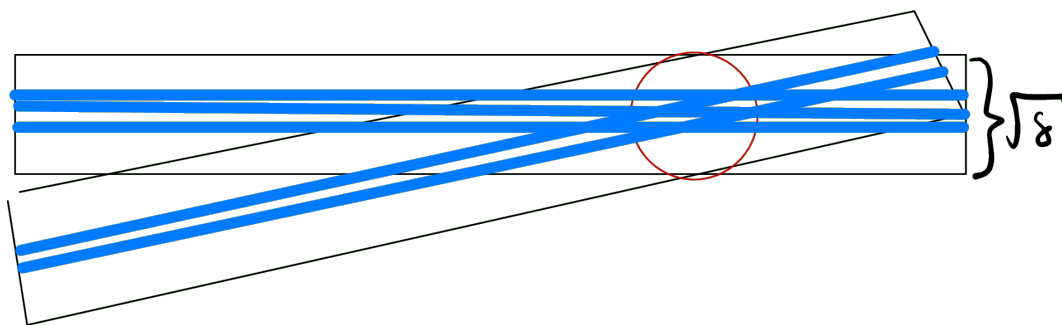


FIGURE 4. Illustration of Lemma 6.4

It is quite difficult to achieve perfect overlap when  $|U(\mathbb{T}[T_{\sqrt{\delta}}]) \cap B_{\sqrt{\delta}}| \ll |B_{\sqrt{\delta}}|$ . In two dimensions, this can only happen in the special case when the tubes thru  $B_{\sqrt{\delta}}$  all lie in a small angular sector, as in the following picture.



In two dimensions, if the fat tubes  $T_{\sqrt{\delta}}$  through  $B_{\sqrt{\delta}}$  are transverse, then the perfect overlap property implies that  $U(\mathbb{T})$  fills  $B_{\rho}$ . We state this result as a lemma.

LEMMA 6.5. — *In two dimensions suppose that  $\rho = \sqrt{\delta}$  and*

- $T_{\rho,1}$  and  $T_{\rho,2}$  pass through  $B_{\rho}$ .
- The tubes  $T_{\rho,1}$  and  $T_{\rho,2}$  are transverse: the angle between them is  $\sim 1$ .
- $U(\mathbb{T}[T_{\rho,1}]) \cap B_{\rho} = U(\mathbb{T}[T_{\rho,2}]) \cap B_{\rho}$ , and these sets are non-empty.

Then  $|U(\mathbb{T}[T_{\rho,1}]) \cap B_{\rho}| \approx |B_{\rho}|$ .

*Proof sketch.* — Let  $T_1 \in \mathbb{T}[T_{\rho,1}]$  be a tube that intersects  $B_{\rho}$ . By hypothesis, each point  $x \in T_1 \cap B_{\rho}$  lies in a tube  $T_{2,x} \in \mathbb{T}[T_{\rho,2}]$ . These tubes are all parallel to each other, and they are roughly perpendicular to  $T_1$ , and so they fill a definite fraction of  $B_{\rho}$ .  $\square$

In three dimensions, there is a more interesting example that satisfies the perfect overlap property.

**Grain example**  $B = B(0, \sqrt{\delta}) \subset \mathbb{R}^3$ .

- $G \subset B$  is a  $\delta \times \sqrt{\delta} \times \sqrt{\delta}$  slab parallel to the  $(x_1, x_2)$ -plane.
- The tubes of  $\mathbb{T}_{\sqrt{\delta}, B}$  are parallel to the  $(x_1, x_2)$ -plane.
- For each  $T_{\sqrt{\delta}} \subset \mathbb{T}_{\sqrt{\delta}}$ ,  $U(\mathbb{T}[T_{\sqrt{\delta}}]) \cap B = G$ .

Figure 5 illustrates the situation. The picture can only show some of the tubes. The tubes running in the  $x_1$  direction should fill the slab  $G$ . These tubes all lie in a single  $T_{\sqrt{\delta}}$  running in the  $x_1$  direction. Similarly, the tubes running in the  $x_2$  direction should fill the slab, and those tubes all lie in a single  $T_{\sqrt{\delta}}$  running in the  $x_2$  direction. We could also add other tubes in any direction parallel to the  $(x_1, x_2)$ -plane.

In three dimensions, any example satisfying the perfect overlap property must either be a union of grains or else have all the fat tubes lie in a small angular sector. It is not hard to check that a worst-case sticky Kakeya set cannot have tubes in such a small angular sector. This leads to the following grain structure lemma.

LEMMA 6.6 (Grain structure). — *Suppose that  $\beta_{\text{sticky}} > 0$  and that  $\mathbb{T}$  is a worst-case sticky Kakeya set. Then for a typical ball  $B = B_{\sqrt{\delta}} \subset U(\mathbb{T}_{\sqrt{\delta}})$ , we can choose coordinates so that*

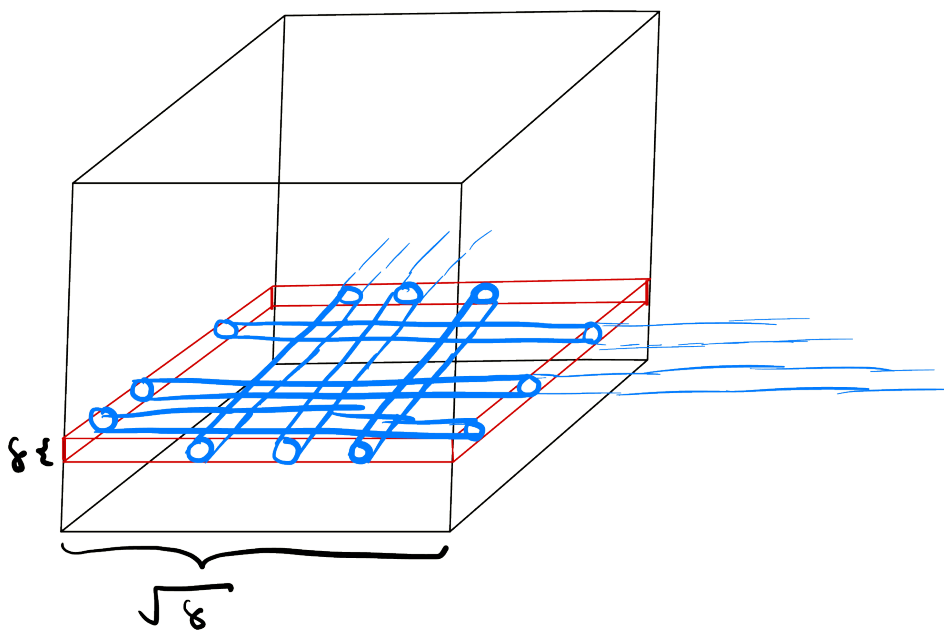


FIGURE 5. Perfect overlap of tubes in a grain

- $U(\mathbb{T}) \cap B$  is a union of slabs  $G$  as in the example above. Therefore,  $U(\mathbb{T}) \cap B$  has the form

$$U(\mathbb{T}) \cap B = [0, \sqrt{\delta}]^2 \times A,$$

where  $A \subset [0, \sqrt{\delta}]$ .

- The tubes of  $\mathbb{T}_{\sqrt{\delta}, B}$  are parallel to the  $(x_1, x_2)$ -plane.

The proof of the grain structure lemma is similar to the proof sketch for Lemma 6.5 above. We omit the details.

We refer to  $B = B_{\sqrt{\delta}} \cap U(\mathbb{T}_{\sqrt{\delta}})$  as a grain box. We call each slab  $G$  a grain. For each grain box  $B$ , we can choose coordinates so that  $B \cap U(\mathbb{T}) = [0, \sqrt{\delta}]^2 \times A$ . That means that the grains within a grain box are parallel. But two grains in different grain boxes need not be parallel.

### 6.3. Fractal structure of $A$

The argument in the perfect overlap section also gives us detailed information about  $|U(\mathbb{T}) \cap B(x, \rho)|$  for any radius  $\rho$ , and this gives us important information about  $A$ .

If  $\mathbb{T}$  is a worst-case sticky Kakeya set, then we know that  $|\mathbb{T}| \approx \delta^{-2}$  and  $\mu(\mathbb{T}) \approx |\mathbb{T}|^{\beta_{\text{sticky}}} = \delta^{-2\beta_{\text{sticky}}}$ . Since  $\mu(\mathbb{T}) = \frac{\sum_{T \in \mathbb{T}} |T|}{|U(\mathbb{T})|} \approx \frac{1}{|U(\mathbb{T})|}$ , we see that

$$(14) \quad |U(\mathbb{T})| \approx \delta^{2\beta_{\text{sticky}}}$$

Now we saw above that if  $\mathbb{T}$  is a worst-case sticky Kakeya set, then  $\mathbb{T}_\rho$  is also a worst-case sticky Kakeya set for every  $\rho \in [\delta, 1]$ . Therefore,  $|U(\mathbb{T}_\rho)| = \rho^{2\beta_{\text{sticky}}}$ . Now

$U(\mathbb{T}_\rho)$  is the  $\rho$ -neighborhood of  $U(\mathbb{T})$ . By uniformity, we will assume that  $|U(\mathbb{T}) \cap B_\rho|$  is the same for each  $B_\rho \subset U(\mathbb{T}_\rho)$ , and then we get

$$|U(\mathbb{T})| \approx \frac{|U(\mathbb{T}) \cap B_\rho|}{|B_\rho|} |U(\mathbb{T}_\rho)|.$$

Plugging in our values for  $|U(\mathbb{T})|$  and  $|U(\mathbb{T}_\rho)|$ , we see that for each  $B_\rho \subset U(\mathbb{T}_\rho)$ ,

$$|U(\mathbb{T}) \cap B_\rho| \approx \left(\frac{\delta}{\rho}\right)^{2\beta_{\text{sticky}}} |B_\rho|.$$

If  $\rho \leq \sqrt{\delta}$ , then  $U(\mathbb{T}) \cap B_\rho$  is described by the grain structure and its geometry depends on the set  $A$ . So for any  $\rho \in [\delta, \sqrt{\delta}]$  and any interval  $I_\rho$  of length  $\rho$  centered at a point of  $A$ , we get

$$(15) \quad |A \cap I_\rho| \approx \left(\frac{\delta}{\rho}\right)^{2\beta_{\text{sticky}}} \rho.$$

It is nicest to rewrite this equation in terms of  $\delta$ -covering numbers. Recall that the  $\delta$ -covering number of a set  $X$ , written  $|X|_\delta$ , is the minimal number of  $\delta$ -balls needed to cover  $X$ . Rewriting (15) in this language gives:

$$(16) \quad |A \cap I_\rho|_\delta \approx \left(\frac{\rho}{\delta}\right)^{1-2\beta_{\text{sticky}}}.$$

This equation describes the way that the set  $A$  is spaced.

Remark. This type of equation appears in the description of fractals like the Cantor set. For instance, if  $A$  was the  $\delta$ -neighborhood of a Cantor set of dimension  $1 - 2\beta_{\text{sticky}}$ , it would satisfy this equation. The technical name for (16) in the literature is that  $A$  is the  $\delta$ -neighborhood of an AD regular set of dimension  $1 - 2\beta_{\text{sticky}}$ .

## 7. Grain structure leads to complex conjugation structure

Each tube of  $\mathbb{T}$  enters many different grain boxes, and the grains in these boxes have different slopes. Let us fix one tube  $T_1 \in \mathbb{T}$  and choose coordinates so that the core line of  $T_1$  is the  $x$ -axis. For each  $x \in [0, 1]$ , the point  $(x, 0, 0)$  lies in a grain box, and the planes in that grain box are parallel to the  $x$ -axis. Therefore, the grain through  $x$  must have the form  $z = s(x)y$ . We call  $s(x)$  the slope function.

In this section, we will study the function  $s(x)$ . Recall that in the Heisenberg group example, in well-chosen coordinates,  $x \in \mathbb{C}$  and  $s(x) = \bar{x}$ . We will see that for a worst-case sticky Keakeya set, the function  $s(x)$  has some special properties analogous to properties of complex conjugation.

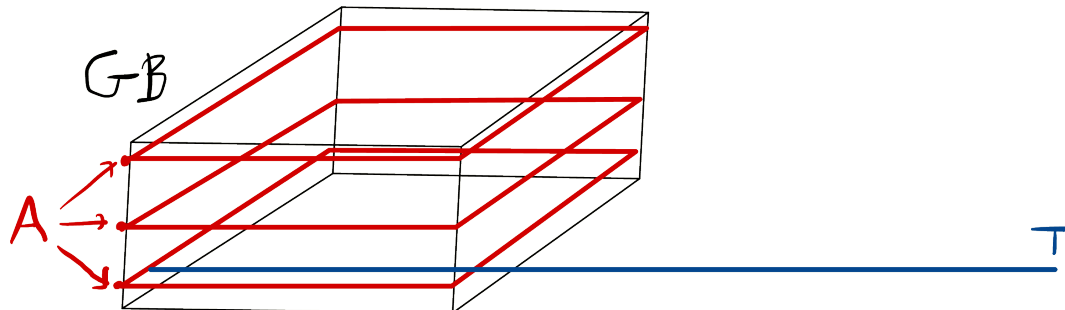
Recall that if  $\mathbb{T}$  is a worst-case sticky Keakeya set, and if  $\rho \gg \delta$ , then  $\mathbb{T}[T_\rho]$  is also a worst-case sticky Keakeya set. So by the grain structure analysis,  $U(\mathbb{T}[T_\rho])$  is also organized into grains. We can compute the dimensions of these grains by a change of

variables argument. There is a linear change of variables that takes  $T_\rho$  to a unit cube and takes  $\mathbb{T}[T_\rho]$  to a set  $\tilde{\mathbb{T}}$  of  $\delta/\rho$ -tubes in the unit cube. According to our grain structure lemma,  $\tilde{\mathbb{T}}$  is organized into grain boxes of side length  $\sqrt{\delta/\rho}$ . Undoing the linear change of variables, we see that the tubes of  $\mathbb{T}[T_\rho]$  are organized into grain boxes of dimensions  $\sqrt{\delta\rho} \times \sqrt{\delta\rho} \times \sqrt{\delta/\rho}$ . We call these long grain boxes, because they are longer and thinner than the regular (cubical) grain boxes. If  $\rho = \sqrt{\delta}$ , then these long grain boxes have dimensions  $\delta^{3/4} \times \delta^{3/4} \times \delta^{1/4}$ . If  $LGB$  is a long grain box, then  $U(\mathbb{T}) \cap LGB$  consists of parallel slabs of dimensions  $\delta \times \delta^{3/4} \times \delta^{1/4}$ . We call these slabs long grains.

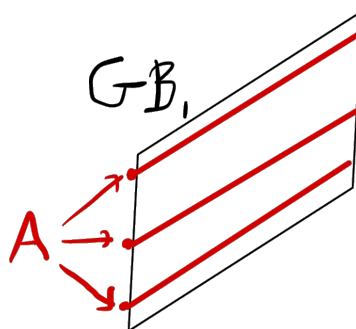
The grains in the grain boxes and the long grains in the long grain boxes have to fit together. In particular, we will fit together two grain boxes and two long grain boxes to form a kind of cycle. Following the grains around this cycle gives interesting information about the slope function  $s(x)$ . Next we explain step by step how the grain boxes and long grain boxes fit together.

We will write  $GB$  for a grain box and  $G$  for a grain. We will write  $LGB$  for a long grain box and  $LG$  for a long grain.

Let's begin with a grain box around some tube  $T_1$ . By choosing coordinates, we can suppose that this box is  $[0, \sqrt{\delta}]^3$ . Because the long grain boxes have height only  $\delta^{3/4}$ , we will focus on only the bottom part of the box, given by  $[0, \sqrt{\delta}]^2 \times [0, \delta^{3/4}]$ . According to the grain structure lemma,  $U(\mathbb{T}) \cap [0, \sqrt{\delta}]^2 \times [0, \delta^{3/4}] = [0, \sqrt{\delta}]^2 \times A$ , where  $A \subset [0, \delta^{3/4}]$ . Here is a picture.

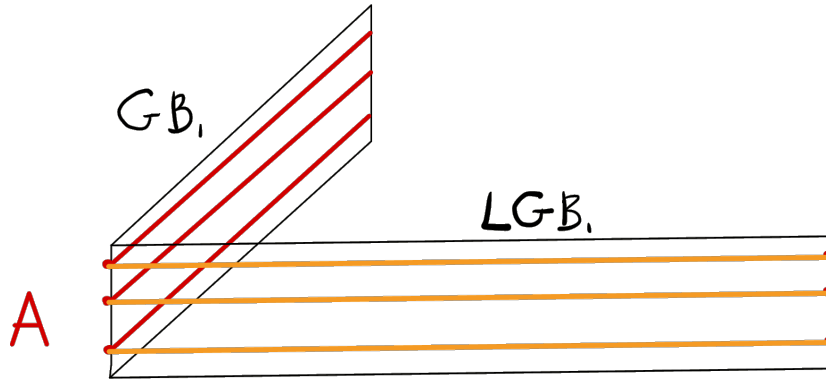


We're going to have to add more objects to the picture, so for simplicity, we only draw one side of our original grain box. Here is the abbreviated picture:



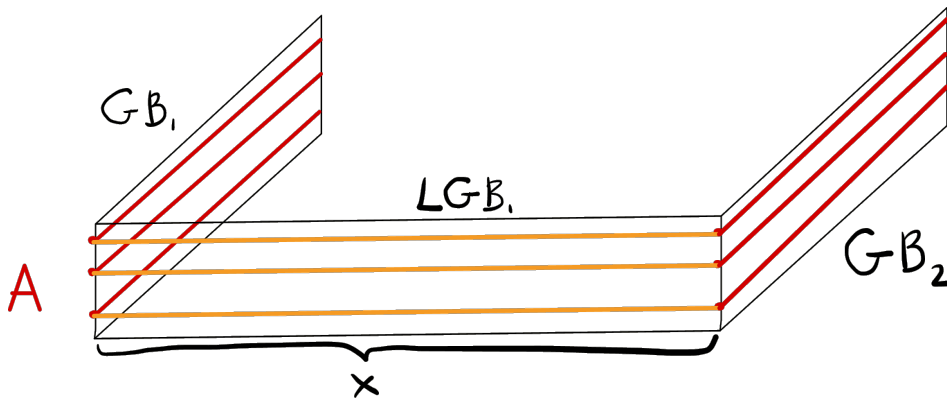
Each line in this picture represents a grain. We label this first grain box  $GB_1$  because we are going to introduce a second grain box soon. The tube  $T_1$  also lies in a long grain

box  $LGB_1$ . The long grain box  $LGB_1$  begins in  $GB_1$  but it is much longer than  $GB_1$ . Let us add  $LGB_1$  to our picture.



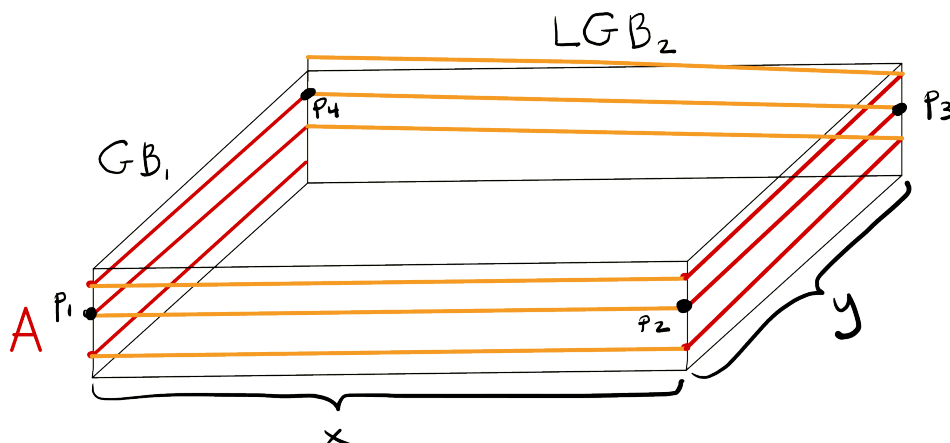
In this picture, the bottom rectangle is  $LGB_1$ , and the horizontal lines in  $LGB_1$  represent the long grains. The long grains have dimensions  $\delta \times \delta^{3/4} \times \delta^{1/4}$ . The line in the picture represents the long axis of the long grain. To keep the picture from being too crowded, we have left out both of the short axes. The  $\delta^{3/4}$  axis of the long grain runs parallel to the grains of  $GB_1$ . (Also, ideally the long grain should be much longer than the grains in  $GB_1$ , but it is hard to get so many scales right in the picture.)

The long grain box  $LGB_1$  enters many other grain boxes. If we start at  $GB_1$  and follow  $LGB_1$  for a distance  $x$ , we end up in a second grain box  $GB_2$ . The grains in these three grain boxes should fit together as shown in the following picture:



Recall that the grains in  $GB_2$  are all parallel to each other, but the slope of the grains in  $GB_2$  may be slightly different from the slope of the grains in  $GB_1$ . We have tried to show this in the picture. The slope of the grains in  $GB_2$  is  $s(x)$ .

Now there are many different long grains running from  $GB_1$  to  $GB_2$ . If we follow  $GB_2$  for a distance  $y$  (in the  $y$  direction), then we end up in a second long grain box  $LGB_2$  parallel to the first long grain box  $LGB_1$ . Both  $LGB_1$  and  $LGB_2$  run from  $GB_1$  to  $GB_2$  and they fit together as in the following picture:



This picture reminds me of the ramps in a parking garage. Starting at some point  $p_1$  in the original grain box, we follow a long grain to  $p_2$ , then a regular grain to  $p_3$ , and then a long grain to  $p_4$ . At the end of this cycle, we have come back to a new “floor” in the original grain box. In this process, no matter what “floor” we start on, the  $z$  coordinate goes up by a constant amount  $\Delta z$ . We can compute  $\Delta z$  in terms of the slopes of the different grains and  $x$  and  $y$ .

Recall that the grains in  $GB_2$  have slope  $s(x)$ , meaning that the planes are defined by equations of the form  $z = s(x)y + c$ . Similarly, let us say that the long grains in the long grain box around  $(0, y, 0)$  have slope  $\tilde{s}(y)$ , meaning that the planes are defined by equations of the form  $z = -\tilde{s}(y)x + c$ . (The negative sign is not important but is convenient.) So the long grains in  $LGB_1$  have slope  $\tilde{s}(0)$  and the long grains in  $LGB_2$  have slope  $\tilde{s}(y)$ . We can choose coordinates so that  $\tilde{s}(0) = 0$ , but  $\tilde{s}(y)$  may not be zero. Now we are ready to compute  $\Delta z$ .

Now, if we start at a point  $p_1 = (0, 0, z)$  with  $z \in A$ , we follow the long grain in  $LGB_1$  to  $p_2 = (x, 0, z + \tilde{s}(0)x) = (x, 0, z)$ . Then we follow the grain in  $GB_2$  to  $p_3 = (x, y, z + s(x)y)$ . Then we follow the long grain in  $LGB_2$  to  $p_4 = (0, y, z + s(x)y + \tilde{s}(y)x)$ . We have now arrived back at a new grain in  $GB_1$ , and so we should have  $z + s(x)y + \tilde{s}(y)x \in A$ . We have arrived at the following key result:

LEMMA 7.1. — *If  $\mathbb{T}$  is a worst-case sticky Kakeya set and  $A, s(x), \tilde{s}(y)$  are defined as above, then for  $x \in [0, \delta^{1/4}]$  and  $y \in [0, \delta^{1/2}]$ ,*

$$(17) \quad A + s(x)y + \tilde{s}(y)x \approx A$$

This result describes a very rigid structure. Even if we just had a single number  $t$  so that  $A + t \approx A$ , this would encode non-trivial arithmetic structure of the set  $A$ . But we have a huge variety of such numbers  $t$ : every number  $t$  that can be written as  $s(x)y + \tilde{s}(y)x$ , with  $x, y$  at the appropriate scales.

In the Heisenberg group example (over  $\mathbb{C}$ ),  $A$  would be the real numbers intersected with a ball of an appropriate size,  $s(x) = \bar{x}$  and  $\tilde{s}(y) = \bar{y}$ . Then Equation (17) would

follow because  $\mathbb{R} + \bar{x}y + \bar{y}x = \mathbb{R}$  for every  $x, y \in \mathbb{C}$ . Equation (17) captures some of the complex conjugation structure in  $\mathbb{C}$  and so we call (17) complex conjugation structure.

In this exposition, we have suppressed the difference between things that are true for all points and things that are true for most/many points. In the full proof this requires technical care. In particular, the precise statement of Lemma 7.1 would contain a lot of quantification. It roughly says that for a large subset of  $z \in A$  and a large subset of pairs  $(x, y) \in [0, \delta^{1/4}] \times [0, \delta^{1/2}]$ ,  $z + s(x)y + \tilde{s}(y)x \in A$ .

## 8. Complex conjugation structure leads to a contradiction

Equation (17) captures some of the structure of complex conjugation and also the way that the real numbers fit in the complex numbers. This is related to several other special features of the way that  $\mathbb{R}$  fits in  $\mathbb{C}$ . The most fundamental feature is that  $\mathbb{R}$  is closed under both addition and multiplication. Starting from Lemma 7.1, the proof of Kakeya shows that there must be a set  $A \subset \mathbb{R}$  which is approximately closed under both addition and multiplication and also obeys spacing estimates as in (16). This area is called sum-product theory, and we give a short introduction to it here.

Suppose that  $A \subset \mathbb{R}$ . We write  $A + A$  for the sumset

$$A + A = \{a_1 + a_2 \mid a_1, a_2 \in A\}$$

and we write  $A \cdot A$  for the product set

$$A \cdot A = \{a_1 a_2 \mid a_1, a_2 \in A\}.$$

If  $A \subset \mathbb{R}$  is a finite set, it is interesting to consider  $|A + A|$  and  $|A \cdot A|$ . If  $A = \{1, \dots, N\}$ , then  $|A| = N$  and  $|A + A| \sim N$ , but  $|A \cdot A| \approx N^2$ . Similarly, if  $A$  is a geometric progression such as  $\{2^{n/N}\}_{n=1}^N$ , then  $|A| = N$  and  $|A \cdot A| \sim N$ , but  $|A + A| \approx N^2$ . Erdős conjectured that for any finite set  $A \subset \mathbb{R}$ ,  $|A + A| + |A \cdot A| \gtrsim |A|^2$ . This conjecture is still open, but we do have some weaker bounds. Erdős and Szemerédi (1983) proved that there is an exponent  $\alpha > 0$  so that  $|A + A| + |A \cdot A| \gtrsim |A|^{1+\alpha}$ . The best known value of  $\alpha$  was improved many times, and it is currently a little more than  $1/3$ .

These results are not exactly what is needed in Kakeya type problems. The set  $A$  that appears in our story is not a finite set of points but a finite set of  $\delta$ -intervals. Instead of measuring the sizes of finite sets by cardinality, we need to measure the sizes of sets by  $\delta$ -covering numbers: if  $A \subset \mathbb{R}$ , we write  $|A|_\delta$  for the minimum number of  $\delta$ -balls needed to cover  $A$ . Switching our point of view to  $\delta$ -covering numbers, it is natural to ask: if  $A \subset \mathbb{R}$ , then is it true that  $|A + A|_\delta + |A \cdot A|_\delta \gtrsim |A|_\delta^{1+\alpha}$  for some  $\alpha > 0$ ? The answer to this question is no. For instance, this inequality fails if  $A = \{1 + j\delta\}_{j=1}^J$  whenever  $1 \ll J \leq \delta^{-1}$ . In order to get a non-trivial sum-product inequality, we need to add an extra assumption saying that  $A$  is not packed too tightly into an interval. The first such theorem was proven by Bourgain (2003), and there is also a closely related result by Edgar and Miller (2003).

**THEOREM 8.1** (Bourgain “discretized” sum-product theorem)

*Suppose that  $0 < s < 1$ , and  $A \subset [0, 1]$  with*

$$- |A|_\delta \approx \delta^{-s}.$$

$$- \text{If } B(x, r) \subset [0, 1], \text{ then } |A \cap B(x, r)|_\delta \lesssim (r/\delta)^s$$

*Then  $|A + A|_\delta + |A \cdot A|_\delta \gtrsim \delta^{-s-\epsilon}$ , where  $\epsilon = \epsilon(s) > 0$ .*

With this background, we can return to discussing the proof of the sticky Kakeya theorem. Recall from (16) that the set  $A$  obeys the hypotheses above with  $s = 1 - 2\beta_{\text{sticky}}$ . We know by earlier arguments that  $\beta_{\text{sticky}} \leq 1/4$  and so if  $\beta_{\text{sticky}} > 0$ , then  $0 < s < 1$ . Roughly speaking, we might hope that the complex conjugation structure forces  $A + A$  and  $A \cdot A$  to be small, which would give a contradiction. I think this is morally correct but the actual proof is somewhat more complicated.

First of all there are some trivial ways that  $s(x)$  and  $\tilde{s}(y)$  can satisfy (17). We could have  $s(x) = \tilde{s}(y) = 0$ . Or we could have  $s(x) = x$  and  $\tilde{s}(y) = -y$ . The proof in Wang and Zahl (2026) first rules out these possibilities. For instance, if  $s(x) = 0$ , then it would mean that all the tubes crossing  $T_1$  must lie in a common planar slab, and this would eventually force too many tubes into a planar slab, violating the hypothesis that  $\Delta_{\max}(\mathbb{T}) \lesssim 1$ .

Starting from the complex conjugation structure, the proof eliminates these trivial possibilities and then shows that some new set related to  $A$  is approximately closed under both addition and multiplication, which contradicts Theorem 8.1. The argument has several steps, and we are not able to sketch all the steps in this survey. The ideas were developed by Katz, Tao, Orponen, Shmerkin, Wang, and Zahl. The ideas are related to major recent progress in the field of projection theory.

At the beginning of our discussion, we mentioned that the Kakeya theorem is hard to prove because the Heisenberg group example shows that the analogous statement over  $\mathbb{C}$  is false. The Heisenberg group example is sticky, and so the complex analogue of the sticky Kakeya theorem is also false. Any proof of the Kakeya theorem or sticky Kakeya theorem must include a step that distinguishes  $\mathbb{R}$  from  $\mathbb{C}$ . The sum-product theorem, Theorem 8.1, is the crucial step that distinguishes  $\mathbb{R}$  from  $\mathbb{C}$ .

Let us quickly note that the analogue of Theorem 8.1 over  $\mathbb{C}$  is false. The counterexample comes because  $\mathbb{R}$  is a subring of  $\mathbb{C}$ . More precisely, we let  $A = \mathbb{R} \cap B(0, 1) \subset \mathbb{C}$ . Then  $|A|_\delta \sim |A + A|_\delta \sim |A \cdot A|_\delta$ , and these are all much smaller than  $|B(0, 1)|_\delta$ .

Since distinguishing  $\mathbb{R}$  from  $\mathbb{C}$  is a crucial part of the proof of Kakeya, let us sketch how it is done. The simplest version of the idea takes place over finite fields. We let  $\mathbb{F}_q$  denote the finite field with  $q$  elements. We write  $q = p^r$  with  $p$  prime. If  $r \geq 2$ , then  $\mathbb{F}_q$  has non-trivial subfields. In particular,  $\mathbb{F}_{p^2}$  is a degree 2 extension of  $\mathbb{F}_p$ , just as  $\mathbb{C}$  is a degree 2 extension of  $\mathbb{R}$ . Proving quantitative bounds that distinguish  $\mathbb{R}$  from  $\mathbb{C}$  is closely related to proving quantitative bounds that distinguish  $\mathbb{F}_p$  from  $\mathbb{F}_{p^2}$ .

We will prove a sum-product type estimate over  $\mathbb{F}_p$  when  $p$  is prime. Our result will involve some sets more complicated than  $A + A$  or  $A \cdot A$ . We will need the following

definitions.

$$(A \cdot A)^{\oplus 3} = \{a_1a_2 + a_3a_4 + a_5a_6 : a_i \in A\}.$$

$$\frac{A - A}{A - A} = \left\{ \frac{a_1 - a_2}{a_3 - a_4} : a_i \in A, a_3 \neq a_4 \right\}.$$

$$\frac{(A \cdot A)^{\oplus 3} - (A \cdot A)^{\oplus 3}}{A - A} = \left\{ \frac{b_1 - b_2}{a_1 - a_2} : b_i \in (A \cdot A)^{\oplus 3}, a_i \in A, a_1 \neq a_2 \right\}.$$

If  $c \in \mathbb{F}_p$ , then

$$A + cA = \{a_1 + ca_2 : a_i \in A\}.$$

LEMMA 8.2. — *Suppose that  $p$  is prime and  $A \subset \mathbb{F}_p$ . Then either*

1.  $\frac{A-A}{A-A} = \mathbb{F}_p$ , or
2.  $\left| \frac{(A \cdot A)^{\oplus 3} - (A \cdot A)^{\oplus 3}}{A-A} \right| \geq |A|^2$ .

We note that this lemma is not true in all finite fields. If  $q = p^2$  and  $A = \mathbb{F}_p \subset \mathbb{F}_q$ , then all the complicated sets in the bullet points are equal to  $A$ , and they all have cardinality much smaller than  $|\mathbb{F}_q|$  or  $|A|^2$ . The proof of this theorem must distinguish  $\mathbb{F}_p$  from  $\mathbb{F}_{p^2}$ .

*Proof.* — First, note that if  $c \notin \frac{A-A}{A-A}$ , then  $|A + cA| = |A|^2$ . Indeed if  $|A + cA| < |A|^2$ , then by the pigeonhole principle, we could find  $a_1, a_2, a'_1, a'_2 \in A$  with  $a_1 + ca_2 = a'_1 + ca'_2$ . But this implies  $c = \frac{a'_1 - a_1}{a_2 - a'_2} \in \frac{A-A}{A-A}$ .

Next, note that if  $\frac{A-A}{A-A} \neq \mathbb{F}_p$  then there is some  $b \in \frac{A-A}{A-A}$  with  $b + 1 \notin \frac{A-A}{A-A}$ . This step is true in  $\mathbb{F}_p$  but it would fail in  $\mathbb{F}_{p^2}$ .

Now, if  $\frac{A-A}{A-A} \neq \mathbb{F}_p$ , then we have  $|A + (b + 1)A| \geq |A|^2$  with  $b \in \frac{A-A}{A-A}$ . Expanding everything out, we see that

$$A + (b + 1)A \subset \frac{(A \cdot A)^{\oplus 3} - (A \cdot A)^{\oplus 3}}{A - A}.$$

This gives the second option above. □

Remark. Some version of the idea of this proof goes back to the work of Edgar and Miller (2003), and the argument was adapted by Bourgain, Katz, and Tao, 2004 and Garaev (2007). This proof can also be adapted to the setting of  $\mathbb{R}$  and  $\mathbb{C}$ , which was done by Guth, Katz, and Zahl (2021). There are multiple proofs of Theorem 8.1. An adapted version of Lemma 8.2 plays a key role in one proof (from Guth, Katz, and Zahl, 2021). The proof of Theorem 8.1 is technically complicated, and there is a good recent exposition of this area in O'Regan, Shmerkin, and Wang (2025).

## 9. Leveraging the sum-product theorem at many scales

Let us pause to digest a surprising feature of the proof of the sticky Kakeya theorem. A key obstacle in proving sticky Kakeya is that the analogue over  $\mathbb{C}$  is false. The sum-product theorem (Theorem 8.1) distinguishes  $\mathbb{R}$  from  $\mathbb{C}$ . However, the bounds in Theorem 8.1 are not sharp, and indeed they are very weak. How can a non-sharp and very weak theorem be used as a key step in the proof of a sharp theorem?

Over the last several years, there have been a number of sharp results proven using the non-sharp Theorem 8.1. This body of work has had a major influence in the field of projection theory. Some of the main contributors are Orponen, Shmerkin, Ren, and Wang. Roughly speaking, it is possible to prove strong and sharp estimates by applying Theorem 8.1 many times at different scales.

The proof we have reviewed can be written in different ways. Let us describe a way to rephrase the proof to highlight the way that we are exploiting Theorem 8.1 many times at different scales.

We define  $\beta_{\text{sticky}}(\delta)$  to be the least exponent so that for every sticky Kakeya set of  $\delta$ -tubes,  $\mu(\mathbb{T}) \leq |\mathbb{T}|^{\beta_{\text{sticky}}(\delta)}$ . In this language, the sticky Kakeya theorem says that  $\lim_{\delta \rightarrow 0} \beta_{\text{sticky}}(\delta) = 0$ . We can organize our proof using a key lemma which says that if  $\delta_2$  is far smaller than  $\delta_1$ , then  $\beta_{\text{sticky}}(\delta_2)$  is smaller than  $\beta_{\text{sticky}}(\delta_1)$  by a definite amount. Here is the precise statement of the lemma.

**LEMMA 9.1.** — *For any  $\beta > 0$ , there are  $\epsilon > 0$  and  $K > 0$  so that if  $\delta_1 < 1/10$  and  $\beta_{\text{sticky}}(\delta_1) \geq \beta$  and  $\delta_2 \leq \delta_1^K$ , then  $\beta_{\text{sticky}}(\delta_2) < \beta_{\text{sticky}}(\delta_1) - \epsilon$ . Also  $\epsilon = \epsilon(\beta)$  and  $K = K(\beta)$  are continuous in  $\beta$ .*

Iterating this key lemma at many different scales gives the sticky Kakeya theorem.

The proof sketched in the sections above can be slightly adapted to give a proof of the key lemma. The proof of the key lemma crucially uses the sum-product theorem, Theorem 8.1. The exponent in Theorem 8.1 is not sharp, and is only a tiny improvement on a trivial bound. Partly for this reason, the exponent  $\epsilon = \epsilon(\beta)$  in the key lemma is not sharp and is only a tiny improvement on a trivial bound. But applying the key lemma many times at different scales, we get essentially sharp bounds. In this process, we are leveraging the sum-product theorem by applying it many times at different scales and getting a small gain each time.

This finishes our sketch of the proof of the sticky Kakeya theorem.

## 10. Sticky vs. not sticky

The proof of the sticky Kakeya theorem shows a remarkable connection between the sticky case of the Kakeya problem and mathematical structures like the Heisenberg group, subrings of  $\mathbb{R}$ , and sum-product inequalities. It is certainly interesting mathematics. But it was not so clear how much progress these results make towards the

general Kakeya conjecture. Is the sticky case a crucial case? Or is it just a rare special case?

Should we expect the “worst” Kakeya set to be sticky? Analysts have considered many cousins of the Kakeya problem. For many years, the worst known example for every cousin problem was sticky. Bourgain (1991) considered a variation of the Kakeya problem for curved tubes in  $\mathbb{R}^3$ . In this curved version, the smallest possible volume of  $|U(\mathbb{T})|$  is  $\sim \delta$ , and the worst-case example is sticky. Katz, Łaba, and Tao (2000) gave the Heisenberg group example, which showed that the analogue of the Kakeya problem with convex Wolff axioms is false in  $\mathbb{C}^3$ . The Heisenberg group example is sticky. It seemed plausible that for a broad class of problems of this type, the worst-case is sticky.

Katz and Tao and others wondered whether the general Kakeya problem could be reduced to the sticky case, but they didn’t see any way to do it. Orponen (2017) proved the sticky case of the Falconer conjecture, another longstanding problem in geometric measure theory which is a kind of cousin of the Kakeya problem. This remarkable proof had a lot of influence in the field, but no one has managed to reduce the general Falconer conjecture to the sticky case.

Then Katz and Zahl (2019) found a new cousin of the Kakeya problem, and gave evidence that the worse case example is not sticky. They considered the Wolff axioms version of the Kakeya problem over a different ring – they replaced  $\mathbb{R}$  by the ring  $A = \mathbb{F}_p[x]/(x^2)$ . The ring  $A$  has a natural notion of distance with two distinct length scales. There is a cousin of the Heisenberg group in  $A^3$  and it leads to a counterexample to the analogue of Theorem 1.2. But unlike in  $\mathbb{C}^3$ , the Heisenberg group cousin in  $A^3$  is *not* sticky. It appears likely that in  $A^3$ , the sticky case of the analogue of Theorem 1.2 is true, but the general case is false.

As of a couple of years ago, I was quite pessimistic about reducing the general case of Kakeya to the sticky case.

The first indication that the sticky case may play a crucial role in problems of this type was the solution of the Furstenberg set conjecture by Orponen–Shmerkin and Ren–Wang in 2024. The Furstenberg set conjecture is a central question in the field called projection theory, which studies the orthogonal projections of sets and measures in  $\mathbb{R}^d$ . In the late 90s, Tom Wolff identified the Kakeya problem, the Falconer problem, and the Furstenberg set problem as cousin problems involving related issues. In particular, all three conjectures have versions over  $\mathbb{C}$  which are false, with counterexamples related to the subfield  $\mathbb{R} \subset \mathbb{C}$ . Orponen and Shmerkin (2023b) proved the sticky case of the Furstenberg set conjecture. Shortly afterward, Ren and Wang (2023) proved the full Furstenberg conjecture.

The proof of the Furstenberg conjecture is a major development in the field, and I think it deserves its own survey article to describe. Some of the key multiscale ideas in the proof of the Kakeya conjecture grew out of this work, developing over multiple papers, including Bourgain (2003, 2010), Demeter and Wang (2025), Orponen (2022), Orponen and Shmerkin (2023a,b), Orponen, Shmerkin, and Wang (2024), Ren and

Wang (2023), Shmerkin and Wang (2025), and Wang and Wu (2024). The proof of the sticky case in Orponen and Shmerkin (2023b) leverages the sum-product theorem 8.1 at many different scales, just like the proof of sticky Kakeya that we saw here. It also has its own unique issues and features. The proof of the general case of the Furstenberg conjecture reduces the problem to two cases: the sticky case and a case which is far from sticky, which they call the semi-well-spaced case. They resolve the semi-well-spaced case using Fourier analytic methods building on Guth, Solomon, and Wang (2019). And they give a short and elegant multiscale argument which reduces the general Furstenberg conjecture to these two cases.

It was quite surprising to me that the general Furstenberg problem reduces to these two cases. This work gave a hint that the sticky case might be a key case in other problems as well. Then Wang and Zahl (2025) reduced the general case of the Kakeya problem to the sticky case. The second main part of our survey describes this reduction.

## 11. The $L^2$ method

Before discussing the reduction to the sticky case, let us briefly recall the classical  $L^2$  method, which we will need in the proof.

If  $T_1$  and  $T_2$  are two  $\delta$ -tubes in  $\mathbb{R}^n$  that intersect at a point and the angle between their core lines is  $\theta \geq \delta$ , then

$$(18) \quad |T_1 \cap T_2| \sim \delta^n \theta^{-1}.$$

If  $\mathbb{T}$  is a set of  $\delta$ -tubes in  $B_1$ , then we can use (18) to upper bound

$$\int_{B_1} \left| \sum_{T \in \mathbb{T}} 1_T \right|^2 = \sum_{T_1, T_2 \in \mathbb{T}} |T_1 \cap T_2|.$$

Assuming that  $|\mathbb{T}| \approx \delta^{-(n-1)}$  and that  $\Delta_{max}(\mathbb{T}) \lesssim 1$ , this method gives the sharp bound

$$\int_{B_1} \left| \sum_{T \in \mathbb{T}} 1_T \right|^2 \lesssim \delta^{-(n-2)}.$$

(This bound is sharp when  $\mathbb{T}$  has one tube in each direction and they all go through the origin.)

Combining this  $L^2$  bound with Cauchy-Schwarz gives a lower bound on  $|U(\mathbb{T})|$ . If  $\mathbb{T}$  is a set of  $\delta$ -tubes in  $\mathbb{R}^2$  with  $|\mathbb{T}| \approx \delta^{-1}$  and  $\Delta_{max}(\mathbb{T}) \lesssim 1$ , this method shows that  $|U(\mathbb{T})| \gtrsim 1$ , which is equivalent to  $\mu(\mathbb{T}) \lesssim 1$ . This method resolves the Kakeya conjecture in two dimensions.

In higher dimensions, while this  $L^2$  estimate is sharp, it does not lead to good information about  $|U(\mathbb{T})|$  or  $\mu(\mathbb{T})$ .

On the other hand, this  $L^2$  method also works well for slabs in  $\mathbb{R}^3$ . For instance, using the same method, we can prove that if  $\mathbb{S}$  is a set of  $\delta \times 1 \times 1$  slabs in  $B_1 \subset \mathbb{R}^3$  with  $|\mathbb{S}| \sim \delta^{-1}$  and  $\Delta_{max}(\mathbb{S}) \lesssim 1$ , then  $\mu(\mathbb{S}) \lesssim 1$  and  $|U(\mathbb{S})| \gtrsim 1$ .

This method can handle many questions about tubes in  $\mathbb{R}^2$  and slabs in  $\mathbb{R}^3$ . In the proof sketch below, we will meet a few problems of this type, and we will mention that they can be handled by the  $L^2$  method.

## 12. The worst case is sticky

We now begin the second big part of the proof ofakeya: showing that a worst-caseakeya set is sticky. This part of the proof was done in Wang and Zahl (2025). Here we will follow the exposition in Guth, Wang, and Zahl (2026), which slightly streamlines the original proof. We describe the proof in Sections 12, 13, and 14.

Recall that  $\beta$  is the infimal number so that, whenever  $\Delta_{max}(\mathbb{T}) \lesssim 1$ , we have

$$(19) \quad \mu(\mathbb{T}) \lesssim |\mathbb{T}|^\beta$$

Recall that  $\mathbb{T}$  is a worst-caseakeya set if  $\Delta_{max}(\mathbb{T}) \lesssim 1$  and  $\mu(\mathbb{T}) \approx |\mathbb{T}|^\beta$ . The hypothesis that  $\Delta_{max}(\mathbb{T}) \lesssim 1$  implies that  $|\mathbb{T}| \lesssim \delta^{-2}$ . For these notes, we focus on the case that  $|\mathbb{T}| \approx \delta^{-2}$ , which shows the main ideas and leaves out some technical issues.

Our goal is to prove that a worst-caseakeya set must be sticky. Assuming that  $\beta > 0$  and that  $\mathbb{T}$  is not sticky, we will prove that

$$(20) \quad \mu(\mathbb{T}) \ll |\mathbb{T}|^\beta$$

(Recall we use the notation  $\mu(\mathbb{T}) \ll |\mathbb{T}|^\beta$  to mean that  $\mu(\mathbb{T})$  is much less than  $|\mathbb{T}|^\beta$ .) This inequality contradicts our assumption that  $\mu(\mathbb{T}) \approx |\mathbb{T}|^\beta$ , and so we conclude that  $\mathbb{T}$  must be sticky. Then the stickyakeya theorem implies that  $\beta = 0$ .

Since  $|\mathbb{T}| \sim \delta^{-2}$ , if  $\mathbb{T}$  is not sticky it means that there is some scale  $\rho \in [\delta, 1]$  so that  $|\mathbb{T}[T_\rho]| \ll (\rho/\delta)^{-2}$  and  $|\mathbb{T}_\rho| \gg \rho^{-2}$ . We will focus on the case that  $\mathbb{T}$  is not-sticky-at-all-scales, meaning that for every  $\rho$  in the range  $\delta \ll \rho \ll 1$ , we have

$$(21) \quad |\mathbb{T}[T_\rho]| \ll (\rho/\delta)^{-2} \text{ and } |\mathbb{T}_\rho| \gg \rho^{-2}.$$

Assuming (21), we will sketch the proof of (20).

In order to use the definition of  $\beta$ , we need to relate  $\mathbb{T}$  with other sets of tubes. We will relate  $\mathbb{T}$  to some other set of tubes  $\mathbb{T}'$  with  $\Delta_{max}(\mathbb{T}') \lesssim 1$  and we use that  $\mu(\mathbb{T}') \lesssim |\mathbb{T}'|^\beta$ . In the not sticky case, it is tricky to find helpful sets of tubes  $\mathbb{T}'$ . Recall that in the sticky case,  $\mathbb{T}_\rho$  and  $\mathbb{T}[T_\rho]$  were both stickyakeya sets. In the not sticky case, we still have  $\Delta_{max}(\mathbb{T}[T_\rho]) \lesssim 1$ , but we don't have  $\Delta_{max}(\mathbb{T}_\rho) \lesssim 1$ , so it is harder to make use of  $\mathbb{T}_\rho$  directly. One of the main new ideas in the proof is a way to find other relevant sets of tubes  $\mathbb{T}'$ . We will see below a couple different clever ways of doing this.

### 12.1. Looking at $\mathbb{T}$ inside a small ball

Let  $\rho$  be an intermediate scale with  $\delta \ll \rho \ll 1$ .

To bound  $\mu(\mathbb{T})$  in the sticky case, we considered  $\mathbb{T}[T_\rho]$  and  $\mathbb{T}_\rho$ . In the non-sticky case, Wang and Zahl also consider a new set of tubes formed by intersecting tubes of  $\mathbb{T}$  with a smaller ball  $B \subset B_1$ .

To set this up, let's first let's think about how the tubes of  $\mathbb{T}[T_\rho]$  intersect each other. If  $T_1, T_2 \in \mathbb{T}$ ,  $T_1$  and  $T_2$  intersect, and the angle between  $T_1, T_2 \approx \rho$ , then  $T_1 \cap T_2$  is approximately a shorter tube of radius  $\delta$  and length  $\delta/\rho$ . Therefore,  $U(\mathbb{T}[T_\rho])$  is a union of shorter tubes of this kind. Each of these short tubes lies in  $\approx \mu(\mathbb{T}[T_\rho])$  long tubes  $T \in \mathbb{T}[T_\rho]$ . Now let  $B$  be a ball of radius  $r = \delta/\rho$  and consider how these shorter tubes overlap inside  $B$ . Let  $\mathbb{T}[T_\rho]_B$  be the set of these shorter tubes in  $B$ . So each tube  $T_B \in \mathbb{T}[T_\rho]_B$  is a  $\delta \times \delta \times \delta/\rho$  tube in  $B$  which lies in  $\approx \mu(\mathbb{T}[T_\rho])$  tubes of  $\mathbb{T}[T_\rho]$ . Figure 6 shows a picture.

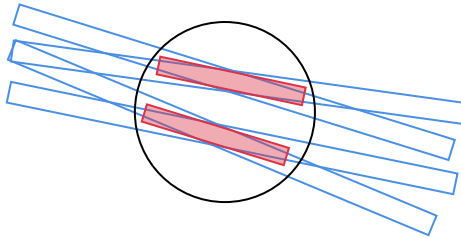


FIGURE 6. Localizing tubes to a ball

In Figure 6, the long tubes belong to  $\mathbb{T}[T_\rho]$ , the short tubes belong to  $\mathbb{T}_{T_\rho, B}$ , and the disk is  $B$ .

Next we define

$$\mathbb{T}_B = \bigcup_{T_\rho \in \mathbb{T}_\rho, T_\rho \cap B \neq \emptyset} \mathbb{T}[T_\rho]_B.$$

Figure 7 is a picture showing  $\mathbb{T}_B$ :

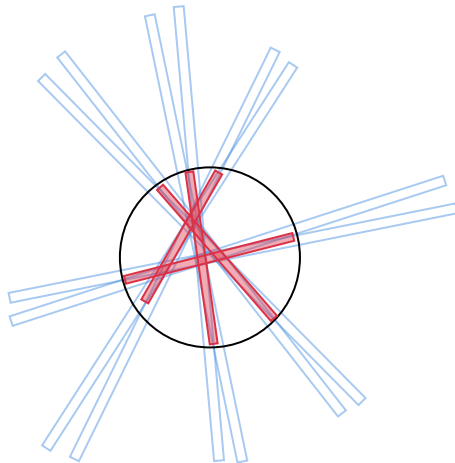
In this picture, the circle is  $B$ , the short tubes belong to  $\mathbb{T}_B$ , and we see that each tube of  $\mathbb{T}_B$  lies in  $\sim \mu(\mathbb{T}[T_\rho])$  longer tubes of  $\mathbb{T}$ . Therefore, we can bound  $\mu(\mathbb{T})$  by

$$(22) \quad \mu(\mathbb{T}) \lesssim \mu(\mathbb{T}[T_\rho])\mu(\mathbb{T}_B).$$

We know that  $\Delta_{\max}(\mathbb{T}[T_\rho]) \lesssim \Delta_{\max}(\mathbb{T}) \lesssim 1$ , and so by the definition of  $\beta$ ,  $\mu(\mathbb{T}[T_\rho]) \lesssim |\mathbb{T}[T_\rho]|^\beta$ . Since we are in the not-sticky-at-all-scales case (see (21)), we also know that  $|\mathbb{T}[T_\rho]| \ll (\rho/\delta)^2$ . So we have

$$(23) \quad \mu(\mathbb{T}[T_\rho]) \ll (\rho/\delta)^{2\beta}.$$

Next, we have to bound  $\mu(\mathbb{T}_B)$ . Here it is much less clear what to do. To get started, let's consider the special case  $\Delta_{\max}(\mathbb{T}_B) \lesssim 1$ . In this special case, we can prove our goal  $\mu(\mathbb{T}) \ll |\mathbb{T}|^\beta$  by a simple induction argument.

FIGURE 7. The set of short tubes  $\mathbb{T}_B$ 

LEMMA 12.1. — *If  $\mathbb{T}$  is a worst-case Kakeya set which is not sticky (as in (21)), and IF  $\Delta_{max}(\mathbb{T}_B) \lesssim 1$ , then  $\mu(\mathbb{T}) \ll |\mathbb{T}|^\beta$ .*

*Proof.* — IF  $\Delta_{max}(\mathbb{T}_B) \lesssim 1$ , then we would get

$$\mu(\mathbb{T}_B) \lesssim |\mathbb{T}_B|^\beta.$$

The tubes of  $\mathbb{T}_B$  have radius  $\delta$  and length  $r = \delta/\rho$ , and so the ratio  $\frac{\text{radius}(\mathbb{T}_B)}{\text{length}(\mathbb{T}_B)} = \rho$ . Since  $\Delta_{max}(\mathbb{T}_B) \lesssim 1$ , it follows that  $|\mathbb{T}_B| \lesssim \rho^{-2}$ . Plugging this bound into the last indented equation, we get

$$\mu(\mathbb{T}_B) \lesssim \rho^{-2\beta}.$$

Combining this bound with (23), we get

$$\mu(\mathbb{T}) \lesssim \mu(\mathbb{T}[T_\rho])\mu(\mathbb{T}_B) \ll (\rho/\delta)^{2\beta} \rho^{-2\beta} = \delta^{-2\beta} \approx |\mathbb{T}|^\beta. \quad \square$$

Now the hypothesis that  $\mathbb{T}_B$  has  $\Delta_{max}(\mathbb{T}_B) \lesssim 1$  is a big IF (that's why I wrote IF in all caps). The fact that  $\Delta_{max}(\mathbb{T}) \lesssim 1$  does NOT imply that  $\Delta_{max}(\mathbb{T}_B) \lesssim 1$ . This makes it a little surprising that Lemma 12.1 plays an important role in our proof.

## 12.2. A surprising induction

This is a key philosophical moment in the proof. We are going to try to control  $\mathbb{T}_B$  using induction. But the set of tubes  $\mathbb{T}_B$  need not obey  $\Delta_{max}(\mathbb{T}_B) \lesssim 1$ . This makes it surprising to try to use  $\mathbb{T}_B$  in an inductive proof of Theorem 1.2.

To put this moment in context, let us recall some more history about work on the Kakeya problem. Bennett, Carbery, and Tao (2006) formulated and proved a multilinear cousin of the Kakeya problem. Their proof was simplified in Guth (2015), and the proof there is only a few pages long. Multilinear Kakeya involves  $n$  sets of tubes  $\mathbb{T}_j$  in  $\mathbb{R}^n$ , where the tubes of  $\mathbb{T}_j$  are approximately parallel to the  $x_j$  axis. Multilinear Kakeya is important because it is much easier than Kakeya but still has many applications - for instance in work of Bourgain and Demeter (2015) on decoupling theory.

The key feature that makes multilinear Kakeya much easier than Kakeya is that if we intersect the tubes of each  $\mathbb{T}_j$  with a small ball  $B$ , then the resulting sets of tubes  $\mathbb{T}_{j,B}$  obey the hypotheses of multilinear Kakeya. Therefore, we can easily apply induction to study the intersections of tubes in each small ball  $B$ . In contrast, the hypotheses of the Kakeya problem do not behave well when we restrict the tubes of  $\mathbb{T}$  to a small ball  $B$ .

People in the field (including me) tried to adapt the inductive proof of multilinear Kakeya to the original Kakeya problem and we all gave up. Since the set  $\mathbb{T}_B$  does not obey the hypotheses of the Kakeya conjecture, how can we control it by induction?

In the Wang–Zahl proof, we assume nothing at all about the set of tubes  $\mathbb{T}_B$ . But no matter how  $\mathbb{T}_B$  behaves, they find a way to take advantage of it and prove that  $\mu(\mathbb{T}) \ll |\mathbb{T}|^\beta$ . There are several scenarios and we will discuss the most important scenarios and how to take advantage of each one. This proof outline reminds me of a game that was popular in my elementary school called “heads I win, tails you lose.”

Before turning to details, let us think through the philosophical issue of how to apply induction to  $\mathbb{T}_B$ . While  $\mathbb{T}_B$  itself does not obey  $\Delta_{max}(\mathbb{T}_B) \lesssim 1$ , Wang and Zahl manage to relate  $\mathbb{T}_B$  to another set of tubes  $\mathbb{T}'$  with  $\Delta_{max}(\mathbb{T}') \lesssim 1$ , and then we can use that  $\mu(\mathbb{T}') \lesssim |\mathbb{T}'|^\beta$ . As we go, we will see how to locate this new set of tubes  $\mathbb{T}'$  in various scenarios.

### 12.3. Organizing the different cases

If  $\Delta_{max}(\mathbb{T}_B) \lesssim 1$ , then Lemma 12.1 gives us our goal:  $\mu(\mathbb{T}) \ll |\mathbb{T}|^\beta$ . So we have to consider the case that  $\Delta_{max}(\mathbb{T}_B) \gg 1$ . By definition, this means that there is some convex set  $K \subset B$  so that  $\Delta(\mathbb{T}_B, K) \gg 1$ . We consider the set  $K$  that maximizes  $\Delta(\mathbb{T}_B, K)$ .

Notice that for any set  $K' \subset K$ ,  $\Delta(\mathbb{T}_B, K') \leq \Delta(\mathbb{T}_B, K)$ . This condition plays an important role in the story, so we give it a name.

**DEFINITION 12.1.** — *If  $\tilde{\mathbb{T}}$  is a set of tubes all contained in a convex set  $K$ , then  $\tilde{\mathbb{T}}$  is Frostman in  $K$  if for any convex  $K' \subset K$ ,*

$$\Delta(\tilde{\mathbb{T}}, K') \lesssim \Delta(\tilde{\mathbb{T}}, K).$$

In other words,  $\tilde{\mathbb{T}}$  is Frostman in  $K$  if the tubes of  $\tilde{\mathbb{T}}$  are contained in  $K$  and  $\Delta_{max}(\tilde{\mathbb{T}}) \approx \Delta(\tilde{\mathbb{T}}, K)$ .

Since we picked the set  $K$  to maximize  $\Delta(\mathbb{T}_B, K)$ , we see that  $\mathbb{T}_B[K]$  is Frostman in  $K$ . (Wang and Zahl picked the name Frostman because the condition is similar to the bound that appears in Frostman’s lemma in geometric measure theory.)

Now the proof divides into cases depending on the shape of  $K$ . Since  $K$  is a convex set which contains some tubes of  $\mathbb{T}_B$ ,  $K$  is essentially a rectangular box of dimensions  $a \times b \times r$ , where  $r$  is the radius of  $B$ . The proof divides into cases according to the values of  $a$  and  $b$ .

One important case is when  $K = B$ . In this section we focus on this important case. Then in Section 13 we consider other possible shapes of  $K$ .

### 12.4. The case $K = B$

We consider the case when  $\Delta_{max}(\mathbb{T}_B) \approx \Delta(\mathbb{T}_B, B) \gg 1$ . In this case,  $\mathbb{T}_B$  is Frostman in  $B$ .

In this case, we first prove a lower bound on  $|U(\mathbb{T}_B)| \approx |U(\mathbb{T}) \cap B|$  which leads to a lower bound for  $|U(\mathbb{T})|$ , which will imply that  $\mu(\mathbb{T}) \ll |\mathbb{T}|^\beta$ .

(Let us pause here to recall that  $U(\mathbb{T})$  and  $\mu(\mathbb{T})$  are closely related to each other:  $\mu(\mathbb{T}) = \frac{|\mathbb{T}||T|}{|U(\mathbb{T})|}$ . Using this equation, we can translate bounds on  $|U(\mathbb{T})|$  into bounds on  $\mu(\mathbb{T})$  and vice versa.)

The key ingredient in the case  $K = B$  is a lower bound for  $|U(\tilde{\mathbb{T}})|$  when  $\tilde{T}$  is a Frostman set of tubes. We state the lemma as a lower bound for a set of tubes in  $B_1$ , but we can apply it to  $\mathbb{T}_B$  by rescaling.

**LEMMA 12.2** (High density lemma). — *Suppose that the exponent  $\beta$  is as defined above. Suppose  $\tilde{\mathbb{T}}$  is a set of  $\delta$  tubes in  $B_1$  which is Frostman in  $B_1$ . Then*

$$|U(\tilde{\mathbb{T}})| \gtrsim (\delta^2 |\tilde{\mathbb{T}}|)^{1-\beta} \delta^{2\beta}.$$

Let us digest this lemma. Recall the definition of Frostman: for any convex set  $K \subset B_1$ ,  $\Delta(\tilde{\mathbb{T}}, K) \lesssim \Delta(\tilde{\mathbb{T}}, B_1)$ . If we let  $K$  be one of the tubes of  $\tilde{\mathbb{T}}$ , then we see that  $\Delta(\tilde{\mathbb{T}}, K) \geq 1$ , and therefore  $\Delta(\tilde{\mathbb{T}}, B_1) \gtrsim 1$ . This implies that  $|\tilde{\mathbb{T}}| \gtrsim \delta^{-2}$ .

In the special case that  $|\tilde{\mathbb{T}}| \approx \delta^{-2}$ , then we have  $\Delta(\tilde{\mathbb{T}}, B_1) \approx 1$ , and so  $\Delta_{max}(\tilde{\mathbb{T}}) \approx \Delta(\tilde{\mathbb{T}}, B_1) \approx 1$ . In this special case, we can apply the definition of  $\beta$  to get  $\mu(\tilde{\mathbb{T}}) \lesssim |\tilde{\mathbb{T}}|^\beta \approx \delta^{-2\beta}$ , and this gives the lower bound  $|U(\tilde{\mathbb{T}})| \gtrsim \delta^{2\beta}$ . To summarize, if  $\tilde{\mathbb{T}}$  is Frostman in  $B_1$  and  $|\tilde{\mathbb{T}}| \approx \delta^{-2}$ , then  $|U(\tilde{\mathbb{T}})| \gtrsim \delta^{2\beta}$ .

The main content of Lemma 12.2 is that if  $|\tilde{\mathbb{T}}| \gg \delta^{-2}$ , then  $|U(\tilde{\mathbb{T}})| \gg \delta^{2\beta}$ . Equivalently, if  $\tilde{\mathbb{T}}$  is Frostman in  $B_1$  and  $\Delta_{max}(\tilde{\mathbb{T}}) \gg 1$ , then  $|U(\tilde{\mathbb{T}})| \gg \delta^{2\beta}$ .

The proof of Lemma 12.2 is complex, and we will discuss it in Section 14.

Now we return to  $\mathbb{T}_B$ . We are considering the case when  $\mathbb{T}_B$  is Frostman in  $B$  and  $\Delta_{max}(\mathbb{T}_B) \gg 1$ . Rescaling and applying Lemma 12.2 to  $\mathbb{T}_B$  gives the bound

$$(24) \quad |U(\mathbb{T}) \cap B| = |U(\mathbb{T}_B)| \gg \left(\frac{\delta}{r}\right)^{2\beta} |B|.$$

In words, the high density lemma tells us that if  $\mathbb{T}_B$  is Frostman and  $\Delta_{max}(\mathbb{T}_B) \gg 1$ , then  $U(\mathbb{T})$  fills a surprisingly large fraction of  $B$ . We will use this fact to show that  $U(\mathbb{T})$  is surprisingly large, which implies that  $\mu(\mathbb{T}) \ll |\mathbb{T}|^\beta$ .

### 12.5. Density of the Kakeya set in small balls

Upper bounds for the multiplicity of the Kakeya set are closely related to lower bounds for the volume of the Kakeya set. Recall that we defined the multiplicity by

$$\mu(\mathbb{T}) = \frac{|\mathbb{T}||T|}{|U(\mathbb{T})|}.$$

Since we are assuming  $|\mathbb{T}| \approx \delta^{-2}$ ,

$$(25) \quad \mu(\mathbb{T}) \ll |\mathbb{T}|^\beta \text{ is equivalent to } |U(\mathbb{T})| \gg \delta^{2\beta}.$$

If  $B_r$  is a “typical” ball of radius  $r$  intersecting the Keakeya set, then we can write

$$(26) \quad |U(\mathbb{T})| \approx |U(\mathbb{T}_r)| \cdot \frac{|U(\mathbb{T}) \cap B_r|}{|B_r|}.$$

We will refer to  $\frac{|U(\mathbb{T}) \cap B_r|}{|B_r|}$  as the density of the Keakeya set in  $B_r$ .

Given that  $\Delta_{max}(\mathbb{T}) \lesssim 1$  and  $|\mathbb{T}| \approx \delta^{-2}$ , it is not hard to show that for all  $\delta \leq r \leq 1$  we have

$$(27) \quad |U(\mathbb{T}_r)| \gtrsim r^{2\beta}.$$

In the special case when  $\mathbb{T}_B$  is Frostman and  $\Delta_{max}(\mathbb{T}_B) \gg 1$ , the high density lemma gives us (24):

$$(28) \quad \frac{|U(\mathbb{T}) \cap B_r|}{|B_r|} \gg \left(\frac{\delta}{r}\right)^{2\beta}.$$

Putting together the last three indented equations gives  $|U(\mathbb{T})| \gg \delta^{2\beta}$ , which is equivalent to  $\mu(\mathbb{T}) \ll |\mathbb{T}|^\beta$ .

Let us take stock. If  $\mathbb{T}$  is not sticky (at all scales), then our goal is to prove that  $\mu(\mathbb{T}) \ll |\mathbb{T}|^\beta$ . So far we have achieved this goal in two cases: the case when  $\Delta_{max}(\mathbb{T}_B) \lesssim 1$  and the case when  $\Delta_{max}(\mathbb{T}_B) \approx \Delta(\mathbb{T}_B, B) \gg 1$ . Recall that we defined  $K \subset B$  to be the set that maximizes  $\Delta(\mathbb{T}_B, K)$ . Our two cases are the case when  $K$  is a tube  $T_B \in \mathbb{T}_B$  and the case when  $K = B$ . In the next section we discuss other shapes of  $K$ .

### 13. Other possible shapes of $K$

Suppose that  $\Delta_{max}(\mathbb{T}_B) = \Delta(\mathbb{T}_B, K)$ . In general  $K$  is a convex set of dimensions  $a \times b \times r$ . So far we have discussed the two extreme cases:  $K = T_B$  and  $K = B$ . Now we turn to various intermediate cases. By doing some pigeonholing, we can assume that there is a set  $\mathbb{K}$  of convex sets  $K \subset B$ , all with the same dimensions, so that  $\Delta(\mathbb{T}_B, K) \approx \Delta_{max}(\mathbb{T})$  for each  $K \in \mathbb{K}$  and so that each  $T_B \in \mathbb{T}_B$  lies in one  $K \in \mathbb{K}$ .

Essentially this set  $\mathbb{K}$  is formed by using the greedy algorithm. First we find the convex set  $K_1$  which maximizes  $\Delta(\mathbb{T}_B, K_1)$ . We call  $K_1$  the densest convex set for  $\mathbb{T}_B$ . Then we remove  $\mathbb{T}_B[K_1]$  from  $\mathbb{T}_B$ , and we let  $K_2$  be the densest convex set for the remaining tubes. We continue in this way until all (or most) of the tubes of  $\mathbb{T}_B$  belong to one of the sets  $K \in \mathbb{K}$ .

From the nature of this procedure, it follows that  $\mu(\mathbb{K}) \lesssim 1$ . The reason is that if many sets  $K \in \mathbb{K}$  pack into a larger convex set  $L$ , then we would have  $\Delta(\mathbb{T}_B, L) \gg \Delta(\mathbb{T}_B, K)$ ,

and so the greedy algorithm would have chosen  $L$  instead of  $K$ . Also, for each  $K \in \mathbb{K}$ ,  $\mathbb{T}_B[K]$  is Frostman.

The argument divides into cases according to the shape of  $K$ .

**Case 1: Fat planks.**

If  $a \gg \delta$ , then we can adapt the high density argument from the last subsection to show that  $U(\mathbb{T}) \cap K$  fills a surprisingly large fraction of  $K$  and hence  $U(\mathbb{T}) \cap B_a$  fills a surprisingly large fraction of  $B_a$ . This requires some technical care, but the high level ideas are the same as in the case  $K = B$  above, and the most ingredient is the high density lemma, Lemma 12.2.

This leaves the case when  $a \approx \delta$ , so  $K$  has dimensions roughly  $\delta \times b \times r$ . When we study how the tubes  $\mathbb{T}_B[K]$  sit inside of  $K$ , we essentially have a two-dimensional Kakeya problem. Using the  $L^2$  method, it follows that  $|U(\mathbb{T}_B[K])| \gtrsim |K|$ , and so  $\mathbb{T}_B[K]$  essentially fills  $K$ .

If  $b \ll r$ , then we call  $K$  a thin plank, and if  $b \approx r$ , then we call  $K$  a thin slab. In the thin plank case, we have to study how the planks of  $\mathbb{K}$  intersect each other. Each plank  $K \in \mathbb{K}$  has a two-dimensional tangent plane, spanned by the two longest axes of  $K$ . When two planks intersect, we call the intersection tangential if the two tangent planes are equal and we call the intersection transverse if the two tangent planes are transverse.

**Case 2: Thin planks intersecting transversely.**

When most intersections are transverse, then the  $L^2$  method gives strong bounds. If we intersect a plank with a ball of radius  $b$ , we get a slab. Inside such a ball  $B_b$ , we would see several slabs intersecting transversely. Using the  $L^2$  method, it follows that  $U(\mathbb{K})$  essentially fills  $B_b$  and so  $U(\mathbb{T})$  essentially fills  $B_b$ . Then we can finish the proof as in Subsection 12.5.

**Case 3: Thin planks intersecting tangentially.**

In this case, we will relate  $\mu(\mathbb{T})$  to  $\mu(\mathbb{K})$ . We will bound  $\mu(\mathbb{K})$  by changing coordinates so that the planks become a set of tubes  $\mathbb{T}'$ , and we bound  $\mu(\mathbb{T}')$  using induction. Here we sketch how to relate  $\mathbb{K}$  to  $\mathbb{T}'$ .

If all intersections are tangential, then all the planks  $K$  that intersect a given plank  $K_0$  lie in a slab  $S$  of dimensions  $\frac{\delta r}{b} \times r \times r$ . In fact, each plank  $K$  that intersects  $K_0$  lies in  $S$  and also has tangent plane close to the tangent plane of  $S$ . We let  $\mathbb{K}_S$  be the set of planks  $K$  so that  $K \subset S$  and the tangent plane of  $K$  is close to that of  $S$ . When most intersections are tangential, then  $\mu(\mathbb{K}) \approx \mu(\mathbb{K}_S)$ .

Now we can change coordinates so that  $S$  becomes the unit cube and so each  $K \in \mathbb{K}_S$  becomes a tube  $T'$  in the unit cube.

$$\text{Planks } K \in \mathbb{K}_S \longleftrightarrow \text{tubes in } B_1.$$

Figure 8 illustrates this correspondence:

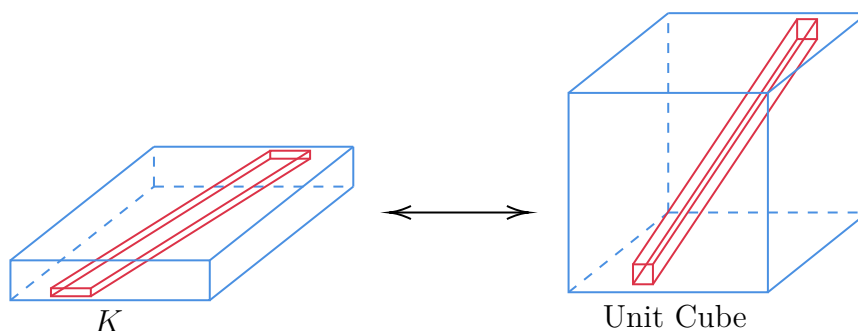


FIGURE 8. Correspondence between planks in  $K$  and tubes in  $B_1$

In Figure 8, the plank on the left is a plank  $K \subset S$ . Under the linear change of variables, the plank  $K$  corresponds to the tube on the right.

We let  $\mathbb{T}'$  be the set of tubes on the right. Then we have  $\mu(\mathbb{K}) \approx \mu(\mathbb{K}_S) \approx \mu(\mathbb{T}')$ . We also have  $\Delta_{max}(\mathbb{T}') \approx \Delta_{max}(\mathbb{K}_S) \leq \Delta_{max}(\mathbb{K}) \lesssim 1$ . Therefore we can bound  $\mu(\mathbb{T}') \lesssim |\mathbb{T}'|^\beta$ .

With some additional computation, this leads to a bound for  $\mu(\mathbb{T})$  which gives the desired improvement  $\mu(\mathbb{T}) \ll |\mathbb{T}|^\beta$ . This computation is similar to the proof of Lemma 12.1, although a little more complicated. We omit the details.

#### Case 4: Thin slabs.

Finally we come to the thin slab case when  $K$  has dimensions roughly  $\delta \times r \times r$ . In the thin slab case, we can get very strong bounds as long as  $r$  is close to 1. If  $r$  is close to 1, then the assumption  $\Delta_{max}(\mathbb{T}) \lesssim 1$  guarantees that not too many tubes of  $\mathbb{T}$  can heavily intersect  $K$ , and so it will follow that there are many thin slabs  $K$ . The intersections of thin slabs with each other are well controlled by the  $L^2$  method. Therefore, if  $r$  is close to 1, then  $|U(\mathbb{K})|$  must be close to 1, and hence  $|U(\mathbb{T})|$  must be close to 1 also.

Recall from the start of Section 12.1 that we chose an angle  $\rho \in [\delta, 1]$  with  $|\mathbb{T}_\rho| \gg \rho^{-2}$ , and the ball  $B$  has radius  $r = \delta/\rho$ . In order to make  $r$  close to 1, we need to choose  $\rho$  close to  $\delta$ . For this reason, we need to know that  $\mathbb{T}$  is not sticky at scales  $\rho$  very close to  $\delta$ . Also, when we fill in the details in Case 3, we actually need to know that  $\mathbb{T}$  is not sticky at all scales  $\rho$  with  $\delta \ll \rho \ll 1$ . This was the reason that we focused on the not-sticky-at-all-scales case above.

### 13.1. How do we reduce to the not-sticky-at-all-scales case?

We saw in the argument above that it was important to reduce to the not-sticky-at-all-scales case. In this subsection, we explain how to do that. We start by recalling the sticky case, the not-sticky case, and the not-sticky-at-all-scales case.

Recall that  $\mathbb{T}$  is a set of  $\delta$ -tubes in  $B_1$  with  $\Delta_{max}(\mathbb{T}) \lesssim 1$  and that  $|\mathbb{T}| \approx \delta^{-2}$ .

The sticky case means that for every  $\rho \in [\delta, 1]$ ,  $\Delta_{max}(\mathbb{T}_\rho) \lesssim 1$ . Since  $|\mathbb{T}| \approx \delta^{-2}$ , this is equivalent to  $|\mathbb{T}[T_\rho]| \approx (\delta/\rho)^{-2}$ .

If  $\mathbb{T}$  is not sticky, it means that there is some  $\rho \in [\delta, 1]$  so that  $|\mathbb{T}[T_\rho]| \ll (\delta/\rho)^{-2}$ . Such a  $\rho$  must lie in the range  $\delta \ll \rho \ll 1$ .

We say that  $\mathbb{T}$  is not-sticky-at-all-scales if  $|\mathbb{T}[T_\rho]| \ll (\delta/\rho)^{-2}$  for every  $\rho$  in the range  $\delta \ll \rho \ll 1$ .

Here is the rough idea how to reduce the not-sticky case to the not-sticky-at-all-scales case. Suppose that there is some scale  $\rho$  so that  $|\mathbb{T}[T_\rho]| \approx (\delta/\rho)^{-2}$ . It follows that  $\Delta_{max}(\mathbb{T}_\rho) \lesssim 1$ . Now we can bound

$$\mu(\mathbb{T}) \lesssim \mu(\mathbb{T}[T_\rho])\mu(\mathbb{T}_\rho),$$

which reduces our original problem to two similar problems at smaller scales. We try to keep reducing in this way. If one of the smaller problems is not-sticky-at-all-scales then we are stuck and we cannot reduce further. Otherwise we can reduce further. If we can keep reducing in this way to very small problems, it means that our original set of tubes  $\mathbb{T}$  was sticky, and we can handle it using the sticky Kakeya theorem. Otherwise, we get stuck with a problem that is not-sticky-at-all-scales and we can handle it using the argument we have sketched in the last two sections.

#### 14. The high density lemma

We now come to the last ingredient in the proof that a worst case Kakeya set must be sticky: the proof of the high density lemma. The Frostman condition plays a key role in the high density lemma, so we recall the Frostman condition and put it in a slightly more general context. Suppose that  $\mathbb{W}$  is a set of convex sets  $W$  all lying in a given convex set  $U$ . We say that  $\mathbb{W}$  is Frostman in  $U$  if  $\Delta_{max}(\mathbb{W}) \approx \Delta(\mathbb{W}, U)$ . In the statement of the high density lemma,  $\mathbb{W}$  will be a set of tubes, but later in our discussion we will see more general convex sets.

LEMMA (High density lemma). — *Suppose that the exponent  $\beta$  is as defined above. Suppose  $\mathbb{T}$  is a set of  $\delta$  tubes which is Frostman in  $B_1$ . Then*

$$|U(\mathbb{T})| \gtrsim (\delta^2|\mathbb{T}|)^{1-\beta} \delta^{2\beta}.$$

*Equivalently,*

$$\mu(\mathbb{T}) \lesssim (\delta^2|\mathbb{T}|)^{1-\beta} \delta^{-2\beta}.$$

To digest this lemma we start with the case when  $|\mathbb{T}| \sim \delta^{-2}$ . If  $|\mathbb{T}| \sim \delta^{-2}$ , then  $\Delta(\mathbb{T}, B_1) \sim 1$ . Since  $\mathbb{T}$  is Frostman,  $\Delta_{max}(\mathbb{T}) \approx \Delta(\mathbb{T}, B_1) \sim 1$ , and so  $|U(\mathbb{T})| \gtrsim \delta^{2\beta}$ . This matches the conclusion of the high-density lemma when  $|\mathbb{T}| \sim \delta^{-2}$ . The content of the high density lemma is that when  $\mathbb{T}$  is Frostman and  $|\mathbb{T}| \gg \delta^{-2}$ ,  $|U(\mathbb{T})|$  is much bigger than  $\delta^{2\beta}$ .

If  $\mathbb{T}$  is Frostman with  $|\mathbb{T}| \gg \delta^{-2}$ , it is easy to prove that  $|U(\mathbb{T})| \gtrsim \delta^{2\beta}$ . Randomly decompose  $\mathbb{T}$  as a disjoint union  $\mathbb{T} = \sqcup_j \mathbb{T}_j$ , where  $|\mathbb{T}_j| \sim \delta^{-2}$ . Since  $\mathbb{T}$  obeys the

Frostman condition, it is not hard to check that  $\mathbb{T}_j$  also obeys the Frostman condition. Since  $|\mathbb{T}_j| \sim \delta^{-2}$ , we checked above that  $|U(\mathbb{T}_j)| \gtrsim \delta^{2\beta}$ , and so

$$(29) \quad |U(\mathbb{T})| \gtrsim |U(\mathbb{T}_j)| \gtrsim \delta^{2\beta}.$$

Even a small improvement on this trivial bound is enough to power the inductive argument in the last sections. If (29) were sharp, it would mean that for each  $j$ ,  $|U(\mathbb{T})| \approx |U(\mathbb{T}_j)|$ . This sounds intuitively unlikely: when  $|\mathbb{T}|$  is far bigger than  $|\mathbb{T}_j|$ , we might expect  $|U(\mathbb{T})|$  to be at least a little bigger than  $|U(\mathbb{T}_j)|$ . However, it is not easy to prove this.

In the proof, we will work with the formulation involving  $\mu(\mathbb{T})$ . The two formulations are equivalent because of the definition  $\mu(\mathbb{T}) = \frac{|\mathbb{T}||U|}{|U(\mathbb{T})|}$ . Let  $\gamma$  be the smallest exponent so that, if  $\mathbb{T}$  is a Frostman set of tubes in  $B_1$ , then

$$(30) \quad \mu(\mathbb{T}) \lesssim (\delta^{-2})^\beta (\delta^2 |\mathbb{T}|)^\gamma.$$

So Lemma 12.2 says that  $\gamma = 1 - \beta$ . To power the inductive argument in the last sections, we just need to prove that  $\gamma < 1$  (assuming that  $\beta > 0$ ). We say that  $\mathbb{T}$  is a worst-case example for Lemma 12.2 if

$$(31) \quad \mu(\mathbb{T}) \approx (\delta^{-2})^\beta (\delta^2 |\mathbb{T}|)^\gamma.$$

### 14.1. Looking for stickyakeya sets

A crucial input to the proof is the stickyakeya theorem. One scenario is that  $\mathbb{T}$  contains a stickyakeya set  $\mathbb{T}'$ . In this case,  $|U(\mathbb{T})| \geq |U(\mathbb{T}')| \gtrsim 1$ , and we are done. When does  $\mathbb{T}$  contain a stickyakeya set? Recall the definition of a stickyakeya set. A set of tubes  $\mathbb{T}$  is a stickyakeya set if:

- $|\mathbb{T}_\rho| \approx \rho^{-2}$  for all  $\rho \in [\delta, 1]$
- $\Delta_{max}(\mathbb{T}) \lesssim 1$ .

To look for a stickyakeya set, it helps to consider a dense sequence of scales  $1 = \rho_0 > \rho_1 > \dots > \rho_N = \delta$ . We say the sequence of scales is dense if each quotient  $\frac{\rho_{j-1}}{\rho_j}$  is very small. We let  $\mathbb{T}_{\rho_j}[T_{\rho_{j-1}}]$  be the set of all  $T_{\rho_j} \in \mathbb{T}_{\rho_j}$  lying in the thicker tube  $T_{\rho_{j-1}} \in \mathbb{T}_{\rho_{j-1}}$ . The definition of sticky can be rephrased in terms of these sets  $\mathbb{T}_{\rho_j}[T_{\rho_{j-1}}]$ . If the sequence of scales is dense enough, then  $\mathbb{T}$  is sticky if and only if for every  $j$ ,

- $|\mathbb{T}_{\rho_j}[T_{\rho_{j-1}}]| \approx \left(\frac{\rho_{j-1}}{\rho_j}\right)^2$
- $\Delta_{max}(\mathbb{T}_{\rho_j}[T_{\rho_{j-1}}]) \lesssim 1$ . Equivalently,  $\mathbb{T}_{\rho_j}[T_{\rho_{j-1}}]$  is Frostman.

Not every Frostman set of tubes in  $B_1$  contains a stickyakeya set  $\mathbb{T}'$ . From the characterization of stickyakeya sets above, we see that if  $\mathbb{T}$  does contain a stickyakeya set, then for each  $j$ ,  $\mathbb{T}_{\rho_j}[T_{\rho_{j-1}}]$  must contain a Frostman set of tubes. This can fail. On the other hand, if each set  $\mathbb{T}_{\rho_j}[T_{\rho_{j-1}}]$  is Frostman, then  $\mathbb{T}$  does contain a sticky subset  $\mathbb{T}'$ . We state this as a lemma.

LEMMA 14.1. — *If  $\mathbb{T}_{\rho_j}[T_{\rho_{j-1}}]$  is Frostman in  $T_{\rho_{j-1}}$  for each  $j$  and each  $T_{\rho_{j-1}} \in \mathbb{T}_{\rho_{j-1}}$ , then  $\mathbb{T}$  contains a subset  $\mathbb{T}'$  which is a sticky Keakeya set. In fact, we can even decompose  $\mathbb{T}$  as  $\mathbb{T} = \sqcup_j \mathbb{T}_j$  where each  $\mathbb{T}_j$  is a sticky Keakeya set.*

*Proof sketch.* — Since  $\mathbb{T}_{\rho_1}$  is Frostman,  $|\mathbb{T}_{\rho_1}| \gtrsim \rho_1^{-2}$ . Choose a random subset  $\mathbb{T}'_{\rho_1} \subset \mathbb{T}_{\rho_1}$  with  $|\mathbb{T}'_{\rho_1}| \sim \rho_1^{-2}$ . Then  $\mathbb{T}'_{\rho_1}$  is also Frostman.

For each  $T_{\rho_1} \in \mathbb{T}'_{\rho_1}$ , note that  $\mathbb{T}_{\rho_2}[T_{\rho_1}]$  is Frostman and so  $|\mathbb{T}_{\rho_2}[T_{\rho_1}]| \gtrsim (\rho_1/\rho_2)^2$ . Choose a random subset  $\mathbb{T}'_{\rho_2}[T_{\rho_1}] \subset \mathbb{T}_{\rho_2}[T_{\rho_1}]$  with  $|\mathbb{T}'_{\rho_2}[T_{\rho_1}]| \sim (\rho_1/\rho_2)^2$ . Then  $\mathbb{T}'_{\rho_2}[T_{\rho_1}]$  is also Frostman. We set  $\mathbb{T}'_{\rho_2} = \cup_{\rho_1 \in \mathbb{T}'_{\rho_1}} \mathbb{T}'_{\rho_2}[T_{\rho_1}]$ .

Proceeding in this way, we define  $\mathbb{T}'$ , and  $\mathbb{T}'$  is sticky because of the criterion above.

If we choose random decompositions instead of just random subsets then we get a decomposition  $\mathbb{T} = \sqcup_j \mathbb{T}_j$  where each  $\mathbb{T}_j$  is sticky. □

This raises the question whether we can find a dense sequence of scales  $1 = \rho_0 > \rho_1 > \dots > \rho_N = \delta$  so that for each  $j$ ,  $\mathbb{T}_{\rho_j}[T_{\rho_{j-1}}]$  is Frostman in  $T_{\rho_{j-1}}$ . We call such a sequence a good sequence of scales. We can try to build a good sequence of scales by adding one scale at a time. The process boils down to asking: is there a scale  $\rho$  with  $1 \gg \rho \gg \delta$  so that  $\mathbb{T}_\rho$  is Frostman and  $\mathbb{T}[T_\rho]$  is Frostman? Call such a  $\rho$  a good scale.

Since  $\mathbb{T}$  is Frostman in  $B_1$ , it follows that  $\mathbb{T}_\rho$  is also Frostman in  $B_1$ . The reason is that if  $\Delta(\mathbb{T}_\rho, K) \gg \Delta(\mathbb{T}_\rho, B_1)$ , then it would follow that  $\Delta(\mathbb{T}, K) \gg \Delta(\mathbb{T}, B_1)$ .

But  $\mathbb{T}[T_\rho]$  is not necessarily Frostman. Since  $\mathbb{T}$  is Frostman in  $B_1$ , we do know that for any convex set  $K$ ,  $\Delta(\mathbb{T}, K) \lesssim \Delta(\mathbb{T}, B_1)$ . But if  $\Delta(\mathbb{T}, T_\rho) \ll \Delta(\mathbb{T}, B_1)$ , then there could be some  $K \subset T_\rho$  with  $\Delta(\mathbb{T}, K) \gg \Delta(\mathbb{T}, T_\rho)$ .

If  $\mathbb{T}[T_\rho]$  is Frostman, then we have a good scale and we can start to build a good sequence of scales which would lead to a sticky Keakeya set  $\mathbb{T}' \subset \mathbb{T}$ . So we need to analyze the case when  $\mathbb{T}[T_\rho]$  is not Frostman and find some useful structure there.

### 14.2. Analyzing the case that $\mathbb{T}[T_\rho]$ is not Frostman

If  $\mathbb{T}[T_\rho]$  is not Frostman, then it means by definition that there is some subset  $W \subset T_\rho$  so that  $\Delta(\mathbb{T}[T_\rho], W) \gg \Delta(\mathbb{T}[T_\rho], T_\rho)$ . As in Section 13, it is helpful to organize  $\mathbb{T}[T_\rho]$  by choosing sets  $W$  that maximize  $\Delta(\mathbb{T}[T_\rho], W)$ . As in the beginning of Section 13, we can find a set  $\mathbb{W}(T_\rho)$  of such maximal  $W$  by choosing them one at a time until each tube  $T \in \mathbb{T}[T_\rho]$  lies in  $\approx 1$   $W \in \mathbb{W}(T_\rho)$ .

We can assume that each  $W \in \mathbb{W}(T_\rho)$  has roughly the same dimensions: each  $W$  has dimensions roughly  $a \times b \times 1$ , where  $\delta \leq a \leq b \leq \rho$ . We will see that if  $\mathbb{T}$  is a worst-case Keakeya set, then it strongly constrains the geometry of  $W$ .

LEMMA 14.2. — *If  $\mathbb{T}$  is a worst-case example for Lemma 12.2 as in (31) and  $\mathbb{W}[T_\rho]$  is defined as above, then  $a \approx b$ . So each  $W \in \mathbb{W}(T_\rho)$  is approximately a tube  $T_a$  of radius  $a$  and length 1.*

*Proof sketch.* — For each  $T_\rho \in \mathbb{T}[T_\rho]$ , we have defined  $\mathbb{W}(T_\rho)$ . We gather all these sets together to form  $\mathbb{W} = \cup_{T_\rho \in \mathbb{T}_\rho} \mathbb{W}(T_\rho)$ . Now each  $T \in \mathbb{T}$  lies in  $\approx 1$   $W \in \mathbb{W}$ . Therefore we have  $|\mathbb{T}| \approx |\mathbb{T}[W]| |\mathbb{W}|$  and we can bound  $\mu(\mathbb{T})$  by

$$(32) \quad \mu(\mathbb{T}) \lesssim \mu(\mathbb{T}[W])\mu(\mathbb{W}).$$

Moreover,  $\mathbb{T}[W]$  and  $\mathbb{W}$  are each Frostman. The set  $\mathbb{T}[W]$  is Frostman because we chose  $W$  to maximize  $\Delta(\mathbb{T}[T_\rho], W)$ . To see that  $\mathbb{W}$  is Frostman, we arrange by some pigeonholing arguments that the sets  $\mathbb{T}[W]$  have roughly the same cardinality. Now if  $K \subset B_1$  has  $\Delta(\mathbb{W}, K) \gg \Delta(\mathbb{W}, B_1)$  it would imply  $\Delta(\mathbb{T}, K) \gg \Delta(\mathbb{T}, B_1)$ . Since  $\mathbb{T}$  is Frostman, it follows that  $\mathbb{W}$  is Frostman too.

Since  $\mathbb{T}[W]$  and  $\mathbb{W}$  are each Frostman, it almost looks like we could bound  $\mu(\mathbb{T}[W])$  and  $\mu(\mathbb{W})$  by induction using (30). That's not quite possible because (30) applies to a set of tubes which is Frostman in a ball, whereas  $\mathbb{T}[W]$  is a set of tubes that is Frostman in a convex set  $W$ , and  $\mathbb{W}$  is a set of convex sets which is Frostman in a ball. In order to use induction, we generalize Lemma 12.2 and (30) to a set of convex sets which is Frostman in another convex set. This more general version then applies to both  $\mathbb{T}[W]$  and  $\mathbb{W}$ .

Suppose that  $\mathbb{V}$  is a set of convex sets which is Frostman in a convex set  $U$ . By a linear change of variables, we can reduce to the case that  $U = B_1$ . Then we examine the dimensions of  $V \in \mathbb{V}$ . Say each  $V \in \mathbb{V}$  has dimensions  $a_{\mathbb{V}} \times b_{\mathbb{V}} \times 1$ . If  $a_{\mathbb{V}} \approx b_{\mathbb{V}}$ , then  $V$  is a tube and we can apply (30). The other extreme example is when  $V$  is a slab of dimensions  $a_{\mathbb{V}} \times 1 \times 1$ . Sharp bounds for intersecting slabs have been known for a long time by the  $L^2$  method. So in the slab case, we get very strong bounds for  $\mu(\mathbb{V})$ . This leaves an intermediate case when  $V$  has dimensions  $a_{\mathbb{V}} \times b_{\mathbb{V}} \times 1$  with  $a_{\mathbb{V}} \ll b_{\mathbb{V}} \ll 1$ . In this case, the shape of  $V$  is intermediate between a tube and a slab. In fact, at large scales  $V$  looks like a tube, but if we intersect  $V$  with a small ball, then it looks like a slab. In this case,  $\mu(\mathbb{V})$  can be bounded by a multiscale argument that combines estimates for tubes from (30) with estimates for slabs from the  $L^2$  method.

We apply this method to bound  $\mu(\mathbb{T}[W])$  and  $\mu(\mathbb{W})$  and then plug those bounds into (32) to bound  $\mu(\mathbb{T})$ . The computation shows that  $\mu(\mathbb{T}) \ll (\delta^{-2})^\beta (\delta^2 |\mathbb{T}|)^\gamma$  unless  $a \approx b$ . In other words, if  $\mathbb{T}$  is a worst-case example for Lemma 12.2, then  $a \approx b$ .

Morally, the reason that the computation works out this way is that the bounds for slabs are sharp. When we unwind the argument above, we bound  $\mu(\mathbb{T})$  by applying (30) at some scales and the  $L^2$  bounds for slabs at other scales. Since the bound for slabs is so strong, this improves on (30) as long as we use the slab bound at some scales. When  $a \approx b$ , then all the convex sets in the argument are tubes, and the slab bound never appears. But if  $a \ll b$ , then each  $W \in \mathbb{W}$  looks like a plank and there are some scales where the slab bounds come into play.  $\square$

### 14.3. Finding a good scale

Lemma 14.2 leads to a general condition for finding a good scale.

LEMMA 14.3. — *Suppose that  $\mathbb{T}$  is a worst-case example for the high density lemma, as in (31). If  $|\mathbb{T}| \gg \delta^{-2}$ , then there is a good scale  $a$ . Recall this means that*

- $\delta \ll a \ll 1$ .
- $\mathbb{T}[T_a]$  is Frostman and  $\mathbb{T}_a$  is Frostman

*Moreover,  $\mathbb{T}[T_a]$  and  $\mathbb{T}_a$  will also be worst-case examples for the high density lemma.*

*Proof sketch.* — Since  $|\mathbb{T}| \gg \delta^{-2}$ , we can choose  $\rho \ll 1$  so that  $|\mathbb{T}[T_\rho]| \gg (\rho/\delta)^2$  and so  $\Delta(\mathbb{T}, T_\rho) \gg 1$ .

We define  $\mathbb{W}(T_\rho)$  as above. Recall that each  $W \in \mathbb{W}(T_\rho)$  maximizes  $\Delta(\mathbb{T}[T_\rho], W)$ , and so  $\mathbb{T}[W]$  is Frostman.

Each  $W$  has dimensions  $a \times b \times 1$ . By Lemma 14.2,  $a \approx b$ , and so  $W$  is essentially a tube of radius  $a$ ,  $T_a$ . We also have  $a \leq \rho \ll 1$ .

Recall that  $\mathbb{W} = \cup_{T_\rho \in \mathbb{T}} \mathbb{W}(T_\rho)$ , and that each  $T \in \mathbb{T}$  lies in  $\approx 1$  set  $W \in \mathbb{W}$ . Since each  $W$  is a tube of radius  $a$ , we must have  $\mathbb{W} = \mathbb{T}_a$ .

Since  $\mathbb{T}[W]$  is Frostman for each  $W \in \mathbb{W}$ , we see that  $\mathbb{T}[T_a]$  is Frostman for each  $T_a \in \mathbb{T}_a$ . On the other hand, since  $\mathbb{T}$  is Frostman it implies that  $\mathbb{T}_a$  is Frostman.

Next we have to check that  $a \gg \delta$ . Recall that we chose  $\rho$  so that  $\Delta(\mathbb{T}[T_\rho], T_\rho) \gg 1$ . But  $\Delta(\mathbb{T}[W], W) \approx \Delta_{\max}(\mathbb{T}[T_\rho]) \geq \Delta(\mathbb{T}[T_\rho], T_\rho) \gg 1$ . But if  $a \approx \delta$ , then  $W = T_a$  would essentially be a  $\delta$ -tube, and then we would have  $\Delta(\mathbb{T}[W], W) \approx 1$ . Therefore, we must have  $a \gg \delta$  as desired. This shows that  $a$  is a good scale.

Finally we sketch the proof that  $\mathbb{T}[T_a]$  and  $\mathbb{T}_a$  are both worst-case examples for the high density lemma. Recall that

$$\mu(\mathbb{T}) \lesssim \mu(\mathbb{T}[T_a])\mu(\mathbb{T}_a).$$

Since  $\mathbb{T}[T_a]$  and  $\mathbb{T}_a$  are both Frostman, we can bound  $\mu(\mathbb{T}[T_a])$  and  $\mu(\mathbb{T}_a)$  using (30), the definition of the exponent  $\gamma$ . When we plug in and simplify the right-hand side, we get  $\delta^{-2\beta} (\delta^2 |\mathbb{T}|)^\gamma$ , and since  $\mathbb{T}$  is worst-case, this is  $\approx \mu(\mathbb{T})$ . Therefore, every step in this chain must be an approximate equality. In particular this means that  $\mu(\mathbb{T}[T_a])$  and  $\mu(\mathbb{T}_a)$  must be worst-case for the high density lemma.  $\square$

### 14.4. Multiscale decomposition

By applying Lemma 14.3 repeatedly, we can choose a sequence of scales  $1 = \rho_0 \gg \rho_1 \gg \rho_2 \gg \dots \gg \rho_N = \delta$  so that  $\mathbb{T}_{\rho_j}[T_{\rho_{j-1}}]$  is always Frostman, and for each  $j$  one of the following holds:

1.  $\rho_{j-1}/\rho_j$  is very small, or
2.  $|\mathbb{T}_{\rho_j}[T_{\rho_{j-1}}]| \approx \left(\frac{\rho_{j-1}}{\rho_j}\right)^2$ .

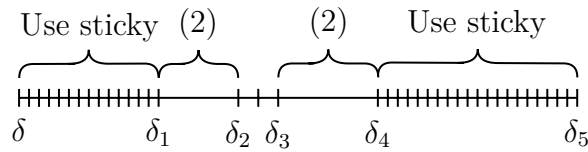


FIGURE 9. Key scales

Figure 9 is a picture showing how this sequence of scales may look:

Here each short vertical line represents a scale  $\rho_j$ . These scales are generally quite close together, except for two significant gaps. Each significant gap must be in case (2), and so we labelled them (2).

If  $|\mathbb{T}| \gg \delta^{-2}$ , not every interval can be in case 2 – a definite fraction of intervals must be in case 1. Since the intervals in case 1 are very small, a definite fraction of scales must consist of many small intervals. On any such block of small intervals, we can apply sticky Kakeya, giving a very strong bound. In our picture, we have drawn two such blocks of small intervals, and they are labelled “use sticky Kakeya”.

To bound  $\mu(\mathbb{T})$ , we begin by factoring  $\mu(\mathbb{T})$  into contributions coming from different scale ranges. For instance, in the scenario illustrated in figure 5, we would bound  $\mu(\mathbb{T})$  by

$$(33) \quad \mu(\mathbb{T}) \lesssim \mu(\mathbb{T}[T_{\delta_1}])\mu(\mathbb{T}_{\delta_1}[T_{\delta_2}])\mu(\mathbb{T}_{\delta_2}[T_{\delta_3}])\mu(\mathbb{T}_{\delta_3}[T_{\delta_4}])\mu(\mathbb{T}_{\delta_4}).$$

The five factors on the right-hand side correspond to five scale ranges in the picture. Note that if we used the trivial bound (29) to bound each factor on the right-hand side, then we would get back the trivial bound. Instead, we use sticky Kakeya on the scale ranges labelled sticky. On the other scale ranges, we use the trivial bound. But since  $\beta > 0$ , the bound in the sticky case is better than the trivial bound, and so our overall bound for  $\mu(\mathbb{T})$  is better than the trivial bound (29). A careful calculation gives the value  $\gamma = 1 - \beta$ .

This finishes the outline of the proof of the high density lemma.

Notice that in this argument, we broke the scales from  $\delta$  to 1 into several ranges in a strategic way that was tailored to the geometry of  $\mathbb{T}$ . The idea of choosing these ranges strategically was introduced by Keleti and Shmerkin (2019), and it has become a major tool in this circle of problems. For instance, it plays an important role in the solution of the Furstenberg set problem in Keleti and Shmerkin (2019) and Ren and Wang (2023).

## 15. Final recap

At the beginning of the survey, we said that the hero of the proof was multiscale analysis. We begin with a worst-case Kakeya set  $\mathbb{T}$ , with  $\mu(\mathbb{T}) \approx |\mathbb{T}|^\beta$ . We then relate  $\mathbb{T}$

to many other sets of tubes  $\mathbb{T}'$  obeying  $\Delta_{max}(\mathbb{T}') \lesssim 1$ . By definition of  $\beta$ , we know that  $\mu(\mathbb{T}') \lesssim |\mathbb{T}'|^\beta$ , and this gives us information about  $\mathbb{T}$ .

As a final recap, let us look back and see how we found all these sets of tubes  $\mathbb{T}'$ . We used several different ways to relate a set of tubes  $\mathbb{T}$  to a new set of tubes, or more generally to a new set of convex sets  $\mathbb{W}$ . If we start with  $\mathbb{T}$ , we can

- Look at thicker tubes  $\mathbb{T}_\rho$ .
- Look at  $\mathbb{T}[T_\rho]$ .
- Intersect tubes of  $\mathbb{T}$  with a smaller ball  $B$  and look at  $\mathbb{T}_B$ .
- Find convex sets  $K$  that maximize  $\Delta(\mathbb{T}, K)$ . Then we can look at  $\mathbb{T}[K]$ . Also, we can form a set  $\mathbb{K}$  of these convex sets  $K$  and look at  $\mathbb{K}$ . Sometimes we can change coordinates to convert  $\mathbb{K}$  to a set of tubes.

These operations can be chained together. For instance, starting with  $\mathbb{T}$  we might first look at  $\mathbb{T}_B$ . Then we might find convex sets  $K$  that maximize  $\Delta(\mathbb{T}_B, K)$  and study  $\mathbb{K}$ . After some coordinate changes, we might be led to a new set of tubes  $\mathbb{T}'$ . Or starting with  $\mathbb{T}$  we might first look at  $\mathbb{T}_B$  and then look at a thicker set of tubes  $\mathbb{T}_{B,\rho}$ . Starting at  $\mathbb{T}$ , we may need to use several operations to arrive at a new set  $\mathbb{T}'$  that obeys  $\Delta_{max}(\mathbb{T}') \lesssim 1$ . By combining information from many such sets of tubes  $\mathbb{T}'$ , the proof shows that if  $\beta > 0$ , the set  $\mathbb{T}$  would have to have special geometric and algebraic structure, closely matching the Heisenberg group.

We also note that the argument is essentially by induction. If we unwind the induction, then the argument effectively uses the above operations many times to get from the initial set of tubes  $\mathbb{T}$  to other sets of tubes  $\mathbb{T}'$ .

At the beginning of the survey, we said that the proof of the Kakeya problem is based on studying the problem at many scales. Looking at a problem at many scales has been a central theme in harmonic analysis for a hundred years, and this proof can be regarded as part of that tradition. But when we said that the proof is based on studying the problem at many scales, we really meant that the proof is based on bringing into play all the sets of tubes  $\mathbb{T}'$  that can be built from  $\mathbb{T}$  by the operations above. This version of multiscale analysis is much newer. It was built over the last twenty-five years, with important parts of the picture appearing just in the last few years.

To finish this essay, let us return to the Katz–Zahl example. As we mentioned in Section 10, Katz and Zahl found a cousin problem to the Kakeya problem where the analogue of Kakeya does not hold but the analogue of the sticky case appears likely to hold. That example made me think it was unlikely that the Kakeya problem could be reduced to the sticky case. Wang and Zahl did reduce the general Kakeya problem to the sticky case, and so it is natural to ask why their method does not apply to the Katz–Zahl example.

The Katz–Zahl example concerns a cousin of the Kakeya problem where  $\mathbb{R}$  is replaced by the ring  $A = \mathbb{F}_p[x]/(x^2)$ . The ring  $A$  has a natural notion of distance with two

distinct length scales. If  $a + bx \in A$  with  $a, b \in \mathbb{F}_p$ , we define

$$\|a + bx\|_A := \begin{cases} 1 & \text{if } a \neq 0 \\ p^{-1} & \text{if } a = 0, b \neq 0 \\ 0 & \text{if } a = b = 0 \end{cases}$$

There is a cousin of the Heisenberg group in  $A^3$  and it leads to a counterexample to the analogue of Theorem 1.2. But unlike in  $\mathbb{C}^3$ , the Heisenberg group cousin in  $A^3$  is *not* sticky. It appears likely that in  $A^3$ , the sticky case of Wolff axiomakey conjecture is true, but the general conjecture is false.

Looking back at the proof of theakey conjecture, the key distinction between  $\mathbb{R}$  and the ring  $A$  is that the ring  $A$  has only two distinct non-zero scales. The proof of theakey conjecture requires discussing many scales in order to run the multiscale analysis. In the ring  $A$ , we do not have the rich range of related sets of tubes  $\mathbb{T}'$  that are used in the proof of theakey conjecture.

## References

- Jonathan Bennett, Anthony Carbery, and Terence Tao (2006). “On the multilinear restriction andakey conjectures”, *Acta Math.* **196** (2), pp. 261–302.
- Jean Bourgain (1991). “ $L^p$ -estimates for oscillatory integrals in several variables”, *Geom. Funct. Anal.* **1** (4), pp. 321–374.
- (2003). “On the Erdős–Volkmann and Katz–Tao ring conjectures”, *Geom. Funct. Anal.* **13** (2), pp. 334–365.
- (2010). “The discretized sum-product and projection theorems”, *J. Anal. Math.* **112**, pp. 193–236.
- Jean Bourgain and Ciprian Demeter (2015). “The proof of the  $l^2$  decoupling conjecture”, *Ann. Math. (2)* **182** (1), pp. 351–389.
- Jean Bourgain, Nets Hawk Katz, and Terence Tao (2004). “A sum-product estimate in finite fields, and applications”, *Geom. Funct. Anal.* **14** (1), pp. 27–57.
- Ciprian Demeter and Hong Wang (2025). “Szemerédi–Trotter bounds for tubes and applications”, *Ars Inven. Anal.* **2025**. Id/No 1, p. 46.
- Gerald Edgar and Chris Miller (2003). “Borel subrings of the reals”, *Proc. Am. Math. Soc.* **131** (4), pp. 1121–1129.
- Paul Erdős and Endre Szemerédi (1983). *On sums and products of integers*. Studies in Pure Mathematics, Mem. of P. Turán, 213–218 (1983).
- Moubariz Garaev (2007). “An explicit sum-product estimate in  $\mathbb{F}_p$ ”, *Int. Math. Res. Not.* **2007** (11). Id/No rnm035, p. 11.
- Larry Guth (2015). “A short proof of the multilinearakey inequality”, *Math. Proc. Camb. Philos. Soc.* **158** (1), pp. 147–153.

- Larry Guth (2025). *Outline of the Wang–Zahl proof of the Kakeya conjecture in  $\mathbb{R}^3$* . arXiv: [2508.05475 \[math.CA\]](#).
- Larry Guth, Nets Hawk Katz, and Joshua Zahl (2021). “On the discretized sum-product problem”, *Int. Math. Res. Not.* **2021** (13), pp. 9769–9785.
- Larry Guth, Noam Solomon, and Hong Wang (2019). “Incidence estimates for well spaced tubes”, *Geom. Funct. Anal.* **29** (6), pp. 1844–1863.
- Larry Guth, Hong Wang, and Joshua Zahl (2026). *A streamlined proof of the Kakeya set conjecture in  $\mathbb{R}^3$* . arXiv: [2601.14411 \[math.CA\]](#).
- Nets Hawk Katz, Izabella Łaba, and Terence Tao (2000). “An improved bound on the Minkowski dimension of Besicovitch sets in  $\mathbb{R}^3$ ”, *Ann. Math. (2)* **152** (2), pp. 383–446.
- Nets Hawk Katz and Joshua Zahl (2019). “An improved bound on the Hausdorff dimension of Besicovitch sets in  $\mathbb{R}^3$ ”, *J. Am. Math. Soc.* **32** (1), pp. 195–259.
- Tamás Keleti and Pablo Shmerkin (2019). “New bounds on the dimensions of planar distance sets”, *Geom. Funct. Anal.* **29** (6), pp. 1886–1948.
- William O’Regan, Pablo Shmerkin, and Hong Wang (2025). *Simple proofs of discretised projection theorems*. arXiv: [2511.21656 \[math.CA\]](#).
- Tuomas Orponen (2017). “On the distance sets of Ahlfors–David regular sets”, *Adv. Math.* **307**, pp. 1029–1045.
- (2022). “On arithmetic sums of Ahlfors-regular sets”, *Geom. Funct. Anal.* **32** (1), pp. 81–134.
- Tuomas Orponen and Pablo Shmerkin (2023a). “On the Hausdorff dimension of Furstenberg sets and orthogonal projections in the plane”, *Duke Math. J.* **172** (18), pp. 3559–3632.
- (2023b). *Projections, Furstenberg sets, and the ABC sum-product problem*. arXiv: [2301.10199 \[math.CA\]](#).
- Tuomas Orponen, Pablo Shmerkin, and Hong Wang (2024). “Kaufman and Falconer estimates for radial projections and a continuum version of Beck’s theorem”, *Geom. Funct. Anal.* **34** (1), pp. 164–201.
- Kevin Ren and Hong Wang (2023). *Furstenberg sets estimate in the plane*. arXiv: [2308.08819 \[math.CA\]](#).
- Pablo Shmerkin and Hong Wang (2025). “On the distance sets spanned by sets of dimension  $d/2$  in  $\mathbb{R}^d$ ”, *Geom. Funct. Anal.* **35** (1), pp. 283–358.
- Terence Tao (2001). “From rotating needles to stability of waves: emerging connections between combinatorics, analysis, and PDE”, *Notices Am. Math. Soc.* **48** (3), pp. 294–303.
- (2014). *Stickiness, graininess, planiness, and a sum-product approach to the Kakeya problem*. Blog post.
- Hong Wang and Shukun Wu (2024). *Restriction estimates using decoupling theorems and two-ends Furstenberg inequalities*. arXiv: [2411.08871 \[math.CA\]](#).
- Hong Wang and Joshua Zahl (2025). *Volume estimates for unions of convex sets, and the Kakeya set conjecture in three dimensions*. arXiv: [2502.17655 \[math.CA\]](#).

- (2026). “Sticky Kakeya sets and the sticky Kakeya conjecture”, *J. Am. Math. Soc.* **39** (2), pp. 515–585.
- Thomas Wolff (1995). “An improved bound for Kakeya type maximal functions”, *Rev. Mat. Iberoam.* **11** (3), pp. 651–674.
- (1999). “Recent work connected with the Kakeya problem”, in: *Prospects in mathematics. Invited talks on the occasion of the 250th anniversary of Princeton University. Papers from the conference, Princeton, NJ, USA, March 17–21, 1996*. Providence, RI: American Mathematical Society, pp. 129–162.

Larry Guth

MIT

*E-mail*: `lguth@math.mit.edu`