Abstract. A Besicovitch set is a set which contains a unit line segment in any direction. It is known that the Minkowski and Hausdorff dimensions of such a set must be greater than or equal to $5/2$ in $\mathbf{R}^3$. In this paper we show that the Minkowski dimension must in fact be greater than $5/2+\varepsilon$ for some absolute constant $\varepsilon > 0$. One observation arising from the argument is that Besicovitch sets of near-minimal dimension have to satisfy certain strong properties, which we call "stickiness", "planiness", and "graininess".

AN IMPROVED BOUND ON THE MINKOWSKI DIMENSION OF BESICOVITCH SETS IN $\mathbf{R}^3$

NETS HAWK KATZ, IZABELLA ŁABA, AND TERENCE TAO

The purpose of this paper is to improve upon the known bounds for the Minkowski dimension of Besicovitch sets in three dimensions. As a by-product of the argument we obtain some strong conclusions on the structure of Besicovitch sets with almost-minimal Minkowski dimension.

Definition 0.1. A *Besicovitch set* (or "Kakeya set") $E \subset \mathbf{R}^n$ is a set which contains a unit line segment in every direction.

Informally, the Kakeya conjecture states that all Besicovitch sets in $\mathbf{R}^n$ have full dimension; this conjecture has been verified for $n = 2$ but is open otherwise. For the purposes of this paper we shall restrict ourselves to the Minkowski dimension, which we now define.

Definition 0.2. If E is in $\mathbf{R}^n$, we define the δ-entropy $\mathcal{E}_\delta(E)$ of E to be the cardinality of the largest δ-separated subset of E, and $N_\delta(E)$ to be the δ-neighbourhood of E.

Definition 0.3. For any set $E \subset \mathbf{R}^n$, the (upper) *Minkowski dimension* $\overline{dim}(E)$ is defined as

$$\overline{dim}(E) = \limsup_{\delta\to 0} \log_{1/\delta} \mathcal{E}_\delta(E) = n - \liminf_{\delta\to 0} \log_\delta |N_\delta(E)|.$$

1991 *Mathematics Subject Classification.* 42B15.

In $\mathbf{R}^n$, Wolff [14] showed the estimate

$$\overline{dim}(E) \geq \frac{1}{2}n + 1, \tag{1}$$

while more recently Bourgain [4] has shown

$$\overline{dim}(E) \geq \frac{13}{25}n + \frac{12}{25}. \tag{2}$$

For further results, generalizations, and applications see [16].

When $n = 3$ Wolff's bound is superior, and thus the best previous result on the three-dimensional problem was $\overline{dim}(E) \geq 5/2$. By combining the ideas of Wolff and Bourgain with some observations on the structure of (hypothetical) extremal counterexamples to the Kakeya problem, we have obtained the following improvement, which is the main result of this paper.

Theorem 0.4. *There exists an $\varepsilon > 0$ such that $\overline{dim}(E) \geq 5/2+\varepsilon$ for all Besicovitch sets E in $\mathbf{R}^3$.*

While the epsilon in this theorem could in principle be computed, we have not tried to optimize our arguments in order to produce an efficient value for ε. The argument in this paper certainly works for $\varepsilon = 10^{-10}$, but this is definitely far from best possible.

Broadly speaking, the argument is a proof by contradiction. A hypothetical counterexample to the theorem is assumed to exist for some small ε. By refining the collection of tubes slightly, and at one point passing from scale δ to scale $\rho = \sqrt{\delta}$, one can impose a surprisingly large amount of structure on "most" of the Besicovitch set. Eventually there will be enough structure that one can apply the techniques of Bourgain[4] efficiently and obtain a contradiction. (Of course, if one applies these techniques directly then one would only obtain (2), which is inferior for $n = 3$).

By the term "most" used in the previous paragraph, we roughly mean that the portion of the Besicovitch set for which our structural assumptions fail only occupies an extremely small fraction of the entire set; we will make this notion precise in Section 5. We shall need this very strong control on the exceptional set, as there is a key stage in the argument in which we need to find an arithmetic progressions of length three in the non-exceptional portion of the Besicovitch set. A discussion of the difficulties of this approach when one only knows that a small portion of each tube is "good" can be found in Bourgain [4]. This means that we will not use methods such as pigeonholing to obtain structural assumptions on our set, as these types of methods usually only give a non-exceptional set which is about $\log(1/\delta)^{-1}$ of the full Besicovitch set.

It may well be that one can use further ideas in [4], such as using triples of points whose *reciprocals* are in arithmetic progression, in order to circumvent this restriction. However, there is an additional obstruction preventing us from obtaining improvements to Theorem 0.4 such as a Hausdorff dimension or maximal function result, as in [14] or [4]. Namely, our argument crucially requires control of the entropy of the Besicovitch set not only at scale δ, but also at many intermediate scales between δ and 1. In particular, the scale $\rho = \sqrt{\delta}$ plays a key role. Such control is readily available in the case when the upper Minkowski dimension is assumed to be small, but not in the other cases just discussed.

We will derive structural properties on our hypothetical low-dimensional Besicovitch set in the following order. Firstly, we follow an observation in Wolff [15], and observe that the Besicovitch set must be "sticky", which roughly states that the map from directions to line segments in the Besicovitch set is almost Lipschitz. To make this observation rigorous we require the x-ray estimate in [15], and also rely crucially on the fact mentioned earlier, that we have control of the Besicovitch set at multiple scales. We will achieve stickiness in Section 3.

Once we have obtained stickiness, it is a fairly routine matter to show that the Besicovitch set must behave in a self-similar fashion. For instance, if one takes a ρ-tube centered around one of the line segments in the Besicovitch set, and dilates it by ρ^{-1} around its axis, one should obtain a new Besicovitch set with almost identical properties. We will obtain quantitative versions of these heuristics in Section 6, after some preliminaries in Sections 4 and 5.

We then combine these self-similarity properties with the following geometric heuristic: if for $i = 1, 2, 3$, we have a vector v_i and a family of tubes $\mathbf{T}_i$ which all approximately point in the direction v_i, then the triple intersection

$$\bigcap_{i=1}^{3} \Big(\bigcup_{T \in \mathbf{T}_i} T \Big) \tag{3}$$

will be fairly small unless v_1, v_2, v_3 are almost co-planar. As a consequence we be able to conclude a remarkable structural property on the Besicovitch set, which we call "planiness". Roughly speaking, it asserts that for most points x in the Besicovitch set, most of the line segments passing through x lie on a plane, or on the union of a small number of planes. At first we shall only derive this property at scale ρ, for reasons which shall become clear, but by changing scale we may easily impose this property at scale δ as well.

One can analyze the derivation of planiness further, and obtain an important additional property which we call "graininess". The key observation is that even if v_1, v_2, v_3 lie in a common plane π, the set (3) will still be small unless the sets $\bigcup_{T \in \mathbf{T}_i} T$ are essentially the union of "squares" parallel to π, providing that the angles between v_1, v_2, and v_3 are fairly large. Inserting this observation into the above argument, one obtains the graininess property, which roughly speaking asserts that the intersection of the Besicovitch set with any ρ-cube will, when studied at scale δ, look like a union of parallel $\delta \times \rho \times \rho$ boxes.

We will define the properties of planiness and graininess in Section 8, after some preliminaries in Section 7.

We remark that the arguments up to this point are not necessarily restricted to sets of dimension close to $5/2$, although for other dimensions one must find an analogue of Wolff's x-ray estimate [15] to begin the argument. The argument below, however, is only effective near dimension $5/2$.

Next, we follow the philosophy of Bourgain [4] and find three ρ-cubes Q_0, Q_1, Q_2 in arithmetic progression which each satisfy certain good properties. To do this it is important that all the properties attained up to this point occur on a very large fraction of the set, which unfortunately causes the arguments in previous sections to be somewhat involved. However, once we have obtained the arithmetic progression then one can be far less stringent, and deal with properties that are satisfied fairly sparsely.

We now apply the ideas in [4], which we now pause to recall. Very roughly, the argument in [4] for the Minkowski dimension runs as follows. Let A, B, C be the intersections of the Besicovitch set with three planes in arithmetic progression. As the Besicovitch set contains many line segments through A, B, C, it follows that there are many pairs (a, b) of points in $A \times B$ whose midpoint is in C. In fact, if the dimension of the Besicovitch set is close to $(n + 1)/2$, then a large fraction of $A \times B$ will have this property. Schematically, we may write this property as

$$(A + B) \subset 2C.$$

In particular, since we expect A, B, C to be of comparable size, we should have

$$|A + B| \approx |A| \approx |B|, \tag{4}$$

where we need to discretize A, B at some scale δ to make sense of the above expressions. From the combinatorial lemmas in [4] relating sums to differences, we should therefore have (in an appropriate sense)

$$|A - B| \ll |A||B|. \tag{5}$$

But this implies that many of the line segments connecting A with B are parallel, which contradicts the definition of a Besicovitch set. Hence we cannot have a Besicovitch set with dimension close to $(n + 1)/2$.

Suppose we applied the above ideas to our situation. If Q_0, Q_1, and Q_2 are three ρ-cubes in arithmetic progression, the Besicovitch set property should imply that there are many pairs (a, b) of points in $Q_0 \times Q_2$ whose midpoint is in Q_1. Unfortunately, if our set has dimension close to $5/2$, this set of pairs of points is very sparse compared to $Q_0 \times Q_2$, and the combinatorial lemmas in [4] are not effective in this context.

However, this can be salvaged by using the planiness and graininess properties of our set, and especially the fact that the planes through x are almost always parallel to the squares through x. These very restrictive properties drastically reduce the possible degrees of freedom of the Besicovitch set, and many pairs in $Q_0 \times Q_2$ can be ruled out *a priori* as being of the form above. The end result is that one can find a well-behaved subset G of $Q_0 \times Q_2$ which is determined by the planes and grains, such that the midpoint of a and b is in Q_1 for a large fraction of pairs (a, b) in G. By applying the lemmas of [4] we find once again that the Besicovitch set contains many tubes which are parallel, if ε is sufficiently small. This is our desired contradiction, and we are done.

For completeness we also give in an Appendix a sketch of the argument from [4] that we use.

The properties of planiness and graininess may seem strange, but there is a simple example of an object which resembles a 5/2-dimensional Kakeya set - albeit in $\mathbf{C}^3$ instead of $\mathbf{R}^3$ - and which does obey these properties. Namely, the Heisenberg group

$$\{(z_1, z_2, z_3) \in \mathbf{C}^3 : Im(z_3) = Im(z_1 \overline{z_2})\}.$$

has real dimension 5, contains a four-parameter family of lines (some of which, though, are parallel), and satisfies the planiness and graininess properties perfectly. We will discuss this example further in Section 13.

The authors are indebted to Jean Bourgain for explaining his recent work to one of the authors, and especially to Tom Wolff for his constant encouragement and mathematical generosity. Many of the "new" ideas in this paper were inspired, directly or indirectly, by the (mostly unpublished) heuristics, computations, and insights which Tom Wolff kindly shared with the authors. The first and third authors are supported by NSF grants DMS-9801410 and DMS-9706764 respectively.

1. Notation and preliminaries I.

Throughout this paper we shall always be working in three dimensions $\mathbf{R}^3$ unless otherwise specified. We use italic letters x, y, z to denote points in $\mathbf{R}^3$, and Roman letters $(\mathrm{x}, \mathrm{y}, \mathrm{z})$ to denote co-ordinates of points in $\mathbf{R}^3$.

Unless otherwise specified, all integrals will be over $\mathbf{R}^3$ with Lebesgue measure.

In this paper δ refers to a number such that $0 < \delta \ll 1$, and ε refers to a fixed number such that $0 < \varepsilon \ll 1$. d will be a number such that $2 < d < 3$; later on we will set $d = 5/2$. The symbol ρ will always denote the quantity $\rho = \sqrt{\delta}$.

We use C, c to denote generic positive constants, varying from line to line (unless subscripted), which are independent of ε, δ, but which may depend on the parameter d. C will denote the large constants and c will denote the small constants.

We will use $X \lesssim Y$, $Y \gtrsim X$, or $X = O(Y)$ to denote the inequality $X \leq AY$, where

A is a positive quantity which may depend on ε. We use $X \gg Y$ to denote the statement $X \geq AY$ for a large constant A. We use $X \sim Y$ to denote the statement that $X \lesssim Y$ and $Y \lesssim X$.

We will use $X \lessapprox Y$, $Y \gtrapprox X$, or "Y majorizes X" to denote the inequality

$$X \leq A\delta^{-C\varepsilon} Y,$$

where A is a positive quantity which may depend on ε, and C is a quantity which does not depend on ε. We use $X \approx Y$ to denote the statement that $X \lessapprox Y$ and $Y \lessapprox X$. In particular we have $\varepsilon \approx 1$.

If E is a subset of $\mathbf{R}^n$, we use $|E|$ to denote its Lebesgue measure; if I is a finite set, we use $\#I$ to denote its cardinality.

For technical reasons, we will require a non-standard definition of a δ-tube. Namely, a δ-tube T is a δ-neighbourhood of a line segment whose endpoints x_0 and x_1 are on the planes $\{(\mathrm{x}, \mathrm{y}, \mathrm{z}) : \mathrm{z} = 0\}$ and $\{(\mathrm{x}, \mathrm{y}, \mathrm{z}) : \mathrm{z} = 1\}$ respectively, and whose orientation is within $\frac{1}{10}$ of the vertical. Note that

$$|T_\sigma| \sim \sigma^2 \tag{6}$$

for any σ-tube T_σ. We define a *direction* to be any quantity of the form $(\mathrm{x}, \mathrm{y}, 1)$ with $|\mathrm{x}|, |\mathrm{y}| \lesssim 1$, and $\mathrm{dir}(T)$ to be the direction $x_1 - x_0$.

If T is a tube, we define CT to be the dilate of T about its axis by a factor C. We say that two tubes T and T' are *equivalent* if $T \subset CT'$ and $T' \subset CT$. If $\mathbf{T}$ is a set of tubes, we say that $\mathbf{T}$ *consists of essentially distinct tubes* if for any $T \in \mathbf{T}$ there are at most $O(1)$ tubes T' which are equivalent to T.

We use the term r-ball to denote a ball of radius r, and use $B(x, r)$ to denote the r-ball centered at x. By a simple covering argument using δ-balls, we see that the quantities $\mathcal{E}_\delta(E)$ and $N_\delta(E)$ defined in Definition 0.2 are related in $\mathbf{R}^3$ by the basic estimate

$$\delta^3 \mathcal{E}_\delta(E) \sim |N_\delta(E)|. \tag{7}$$

If μ is a function, we use $\mathrm{supp}(\mu)$ to denote the support of μ.

If $1 < d \leq \infty$ is an exponent, we define the dual exponent by $d' = d/(d-1)$, with $\infty' = 1$ of course. The support, L^1 norm, $L^{d'}$ norm, and size of a function are all related of course by such standard inequalities as Hölder and Chebyshev. We will rely on these inequalities extremely often, and so we write them down for future reference.

Lemma 1.1. *For any non-negative function μ on a measure space (continuous or discrete), any $1 < d \leq \infty$, and any $\lambda > 0$, we have*

$$\|\mu\|_1 \lessapprox \|\mu\|_{d'} |\mathrm{supp}(\mu)|^{1/d} \tag{8}$$

$$\|\mu\|_{d'} \gtrapprox \|\mu\|_1 |\mathrm{supp}(\mu)|^{-1/d} \tag{9}$$

$$\int_{\mu \gtrapprox \lambda} \mu \lessapprox \lambda^{1-d'} \|\mu\|_{d'}^{d'} \tag{10}$$

$$\int_{\mu \lessapprox \lambda} \mu \lessapprox \lambda |\mathrm{supp}(\mu)| \tag{11}$$

$$|\{\mu \gtrapprox \lambda\}| \lessapprox \lambda^{-d'} \|\mu\|_{d'}^{d'}. \tag{12}$$

We remark that all the above quantites will automatically be finite in our applications.

Finally, we observe the following trivial uniformity lemma.

Lemma 1.2. *Let μ be a non-negative function on a measure space (continuous or discrete), such that*

$$|\mathrm{supp}(\mu)| \leq A$$
$$\|\mu\|_\infty \leq B$$
$$\|\mu\|_1 \gtrapprox AB$$

for some $A, B > 0$. Then there exists a non-empty set $E \subset \mathrm{supp}(\mu)$ such that

$$|E| \approx |\mathrm{supp}(\mu)| \approx A$$
$$\mu \approx B \text{ on } E$$
$$\|\mu\|_{L^1(E)} \approx AB$$

Proof Define λ by

$$\lambda = (AB)^{-1} \|\mu\|_1;$$

from hypothesis we have $\lambda \approx 1$. We define E to be the set

$$E = \{\mu > \frac{1}{2}\lambda B\}.$$

From the estimate

$$\lambda AB = \|\mu\|_{L^1(E)} + \|\mu\|_{L^1(E^c)} \leq B|E| + \frac{1}{2}\lambda BA$$

which follows from the hypotheses, we see that $|E| \gtrapprox A$. The verification of the remainder of the properties are now routine. ∎

It is of course possible to make this lemma more precise (e.g. by the pigeonhole principle), but we shall not do so here.

2. Kakeya estimates

In this section we summarize the Kakeya and x-ray estimates which we shall need. In the following σ, θ are quantities such that $\delta \leq \sigma \leq \theta \leq 1$.

Definition 2.1. If $\mathbf{T}_\sigma$ is a collection of σ-tubes, we define the *directional multiplicity* $m = m(\mathbf{T}_\sigma)$ to be the largest number of tubes in $\mathbf{T}_\sigma$ whose directions all lie in a cap of radius σ. If $m \approx 1$, we say that $\mathbf{T}_\sigma$ is *direction-separated.*

Definition 2.2. Let $2 < d < 3$. We say that there is an *x-ray estimate at dimension* d if there exist $0 < \alpha, \beta < 1$ for which the following statement holds: For any δ-separated set $\mathcal{E}$ of directions and any collection $\mathbf{T}$ of essentially distinct tubes pointing in directions in $\mathcal{E}$, we have

$$\| \sum_{T \in \mathbf{T}} \chi_T \|_{d'} \lesssim \delta^{1-\frac{3}{d}} m^{1-\beta} (\delta^2 \# \mathcal{E})^\alpha, \tag{13}$$

where m is the directional multiplicity of $\mathbf{T}$. If we only assume that (13) holds with $0 < \alpha < 1$ and $\beta = 0$, then we say that there is a *Kakeya estimate at dimension* d.

Clearly, an x-ray estimate is stronger than a Kakeya estimate at the same dimension.

Theorem 2.3. [14, 15] *There is a x-ray estimate at dimension* $5/2$*.*

Indeed, (13) is proven in [14] for $(\alpha, \beta, d) = (7/10, 0, 5/2)$, while in [15] this is improved to $(\alpha, \beta, d) = (7/10, 1/4, 5/2)$. Although these values of α and β are sharp for this value of d, their exact values are not particularly important for our purposes. We remark that an estimate with $\alpha = 0$ can automatically be improved to an estimate for some positive α thanks to Nikishin-Pisier factorization theory; a discussion of this phenomenon may be found in [1].

We will usually only rely on the following variants of the above estimate:

Lemma 2.4. *Let $\delta \leq \sigma \leq \theta \ll 1$, and let $\boldsymbol{T}_\sigma$ be a collection of σ-tubes whose set of directions all lie in a cap of radius θ. Let $2 < d < 3$ be fixed.*

- *If we have a Kakeya estimate at some dimension d, and if the collection $\boldsymbol{T}_\sigma$ is direction-separated, then*

$$\| \sum_{T_\sigma \in \boldsymbol{T}_\sigma} \chi_{T_\sigma} \|_{d'} \lessapprox \sigma^{\frac{d-3}{d}} \theta^{\frac{d+1}{d}}. \tag{14}$$

- *If we have an x-ray estimate at some dimension d, and if $\boldsymbol{T}_\sigma$ consists of essentially distinct tubes, then*

$$\| \sum_{T_\sigma \in \boldsymbol{T}_\sigma} \chi_{T_\sigma} \|_{d'} \lessapprox \sigma^{\frac{d-3}{d}} \theta^{\frac{d+1}{d}} m^{1-\beta} \tag{15}$$

for some $\beta > 0$, where m is the directional multiplicity of $\boldsymbol{T}_\sigma$.

Proof We prove only the second claim, as the first follows by setting $m \approx 1$ and $\beta = 0$. By an affine transformation we may assume that $\omega = (0,0,1)$. Apply the non-isotropic dilation $(\mathrm{x},\mathrm{y},\mathrm{z}) \mapsto (\theta^{-1}\mathrm{x}, \theta^{-1}\mathrm{y}, \mathrm{z})$. This transforms $\mathbf{T}_\sigma$ to a collection $\mathbf{T}'$ of $\sigma\theta^{-1}$ tubes pointing in a $\sigma\theta^{-1}$-separated set of directions, without significantly affecting the directional multiplicity. From (13) we have

$$\| \sum_{T \in \mathbf{T}'} \chi_{T'} \|_{d'} \lessapprox \sigma^{\frac{d-3}{d}} \theta^{\frac{3-d}{d}} m^{1-\beta}.$$

The claim then follows by undoing the dilation. ∎

In the specific case $d = 5/2$, $\alpha = 7/10$ we can also obtain the above Lemma directly from (13).

3. The sticky reduction

In the rest of the paper, $2 < d < 3$ will be a number such that there is an x-ray estimate at dimension d. In particular, by the results in [15] we may choose $d = 5/2$.

It is well known that Besicovitch sets must have Hausdorff and Minkowski dimensions $\geq d$. The purpose of this section is to show that one can push this observation a bit further, and conclude that sets whose Minkowski dimension are close to d have a certain "sticky" structure. Our arguments crucially rely on the fact that sets with small Minkowski dimension are under control at several scales simultaneously (in particular, at the scales δ and $\rho = \sqrt{\delta}$); there does not appear to be any obvious way to apply these heuristics to a set with Hausdorff dimension close to d, for instance.

Definition 3.1. Let $\mathbf{T}$ be a collection of δ-tubes. We say that $\mathbf{T}$ is *sticky* if it is direction-separated and there exists a collection $\mathbf{T}_\rho$ of direction-separated ρ-tubes and a partition of $\mathbf{T}$ into disjoint sets $\mathbf{T}[T_\rho]$ for $T_\rho \in \mathbf{T}_\rho$ such that

$$T \subset T_\rho \text{ for all } T_\rho \in \mathbf{T}_\rho \text{ and } T \in \mathbf{T}[T_\rho] \tag{16}$$

and we have the cardinality estimates

$$\#\mathbf{T} \approx \delta^{-2} \tag{17}$$

$$\#\mathbf{T}_\rho \approx \delta^{-1} \tag{18}$$

$$\#\mathbf{T}[T_\rho] \approx \delta^{-1} \text{ for all } T_\rho \in \mathbf{T}_\rho. \tag{19}$$

For technical reasons we shall also need an iterated version of stickiness:

Definition 3.2. Let $\mathbf{T}$ be a collection of direction-separated δ-tubes of cardinality $\approx \delta^{-2}$. We say that $\mathbf{T}$ is *doubly sticky* if it is sticky with an associated collection $\mathbf{T}_\rho$ of ρ-tubes, and the collection $\mathbf{T}_\rho$ is also sticky (with δ replaced throughout by ρ in the previous definition), with an associated collection $\mathbf{T}_{\delta^{1/4}}$ of $\delta^{1/4}$-tubes.

Proposition 3.3. *Suppose that there is an x-ray estimate at dimension d, and that there exists a Besicovitch set E with $\overline{dim}(E) < d + \varepsilon$. Then for any sufficiently small δ, there exists a doubly sticky collection $\boldsymbol{T}$ of tubes at scale δ with associated collections $\boldsymbol{T}_\rho$ and $\boldsymbol{T}_{\delta^{1/4}}$ of ρ-tubes and $\delta^{1/4}$-tubes respectively, such that*

$$|\bigcup_{T \in \boldsymbol{T}} T| \lessapprox \delta^{3-d} \tag{20}$$

$$|\bigcup_{T_\rho \in \boldsymbol{T}_\rho} T_\rho| \lessapprox \rho^{3-d} \tag{21}$$

$$|\bigcup_{T_{\delta^{1/4}} \in \boldsymbol{T}_{\delta^{1/4}}} T_{\delta^{1/4}}| \lessapprox \delta^{(3-d)/4}. \tag{22}$$

Proof Let E be a Besicovitch set with Minkowski dimension at most $d + \varepsilon$, and let $\delta \ll 1$ be a fixed. We may assume without loss of generality that E is contained in a fixed ball $B(0, C)$. Then by Definition 0.3 we have

$$|N_\sigma(E)| \lessapprox \sigma^{3-d} \text{ for all } \delta \leq \sigma \leq 1. \tag{23}$$

By taking a δ-separated set of directions oriented within $\frac{1}{10}$ of the vertical, and looking at the associated line segments in E, we can find (possibly after rescaling E slightly) a direction-separated set $\mathbf{T}$ of δ-tubes satisfying (17) such that each tube is contained in $N_{C\delta}(E)$. Unfortunately this collection need not be sticky, let alone doubly sticky. To remedy this we shall prune $\mathbf{T}$ of its non-sticky components.

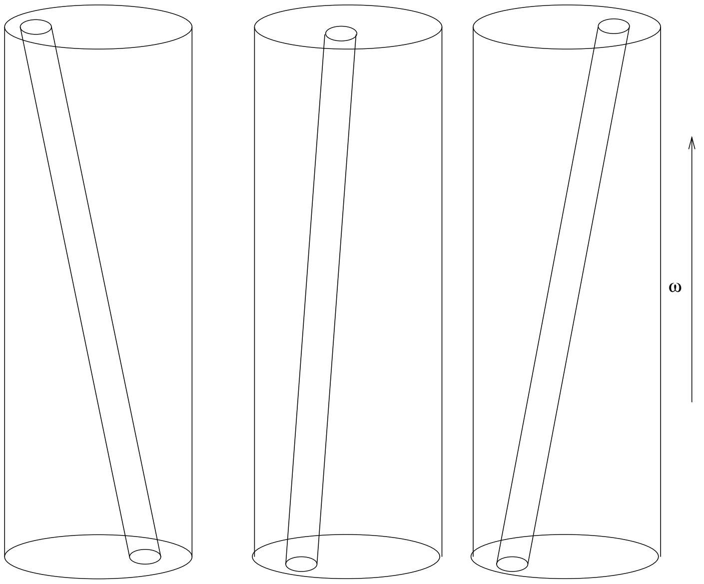

FIGURE 1. An example of a non-sticky direction ω. The thin tubes are δ-tubes in $N_{C\delta}(E)$, the fat tubes are ρ-tubes in $N_{C\rho}(E)$. Note that $N_{C\rho}(E)$ is rather large.

Let $\mathcal{E}$ be a maximal ρ-separated set of directions. Call a direction $\omega \in \mathcal{E}$ *sticky* if the tubes $T \in \mathbf{T}$ such that $\mathrm{dir}(T) \in B(\omega, \rho)$ can be covered by $O(\delta^{-C_1\varepsilon})$ ρ-tubes pointing in the direction ω; here C_1 is a constant to be chosen later.

We now observe that only a small number of directions ω are non-sticky. More precisely, let $\mathcal{E}_1$ be the sub-collection of directions in $\mathcal{E}$ which are non-sticky. By definition, we can find for each $\omega \in \mathcal{E}_1$ a collection of $\gtrapprox \delta^{-C_1\varepsilon}$ disjoint ρ-tubes which are contained in $N_{C\rho}(E)$. Call the union of all these collections $\mathbf{T}'_\rho$. Then by (6)

$$\| \sum_{T \in \mathbf{T}'_\rho} \chi_T \|_1 \sim \rho^2 \# \mathbf{T}'_\rho \gtrapprox \delta^{-C_1\varepsilon} (\rho^2 \# \mathcal{E}_1).$$

From (9) and (23) we thus have

$$\| \sum_{T \in \mathbf{T}'_\rho} \chi_T \|_{d'} \gtrapprox \delta^{-C_1\varepsilon} (\rho^2 \# \mathcal{E}_1) \rho^{-\frac{3-d}{d}}.$$

On the other hand, from (13) we have

$$\| \sum_{T \in \mathbf{T}'_\rho} \chi_T \|_{d'} \lessapprox \rho^{-\frac{3-d}{d}} \delta^{-C_1\varepsilon(1-\beta)} (\rho^2 \# \mathcal{E}_1)^\alpha.$$

Combining the two estimates we obtain

$$\# \mathcal{E}_1 \lessapprox \delta^{-1} \delta^{cC_1\varepsilon},$$

where c depends on α, β. Since the tubes in $\mathbf{T}$ are direction-separated, we thus see that at most $\lessapprox \delta^{-2}\delta^{cC_1\varepsilon}$ tubes in $\mathbf{T}$ that point within $O(\delta)$ of a non-sticky direction.

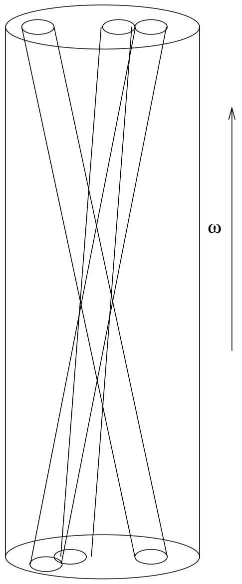

FIGURE 2. An example of a sticky direction ω.

We may therefore remove these tubes from $\mathbf{T}$ without significantly affecting (17), if C_1 was chosen sufficiently large.

The collection $\mathbf{T}$ now has no non-sticky directions. Thus, we may cover the tubes in $\mathbf{T}$ by a family of ρ-tubes $\mathbf{T}_\rho$ which are direction-separated. In particular, we have the upper bound in (18). Since $\mathbf{T}$ is direction-separated, each $\mathbf{T}_\rho$ can cover at most $\approx \delta^{-1}$ tubes in $\mathbf{T}$, and so we have the lower bound in (18). Let $\{\mathbf{T}[T_\rho]\}$ be any partition of $\mathbf{T}$ for which (16) always holds. As we have just observed, we have

$$\#\mathbf{T}[T_\rho] \lessapprox \delta^{-1}$$

for each $T_\rho \in \mathbf{T}_\rho$. On the other hand, from (17) we have

$$\sum_{T_\rho \in \mathbf{T}_\rho} \#\mathbf{T}[T_\rho] \approx \delta^{-2}.$$

From these estimates, (18), and Lemma 1.2, we can find a subset $\mathbf{T}_\rho{}'$ of $\mathbf{T}_\rho$ such that $\#\mathbf{T}_\rho{}' \approx \delta^{-1}$ and (19) holds for all $T_\rho \in \mathbf{T}_\rho{}'$. If we now replace $\mathbf{T}$ with $\bigcup_{T_\rho \in \mathbf{T}_\rho{}'} \mathbf{T}[T_\rho]$ and $\mathbf{T}_\rho$ with $\mathbf{T}_\rho{}'$ we see that we have obtained (19) without significantly affecting any of the other properties just derived.

We now repeat the above procedure but with δ replaced by ρ, and $\mathbf{T}$ replaced by $\mathbf{T}_\rho$. This allows us to refine the collection $\mathbf{T}_\rho$ so that it is also sticky, with an associated collection $T_{\delta^{1/4}}$ of $\delta^{1/4}$-tubes. Of course, to maintain consistency we have to remove the tubes $\mathbf{T}[T_\rho]$ from $\mathbf{T}$ every time we remove a tube T_ρ from $\mathbf{T}_\rho$, but this does not cause any difficulty, and one may verify that all the claims in the Proposition hold. Note that the claims (20), (21), and (22) will follow from (23). ∎

Henceforth we assume that C_1 has been set to an absolute constant, and allow all future constants to implicitly depend on C_1. We will continue this convention when we choose C_2, C_3, etc.

The above result established stickiness at the scales ρ and $\delta^{1/4}$. In fact, one could quite easily establish stickiness at every scale $\delta \leq \sigma \leq 1$ from the Minkowski dimension hypothesis, but we shall not need to do so here. We also remark that the above argument requires no special numerology on d, and would work perfectly well for $d > 5/2$, providing of course that we had an x-ray estimate at d.

Mostly we will work with the collections $\mathbf{T}$ and $\mathbf{T}_\rho$ of tubes, and make no reference to the collection $\mathbf{T}_{\delta^{1/4}}$ or (22). In fact, the only purpose of $\mathbf{T}_{\delta^{1/4}}$ and (22) is so that we may pass from scale δ to scale ρ at a crucial step in the argument (Proposition 9.2), without losing stickiness.

4. Notation and preliminaries II.

In the rest of the paper, $0 < \varepsilon \ll 1$ will be a fixed small number (say $\varepsilon = 10^{-10}$), and $\mathbf{T}$ will be a sticky collection of tubes satisfying (20) and (21). With the exception of Proposition 5.10, we will not use the double-stickiness property, and will not change the value of δ.

For future reference we shall set out some notation and estimates which we shall use frequently.

We use T_σ to denote a tube of thickness σ; if T is unsubscripted, we assume it to have thickness δ.

Definition 4.1. For any $x \in \mathbf{R}^3$ and $T_\rho \in \mathbf{T}_\rho$, we define the sets $\mathbf{T}(x)$, $\mathbf{T}_\rho(x)$, and $\mathbf{T}[T_\rho](x)$ by

$$\mathbf{T}(x) = \{T \in \mathbf{T} : x \in T\}$$
$$\mathbf{T}_\rho(x) = \{T_\rho \in \mathbf{T}_\rho : x \in T_\rho\}$$
$$\mathbf{T}[T_\rho](x) = \{T \in \mathbf{T}[T_\rho] : x \in T\}.$$

Definition 4.2. We define the sets E_δ, E_ρ, and $E_\delta[T_\rho]$ for all $T_\rho \in \mathbf{T}_\rho$ by

$$E_\delta = \bigcup_{T \in \mathbf{T}} T$$
$$E_\rho = \bigcup_{T_\rho \in \mathbf{T}_\rho} T_\rho$$
$$E_\delta[T_\rho] = \bigcup_{T \in \mathbf{T}[T_\rho]} T.$$

We similarly define the multiplicity functions μ_δ, μ_ρ, and $\mu_\delta[T_\rho]$ by

$$\mu_\delta(x) = \sum_{T \in \mathbf{T}} \chi_T(x) = \#\mathbf{T}(x)$$
$$\mu_\rho(x) = \sum_{T_\rho \in \mathbf{T}_\rho} \chi_{T_\rho}(x) = \#\mathbf{T}_\rho(x)$$
$$\mu_\delta[T_\rho](x) = \sum_{T \in \mathbf{T}[T_\rho]} \chi_{T_\rho}(x) = \#\mathbf{T}[T_\rho](x).$$

The size of E_δ and E_ρ is controlled by (20) and (21), while the $L^{d'}$ norms of μ_δ, μ_ρ, and $\mu_\delta[T_\rho]$ can be controlled by (14). To complement these bounds we have the following precise control on the L^1 norms of μ_δ, $\mu_s d$, and $\mu_\delta[T_\rho]$:

Lemma 4.3. *Let T_ρ be an element of $\mathbf{T}_\rho$, and x_0 be a point in T_ρ. Then we have the L^1 estimates*

$$||\mu_\delta||_1 \approx 1 \tag{24}$$
$$||\mu_\rho||_1 \approx 1 \tag{25}$$
$$||\mu_\delta[T_\rho]||_1 \approx \delta \tag{26}$$
$$||\mu_\delta[T_\rho]||_{L^1(B(x_0,C\rho))} \approx \delta^{3/2} \tag{27}$$

Proof The estimates (24), (25), (26) follow from (17), (18), (19), and (6). The proof of (27) is similar and relies on the geometrical observation (cf. Figure 3)

$$\int_{B(x_0,C\rho)} \chi_T \approx \delta^{5/2} \tag{28}$$

for any $T \in \mathbf{T}[T_\rho]$. ∎

5. Properties which occur with probability close to 1

In the sequel we shall frequently need a quantitative version of the statement "The property $P(x)$ implies the property $Q(x)$ with probability close to 1". This motivates

Definition 5.1. Let $P(x)$ and $Q(x)$ be logical statements with free parameters $x = (x_1, \dots, x_n)$, where each of the variables x_i range either over a subset of Euclidean space, or over a discrete set. We use

$$\tilde{\forall}_{x} Q(x) : P(x) \tag{29}$$

to denote the statement that

$$|\{x : Q(x) \text{ holds, but } P(x) \text{ fails}\}| \lessapprox \delta^{c\sqrt{\epsilon}} |\{x : Q(x) \text{ holds}\}| \tag{30}$$

for some absolute constant $c > 0$, where the sets are measured with respect to the measure $dx = \prod_{i=1}^n dx_i$, and dx_i is Lebesgue measure if the x_i range over a subset of Euclidean space, or counting measure if they range over a discrete set.

In practice our variables x_i will either be points in $\mathbf{R}^3$ (and thus endowed with Lebesgue measure), or tubes in $\mathbf{T}$ or $\mathbf{T}_\rho$ (and thus endowed with counting measure). Thus, for instance,

$$\tilde{\forall}_{T,x}\, T \in \mathbf{T}, x \in T : P(x, T)$$

denotes the statement that

$$\sum_{T \in \mathbf{T}} |\{x \in T : P(x,T) \text{ fails}\}| \lessapprox \delta^{c\sqrt{\epsilon}} \sum_{T \in \mathbf{T}} |T|.$$

The right-hand side of (30) will always be automatically finite in our applications. Note that (29) vacuously holds if $P(x)$ is never satisfied.

Of course, (30) is trivial if $c = 0$. The reason why we choose the factor $\delta^{c\sqrt{\epsilon}}$ is that it is much smaller than any quantity of the form $\delta^{C\varepsilon}$, but much larger than anything of the form δ^c. In particular, (30) implies

$$|\{x : Q(x), P(x)\}| \sim |\{x : Q(x)\}|. \tag{31}$$

Here and in the rest of the paper, the expression "$Q(x), P(x)$" is an abbreviation for "$Q(x)$ and $P(x)$ both hold".

The symbol $\tilde{\forall}_{x}$ should be read as "for most x such that", where "most" means that the event occurs with probability very close to 1.

We now develop some machinery to manipulate expressions of the form (29), which are analogous to the laws of formal logic. We first observe some trivial lemmas.

Lemma 5.2. *Suppose that $m \approx 1$ is an integer and Q and $P_1, \ldots, P_m$ are properties depending on some free parameters $x_1, \ldots, x_n$, such that*

$$\underset{x}{\tilde{\forall}} Q(x) : P_i(x) \text{ holds for all } i = 1, \ldots, m$$

where the implicit constants are independent of i. Then we have $\underset{x}{\tilde{\forall}} Q(x) : P_1(x), \ldots, P_m(x)$.

Proof Apply (30) for each $P_i(x)$ and sum in i. ∎

Let $Q(x) \implies P(x)$ denote the statement "$Q(x)$ fails, or $P(x)$ holds".

Corollary 5.3. *(Modus Ponens) If $R(x)$, $Q(x)$, $P(x)$ are properties such that*

$$\underset{x}{\tilde{\forall}} R(x) : Q(x) \text{ and } \underset{x}{\tilde{\forall}} R(x) : (Q(x) \implies P(x))$$

hold, then we have

$$\underset{x}{\tilde{\forall}} R(x) : P(x).$$

Lemma 5.4. *Let P, Q, R be properties with free parameters $x = (x_1, \ldots, x_n)$ such that*

$$\underset{x}{\tilde{\forall}} R(x) : Q(x) \text{ and } \underset{x}{\tilde{\forall}} R(x), Q(x) : P(x)$$

hold. Then we have $\underset{x}{\tilde{\forall}} R(x) : P(x), Q(x)$.

Proof Applying (31) to the first hypothesis, and (30) to the second, we obtain

$$\underset{x}{\tilde{\forall}} R(x) : (P(x) \implies Q(x)),$$

and the claim follows from Corollary 5.3. ∎

Lemma 5.5. *Let $P(x)$, $Q(x,y)$, $R(x,y)$ be properties with free parameters $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_m)$ such that*

$$\underset{y}{\tilde{\forall}} Q(x,y) : R(x,y)$$

uniformly for all x satisfying $P(x)$. Then we have $\underset{x,y}{\tilde{\forall}} P(x), Q(x,y) : R(x,y)$.

Proof Apply (30) to the hypothesis, and integrate over all x that satisfy $P(x)$. ∎

Unlike the case with the more familiar $\forall$ quantifier, some care must be taken with $\tilde{\forall}$ when adding or removing dummy variables if this significantly changes the underlying measure. For instance,

$$\underset{x}{\tilde{\forall}} x \in \bigcup_{T \in \mathbf{T}} T : P(x)$$

is not necessarily equivalent to

$$\underset{T,x}{\tilde{\forall}}\, T \in \mathbf{T}, x \in T : P(x)$$

because the underlying measures are quite different. In the former case all points x in $\bigcup_{T \in \mathbf{T}}$ have equal weight, whereas in the latter case the weight of a point x is proportional to the multiplicity $\mu_\delta(x)$.

On the other hand, it is legitimate to add or remove dummy variables when the multiplicity is approximately constant. More precisely:

Lemma 5.6. *Let $Q_1(x)$, $Q_2(x,y)$ be properties such that*

$$|\{y : Q_2(x,y)\}| \approx M \text{ whenever } Q_1(x) \text{ holds} \tag{32}$$

for some quantity M independent of x. Then for any property $P(x)$, the statements

$$\underset{x}{\tilde{\forall}} Q_1(x) : P(x)$$

and

$$\underset{x,y}{\tilde{\forall}} Q_1(x), Q_2(x,y) : P(x)$$

are equivalent.

Proof From (32) we see that

$$|\{(x,y) : Q_1(x), Q_2(x,y)\}| \approx M|\{x : Q_1(x)\}|$$

and

$$|\{(x,y) : Q_1(x), Q_2(x,y) \text{ hold, } P(x) \text{ fails}\}| \approx M|\{x : Q_1(x) \text{ holds, } P(x) \text{ fails}\}|.$$

The claim then follows by expanding the hypothesis and conclusion using (30). ■

Corollary 5.7. *Suppose the set Y is partitioned into disjoint subsets $Y[z]$ as z ranges over an index set Z. Then for any properties $Q(x,y)$, $P(x,y)$ and any free variables x, the statements*

$$\underset{x,y}{\tilde{\forall}}\, y \in Y, Q(x,y) : P(x,y)$$

and

$$\underset{x,y,z}{\tilde{\forall}}\, z \in Z, y \in Y[z], Q(x,y) : P(x,y).$$

are equivalent.

Proof Apply Lemma 5.6 with x, y equal to (x,y), z respectively, M equal to 1, and $Q_2((x,y),z)$ equal to the property that $z \in Z$ and $y \in Y[z]$. ∎

We will usually apply this Corollary with $Y = \mathbf{T}$, $Z = \mathbf{T}_\rho$.

Next, we show how a compound $\tilde\forall$ quantifier can be split up into two simpler quantifiers.

Lemma 5.8. *Suppose that $Q(x,y)$, $P(x,y)$ are properties depending on some free parameters $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_m)$. Then the statements*

$$\tilde\forall_{x,y} Q(x,y) : P(x,y) \tag{33}$$

and

$$\tilde\forall_{x,y'} Q(x,y') : [\tilde\forall_y Q(x,y) : P(x,y)] \tag{34}$$

are equivalent (up to changes of constants).

Whenever we have nested $\tilde\forall$ statements as in (34), we always assume the implicit constants in the inner $\tilde\forall$ to be independent of the variables in the outer $\tilde\forall$.

Proof Let $R(x)$ denote the property that

$$\tilde\forall_y Q(x,y) : P(x,y)$$

holds.

Assume first that (33) held. If $R(x)$ failed, then

$$|\{y : Q(x,y) \text{ holds, } P(x,y) \text{ fails}\}| \gtrapprox \delta^{c'\sqrt{\epsilon}} |\{y : Q(x,y)\}|$$

for some c' which we shall choose later. Integrating this over all x for which $R(x)$ failed, we obtain

$$|\{(x,y) : Q(x,y) \text{ holds, } P(x,y) \text{ fails}\}| \gtrapprox \delta^{c'\sqrt{\epsilon}} |\{(x,y) : Q(x,y) \text{ holds, } R(x) \text{ fails}\}|.$$

From (33) we therefore have (if c' is chosen sufficiently small)

$$|\{(x,y) : Q(x,y) \text{ holds, } R(x) \text{ fails}\}| \lessapprox \delta^{c\sqrt{\epsilon}} |\{(x,y) : Q(x,y)\}|,$$

and (34) follows by replacing y with y'.

Now suppose that (34) held. We need to show that

$$\int |\{y : Q(x,y) \text{ holds, } P(x,y) \text{ fails}\}| \, dx \lessapprox \delta^{c\sqrt{\epsilon}} \int |\{y : Q(x,y)\}| \, dx. \tag{35}$$

The contribution to the left-hand side of (35) when $R(x)$ holds is acceptable, by the definition of $R(x)$. The contribution when $R(x)$ fails is majorized by

$$\int_{R(x) \text{ fails}} |\{y : Q(x,y)\}| = \int |\{y' : Q(x,y') \text{ holds, } R(x) \text{ fails}\}| \, dx$$

and this is acceptable by (34). ∎

Corollary 5.9. *Suppose that $Q_1(x)$, $Q_2(x,y)$, and $P(x,y)$ are properties depending on some free parameters $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_m)$ which obey (32). Then, the statements*

$$\tilde{\forall}_{x,y} Q_1(x), Q_2(x,y) : P(x,y) \tag{36}$$

and

$$\tilde{\forall}_{x} Q_1(x) : [\tilde{\forall}_{y} Q_2(x,y) \text{ holds } : P(x,y)] \tag{37}$$

are equivalent (up to changes of constants).

Proof Assume that (36) held. By Lemma 5.8 we have

$$\tilde{\forall}_{x,y'} Q_1(x), Q_2(x,y') : [\tilde{\forall}_{y} Q_1(x), Q_2(x,y) : P(x,y)].$$

The second $Q_1(x)$ is redundant. Eliminating the y' variable using Lemma 5.6, one obtains (37). The converse implication follows by reversing the above steps. ∎

We now apply the above machinery to our specific setting, in which we have a sticky collection of tubes.

Lemma 5.10. *Let $\boldsymbol{T}$ be a sticky collection of tubes, and let $P(y, T_\rho, T)$ be a property. Then the statements*

$$\tilde{\forall}_{T_\rho, T, y} T_\rho \in \boldsymbol{T}_\rho, T \in \boldsymbol{T}[T_\rho], y \in T : P(y, T_\rho, T) \tag{38}$$

$$\tilde{\forall}_{T_\rho, T, y, x} T_\rho \in \boldsymbol{T}_\rho, T \in \boldsymbol{T}[T_\rho], y \in T, x \in T_\rho \cap B(y, C\rho) : P(y, T_\rho, T) \tag{39}$$

and

$$\tilde{\forall}_{T_\rho, x} T_\rho \in \boldsymbol{T}_\rho, x \in T_\rho : [\tilde{\forall}_{T, y} T \in \boldsymbol{T}[T_\rho], y \in T \cap B(x, C\rho) : P(y, T_\rho, T)] \tag{40}$$

are equivalent (up to changes of constants).

Proof The equivalence of (38) and (39) follows from Lemma 5.6, since (32) follows from the trivial estimate

$$|T_\rho \cap B(y, C\rho)| \approx \rho^3 \text{ whenever } y \in T \in \mathbf{T}[T_\rho].$$

But if we re-arrange (39) as

$$\tilde{\forall}_{T_\rho, x, T, y} \; T_\rho \in \mathbf{T}_\rho, x \in T_\rho, T \in \mathbf{T}[T_\rho], y \in T \cap B(x, C\rho) : P(y, T_\rho, T)$$

the equivalence of (39) and (40) follows from Corollary 5.9, since (32) follows from (27). ■

6. Uniformity and self-similarity

In this section we shall investigate how the stickiness hypothesis, combined with the Kakeya estimate at scale d, implies certain self-similarity properties of the set E_δ. We will then show in later sections how these properties imply planiness and graininess properties of the set, with these efforts culminating in Corollary 9.5.

The notation will be as in the previous section. Informally, the main result of this section shall be

Heuristic 6.1. *The set E_δ can be mostly covered by about ρ^{-d} ρ-balls B, such that the intersection of E_δ with most of these balls has volume about $\delta^{3-d}\rho^d$. Most points $x \in \bigcup_{T \in \boldsymbol{T}} T$ are contained in about $\delta^{-(3-d)}$ tubes $T \in \boldsymbol{T}$, which in turn are mostly contained in about $\rho^{-(3-d)}$ families $\boldsymbol{T}[T_\rho]$ in such a way that each family contains $\rho^{-(3-d)}$ of the tubes. Also, the set E_δ is not "lumpy" at any scale greater than δ, in the sense that*

$$|E_\delta \cap B_\sigma| \ll |B_\sigma| \tag{41}$$

for all $\sigma \gg \delta$ and all σ-balls B_σ. Finally, the angle subtended by most pairs of intersecting tubes is ≈ 1.

Note that the numerology in this heuristic is consistent with (20), (21), (14), and Lemma 4.3. Moreover, it is essentially the only such numerology which is consistent with these estimates.

The main purpose of this section will be to prove a rigorous version of the Heuristic 6.1; this will be done in Proposition 6.6. Of course, in order to interpret the term "most" in the above Heuristic, the $\tilde{\forall}$ machinery in the previous section will be used heavily.

We begin with a more accurate version of (41), which may be thought of as a dual to the Minkowski dimension estimate (23).

Proposition 6.2. *For each $\delta \leq \sigma \ll 1$ and $x \in \boldsymbol{R}^3$, consider the statement*

$$|E_\delta \cap B(x, C\sigma)| \lessapprox \delta^{-\sqrt{\epsilon}} \delta^{3-d} \sigma^d. \tag{42}$$

Then, if the constants in (42) are chosen appropriately, we have

$$\underset{T,x}{\tilde{\forall}}\ T \in \boldsymbol{T}, x \in T : (42) \textit{ holds for all } \delta \leq \sigma \ll 1.$$

Proof If (42) holds for σ, and $\sigma' \approx \sigma$, then (42) holds for σ' (with slightly worse constants). Thus in order to verify (42) for all $\delta \leq \sigma \ll 1$, it suffices to verify (42) for all σ of the form $\sigma = \delta^{k\varepsilon}$, where $k = 0, \dots, \varepsilon^{-1}$. Thus it suffices to show that

$$\underset{T,x}{\tilde{\forall}}\ T \in \mathbf{T}, x \in T : (42) \text{ holds for all } \sigma = \delta^{k\varepsilon}, k = 0, \dots, \varepsilon^{-1}.$$

Since $\varepsilon^{-1} \approx 1$, it thus suffices by Lemma 5.2 to show that

$$\underset{T,x}{\tilde{\forall}}\ T \in \mathbf{T}, x \in T : (42) \text{ holds for } \sigma$$

uniformly in σ.

Fix σ, and let X denote the set

$$X = \{x \in \mathbf{R}^3 : (42) \text{ fails for } \sigma.\}$$

Then by (30) and (24) it suffices to show

$$\int_X \mu_\delta \lessapprox \delta^{c\sqrt{\epsilon}}.$$

From (20) and the definition of X we may cover X by a collection $\mathbf{B}$ of σ-balls with cardinality

$$\#\mathbf{B} \lessapprox \delta^{\sqrt{\epsilon}} \sigma^{-d}. \tag{43}$$

From elementary geometry we have

$$\int_B \chi_T \lessapprox \sigma^{-2}\delta^2 \int_B \chi_{N_{C\sigma}(T)}$$

for all $B \in \mathbf{B}$ and $T \in \mathbf{T}$ (cf. (28)). Summing this in B and T we see that

$$\int_X \mu_\delta \lessapprox \sigma^{-2}\delta^2 \int_{\bigcup_{B \in \mathbf{B}} B} \sum_{T \in \mathbf{T}} \chi_{N_{C\sigma}(T)}.$$

By (8) and (43) this is majorized by

$$\sigma^{-2}\delta^2 [\sigma^3 \delta^{\sqrt{\epsilon}} \sigma^{-d}]^{1/d} \|\sum_{T \in \mathbf{T}} \chi_{N_{C\sigma(T)}}\|_{d'}.$$

The collection of $C\sigma$-tubes $N_{C\sigma(T)}$ can be partitioned into about $\sigma^2\delta^{-2}$ subcollections, each of which are direction-separated. Thus from the triangle inequality and (14), we can majorize the above by

$$\sigma^{-2}\delta^2 (\sigma^3 \delta^{\sqrt{\epsilon}} \sigma^{-d})^{1/d} \sigma^2 \delta^{-2} \sigma^{\frac{d-3}{d}} = \delta^{\sqrt{\epsilon}/d}$$

as desired. ∎

We combine this property with two others, which are also related to Heuristic 6.1.

Definition 6.3. If $x_0 \in \mathbf{R}^3$ and $T_\rho \in \mathbf{T}_\rho[x_0]$, we say that $P_1(x_0, T_\rho)$ holds if the three statements

$$|E_\delta[T_\rho] \cap B(x_0, C\rho)| \gtrapprox \delta^{\sqrt{\epsilon}} \delta^{3-d} \rho^d \tag{44}$$

$$\tilde{\forall}_{T,x}\ T \in \mathbf{T}[T_\rho], x \in T \cap B(x_0, C\rho) : (42) \text{ holds for all } \delta \leq \sigma \ll 1 \tag{45}$$

$$\tilde{\forall}_{T,x}\ T \in \mathbf{T}[T_\rho], x \in T \cap B(x_0, C\rho) : (47) \text{ holds.} \tag{46}$$

hold, where (47) is the estimate

$$\mu_\delta[T_\rho](x) \lessapprox \delta^{-\sqrt{\epsilon}} \rho^{-(3-d)}. \tag{47}$$

Proposition 6.4. *If the constants in Definition 6.3 are chosen appropriately, then we have*

$$\tilde{\forall}_{T_\rho, x_0}\ T_\rho \in \boldsymbol{T}_\rho, x_0 \in T_\rho : P_1(x_0, T_\rho).$$

Proof If we re-arrange Proposition 6.2 as

$$\tilde{\forall}_{T_\rho, T, x}\ T_\rho \in \mathbf{T}_\rho, T \in \mathbf{T}[T_\rho], x \in T : (42) \text{ holds for all } \delta \leq \sigma \leq 1$$

and apply Lemma 5.10, we obtain

$$\tilde{\forall}_{T_\rho, x_0}\ T_\rho \in \mathbf{T}_\rho, x_0 \in T_\rho : (45) \text{ holds.}$$

Next, we show

$$\tilde{\forall}_{T_\rho, x_0}\ T_\rho \in \mathbf{T}_\rho, x_0 \in T_\rho : (46) \text{ holds.}$$

By Lemma 5.10 again, it suffices to show that

$$\tilde{\forall}_{T_\rho, T, x}\ T_\rho \in \mathbf{T}_\rho, T \in \mathbf{T}[T_\rho], x \in T : (47) \text{ holds.}$$

By Lemma 5.5 it suffices to show

$$\tilde{\forall}_{T, x}\ T \in \mathbf{T}[T_\rho], x \in T : (47) \text{ holds}$$

uniformly in T_ρ. Fixing T_ρ, we rewrite this as

$$\int_{\mu_\delta[T_\rho] \gtrapprox \delta^{-\sqrt{\epsilon}} \rho^{-(3-d)}} \mu_\delta[T_\rho] \lessapprox \delta^{c\sqrt{\epsilon}} \int \mu_\delta[T_\rho].$$

However, from (14) we have

$$\|\mu_\delta[T_\rho]\|_{d'} \lessapprox \delta^{\frac{d-3}{d}} \rho^{\frac{d+1}{d}}. \tag{48}$$

The claim then follows (with $c = d' - 1$) from (10), (26), and some algebra.

We now show that

$$\tilde{\forall}_{x_0} x_0 \in T_\rho : (44) \text{ holds}$$

for all $T_\rho \in \mathbf{T}_\rho$; the Proposition then follows from Lemma 5.5 and Lemma 5.2.

Fix T_ρ, and let X denote the set

$$X = \{x_0 \in T_\rho : (44) \text{ fails}\}.$$

By (6), we must show that $|X| \lessapprox \delta^{c\sqrt{\epsilon}}\delta$.

If $x_0 \in X$, then from (9), (27), and the failure of (44), we have

$$||\mu_\delta[T_\rho]||_{L^{d'}(B(x_0,C\rho))} \gtrapprox \delta^{3/2}(\delta^{\sqrt{\epsilon}}\delta^{3-d}\rho^d)^{-1/d}.$$

Comparing this with (48) and applying a covering argument one obtains the entropy estimate

$$\mathcal{E}_\rho(X) \lessapprox \left[\delta^{\frac{d-3}{d}}\rho^{\frac{d+1}{d}}\delta^{-3/2}(\delta^{\sqrt{\epsilon}}\delta^{3-d}\rho^d)^{1/d}\right]^{d'}$$

which simplifies to

$$\mathcal{E}_\rho(X) \lessapprox \delta^{\sqrt{\epsilon}d'/d}\delta^{-1/2}.$$

The desired bound on X then follows from (7). ∎

Definition 6.5. Let x_0 be a point in $\mathbf{R}^3$. We say that $P_2(x_0)$ holds if one has

$$\delta^{C\sqrt{\epsilon}}\rho^{-(3-d)} \lessapprox \#\mathbf{T}_\rho(x_0) \lessapprox \delta^{-C\sqrt{\epsilon}}\rho^{-(3-d)} \tag{49}$$

$$\tilde{\forall}_{T_\rho} T_\rho \in \mathbf{T}_\rho(x_0) : P_1(x_0, T_\rho) \tag{50}$$

$$|E_\delta \cap B(x_0, C\rho)| \lessapprox \delta^{-C\sqrt{\epsilon}}\delta^{3-d}\rho^d, \text{ and} \tag{51}$$

$$\#\{T_\rho \in \mathbf{T}_\rho(x_0) : \mathrm{dir}(T_\rho) \in B(\omega, \theta)\} \lessapprox \delta^{-C\sqrt{\epsilon}}\theta^c\rho^{-(3-d)} \tag{52}$$

for all directions ω and all $\delta \le \theta \ll 1$.

The property (52) basically states that the tubes in $\mathbf{T}_\rho(x_0)$ are not clustered in a narrow angular band.

Proposition 6.6. *If the constants in the above definition are chosen appropriately, then we have*

$$\tilde{\forall}_{T_\rho} T_\rho \in \boldsymbol{T}_\rho, x \in T_\rho : P_2(x). \tag{53}$$

The reader should compare this rigorous proposition with Heuristic 6.1.

Proof By Lemma 5.2 it suffices to handle each of the properties in $P_2(x)$ separately.

We first observe from Proposition 6.4 and Lemma 5.8 that

$$\tilde{\forall}_{T_\rho{}', x} T_\rho{}' \in \mathbf{T}_\rho, x \in T_\rho{}' : [\tilde{\forall}_{T_\rho} T_\rho \in \mathbf{T}_\rho, x \in T_\rho : P_1(x, T_\rho)].$$

But this is just a rephrasing of

$$\tilde{\forall}_{T_\rho, x} T_\rho \in \mathbf{T}_\rho, x \in T_\rho : (50) \text{ holds.}$$

Next, we show the upper bound in (49):

$$\tilde{\forall}_{T_\rho, x} T_\rho \in \mathbf{T}_\rho, x \in T_\rho : \mu_\rho(x) \lessapprox \delta^{-C\sqrt{\epsilon}} \rho^{-(3-d)}.$$

We re-arrange this using (25) as

$$\int_{\mu_\rho \gtrapprox \delta^{-C\sqrt{\epsilon}} \rho^{-(3-d)}} \mu_\rho \lessapprox \delta^{c\sqrt{\epsilon}}.$$

On the other hand, from (14) we have

$$\|\mu_\rho\|_{d'} \lessapprox \rho^{\frac{d-3}{d}}, \tag{54}$$

and the claim then follows from (10).

We now turn to the lower bound in (49):

$$\tilde{\forall}_{T_\rho, x} T_\rho \in \mathbf{T}_\rho, x \in T_\rho : \mu_\rho(x) \gtrapprox \delta^{C\sqrt{\epsilon}} \rho^{-(3-d)}.$$

We may re-arrange this using (25) as

$$\int_{\mu_\rho \lessapprox \delta^{C\sqrt{\epsilon}} \rho^{-(3-d)}} \mu_\rho \lessapprox \delta^{c\sqrt{\epsilon}}. \tag{55}$$

The claim then follows from (11), (21), and the fact that μ_ρ is supported on E_ρ.

Next, we show

$$\tilde{\forall}_{T_\rho, x} T_\rho \in \mathbf{T}_\rho, x \in T_\rho : (52) \text{ holds for all directions } \omega \text{ and all } \delta \leq \theta \ll 1.$$

By repeating the argument in Proposition 6.2, it suffices to obtain

$$\tilde{\forall}_{T_\rho, x} T_\rho \in \mathbf{T}_\rho, x \in T_\rho : (52) \text{ holds for all directions } \omega$$

uniformly in θ.

Fix θ; by (25), it suffices to show that

$$\int_X \mu_\rho \lessapprox \delta^{c\sqrt{\epsilon}}$$

where X is the set of points in $\mathbf{R}^3$ for which (52) fails for some ω.

By (8) and (54) it suffices to show that

$$\rho^{\frac{d-3}{d}}|X|^{1/d} \lessapprox \delta^{c\sqrt{\epsilon}},$$

or in other words

$$|X| \lessapprox \delta^{c\sqrt{\epsilon}}\rho^{3-d}.$$

If (52) fails for some ω, it also fails (with slightly different constants) for some ω in a fixed $c\theta$-separated set. Since this set has cardinality $O(\theta^{-2})$, it suffices to show that

$$|\{x \in \mathbf{R}^3 : (52) \text{ fails for } \omega\}| \lessapprox \delta^{c\sqrt{\epsilon}}\rho^{3-d}\theta^2$$

uniformly in ω.

Fix ω. We must show

$$|\{\sum_{T_\rho \in \mathbf{T}_\rho{}'} \chi_{T_\rho} \gtrapprox \delta^{-C\sqrt{\epsilon}}\theta^c\rho^{-(3-d)}\}| \lessapprox \delta^{c\sqrt{\epsilon}}\rho^{3-d}\theta^2 \tag{56}$$

where $\mathbf{T}_\rho{}'$ is the set of tubes in T_ρ whose directions lie in $B(\omega, \theta)$.

By (12) and (14), the left-hand side of (56) is majorized by

$$[\delta^{-C\sqrt{\epsilon}}\theta^c\rho^{-(3-d)}]^{-d'}[\rho^{1-\frac{3}{d}}\theta^{1+\frac{1}{d}}]^{d'},$$

which we rearrange as

$$\delta^{Cd'\sqrt{\epsilon}}\theta^{\frac{3-d}{d-1}-cd'}\rho^{3-d}\theta^2.$$

Since $d < 3$, (56) follows if c is chosen sufficiently small.

Finally, we address (51). This is easily dealt with because (51) follows from the other properties in $P_2(x_0)$. To see this, first observe from (49) and (50) that $P_1(x_0, T_\rho)$ holds for at least one tube $T_\rho \in \mathbf{T}_\rho$. From (44) and (45) we thus see that there exists an $x \in B(x_0, C\rho)$ such that (42) holds for all $\delta \leq \sigma \ll 1$. Applying this with $\sigma = C\rho$ gives the result. (Alternatively, one can modify the proof of Proposition 6.2 to obtain (51)). ■

In the next few sections we will explore some surprising consequences of the property $P_2(x_0)$, namely planiness and graininess.

7. Triple intersections

In this section we fix x_0 to be a point in $\mathbf{R}^3$ such that $P_2(x_0)$ holds.

Let T_ρ be a tube in $\mathbf{T}_\rho(x_0)$ such that $P_1(x_0, T_\rho)$ holds. By Proposition 6.6 this situation occurs fairly often.

Let $A(x_0, T_\rho)$ denote the set

$$A(x_0, T_\rho) = E_\delta[T_\rho] \cap B(x_0, C\rho). \tag{57}$$

From (44) we have

$$|A(x_0, T_\rho)| \gtrapprox \delta^{C\sqrt{\epsilon}} \delta^{3-d} \rho^d.$$

On the other hand, from (51) we have

$$|\bigcup_{T_\rho \in \mathbf{T}_\rho(x_0)} A(x_0, T_\rho)| \lessapprox \delta^{-C\sqrt{\epsilon}} \delta^{3-d} \rho^d.$$

Thus we expect a lot of overlap between the $A(x_0, T_\rho)$. In particular, we expect the size of the triple intersection

$$A(x_0, T^1_\rho) \cap A(x_0, T^2_\rho) \cap A(x_0, T^3_\rho) \tag{58}$$

to be quite large for many triples T^i_ρ, $i = 1, 2, 3$.

However, it turns out that we can get a non-trivial bound on the size of (58) if tubes T^i_ρ are not coplanar. This is basically because (42) forces the projections of (58) along the directions of T^i_ρ to be small. (For a discussion of the relationship between the volume of a set and the areas of its projections, see [7]).

When the tubes T^i_ρ all lie close to a single plane π, then one cannot directly improve the trivial bound on (58) given by (44). However, we can show that this trivial bound is only attained when the sets $A(x_0, T^i_\rho)$ are "grainy", in the sense that they consist of a union of squares parallel to π.

More precisely, we have

Definition 7.1. Let π be a plane. We define a *square parallel to* π to be any rectangular box Q of dimensions $\delta \times \rho \times \rho$ whose two long sides are parallel to π.

Lemma 7.2. *Let x_0 be a point in $\mathbf{R}^3$ such that $P_2(x_0)$ holds, and let $T^1_\rho, T^2_\rho, T^3_\rho$ be three tubes in $\mathbf{T}_\rho(x_0)$. Let v_i be any vectors such that*

$$v_i = \rho \ \mathrm{dir}(T^i_\rho) + O(\delta). \tag{59}$$

Let $\pi(v_1, v_2)$ denote the plane spanned by v_1 and v_2, $1 < M \ll \delta^{-1/4}$ be an arbitrary number, and F be an arbitrary subset of $\mathbf{R}^3$.

- *(Non-planar tubes) If $\angle v_1, v_2 \geq M^{-1}$, $\angle \pi(v_1, v_2), v_3 \geq \rho M^2$, and (42) holds for all $x \in F$ and $\delta \leq \sigma \leq 1$, then*

  $$|A(x_0, T^1_\rho) \cap A(x_0, T^2_\rho) \cap A(x_0, T^3_\rho) \cap F| \lessapprox \delta^{-C\sqrt{\epsilon}} \delta^{3-d} \rho^d M^{-c}. \tag{60}$$

- *(Planar tubes) If $\angle v_1, v_2 \geq M^{-1}$ and $\angle \pi(v_1, v_2), v_3 \leq \rho M^2$ then*

  $$|A(x_0, T^1_\rho) \cap A(x_0, T^2_\rho) \cap A(x_0, T^3_\rho) \cap F| \lessapprox \delta^{-C\sqrt{\epsilon}} \delta^{3-d} \rho^d M^C (\sup_Q \frac{|F \cap Q|}{|Q|})^c, \tag{61}$$

where Q ranges over all squares parallel to $\pi(v_1, v_2)$.

Here $c > 0$ and $C > 0$ are absolute constants.

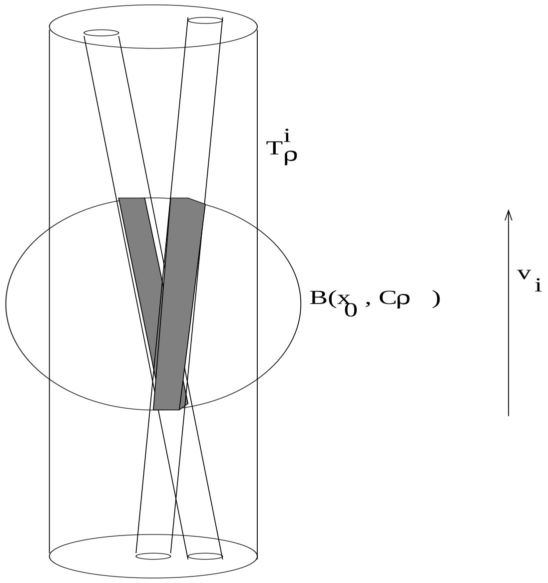

FIGURE 3. The shaded region is a portion of a typical A_i. Note that A_i essentially consists of $\delta \times \rho$ tubes oriented in the direction v_i.

Proof Write $A_i = A(x_0, T^i_\rho)$ for short. The key observation is that for every $T \in \mathbf{T}[T^i_\rho]$, the set $T \cap B(x_0, C\rho)$ is essentially constant in the direction v_i. More precisely, we have the elementary pointwise estimate (see Figure 3)

$$\chi_{B(x_0,C\rho)\cap T} \lesssim \mathbf{E}_{v_i}(\chi_{B(x_0,2C\rho)\cap CT}) \tag{62}$$

where E_v is the averaging operator

$$\mathbf{E}_v(f(x)) = \int_{|t|\lesssim 1} f(x+vt)\, dt.$$

Summing this in T we obtain

$$\chi_{A_i} \lesssim \mathbf{E}_{v_i}(\chi_{\tilde{A}_i}) \tag{63}$$

where

$$\tilde{A}_i = B(x_0, 2C\rho) \cap \bigcup_{T\in\mathbf{T}[T^i_\rho]} CT.$$

It is easy to see that $|\tilde{A}_i| \sim |A_i|$. In particular, from (51) we have

$$|\tilde{A}_i| \lessapprox \delta^{-C\sqrt{\epsilon}}\delta^{3-d}\rho^d. \tag{64}$$

To utilize (63) we invoke the following estimate, which is a variant of the argument in [7].

Lemma 7.3. *Let v_1, v_2, v_3 be any three linearly independent vectors, and let F be a subset of $\boldsymbol{R}^3$. Then for any functions f_1, f_2, f_3, we have*

$$\int_F \boldsymbol{E}_{v_1}(f_1)\boldsymbol{E}_{v_2}(f_2)\boldsymbol{E}_{v_3}(f_3) \lesssim \sup_P \left(\frac{|F\cap P|}{|P|}\right)^{1/3} \prod_{i=1}^3 \|f_i\|_3,$$

where $c > 0$ is an absolute constant, and P ranges over all parallelopipeds with edge vectors v_1, v_2, v_3.

We remark that the factor $\sup_P(\frac{|F\cap P|}{|P|})^{1/3}$ represents a gain over the trivial estimate obtained via Hölder's inequality.

Proof The statement of the lemma is invariant under affine transformations, so we may rescale v_1, v_2, v_3 to be the cardinal directions e_1, e_2, e_3 respectively. It suffices to show that

$$\int_{F\cap P} \mathbf{E}_{v_1}(f_1)\mathbf{E}_{v_2}(f_2)\mathbf{E}_{v_3}(f_3) \lesssim (|F\cap P|)^{1/3} \prod_{i=1}^3 \|f_i\|_{L^3(CP)},$$

for all unit cubes P, since the claim follows by summing over a partition of $\mathbf{R}^3$ and using Hölder's inequality. We may assume that P is centered at the origin, that $F \subset P$ and that f_i are supported on CP.

We have the pointwise estimate

$$\mathbf{E}_{v_1}(f_1)(x_1, x_2, x_3) \lesssim \overline{f}_1(x_2, x_3),$$

where $x = (x_1, x_2, x_3)$ and

$$\overline{f}_1(x_2, x_3) = \int_{|x_1|\lesssim 1} f_1(x_1, x_2, x_3)\ dx_1.$$

Similarly for cyclic permutations of $1, 2, 3$. Since $\|\overline{f}_i\|_3 \lesssim \|f_i\|_3$ by Young's inequality, we are thus reduced to showing that

$$\int_P \chi_F(x)\overline{f}_1(x_2, x_3)\overline{f}_2(x_3, x_1)\overline{f}_3(x_1, x_2)\ dx_1 dx_2 dx_3 \lesssim |F|^c \|\overline{f}_1\|_3 \|\overline{f}_2\|_3 \|\overline{f}_3\|_3$$

We rewrite this as

$$\int_{|x_1|\lesssim 1} \left(\int_{|x_2|,|x_3|\lesssim 1} \chi_F^{x_1}(x_2, x_3)\overline{f}_1(x_2, x_3)\overline{f}_2^{x_1}(x_3)\overline{f}_3^{x_1}(x_2)\ dx_2 dx_3\right)\ dx_1,$$

where $\chi_F^{x_1}(x_2, x_3) = \chi_F(x_1, x_2, x_3)$, etc. By Hölder's inequality this is majorized by

$$\int_{|x_1|\lesssim 1} \|\chi_F^{x_1}\|_3 \|\overline{f}_1\|_3 \|\overline{f}_3^{x_1} \otimes \overline{f}_2^{x_1}\|_3\ dx_1,$$

where the norms are taken in the x_2, x_3 variables only. We can simplify this to

$$\|\overline{f}_1\|_3 \int_{|x_1| \lesssim 1} \|\chi_F^{x_1}\|_3 \|\overline{f}_3^{x_1}\|_3 \|\overline{f}_2^{x_1}\|_3 \ dx_1,$$

and the claim follows from applying Hölder in the x_1 variable. ∎

Combining this estimate with (63) and (64) we obtain

$$|A_1 \cap A_2 \cap A_3 \cap F| \lessapprox \delta^{-C\sqrt{\epsilon}} \delta^{3-d} \rho^d (\sup_P \frac{|F \cap P|}{|P|})^{1/3} \tag{65}$$

where P ranges over all parallelopipeds with edge vectors v_1, v_2, v_3.

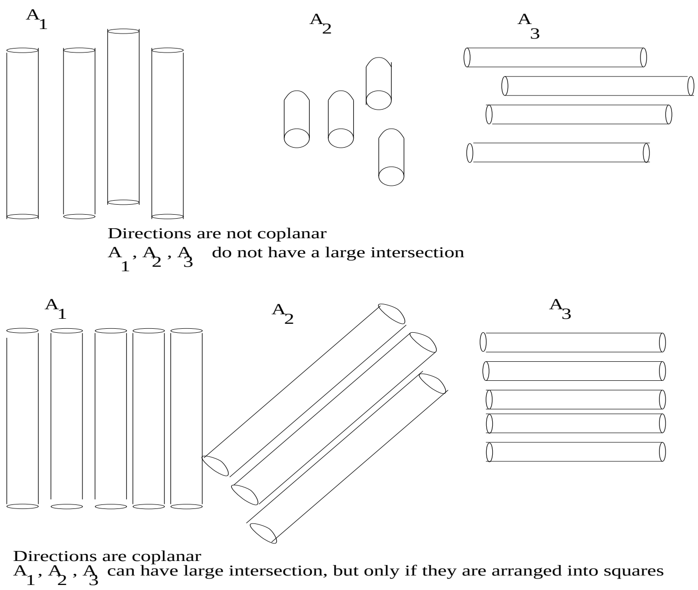

FIGURE 4. The two cases of Lemma 7.2.

We now prove the two claims of Lemma 7.2.

Proof of (60). Since $A_i \subset E_\delta$, we may assume without loss of generality that $F \subset E_\delta$. Since we are assuming (42) to hold on F, we thus see that

$$\frac{|F \cap B|}{|B|} \lessapprox \delta^{-C\sqrt{\epsilon}} (\frac{\sigma}{\delta})^{-(3-d)}$$

for any ball of radius σ.

Let P be any parallelopiped with edge vectors v_1, v_2, v_3. From the hypotheses on v_1, v_2, v_3 and some elementary geometry we see that

$$|P| \sim N_{M\delta}(P),$$

so that P can be efficiently covered by a finitely overlapping collection of $M\delta$-balls. From a simple covering argument we thus obtain

$$\frac{|F \cap P|}{|P|} \lessapprox \delta^{-C\sqrt{\epsilon}} M^{-(3-d)}$$

uniformly in P, and (60) follows from (65).

Proof of (61). By perturbing v_i if necessary we may assume that $\angle \pi(v_1, v_2), v_3 \gtrsim \delta$. The claim then follows from (65) and the geometric observation that, if P is any parallelopiped with edge vectors v_1, v_2, v_3, then P can be covered by $O(M^2)$ squares Q parallel to $\pi(v_1, v_2)$, and $|P| \gtrsim M^{-1}|Q|$ for any one of these squares.

∎

8. Consequences of Lemma 7.2: planiness and graininess

In this section we shall utilize the first conclusion Lemma 7.2 to enforce a "planiness" property on the tubes in $\mathbf{T}_\rho$; roughly speaking, we require for all x_0 satisfying the property P_2 that most of the tubes in $\mathbf{T}_\rho(x_0)$ to be coplanar. Then we will shift scales in order to also enforce this property on the δ-tubes $\mathbf{T}$. Later in this section we will also use the second conclusion of Lemma 7.2 to also enforce a certain "graininess" property on the set E_δ.

We begin this program with

Proposition 8.1. *Let x_0 be such that $P_2(x_0)$ holds. Then one can find a family of planes $\Pi_\rho = \Pi_\rho(x_0)$ through x_0 and subcollections $\boldsymbol{T}_\rho(x_0, \pi) \subset \boldsymbol{T}_\rho(x_0)$ for all $\pi \in \Pi_\rho(x_0)$ such that*

- *For all $\pi \in \Pi_\rho(x_0)$ and $T_\rho \in \boldsymbol{T}_\rho(x_0, \pi)$ we have*

$$T_\rho \subset N(\pi), \tag{66}$$

  where $N(\pi)$ is the $\delta^{-C_2\sqrt{\epsilon}}\rho$-neighbourhood of π, and C_2 is a constant to be chosen later.
- *The $\boldsymbol{T}_\rho(x_0, \pi)$ are disjoint as π varies, and*

$$\#\boldsymbol{T}_\rho(x_0, \pi) \gtrapprox \delta^{C\sqrt{\epsilon}} \rho^{-(3-d)} \tag{67}$$

  for all $\pi \in \Pi_\rho(x_0)$.
- *We have*

$$1 \leq \#\Pi_\rho(x_0) \lessapprox \delta^{-C\sqrt{\epsilon}}. \tag{68}$$

- *We have*

$$\tilde{\forall}_{T_\rho} T_\rho \in \boldsymbol{T}_\rho(x_0) : T_\rho \in \bigcup_{\pi \in \Pi_\rho(x_0)} \boldsymbol{T}_\rho(x_0, \pi) \tag{69}$$

- *For every $\pi \in \Pi_\rho(x_0)$, the set $\boldsymbol{T}_\rho(x_0, \pi)$ is a subset of*

$$\{T_\rho \in \boldsymbol{T}_\rho(x_0) : P_1(x_0, T_\rho) \text{ holds}\}. \tag{70}$$

For the purposes of visualization we recommend thinking of Π_ρ as consisting of a single plane π, and $\mathbf{T}_\rho(x_0, \pi)$ as equal to $\mathbf{T}_\rho(x_0)$.

Proof The idea of the proof is to iterate the following lemma.

Lemma 8.2. *Let x_0 be such that $P_2(x_0)$ holds. Suppose that $\boldsymbol{T}'_\rho$ is a subset of* (70) *such that*

$$\# \boldsymbol{T}'_\rho \geq \delta^{\sqrt{\epsilon}} \# \boldsymbol{T}_\rho(x_0). \tag{71}$$

Then one can find a plane π through x_0 such that

$$\#\{T_\rho \in \boldsymbol{T}'_\rho : T_\rho \subset N(\pi)\} \gtrapprox \delta^{C_2\sqrt{\epsilon}} \rho^{-(3-d)}, \tag{72}$$

if C_2 is chosen sufficiently large.

Proof Suppose for contradiction that (72) failed for all planes π through x_0.

Let F denote the set

$$F = \{x \in E_\delta \cap B(x_0, C\rho) : (42) \text{ holds for all } \delta \leq \sigma \ll 1\}. \tag{73}$$

Let T_ρ be any element of $\mathbf{T}'_\rho$. From (45), (46), and Lemma 5.2 we have

$$\tilde{\forall}_{T,x}\, T \in \mathbf{T}[T_\rho], x \in T \cap B(x_0, C\rho) : x \in F, \mu_\delta[T_\rho](x) \leq \delta^{-\sqrt{\epsilon}} \rho^{-(3-d)}. \tag{74}$$

From (31) we thus obtain

$$\int_{F \cap \{\mu_\delta[T_\rho] \leq \delta^{-\sqrt{\epsilon}} \rho^{-(3-d)}\}} \mu_\delta[T_\rho] \approx \int_{B(x_0, C\rho)} \mu_\delta[T_\rho].$$

By (27) and (11) we thus have

$$\delta^{-\sqrt{\epsilon}} \rho^{-(3-d)} |A(x_0, T_\rho) \cap F| \gtrapprox \delta^{-3/2},$$

where $A(x_0, T_\rho)$ was defined in (57). We re-arrange this as

$$\int_F \chi_{A(x_0, T_\rho)} \gtrapprox \delta^{\sqrt{\epsilon}} \delta^{3-d} \rho^d.$$

Summing this over all $T_\rho \in \mathbf{T}'_\rho$ and using (71) we obtain

$$\int_F \sum_{T_\rho \in \mathbf{T}'_\rho} \chi_{A(x_0,T_\rho)} \gtrapprox \delta^{C\sqrt{\epsilon}} \# \mathbf{T}_\rho(x_0)[\delta^{3-d}\rho^d]. \tag{75}$$

On the other hand, we will shortly prove the estimate

$$\sum_{T^1_\rho}\sum_{T^2_\rho}\sum_{T^3_\rho} |A(x_0,T^1_\rho) \cap A(x_0,T^2_\rho) \cap A(x_0,T^3_\rho) \cap F| \lessapprox \delta^{-C\sqrt{\epsilon}}\delta^{cC_2\sqrt{\epsilon}}\delta^{3-d}\rho^d \# \mathbf{T}_\rho(x_0)^3, \tag{76}$$

where the tubes T^i_ρ will always be assumed to range over $\mathbf{T}'_\rho$, and c denotes a constant independent of C_2 but which can vary from line to line. Rewriting the left-hand side of (76) as

$$\| \sum_{T_\rho \in \mathbf{T}'_\rho} \chi_{A(x_0,T_\rho)} \|^3_{L^3(F)}$$

and using (8) (with d replaced by 3/2) and (51), we thus obtain

$$\| \sum_{T_\rho \in \mathbf{T}'_\rho} \chi_{A(x_0,T_\rho)} \|_{L^1(F)} \lessapprox \delta^{-C\sqrt{\epsilon}}\delta^{cC_2\sqrt{\epsilon}}\delta^{3-d}\rho^d \# \mathbf{T}_\rho(x_0)$$

which contradicts (75) if C_2 is chosen sufficiently large.

It remains to prove (76). Fix $T^1_\rho, T^2_\rho, T^3_\rho$, and choose v_i so that (59) holds. We divide the contributions to (76) into three cases.

Firstly, we consider the contribution when v_1 and v_2 are close together, or more precisely when

$$\angle v_1, v_2 \lessapprox \delta^{C_2\sqrt{\epsilon}/2}. \tag{77}$$

Estimating the summand crudely by (51), we may majorize the contribution of this case by

$$\delta^{-C\sqrt{\epsilon}}\delta^{3-d}\rho^d \#\{(T^1_\rho, T^2_\rho, T^3_\rho) \in {\mathbf{T}'_\rho}^3 : (77) \text{ holds.}\}. \tag{78}$$

However, by (52) we see that for any fixed T^1_ρ there are at most

$$\lessapprox \delta^{-C\sqrt{\epsilon}}(\delta^{C_2\sqrt{\epsilon}/2})^c \rho^{-(3-d)}$$

tubes T^2_ρ for which (77) holds; by (49) this is majorized by

$$\lessapprox \delta^{-C\sqrt{\epsilon}}\delta^{cC_2\sqrt{\epsilon}} \# \mathbf{T}_\rho(x_0).$$

Thus the contribution of this case to (76) is acceptable.

Next we estimate the contribution for which (77) fails, but for which

$$\angle \pi(v_1, v_2), v_3 \lessapprox \delta^{-C_2\sqrt{\epsilon}}\rho. \tag{79}$$

where $\pi(v_1, v_2)$ is the plane spanned by v_1, v_2. By crudely estimating the summand by (51) again, the contribution to (76) is at most

$$\delta^{-C\sqrt{\epsilon}}\delta^{3-d}\rho^d \#\{(T_\rho^1, T_\rho^2, T_\rho^3) \in (\mathbf{T}'_\rho)^3 : (79) \text{ holds.}\}.$$

However, we are assuming (72) to fail for all planes π. This implies that for any fixed T_ρ^1, T_ρ^2, there are at most $\delta^{C_2\sqrt{\epsilon}}\rho^{-(3-d)}$ tubes T_ρ^3 for which (79) holds. By repeating the previous calculation we see that the contribution of this case to (76) is also acceptable.

Finally, we consider the contribution for which (77) and (79) both fail. In this case the first part Lemma 7.3 applies with $M = \delta^{-C_2\sqrt{\epsilon}/2}$. Using (60) we therefore see that this contribution is also acceptable. Combining the three contributions together we obtain the desired estimate. ∎

With this lemma it is easy to prove the Proposition by the following construction. Initialize $\Pi_\rho(x_0)$ to be the empty set and $\mathbf{T}'_\rho$ to be (70). From (50) we see that (71) currently holds.

We now perform the following iteration. We use Lemma 8.2 to extract a plane π which obeys (72) with respect to $\mathbf{T}'_\rho$. We add this plane π to $\Pi_\rho(x_0)$ and define

$$\mathbf{T}_\rho(x_0, \pi) = \{T_\rho \in \mathbf{T}'_\rho : T_\rho \subset N(\pi)\}.$$

We now remove all the elements of $\mathbf{T}_\rho(x_0, \pi)$ from $\mathbf{T}'_\rho$ and repeat the iteration until (71) fails, at which point we halt the construction.

It is easy to verify that this iteration gives all the properties claimed in the Proposition. The bound (69) follows by construction. The upper bound on $\#\Pi_\rho$ follows from (72) and (49), while the lower bound on $\#\Pi_\rho$ follows from (69). ∎

Henceforth we assume that the collections $\Pi_\rho(x_0)$ and $\mathbf{T}_\rho(x_0, \pi)$ have been chosen for all x_0 such that $P_2(x_0)$ held.

Having obtained planiness, we now turn to the companion property of graininess. We remark that the heuristic that one should be able to reduce the study of Besicovitch sets to the grainy case is due to Tom Wolff (private communication). Much of the work in this paper was motivated by a desire to make this heuristic rigorous.

Definition 8.3. Let π be a plane, and x_0 be a point in $\mathbf{R}^3$. A set S is said to be *grainy* with respect to (x_0, π) if $S \subset B(x_0, C\rho)$ and

$$S = \bigcup_{Q \in \mathbf{Q}} Q \tag{80}$$

for some collection $\mathbf{Q}$ of squares Q parallel to π such that

$$\#\mathbf{Q} \lessapprox \delta^{-C\sqrt{\epsilon}}\delta^{(2-d)/2}. \tag{81}$$

If S is a grainy set, we define its dilate CS to be

$$CS = \bigcup_{Q \in \mathbf{Q}} CQ$$

where CQ is the dilate of Q by C around its center.

Note if S is grainy, then CS is also grainy with slightly larger constants. Also, from (81) we have

$$|S|, |CS| \lesssim \delta\rho^2 \#\mathbf{Q} \lessapprox \delta^{-C\sqrt{\epsilon}} \delta^{3-d} \rho^d. \tag{82}$$

Proposition 8.4. *Let the hypotheses and notation be as in Proposition* 8.1. *Then for each* $\pi \in \Pi_\rho(x_0)$ *there exists a set* $S(x_0, \pi)$ *which is grainy with respect to* (x_0, π)*, such that*

$$\tilde{\forall}_{T_\rho, T, x} \; T_\rho \in \boldsymbol{T}_\rho(x_0, \pi), T \in \boldsymbol{T}[T_\rho], x \in T \cap B(x_0, C\rho) : x \in S(x_0, \pi). \tag{83}$$

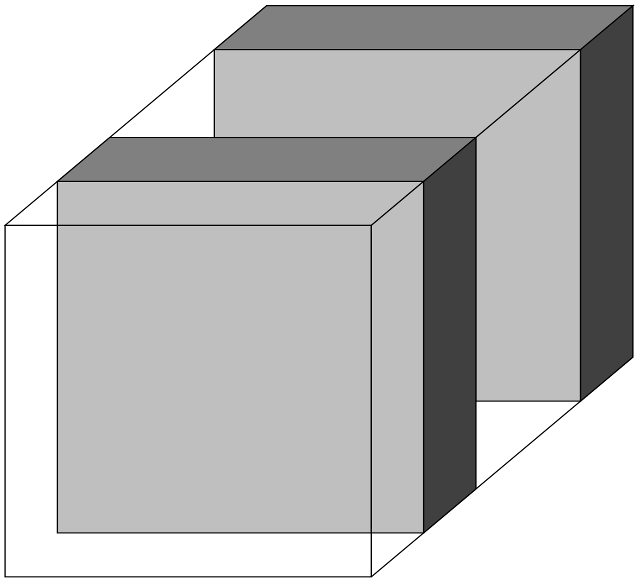

FIGURE 5. A (simplified) typical intersection of E_δ with a ρ-cube. In this case the plane π is oriented in the xz direction.

Proof We will modify the argument in Proposition 8.1.

Fix x_0, π. Cover $B(x_0, C\rho)$ by a finitely overlapping collection of squares Q parallel to π. Define $\mathbf{Q}(x_0, \pi)$ to be the subcollection of these squares Q for which

$$\frac{|E_\delta \cap Q|}{|Q|} \gtrapprox \delta^{C_3\sqrt{\epsilon}}. \tag{84}$$

where $C_3 > 0$ is a constant we will choose later. Define $S(x_0, \pi)$ by

$$S(x_0, \pi) = \bigcup_{Q \in \mathbf{Q}(x_0,\pi)} Q.$$

From (84), (51) and the estimate $|Q| \sim \delta^2$ we obtain (81), and so $S(x_0, \pi)$ is grainy with respect to (x_0, π).

Define the set X by

$$X = (E_\delta \cap B(x_0, C\rho)) \backslash S(x_0, \pi).$$

Then by construction

$$\sup_Q \frac{|X \cap Q|}{|Q|} \lessapprox \delta^{C_3 \sqrt{\epsilon}} \tag{85}$$

where Q ranges over all squares parallel to π.

To finish the proof we must show (83). By summing (46) for all $T_\rho \in \mathbf{T}_\rho(x_0, \pi)$ we have

$$\tilde{\forall}_{T_\rho, T, x} \; T_\rho \in \mathbf{T}_\rho(x_0, \pi), T \in \mathbf{T}[T_\rho], x \in T \cap B(x_0, C\rho) : (47) \text{ holds.}$$

Thus by Lemma 5.2 it suffices to show that

$$\tilde{\forall}_{T_\rho} T_\rho \in \mathbf{T}_\rho(x_0, \pi), T \in \mathbf{T}[T_\rho], x \in T \cap B(x_0, C\rho) : ((47) \implies x \notin X).$$

We rewrite this as

$$\int_X \sum_{T_\rho \in \mathbf{T}_\rho(x_0,\pi) : \mu_\delta[T_\rho] \leq \delta^{-\sqrt{\epsilon}} \rho^{-(3-d)}} \mu_\delta[T_\rho] \lessapprox \delta^{c\sqrt{\epsilon}} \int_{B(x_0, C\rho)} \sum_{T_\rho \in \mathbf{T}_\rho(x_0,\pi)} \mu_\delta[T_\rho]. \tag{86}$$

We majorize the left-hand side of (86) as

$$\int_X \sum_{T_\rho \in \mathbf{T}_\rho(x_0,\pi)} \delta^{-\sqrt{\epsilon}} \rho^{-(3-d)} \chi_{A(x_0, T_\rho)}.$$

The right-hand side of (86) can be estimated using (67) and (27). Thus (86) reduces to showing that

$$\int_X \sum_{T_\rho \in \mathbf{T}_\rho(x_0,\pi)} \delta^{-\sqrt{\epsilon}} \rho^{-(3-d)} \chi_{A(x_0, T_\rho)} \lessapprox \delta^{C\sqrt{\epsilon}} \rho^{-(3-d)} \delta^{3/2}$$

for some sufficiently large C. This simplifies by (49) to

$$\int_X \sum_{T_\rho \in \mathbf{T}_\rho(x_0,\pi)} \chi_{A(x_0, T_\rho)} \lessapprox \delta^{C\sqrt{\epsilon}} \delta^{3-d} \rho^d \# \mathbf{T}_\rho(x_0).$$

We will shortly prove

$$\sum_{T_\rho^1} \sum_{T_\rho^2} \sum_{T_\rho^3} |A(x_0, T_\rho^1) \cap A(x_0, T_\rho^2) \cap A(x_0, T_\rho^3) \cap X| \lessapprox \delta^{-C\sqrt{\epsilon}} \delta^{cC_3\sqrt{\epsilon}} \delta^{3-d} \rho^d \# \mathbf{T}_\rho(x_0)^3, \tag{87}$$

where the T_ρ^i are assumed to range over $\mathbf{T}_\rho(x_0,\pi)$; the desired estimate then follows as before from (8) with d replaced by $3/2$ and then (51), if C_3 is chosen sufficiently large.

As in the proof of Proposition 8.1, the contributions to (87) when

$$\angle v_1, v_2 \lessapprox \delta^{C_3\sqrt{\epsilon}} \tag{88}$$

or when

$$\angle \pi(v_1,v_2), v_3 \gtrapprox \delta^{-2C_3\sqrt{\epsilon}}\rho. \tag{89}$$

are acceptable if C_3 is sufficiently large. Thus it only remains to consider the contribution for which (88) and (89) both fail. However, in this case the second part of Lemma 7.2 applies. From (61) and (85) we have

$$|A(x_0,T_\rho^1)\cap A(x_0,T_\rho^2)\cap A(x_0,T_\rho^3)\cap X| \lessapprox \delta^{-C\sqrt{\epsilon}}\delta^{3-d}\rho^d\delta^{cC_3\sqrt{\epsilon}}.$$

Thus the contribution from this case to (87) are also acceptable, as desired. ■

Henceforth we assume that the sets $S(x_0,\pi)$ have been chosen for all x_0 such that $P_2(x_0)$ holds, and all $\pi \in \Pi_\rho(x_0)$. The constants C_2 and C_3 are now fixed, and future constants C may implicitly depend on these two.

9. Consistency of planes and squares

In Proposition 6.6 we showed that any sticky collection of tubes satisfying (20) and (21) had to satisfy the property $P_2(x)$ extremely often. This property in turn implied two strong properties on the collection, namely planiness and graininess, by Propositions 8.1 and 8.4.

The planiness property is really a property about the ρ-tubes $\mathbf{T}_\rho$ rather than the δ-tubes $\mathbf{T}$. By taking advantage of the double-sticky assumption and changing scale from δ to ρ, we shall be able to also impose a planiness property in the δ-tubes. More precisely, we show

Definition 9.1. Define $P_3(x)$ to be property that one can find a family of planes $\Pi_\delta(x)$ through x and subcollections $\mathbf{T}(x,\pi)\subset\mathbf{T}(x)$ for all $\pi\in\Pi_\delta(x)$ such that

- For all $\pi\in\Pi_\delta(x)$ and $T\in\mathbf{T}(x,\pi)$ we have
$$T\subset N_{\delta^{-C\sqrt{\epsilon}}\delta}(\pi). \tag{90}$$
- The $\mathbf{T}(x,\pi)$ are disjoint as π varies, and $\#\mathbf{T}(x,\pi)\gtrapprox\delta^{C\sqrt{\epsilon}}\delta^{d-3}$ for all $\pi\in\Pi_\delta(x)$.
- We have
$$1\le\#\Pi_\delta(x)\lessapprox\delta^{-C\sqrt{\epsilon}}. \tag{91}$$

- We have

$$\tilde{\forall}_{T} T \in \mathbf{T}(x) : T \in \mathbf{T}(x, \pi) \text{ for some } \pi \in \Pi_\delta(x) \tag{92}$$

- We have

$$\#\{T \in \mathbf{T}(x_0) : \mathrm{dir}(T) \in B(\omega, \theta)\} \lessapprox \delta^{-C\sqrt{\varepsilon}} \theta^c \delta^{d-3} \tag{93}$$

for all directions ω and all $\delta \leq \theta \ll 1$.

Proposition 9.2. *Suppose that there is an x-ray estimate at dimension d, and that there exists a Besicovitch set E with* $\overline{dim}(E) < d + \varepsilon$. *Then there exists a sticky collection* $\boldsymbol{T}$ *of tubes such that* (20) *and* (21) *hold, and such that*

$$\tilde{\forall}_{T,x}\, T \in \boldsymbol{T}, x \in T : P_3(x). \tag{94}$$

Proof By Proposition 3.3 we can find a doubly sticky collection $\mathbf{T}$ of tubes which obey (20), (21), and (22). By Proposition 6.6 the property P_2 occurs almost always, and the conclusions of Proposition 8.1 obtain whenever P_2 is satisfied. We now replace δ by ρ, $\mathbf{T}$ with $\mathbf{T}_\rho$, and $\mathbf{T}_\rho$ with $\mathbf{T}_{\delta^{1/4}}$. The verification of all the required properties now follows from Proposition 8.1 and (52). ■

Henceforth we fix $\mathbf{T}$ to be a collection of tubes satisfying all the conclusions of Proposition 9.2, and assume that we have chosen the collections $\Pi_\delta(x)$ and $\mathbf{T}(x, \pi)$ for all x such that $P_2(x)$ holds, and all $\pi \in \Pi_\delta(x)$. This changing of scale from δ to ρ is intruigingly reminiscent of the techniques employed in proving restriction theorems (see e.g. [1], [2], [3], [11], [12]).

Since $\mathbf{T}$ is sticky and satisfies (20) and (21), we may apply Proposition 6.6 again and obtain (53). In particular, the conclusions of Proposition 8.1 and Proposition 8.4 obtain for a large number of x_0.

We thus have two scales of structure on our set, one at scale δ arising from P_3, and one arising from the Propositions in the previous section. Fortunately, there is a close relationship between the two structures, namely that the planes in $\Pi_\delta(x)$ and the planes in $\Pi_\rho(x_0)$ are mostly parallel if $x \in B(x_0, C\rho)$. In particular, the planes in $\Pi_\delta(x)$ are mostly parallel to the squares Q that contain x. More precisely:

Definition 9.3. We define $P_4(x_0, x, T_\rho, T)$ to be the property that $P_2(x_0)$ and $P_3(x)$ hold, that $x \in B(x_0, C\rho)$, and there exist planes $\pi \in \Pi_\delta(x)$, $\pi' \in \Pi_\rho(x_0)$ such that $T \in \mathbf{T}(x, \pi) \cap \mathbf{T}[T_\rho]$, $T_\rho \in \mathbf{T}_\rho(x_0, \pi')$, $x \in S(x_0, \pi')$, and we have the consistency condition

$$\angle \pi, \pi' \lessapprox \delta^{-C\sqrt{\epsilon}} \rho. \tag{95}$$

Proposition 9.4. *If the constants in the above definition are chosen appropriately, then we have*

(96)
$$\tilde{\forall}_{T_\rho, x_0, T, x} T_\rho \in \boldsymbol{T}_\rho, x_0 \in T_\rho, T \in \boldsymbol{T}[T_\rho], x \in T \cap B(x_0, C\rho) : P_4(x_0, x, T_\rho, T).$$

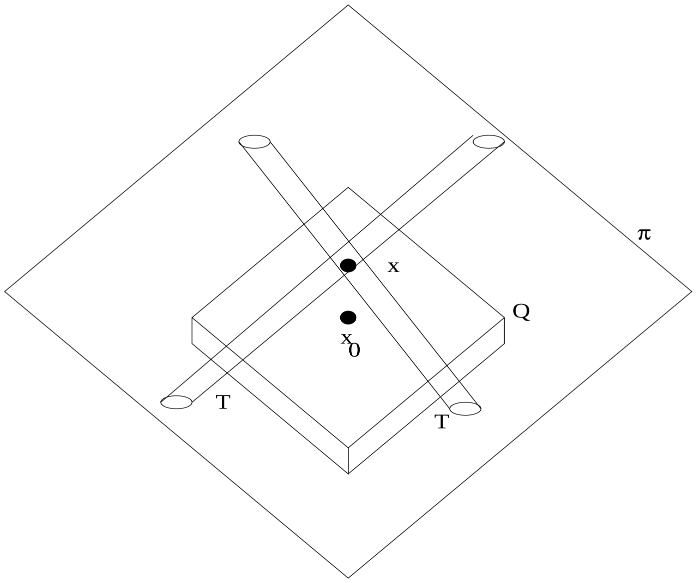

FIGURE 6. A typical situation in which $P_4(x_0, x, T_\rho, T)$ holds. The tubes through x lie on the plane π, which is essentially parallel to the square $Q \subset S(x_0, \pi')$ which contains x. The plane π' (not pictured) differs in angle from π by about ρ.

Proof We need to show that $\tilde{\forall}_{T_\rho, x_0, T, x} Q : P_4$, where $Q = Q(x_0, x, T_\rho, T)$ denotes the property that

$$T_\rho \in \mathbf{T}_\rho, x_0 \in T_\rho, T \in \mathbf{T}[T_\rho], x \in T \cap B(x_0, C\rho).$$

Apart from the condition (95), this can be achieved simply by applying the machinery of Section 5 to the previous propositions, as follows.

From (69) and Lemma 5.5 we have

$$\tilde{\forall}_{x_0, T_\rho} P_2(x_0), T_\rho \in \mathbf{T}_\rho(x_0) : T_\rho \in \bigcup_{\pi' \in \Pi_\rho(x_0)} \mathbf{T}_\rho(x_0, \pi').$$

Clearly this implies

$$\tilde{\forall}_{x_0, T_\rho} P_2(x_0), T_\rho \in \mathbf{T}_\rho(x_0) : [\tilde{\forall}_{T, x} T \in \mathbf{T}[T_\rho], x \in T : T_\rho \in \bigcup_{\pi' \in \Pi_\rho(x_0)} \mathbf{T}_\rho(x_0, \pi')].$$

By Lemma 5.10 we therefore have

$$\tilde{\forall}_{x_0, x, T_\rho, T} P_2(x_0), Q : T_\rho \in \bigcup_{\pi' \in \Pi_\rho(x_0)} \mathbf{T}_\rho(x_0, \pi'). \tag{97}$$

On the other hand, from (53) and another application of Lemma 5.6 we have

$$\tilde{\forall}_{x_0, x, T_\rho, T} Q : P_2(x_0).$$

Combining this with (97) using Lemma 5.4 we obtain

$$\tilde{\forall}_{x_0, x, T_\rho, T} Q : P_2(x_0), T_\rho \in \bigcup_{\pi' \in \Pi_\rho(x_0)} \mathbf{T}_\rho(x_0, \pi'). \tag{98}$$

Now from (83) and Lemma 5.5 we have

$$\tilde{\forall}_{x_0, \pi', T_\rho, T, x} \begin{array}{l} P_2(x_0), \pi' \in \Pi_\rho(x_0), T_\rho \in \mathbf{T}_\rho(x_0, \pi'), \\ \qquad T \in \mathbf{T}[T_\rho], x \in T \cap B(x_0, C\rho) : x \in S(x_0, \pi'), \end{array}$$

which we may re-arrange using Corollary 5.7 as

$$\tilde{\forall}_{x_0, x, T_\rho, T} Q, P_2(x_0), T_\rho \in \bigcup_{\pi' \in \Pi_\rho(x_0)} \mathbf{T}_\rho(x_0, \pi') : x \in S(x_0, \pi')$$

with the understanding that π' is the unique plane in $\Pi_\rho(x_0)$ such that $T_\rho \in \mathbf{T}_\rho(x_0, \pi')$. Combining this with (98) via Lemma 5.4, we obtain

$$\tilde{\forall}_{x_0, x, T_\rho, T} Q : R \tag{99}$$

where $R = R(x_0, x, T_\rho, T)$ denotes the property that

$P_2(x_0)$ holds, and for some $\pi' \in \Pi_\rho(x_0)$ we have $T_\rho \in \mathbf{T}_\rho(x_0, \pi')$ and $x \in S(x_0, \pi')$.

From (92) and Lemma 5.5 we have

$$\tilde{\forall}_{x, T} P_3(x), T \in \mathbf{T}(x) : T \in \bigcup_{\pi \in \Pi_\delta(x)} \mathbf{T}(x, \pi).$$

We re-arrange this using Corollary 5.7 as

$$\tilde{\forall}_{T_\rho, T, x} T_\rho \in \mathbf{T}_\rho, T \in \mathbf{T}[T_\rho], x \in T, P_3(x) : T \in \bigcup_{\pi \in \Pi_\delta(x)} \mathbf{T}(x, \pi).$$

By (94) and Lemma 5.4 we thus have

$$\tilde{\forall}_{T_\rho, T, x} T_\rho \in \mathbf{T}_\rho, T \in \mathbf{T}[T_\rho], x \in T : R'$$

where $R' = R'(x, T)$ denotes the property that

$$P_3(x), \text{ and } T \in \bigcup_{\pi \in \Pi_\delta(x)} \mathbf{T}(x, \pi).$$

By Lemma 5.10 we then obtain

$$\tilde{\forall}_{x_0, x, T_\rho, T} Q : R'. \tag{100}$$

The property $P_4(x_0, x, T_\rho, T)$ is equivalent to

$$R, R', (95)$$

where C_4 is a constant to be chosen later, and we adopt the convention that π is the unique plane in $\Pi_\delta(x)$ such that $T \in \mathbf{T}(x, \pi)$. Thus by (99), (100), and Corollary 5.3 it only remains to show that

$$\tilde{\forall}_{x_0, x, T_\rho, T} Q : ((R, R') \implies (95)),$$

or in other words that

(101)
$$|\{(x_0, x, T_\rho, T) : Q, R, R', \angle\pi, \pi' \gtrapprox \delta^{-C_4\sqrt{\epsilon}}\rho\}| \lessapprox \delta^{c\sqrt{\epsilon}}|\{(x_0, x, T_\rho, T) : Q\}|.$$

The left-hand side of (101) is majorized by

(102)
$$\begin{aligned}|\{(x_0, x, T_\rho, T, \pi, \pi') :&P_2(x_0), P_3(x), \pi' \in \Pi_\rho(x_0), \pi \in \Pi_\delta(x),\\ &T_\rho \in \mathbf{T}_\rho(x_0, \pi'), T \in \mathbf{T}[T_\rho] \cap \mathbf{T}(x, \pi), \angle\pi, \pi' \gtrapprox \delta^{-C_4\sqrt{\epsilon}}\rho\}|.\end{aligned}$$

Of course, T_ρ, T, π, π' are measured by counting measure and x_0, x by Lebesgue measure.

Fix x_0, x, π, π' such that

$$P_2(x_0), P_3(x), \pi' \in \Pi_\rho(x_0), \pi \in \Pi_\delta(x), \angle\pi, \pi' \gtrapprox \delta^{-C_4\sqrt{\epsilon}}\rho,$$

and suppose that T_ρ and T be such that $T_\rho \in \mathbf{T}_\rho(x_0, \pi')$ and $T \in \mathbf{T}[T_\rho] \cap \mathbf{T}(x, \pi)$. From (66) and (90) we see that T is contained in the $\delta^{-C\sqrt{\epsilon}}\rho$-neighbourhood of π', and the $\delta^{-C\sqrt{\epsilon}}\delta$-neighbourhood of π. From the angular separation assumption on π and π' and elementary geometry this implies that the direction of T is constrained to a cap of diameter $O(\delta^{-C\sqrt{\epsilon}}\delta^{C_4\sqrt{\epsilon}})$, whose center depends only on π and π'. From (93) we thus see that the contribution of (x_0, x, π, π') to (102) is majorized by

$$\delta^{-C\sqrt{\epsilon}}(\delta^{-C\sqrt{\epsilon}}\delta^{C_4\sqrt{\epsilon}})^c\delta^{d-3}.$$

We may therefore majorize (102) by

$$\int\int_{P_2(x_0), P_3(x), |x-x_0| \leq C\rho} \sum_{\pi' \in \Pi_\rho(x_0)} \sum_{\pi \in \Pi_\delta(x)} \delta^{-C\sqrt{\epsilon}}(\delta^{-C\sqrt{\epsilon}}\delta^{C_4\sqrt{\epsilon}})^c\delta^{d-3}\ dxdx_0.$$

From (68) and (91) this is majorized by

$$\delta^{-C\sqrt{\epsilon}}\delta^{cC_4\sqrt{\epsilon}}\delta^{d-3}\int_{P_2(x_0)}(\int_{E_\delta \cap B(x_0, C\rho)} dx)dx_0.$$

By (51), the inner integral is $\lessapprox \delta^{-C\sqrt{\epsilon}}\delta^{3-d}\rho^d$. From (21) we can therefore majorize this as

(103)
$$\delta^{-C\sqrt{\epsilon}}\delta^{cC_4\sqrt{\epsilon}}\delta^{3/2}.$$

We now turn our attention to the right-hand side of (101), which we can rewrite as

$$\delta^{c\sqrt{\epsilon}} \int \int_{|x-x_0| \leq C\rho} \sum_{T_\rho \in \mathbf{T}_\rho} \sum_{T \in \mathbf{T}[T_\rho]} \chi_{T_\rho}(x_0) \chi_T(x) \; dx dx_0$$

for some $c > 0$. Evaluating the x-integral using (27), we see that this is

$$\approx \delta^{c\sqrt{\epsilon}} \int \sum_{T_\rho \in \mathbf{T}_\rho} \chi_{T_\rho}(x_0) \delta^{3/2} \; dx_0,$$

which by (25) is

$$\approx \delta^{c\sqrt{\epsilon}} \delta^{3/2}.$$

Comparing this with (103) we obtain (101) as desired, if C_4 is chosen sufficiently large. ∎

Applying Lemma 5.10 to this proposition, we obtain the main result of this section, which is phrased in terms of a rather complicated property P_5.

Corollary 9.5. *If $P_5(x_0, T_\rho)$ denotes the property that*

$$\tilde{\forall}_{T,x} \; T \in \boldsymbol{T}[T_\rho], x \in T \cap B(x_0, C\rho) : P_4(x_0, x, T), \tag{104}$$

then one has

$$\tilde{\forall}_{T_\rho, x_0} \; T_\rho \in \boldsymbol{T}_\rho, x_0 \in T_\rho : P_5(x_0, T_\rho). \tag{105}$$

10. An arithmetic progression of length three

In the last few sections we have been steadily accumulating structural information on a collection $\mathbf{T}$ of tubes. This program has culminated in Proposition 9.5, in which we have shown that a rather complicated property P_5 holds almost always.

We have taken great pains to ensure that this property P_5 almost never failed. We will now see the payoff for this carefulness, in that we will be easily able to find three widely spaced points in arithmetic progression for which P_5 holds. Furthermore we shall be able to impose an additional technical condition, namely that the planes associated to these points are not too parallel. More precisely:

Proposition 10.1. *Let* $\boldsymbol{T}$ *be a sticky collection of tubes satisfying* (105). *Then there exists a tube* $T_\rho \in \boldsymbol{T}_\rho$ *and three points* x_0, x_0+s, x_0+2s *in* T_ρ *such that* $P_5(x_0+si, T_\rho)$ *holds for* $i = 0,1,2$. *Furthermore,* $|s| \sim 1$ *and* $\angle \pi(x_0+si), \pi(x_0+sj) \gtrapprox \delta^{C_5\sqrt{\epsilon}}$ *for distinct* $i,j \in \{0,1,2\}$, *where* $\pi(x)$ *denotes the plane for which* $T_\rho \in \boldsymbol{T}_\rho(x, \pi(x))$, *and* $C_5 > 0$ *is a suitable constant.*

Note that the plane $\pi(x)$ is well defined if $P_5(x, T_\rho)$ holds.

Proof From (105) and Lemma 5.9 we have

$$\tilde{\forall}_{T_\rho} T_\rho \in \mathbf{T}_\rho : [\tilde{\forall}_x x \in T_\rho : P_5(x, T_\rho)].$$

Thus by (31) we can find a tube $T_\rho \in \mathbf{T}_\rho$ such that

$$\tilde{\forall}_x x \in T_\rho : P_5(x, T_\rho). \tag{106}$$

Fix such a tube T_ρ, and let A denote the set

$$A = \{(x_0, s) \in \mathbf{R}^3 \times \mathbf{R}^3 : |s| \sim 1, \quad x_0 + si \in T_\rho \text{ for } i = 0,1,2\}.$$

Note that x_0 lives in T_ρ, and s lives in a translate of CT_ρ. We have to show that the set

$$\{(x_0, s) \in A : P_5(x_0+si, T_\rho), \angle \pi(x_0+si), \pi(x_0+sj) \gtrapprox \delta^{C_5\sqrt{\epsilon}} \text{ for all distinct } i,j = 0,1,2\}$$

is non-empty. On the other hand, from (106) we have

$$|\{(x_0, s) \in A : P_5(x_0 + si, T_\rho) \text{ fails}\}| \lessapprox \delta |\{x \in T_\rho : P_5(x, T_\rho) \text{ fails}\}| \lessapprox \delta^{c\sqrt{\epsilon}}\delta^2$$

for all $i = 0,1,2$. Since $|A| \sim \delta^2$, it will therefore suffice to show that

$$|\{(x_0, s) \in A : P_5(x_0+si, T_\rho), P_5(x_0+sj, T_\rho), \angle \pi(x_0+si), \pi(x_0+sj) \gtrapprox \delta^{C_5\sqrt{\epsilon}}\}| \lessapprox \delta^{c\sqrt{\epsilon}}\delta^2$$

for all distinct $i,j = 0,1,2$. Making the change of variables $x = x_0 + si$ and $y = x_0 + sj$, it therefore suffices to show that

$$|\{(x,y) \in F \times F : \angle(\pi(x), \pi(y)) \lessapprox \delta^{C_5\sqrt{\epsilon}}\}| \lessapprox \delta^{c\sqrt{\epsilon}}\delta^2 \tag{107}$$

for some $c > 0$, where F denotes the set

$$F = \{x \in T_\rho : P_5(x, T_\rho)\}.$$

The idea is as follows. By using the Kakeya estimate (14) one can easily recover the trivial estimate for (107) with $c = 0$. To improve upon this we use the x-ray estimate (15) instead, observing that if the squares corresponding to different x are almost parallel then there will be many parallel tubes in each direction.

We now turn to the details. Let Π be a maximal $\delta^{C_5\sqrt{\epsilon}}$-angular-separated set of planes containing the axis of T_ρ. If $x \in F$, then there exists a $\pi \in \Pi$ such that $\angle(\pi(x), \pi) \lessapprox \delta^{C_5\sqrt{\epsilon}}$. We can therefore majorize the left-hand side of (107) by

$$(107) \lessapprox \sum_{\pi \in \Pi} |\{(x,y) \in F \times F : \angle(\pi(x), \pi), \angle(\pi(y), \pi) \lessapprox \delta^{C_5\sqrt{\epsilon}}\}|$$

which we rewrite as

$$(107) \lessapprox \sum_{\pi \in \Pi} |\{x \in F : \angle(\pi(x), \pi) \lessapprox \delta^{C_5\sqrt{\epsilon}}\}|^2 \tag{108}$$

On the other hand, we clearly have

$$\sum_{\pi \in \Pi} |\{x \in F : \angle(\pi(x), \pi) \lessapprox \delta^{C_5\sqrt{\epsilon}}\}| \lessapprox |F| \lessapprox \delta.$$

It thus suffices to show

$$|\{x \in F : \angle(\pi(x), \pi) \lessapprox \delta^{C_5\sqrt{\epsilon}}\}| \lessapprox \delta^{cC_5\sqrt{\epsilon}}\delta^{-C\sqrt{\epsilon}}\delta \tag{109}$$

uniformly in π, since (108) clearly follows from the above two statements, for a sufficiently large choice of C_5.

We now prove (109). Fix π, and let F' denote the set

$$F' = \{x \in F : \angle(\pi(x), \pi) \lessapprox \delta^{C_5\sqrt{\epsilon}}\}.$$

Suppose that x was an element of F'. From (104) and the definition of P_4 we have

$$\tilde{\forall}_{y,T}\, T \in \mathbf{T}[T_\rho], y \in B(x, C\rho) : y \in S(x, \pi(x)).$$

By (31) and (27) we therefore have

$$\int_{S(x,\pi(x))} \sum_{T \in \mathbf{T}[T_\rho]} \chi_T(y)\ dy \approx \delta^{3/2}. \tag{110}$$

Let I denote the one-dimensional interval in $\mathbf{R}^3$ centered at the origin, parallel to π, orthogonal to the axis of T_ρ, and having length $\delta^{-C_5\sqrt{\epsilon}}\delta$, and let V be a maximal δ-separated subset of I. If Q is a square in $S(x, \pi(x))$ and T is a tube in $\mathbf{T}[T_\rho]$, then (66), the angular separation between $\pi(x)$ and π, and some elementary geometry shows that

$$|T \cap Q| \lesssim \delta^{C_5\sqrt{\epsilon}} \sum_{v \in V} |(T + v) \cap CQ|,$$

where $T + v$ is the translate of T by v, and CQ is a dilate of Q. Therefore

$$\int_{S(x,\pi(x))} \chi_T(y)\ dy \lesssim \delta^{C_5\sqrt{\epsilon}} \sum_{v \in V} \int_{CS(x,\pi(x))} \chi_{T+v}(y)\ dy$$

where $CS(x, \pi(x))$ was defined in Definition 8.3. Combining this with (110) we obtain

$$\int_{CS(x,\pi(x))} \sum_{v \in V} \sum_{T \in \mathbf{T}[T_\rho]} \chi_{T+v}(y)\ dy \gtrapprox \delta^{-C_5\sqrt{\epsilon}}\delta^{3/2}.$$

From (82), (9) and some algebra we therefore obtain

$$\int_{B(x,C\rho)} (\sum_{v \in V} \sum_{T \in \mathbf{T}[T_\rho]} \chi_{T+v}(y))^{d'}\ dy \gtrapprox \delta^{C\sqrt{\epsilon}}\delta^{-d'C_5\sqrt{\epsilon}}\delta^{3-d'}.$$

Integrating this over F' we obtain

$$\int |F' \cap B(y, C\rho)|(\sum_{v \in V} \sum_{T \in \mathbf{T}[T_\rho]} \chi_{T+v}(y))^{d'}\ dy \gtrapprox \delta^{C\sqrt{\epsilon}} \delta^{-d' C_5 \sqrt{\epsilon}} \delta^{3-d'} |F'|.$$

Since $|F' \cap B(y, C\rho)| \lessapprox \delta^{3/2}$, we thus obtain

$$\| \sum_{v \in V} \sum_{T \in \mathbf{T}[T_\rho]} \chi_{T+v}(y)\|_{d'}^{d'} \gtrapprox \delta^{C\sqrt{\epsilon}} \delta^{-d' C_5 \sqrt{\epsilon}} \delta^{\frac{3}{2}-d'} |F'|. \tag{111}$$

On the other hand, the directional multiplicity of the collection $\{T + v : T \in \mathbf{T}, v \in V\}$ is at most $\approx \# V \approx \delta^{-C_5\sqrt{\epsilon}}$. Thus by (15), we have

$$\| \sum_{v \in V} \sum_{T \in \mathbf{T}[T_\rho]} \chi_{T+v}(y)\|_{d'} \lessapprox \delta^{\frac{5}{2d'}-1} (\delta^{-C_5\sqrt{\epsilon}})^{1-\beta}$$

for some $\beta > 0$. Combining the two estimates gives (109) with $c = \beta d'$, as desired. ∎

The above proposition yielded a single arithmetic progression x_0^0, x_1^0, x_2^0 in a single ρ-tube T_ρ of length three which had the property P_5. By definition of P_5, there are many points in $B(x_i^0, C\rho)$, $i = 0, 1, 2$, with the property P_4. We therefore expect a large number (about δ^{-1}) of tubes T in $\mathbf{T}[T_\rho]$ to satisfy property P_4 when they pass near x_i^0.

The property P_4 is of course very complicated. However, we shall need only a few of the components of P_4 for the purposes of our argument, namely the graininess property, the planiness property at scale δ, and the consistency condition (95). More precisely, the proposition we will need for later sections is as follows.

Proposition 10.2. *Suppose we have an x-ray estimate at dimension d, and there exists a Besicovitch set of Minkowski dimension at most $d + \varepsilon$. Then for any $0 < \delta \ll 1$ we can find a ρ-tube T_ρ, a direction-separated set $\mathbf{T}[T_\rho]$ of δ-tubes in T_ρ with cardinality $\gtrapprox \delta^{C\sqrt{\epsilon}}\delta^{-1}$, three points x_i^0 in arithmetic progression in T_ρ with spacing ~ 1, and three planes π_0^0, π_1^0, π_2^0 each containing the axis of T_ρ, such that*

$$\angle \pi_i^0, \pi_j^0 \gtrapprox \delta^{C\sqrt{\epsilon}} \tag{112}$$

for all $i \neq j$. Furthermore, for each $i = 0, 1, 2$ we can find sets S_i which are grainy with respect to (x_i^0, π_i^0), and to each $x_i \in S_i$ we can associate a plane $\pi_i(x_i)$ containing x_i, such that

$$\angle \pi_i(x_i), \pi_i^0 \lessapprox \delta^{-C\sqrt{\epsilon}} \rho \tag{113}$$

holds for all i and $x_i \in S_i$, and

(114)
$$\int_X \#\{T \in \boldsymbol{T}[T_\rho] \cap \boldsymbol{T}(x_0) \cap \boldsymbol{T}(x_2) : T \in N(\pi_0(x_0)) \cap N(\pi_2(x_2))\}\ dx_0 dx_2 \gtrapprox \delta^{C\sqrt{\epsilon}}\delta^4,$$

where

$$N(\pi) = N_{\delta^{-C\sqrt{\epsilon}}\delta}(\pi)$$

and

$$X = \{(x_0, x_2) \in S_0 \times S_2 : \frac{1}{2}(x_0 + x_2) \in S_1\}.$$

Note that this Proposition differs from the previous ones, in that the $\tilde{\forall}$ symbol is completely absent. That stringent notation was necessary only in order to obtain the arithmetic progression in Proposition 10.1. Henceforth we shall be content with properties which are satisfied[1] only about $\delta^{C\sqrt{\epsilon}}$ of the time.

This Proposition is not so much a statement on the collection $\mathbf{T}$, but rather on the collection $\mathbf{T}[T_\rho]$, which can be regarded as a rescaled Besicovitch set. We are restricting ourselves to this portion of $\mathbf{T}$ in order to take full advantage of graininess, and will no longer concern ourselves with the remainder of $\mathbf{T}$.

Proof Fix T_ρ, x_0, s so that the conclusions of Proposition 10.1 hold. Set x_i^0 equal to $x_0 + si$ and S_i equal to the set $CS(x_i^0, \pi(x_i^0))$. Set π_i^0 to be any plane within an angle of $\delta^{-C\sqrt{\epsilon}}\rho$ to $\pi(x_i^0)$ which contains the axis of T_ρ; such a plane exists from (66). Note that since S_i is grainy with respect to $(x_i^0, \pi(x_i^0))$, it is also grainy with respect to (x_i^0, π_i^0) with slightly worse constants.

Let $i = 0, 1, 2$. From the definition of $P_5(x_i^0, T_\rho)$ we have

$$\tilde{\forall}_{T, x_i}\ T \in \mathbf{T}[T_\rho], x_i \in T \cap B(x_i^0, C\rho) : P_4(x_i^0, x_i, T, T_\rho).$$

By Lemma 5.6 we therefore have

$$\tilde{\forall}_{T, x_0, x_1, x_2}\ T \in \mathbf{T}[T_\rho], x_j \in T \cap B(x_j^0, C\rho) \text{ for } j = 0, 1, 2 : P_4(x_i^0, x_i, T, T_\rho).$$

From Lemma 5.2 we thus have

$$\tilde{\forall}_{T, x_0, x_1, x_2}\ T \in \mathbf{T}[T_\rho], x_j \in T \cap B(x_j^0, C\rho) \text{ for } j = 0, 1, 2 : P_4(x_i^0, x_i, T, T_\rho) \text{ holds for all } i.$$

Applying (31), we obtain

$$\sum_{T \in \mathbf{T}[T_\rho]} \prod_{i=1}^{3} |\{x_i \in T \cap B(x_i^0, C\rho) : P_4(x_i^0, x_i, T, T_\rho)\}| \approx \sum_{T \in \mathbf{T}[T_\rho]} \prod_{i=1}^{3} |T \cap B(x_i^0, C\rho)|.$$

[1] See however Proposition 14.5 in the Appendix.

Computing the right-hand side using (28) and (19) we thus obtain

$$\sum_{T\in \mathbf{T}[T_\rho]} \prod_{i=1}^{3} |\{x_i \in T \cap B(x_i^0, C\rho) : P_4(x_i^0, x_i, T, T_\rho)\}| \gtrapprox \delta^{13/2}. \tag{115}$$

For $i = 0, 2$, define the sets

$$F_i = \{x_i \in S_i : P_3(x_i)\}.$$

For any x_0, x_2, T, the contribution to the above sum is $O(\delta^{5/2})$. Furthermore, from elementary geometry and the definition of P_4 and S_i, we see that this contribution vanishes unless $x_0 \in T \cap F_0$, $x_2 \in T \cap F_2$, and $(x_0, x_2) \in X$. Thus we have

$$\sum_{T\in \mathbf{T}[T_\rho]} |\{(x_0, x_2) \in X \cap ((T \cap F_0) \times (T \cap F_2)) : P_4(x_i^0, x_i, T, T_\rho) \text{ holds for } i = 0, 2\}| \gtrapprox \delta^4.$$

We may rewrite the above as

$$\int_{X\cap(F_0\times F_2)} \#\{T \in \mathbf{T}[T_\rho] : P_4(x_0^0, x_0, T, T_\rho), P_4(x_2^0, x_2, T, T_\rho)\}\; dx_0 dx_2 \gtrapprox \delta^4. \tag{116}$$

From the definition of P_4, we have

$$\begin{aligned} &\#\{T \in \mathbf{T}[T_\rho] : P_4(x_i^0, x_i, T, T_\rho) \text{ holds for } i = 0, 2\} \\ &\lessapprox \sum_{\pi_0 \in \Pi_\delta^0(x_0)} \sum_{\pi_2 \in \Pi_\delta^2(x_2)} \#\{T \in \mathbf{T}[T_\rho] \cap \mathbf{T}(x_0) \cap \mathbf{T}(x_2) : T \in N(\pi_0) \cap N(\pi_2)\}, \end{aligned}$$

where

$$\Pi_\delta^i(x_i) = \{\pi \in \Pi_\delta(x_i) : \angle \pi, \pi_i^0 \lessapprox \delta^{-C\sqrt{\epsilon}}\rho\}.$$

It therefore follows that we may find a way of assigning to each $x_i \in F_i$, $i = 0, 2$ a plane $\pi_i(x_i) \in \Pi_\delta^i(x_i)$ such that

$$\begin{aligned} \int_{X\cap(F_0\times F_2)} &\#\{T \in \mathbf{T}[T_\rho] \cap \mathbf{T}(x_0) \cap \mathbf{T}(x_2) : \\ &\quad T \in N(\pi_0(x_0)) \cap N(\pi_2(x_2))\}\; dx_0 dx_2 \gtrapprox \delta^{C\sqrt{\epsilon}}\delta^4 \end{aligned} \tag{117}$$

for some large C. For, if (117) failed for all choices of assignments $\pi_0(x_0)$ and $\pi_2(x_2)$, then we could average over all such assignments using (91) and contradict (116) if C was large enough.

We thus choose planes $\pi_i(x_i) \in \Pi_\delta^i(x_i)$ so that (117) holds; from (95) we have (113) for all $x_i \in F_i$. We then extend $\pi_i()$ to all of S_i so that (113) holds for all $x_i \in S_i$. It is then an easy matter to check all the properties in the Proposition. ∎

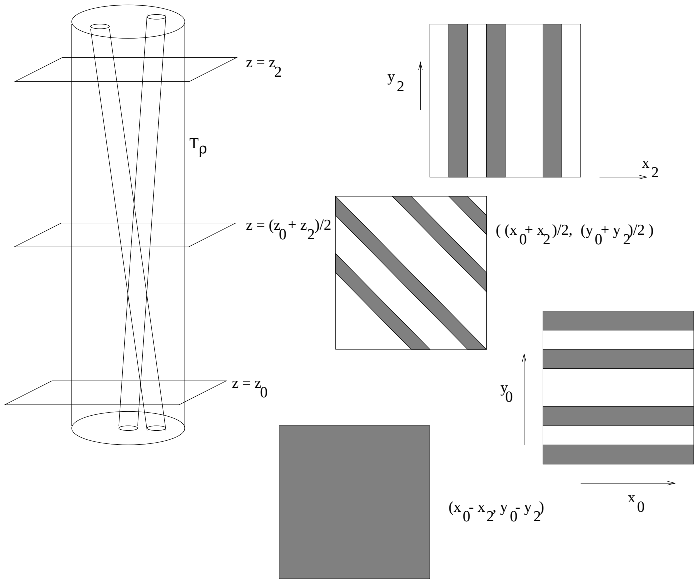

FIGURE 7. The set $E_\delta[T_\rho]$, and its (grainy) slices at the planes $\mathrm{z} = \mathrm{z}_0, \mathrm{z} = \mathrm{z}_2, \mathrm{z} = (\mathrm{z}_0+\mathrm{z}_2)/2$. The direction-separated assumption on $\mathbf{T}[T_\rho]$ means that the vectors $(\mathrm{x}_0 - \mathrm{x}_2, \mathrm{y}_0 - \mathrm{y}_2)$ must essentially fill out a $\rho \times \rho$ square.

11. Reduction to an additive setting

The purpose of this section is to convert the conclusions of Proposition 10.2 into pure additive combinatorics via some affine transformations and a discretization. This follows the philosophy of Bourgain [4], in which the properties of the Besicovitch set were converted into statements about sums and differences of sets; our main innovation is that we shall also encode the planiness and graininess properties into this additive combinatorial setting. After some technical preparations, the desired contradiction will then follow from the lemmas in [4].

For reasons that shall become clear, we shall parameterize the integer lattice $\mathbf{Z}^4$ by $(\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2)$.

Proposition 11.1. *Suppose there exists a Besicovitch set of Minkowski dimension at most $d + \varepsilon$ for some $0 \leq \varepsilon \ll 1$, and let $M \gg 1$ be a large number. Then there exists a subset $\Omega \subset \boldsymbol{Z}^4$ with cardinality*

$$\#\Omega \gtrsim M^{-C\sqrt{\varepsilon}} M^2 \tag{118}$$

such that

- *(Graininess) If we define the projections* $X_0, Y_0, X_2, Y_2 \in \boldsymbol{Z}$ *of* Ω *by*

$$X_0 = \{\mathrm{x}_0 : (\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2) \in \Omega \textit{ for some } \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2\}$$
$$Y_0 = \{\mathrm{y}_0 : (\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2) \in \Omega \textit{ for some } \mathrm{x}_0, \mathrm{x}_2, \mathrm{y}_2\}$$
$$X_2 = \{\mathrm{x}_2 : (\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2) \in \Omega \textit{ for some } \mathrm{x}_0, \mathrm{y}_0, \mathrm{y}_2\}$$
$$Y_2 = \{\mathrm{y}_2 : (\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2) \in \Omega \textit{ for some } \mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2\}$$

  then we have

$$\#X_0 \lesssim M^{C\sqrt{\epsilon}} M \tag{119}$$
$$\#Y_0 \lesssim M^{C\sqrt{\epsilon}} M^{d-2} \tag{120}$$
$$\#X_2 \lesssim M^{C\sqrt{\epsilon}} M^{d-2} \tag{121}$$
$$\#Y_2 \lesssim M^{C\sqrt{\epsilon}} M \tag{122}$$

  Furthermore, we have

$$\#\{(\mathrm{x}_2 + \mathrm{y}_0) + (\mathrm{x}_0 + \mathrm{y}_2) : (\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2) \in \Omega\} \lesssim M^{C\sqrt{\epsilon}} M^{d-2}. \tag{123}$$

- *(Planiness) There exist functions* $g : X_0 \times Y_0 \to Y_2$ *and* $h : X_2 \times Y_2 \to X_0$ *such that*

$$\mathrm{y}_2 = g(\mathrm{x}_0, \mathrm{y}_0) \tag{124}$$

  and

$$\mathrm{x}_0 = h(\mathrm{x}_2, \mathrm{y}_2) \tag{125}$$

  for all $(\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2) \in \Omega$.
- *(Direction-separated tubes). We have*

$$\#\{(\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2) \in \Omega : (\mathrm{x}_2 + \mathrm{y}_0) - (\mathrm{x}_0 + \mathrm{y}_2) = w\} \lesssim M^{C\sqrt{\epsilon}} M \tag{126}$$

  for all $w \in \boldsymbol{Z}$.

One should view (126) and (123) as a statement about the difference set and sum set of X_2+Y_0 and X_0+Y_2. The conflict between these two estimates will eventually cause a contradiction if ε is sufficiently small.

Proof Fix M, and define $\delta = M^{-2}$, so that $M = \delta^{-1/2}$. Note that it suffices to prove the above estimates with $\lesssim$ replaced by $\lessapprox$.

For technical reasons we shall prove the Proposition with the lattice $\mathbf{Z}$ replaced by $\delta\mathbf{Z}$, and $\mathbf{Z}^4$ by $(\delta\mathbf{Z})^4$. Since the two lattices are isomorphic, this will not make any difference.

By the results of the previous section, we can find T_ρ, $\mathbf{T}(T_\rho)$, x_i^0, π_i^0, S_i, and $\pi_i()$ which obey the conclusions of Proposition 10.2.

We now apply some affine transformations to place these objects into a standardized form. Firstly we may apply an affine transformation (preserving the z co-ordinate) so that T_ρ is the vertical tube

$$T_\rho = \{(\mathrm{x}, \mathrm{y}, \mathrm{z}) : |\mathrm{z}| \lesssim 1, \quad |\mathrm{x}|, |\mathrm{y}| \lesssim \rho\}.$$

By construction and (112), the planes π_i^0 all contain the z-axis, and make an angle of $\gtrapprox \delta^{C\sqrt{\epsilon}}$ with respect to each other.

By a rotation around the z axis we may make π_0^0 the xz-plane:

$$\pi_0^0 = \{(\mathrm{x}, \mathrm{y}, \mathrm{z}) : \mathrm{y} = 0\}. \tag{127}$$

By (112), we may apply a linear transformation $(\mathrm{x}, \mathrm{y}, \mathrm{z}) \mapsto (\mathrm{x} + k\mathrm{y}, \mathrm{y}, \mathrm{z})$ for some $|k| \lesssim \delta^{-C\sqrt{\epsilon}}$ and make π_2^0 the yz-plane without affecting π_0^0:

$$\pi_2^0 = \{(\mathrm{x}, \mathrm{y}, \mathrm{z}) : \mathrm{x} = 0\}. \tag{128}$$

This transformation may distort the tube T_ρ and the tubes T in $\mathbf{T}(T_\rho)$ slightly by a factor of $O(\delta^{-C\sqrt{\epsilon}})$, but this will not significantly affect the argument which follows.

Using (112) again, we see that by applying the linear transformation $(\mathrm{x}, \mathrm{y}, \mathrm{z}) \mapsto (k'\mathrm{x}, k''\mathrm{y}, \mathrm{z})$ for some $\delta^{C\sqrt{\epsilon}} \lesssim |k'|, |k''| \lesssim \delta^{-C\sqrt{\epsilon}}$ we may make π_1^0 equal to the $\mathrm{x} + \mathrm{y} = 0$ plane without affecting π_0^0 or π_2^0:

$$\pi_1^0 = \{(\mathrm{x}, \mathrm{y}, \mathrm{z}) : \mathrm{x} + \mathrm{y} = 0\}. \tag{129}$$

Again this may distort the tubes T_ρ and $\mathbf{T}(T_\rho)$ slightly, but this will not affect the argument.

For any $x_0, x_2 \in \mathbf{R}^3$ let $\mathbf{T}(x_0, x_2)$ denote the set

$$\mathbf{T}(x_0, x_2) = \{T \in \mathbf{T}[T_\rho] \cap \mathbf{T}(x_0) \cap \mathbf{T}(x_2) : T \subset N(\pi_0(x_0)) \cap N(\pi_2(x_2))\}.$$

From (114) we have

$$\int\int [\int_{X(\mathrm{z}_0,\mathrm{z}_2)} \#\mathbf{T}(x_0, x_2) d\mathrm{x}_0 d\mathrm{y}_0 d\mathrm{x}_2 d\mathrm{y}_2]\ d\mathrm{z}_0 d\mathrm{z}_2 \gtrapprox \delta^{C\sqrt{\epsilon}}\delta^4,$$

where $x_i = (\mathrm{x}_i, \mathrm{y}_i, \mathrm{z}_i)$ for $i = 0, 2$ and

$$X(\mathrm{z}_0, \mathrm{z}_2) = \{(\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2) : ((\mathrm{x}_0, \mathrm{y}_0, \mathrm{z}_0), (\mathrm{x}_2, \mathrm{y}_2, \mathrm{z}_2)) \in X\}.$$

The quantities z_0, z_2 both range inside intervals of length $O(\rho)$. Thus one can find $\mathrm{z}_0, \mathrm{z}_2 \in \mathbf{R}$ such that

$$\int_{X(\mathrm{z}_0,\mathrm{z}_2)} \#\mathbf{T}(x_0, x_2) d\mathrm{x}_0 d\mathrm{y}_0 d\mathrm{x}_2 d\mathrm{y}_2 \gtrapprox \delta^{C\sqrt{\epsilon}}\delta^3.$$

Fix z_0, z_2 so that the above estimate holds. By factoring the $(\mathrm{x}_0, \mathrm{y}_0)$ and $(\mathrm{x}_2, \mathrm{y}_2)$ planes into translates of the lattice $(\delta\mathbf{Z})^2$, we see that there must exist some translates L_0 and L_2 of the lattice $(\delta\mathbf{Z})^2$ such that

$$\sum_{(\mathrm{x}_0,\mathrm{y}_0,\mathrm{x}_2,\mathrm{y}_2)\in X(\mathrm{z}_0,\mathrm{z}_2)\cap(L_0\times L_2)} \#\mathbf{T}(x_0, x_2) \gtrapprox \delta^{-C\sqrt{\epsilon}}\delta^{-1}.$$

By a very mild affine transformation we may make $L_0 = L_2 = (\delta\mathbf{Z})^2$.

Define Ω to be the set of all quadruplets $(\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2)$ which give a non-zero contribution to the above sum. From the direction-separated nature of $\mathbf{T}[T_\rho]$, each such quadruplet contributes at most $\lessapprox \delta^{-C\sqrt{\epsilon}}$ to the sum, and so we see that Ω satisfies (118).

Since $(\mathrm{x}_0, \mathrm{y}_0, \mathrm{z}_0)$ lies in S_0, and x_0, y_0 live in $\delta\mathbf{Z}$, the bounds (119) and (120) follow from the graininess of S_0 and (127). Similarly the bounds (121) and (122) follow from the graininess of S_2 and (128).

Also, since $(\frac{\mathrm{x}_0+\mathrm{x}_2}{2}, \frac{\mathrm{y}_0+\mathrm{y}_2}{2}, \frac{\mathrm{z}_0+\mathrm{z}_2}{2})$ lies in S_0, the bound (123) follows from (129) and the graininess of S_1.

From the direction-separated nature of the tubes in $\mathbf{T}[T_\rho]$ we see that

$$\#\{(\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2) \in \Omega : (\mathrm{x}_2 - \mathrm{x}_0, \mathrm{y}_2 - \mathrm{y}_0) = (x, y)\} \lessapprox M^{C\sqrt{\epsilon}}$$

for all x, y. The estimate (126) then follows from this and the fact that $\mathrm{x}_2 - \mathrm{x}_0$ ranges over a δ-separated subset of a $O(\delta^{-C\sqrt{\epsilon}}\delta^{1/2})$-interval, which therefore has cardinality $O(M^{C\sqrt{\epsilon}}M)$.

Finally, we turn to (125) and (124). Fix $(\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2) \in \Omega$, and let T be a tube in $\mathbf{T}(x_0, x_2)$, where $x_i = (\mathrm{x}_i, \mathrm{y}_i, \mathrm{z}_i)$. Then by definition of $\mathbf{T}(x_0, x_2)$ we see that T is contained in a $\delta^{-C\sqrt{\epsilon}}\delta$-neighbourhood of the plane $\pi_0(\mathrm{x}_0, \mathrm{y}_0, \mathrm{z}_0)$. From (113), (127), and elementary geometry we therefore see that y_2 is contained in a set of cardinality $O(M^{C\sqrt{\epsilon}})$ which is determined only by $\pi_0(\mathrm{x}_0, \mathrm{y}_0, \mathrm{z}_0)$ (so it is independent of x_2 or y_2). By the pigeonhole principle we may therefore refine Ω so that (124) holds for some function g, without affecting (118) (other than a worsening of constants). Similarly we can refine Ω further to obtain (125) for some function h. ∎

12. Sums, differences, and functional relationships

In the previous arguments, the parameter d was not fixed to be any particular value. However, the arguments we shall now use work for $d = 5/2$ only, and we shall fix this value accordingly. Of course, this value of d is admissible by Theorem 2.3.

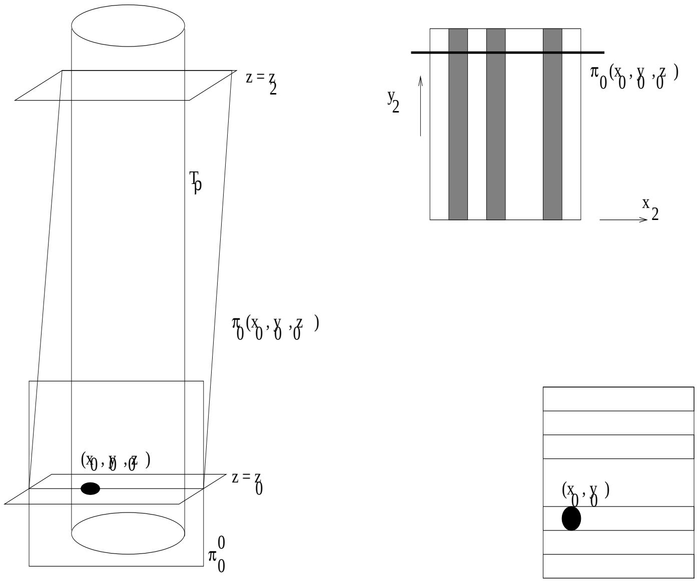

FIGURE 8. Since the plane $\pi_0(x_0, y_0, z_0)$ must be close to π_0^0, its intersection with $\{z = z_2\}$ is essentially a horizontal line, which explains (124).

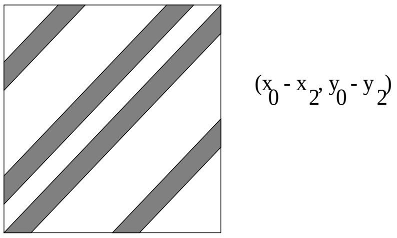

FIGURE 9. The shape of (130) which is dictated by the all the properties of Proposition 11.1, except for (126), with which it is incompatible.

Having obtained Proposition 11.1, we shall now use pure combinatorics and the lemmas in [4] to show that the conclusions of this Proposition lead to a contradiction if ε is sufficiently small.

The idea is to show that (126) is incompatible with the rest of the properties in Proposition 11.1. In fact, those properties have a strong tendency to make the set

$$\{(\mathrm{x}_2 - \mathrm{x}_0, \mathrm{y}_2 - \mathrm{y}_0) : (\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2) \in \Omega\} \tag{130}$$

resemble the set depicted in Figure 9; this is due to the heuristic implication of (5) from (4), combined with the functional relationships (125) and (124). Of course, such a set is highly incompatible with (126), and so we can attain the desired contradiction.

The property of planiness has evolved into the functional relationships (125) and (124), which state that x_0 is independent of y_0 and y_2 is independent of x_2. It turns out that we can strengthen these relationships further in the case $d = 5/2$, and conclude that y_2 is also independent of y_0, and x_0 is independent of x_2. In other words, there is a direct functional relationship between y_2 and x_0.

More precisely:

Proposition 12.1. *Let the notation be as in Proposition* 11.1, *and suppose that* $d = 5/2$. *Then there exists a subset* Ω_1 *of* Ω *with cardinality*

$$\#\Omega_1 \gtrsim M^{-C\sqrt{\epsilon}} M^2 \tag{131}$$

and functions $\tilde{g} : X_0 \to Y_2$ *and* $\tilde{h} : Y_2 \to X_0$ *such that*

$$\mathrm{y}_2 = \tilde{g}(\mathrm{x}_0) \tag{132}$$

and

$$\mathrm{x}_0 = \tilde{h}(\mathrm{y}_2) \tag{133}$$

for all $(\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2) \in \Omega_1$. *In particular, we have*

$$\#\{(\mathrm{x}_2 + \mathrm{y}_0) + f(\mathrm{y}_2) : (\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2) \in \Omega_1\} \lesssim M^{C\sqrt{\epsilon}} M^{1/2} \tag{134}$$

and

$$\#\{(\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2) \in \Omega : (\mathrm{x}_2 + \mathrm{y}_0) - f(\mathrm{y}_2) = w\} \lesssim M^{C\sqrt{\epsilon}} M \tag{135}$$

for all $w \in \boldsymbol{Z}$, *where* $f : Y_2 \to \boldsymbol{Z}$ *is the function* $f(\mathrm{y}_2) = \mathrm{y}_2 + \tilde{h}(\mathrm{y}_2)$.

Proof The argument will be a heavily disguised variant of the "hairbrush" constructions in [14].

For each $(\mathrm{x}_0, \mathrm{y}_0) \in X_0 \times Y_0$, define the multiplicity

$$\mu_0(\mathrm{x}_0, \mathrm{y}_0) = \#\{(\mathrm{x}_2, \mathrm{y}_2) : (\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2) \in \Omega\}.$$

From (121) and (124) we have

$$\|\mu_0\|_\infty \lessapprox M^{C\sqrt{\epsilon}} M^{1/2}.$$

From (118) we have

$$\|\mu_0\|_1 \gtrapprox M^{-C\sqrt{\epsilon}} M^2.$$

Finally, from (119) and (120) we have

$$\#\mathrm{supp}(\mu_0) \lessapprox M^{C\sqrt{\epsilon}} M^{3/2}.$$

Combining these facts by Lemma 1.2, we may therefore find a subset $S \subset X_0 \times Y_0$ with cardinality

$$\#S \gtrapprox M^{-C\sqrt{\epsilon}} M^{3/2} \tag{136}$$

such that

$$\mu_0(\mathrm{x}_0, \mathrm{y}_0) \gtrapprox M^{-C\sqrt{\epsilon}} M^{1/2} \text{ for all } (\mathrm{x}_0, \mathrm{y}_0) \in S. \tag{137}$$

In particular, if we define $\Omega' \subset \Omega$ to be the set

$$\{(\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2) \in \Omega : (\mathrm{x}_0, \mathrm{y}_0) \in S\}$$

then we have

$$\#\Omega' \gtrapprox M^{-C\sqrt{\epsilon}} M^2. \tag{138}$$

Now for each $(\mathrm{x}_2, \mathrm{y}_2) \in X_2 \times Y_2$, define the multiplicity

$$\mu_2(\mathrm{x}_2, \mathrm{y}_2) = \#\{(\mathrm{x}_0, \mathrm{y}_0) : (\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2) \in \Omega'\}.$$

From (138) we have

$$\sum_{\mathrm{x}_2} \sum_{\mathrm{y}_2} \mu_2(\mathrm{x}_2, \mathrm{y}_2) \gtrapprox M^{-C\sqrt{\epsilon}} M^2.$$

From (121), we can therefore find a number $\mathrm{x}_2^0 \in X_2$ such that

$$\sum_{\mathrm{y}_2} \mu_2(\mathrm{x}_2^0, \mathrm{y}_2) \gtrapprox M^{-C\sqrt{\epsilon}} M^{3/2}.$$

Fix this x_2^0. From (120) and (125) we have

$$\sup_{\mathrm{y}_2} \mu_2(\mathrm{x}_2^0, \mathrm{y}_2) \lessapprox M^{C\sqrt{\epsilon}} M^{1/2}.$$

Finally, from (122) we have

$$\#\mathrm{supp}(\mu_2(\mathrm{x}_2^0, \cdot)) \lessapprox M^{C\sqrt{\epsilon}} M.$$

Combining these facts by Lemma 1.2, we may therefore find a set $Y_2' \subset Y_2$ of cardinality

$$\#Y_2' \gtrapprox M^{-C\sqrt{\epsilon}} M \tag{139}$$

such that

$$\mu_2(\mathrm{x}_2^0, \mathrm{y}_2) \gtrsim M^{-C\sqrt{\epsilon}} M^{1/2} \text{ for all } \mathrm{y}_2 \in Y_2'. \tag{140}$$

For each y_2 in Y_2, define the set

$$A(\mathrm{y}_2) = \{(\mathrm{x}_0, \mathrm{y}_0) \in Y_0 : (\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2^0, \mathrm{y}_2) \in \Omega'\}.$$

By (124) these sets are disjoint as y_2 varies. By (140) these sets have cardinality

$$\# A(\mathrm{y}_2) \gtrsim M^{-C\sqrt{\epsilon}} M^{1/2}. \tag{141}$$

By (125) these sets are contained in the row $\{h(\mathrm{x}_2^0, y_2)\} \times Y_0$. Combining these three facts with (120) we thus see that

$$\#\{\mathrm{y}_2 \in Y_2' : h(\mathrm{x}_2^0, \mathrm{y}_2) = \mathrm{x}_0\} \lesssim M^{C\sqrt{\epsilon}} \text{ for all } x_0.$$

We may therefore find a subset Y_2'' of Y_2' with cardinality

$$\# Y_2'' \gtrsim M^{-C\sqrt{\epsilon}} M \tag{142}$$

such that the map $h(\mathrm{x}_2^0, \cdot)$ is injective on Y_2''.

We now define the set Ω_1 by

$$\Omega_1 = \{(\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2) : (\mathrm{x}_0, \mathrm{y}_0) \in \bigcup_{\mathrm{y}_2 \in Y_2''} A(\mathrm{y}_2).\}$$

By construction we have $A(\mathrm{y}_2) \subset S$ for all $\mathrm{y}_2 \in Y_2$. Thus from (137), (141), (142) and the disjointness of the $A(\mathrm{y}_2)$ we have

$$\#\Omega_1 \gtrsim M^{-C\sqrt{\epsilon}} M^{1/2} \# \bigcup_{\mathrm{y}_2 \in Y_2''} A(\mathrm{y}_2) \gtrsim M^{-C\sqrt{\epsilon}} M^{1/2} M^{-C\sqrt{\epsilon}} M M^{-C\sqrt{\epsilon}} M^{1/2}$$

which is (131).

We now show (132). Let $(\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2)$ be an element of Ω_1. Then by construction there exists a y_2' such that $(\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2^0, \mathrm{y}_2') \in \Omega'$. By (125) we have

$$\mathrm{x}_0 = h(\mathrm{x}_2, \mathrm{y}_2) = h(\mathrm{x}_2^0, \mathrm{y}_2').$$

But by (124) we have $\mathrm{y}_2 = \mathrm{y}_2'$. Thus (133) follows by setting $\tilde{h}(\mathrm{y}_2) = h(\mathrm{x}_2^0, \mathrm{y}_2)$. The claim (132) then follows since $\tilde{h}$ is injective by construction. Finally, the claims (134) and (135) follow from (123) and (126) respectively after applying (133). ∎

Having strengthened the functional relationship, the next step is to reduce the number of values that $\mathrm{x}_2 + \mathrm{y}_0$ and $f(\mathrm{y}_2)$ can take. More precisely:

Proposition 12.2. *Let the notation be as in the previous Propositions. Then one can find a subset Ω_2 of Ω_1 with cardinality*

$$\#\Omega_2 \gtrsim M^{-C\sqrt{\epsilon}} M^2 \tag{143}$$

such that, if A and B denote the sets

$$A = \{\mathrm{x}_2 + \mathrm{y}_0 : (\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2) \in \Omega_2\}$$

and

$$B = \{f(\mathrm{y}_2) : (\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2) \in \Omega_2\},$$

then

$$\#A, \#B \lesssim M^{C\sqrt{\epsilon}} M^{1/2}. \tag{144}$$

Proof Define the multiplicity function $\mu_1(a)$ on $\mathbf{Z}$ by

$$\mu_1(a) = \{(\mathrm{x}_2, \mathrm{y}_0) \in X_2 \times Y_0 : \mathrm{x}_2 + \mathrm{y}_0 = a\}.$$

We consider the problem of estimating the cardinality of

$$\{(\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2) \in \Omega_1 : \mu_1(\mathrm{x}_2 + \mathrm{y}_0) \leq M^{-C_7\sqrt{\epsilon}} M^{1/2}\} \tag{145}$$

where C_7 is a constant to be chosen later.

First consider the contribution of a single $\mathrm{y}_2 \in Y_2$ to (145). The quantity x_0 is fixed by (133). By (134) there are $\lesssim M^{C\sqrt{\epsilon}} M^{1/2}$ possible values of $\mathrm{x}_2 + \mathrm{y}_0$ which contribute. Thus the total contribution of y_2 to the cardinality of (145) is at most

$$M^{C\sqrt{\epsilon}} M^{1/2} M^{-C_7\sqrt{\epsilon}} M^{1/2}. \tag{146}$$

By (122) we therefore have

$$\#(145) \lesssim M^{C\sqrt{\epsilon}} M^{-C_7\sqrt{\epsilon}} M^2.$$

Similarly, define the multiplicity function $\mu_2(b)$ on $\mathbf{Z}$ by

$$\mu_2(b) = \{y_2 \in Y_2 : f(y_2) = b\},$$

and consider the cardinality of

$$\{(\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2) \in \Omega_1 : \mu_2(f(\mathrm{y}_2)) \leq M^{-C_7\sqrt{\epsilon}} M^{1/2}\}. \tag{147}$$

We consider the contribution of a single $(\mathrm{x}_2, \mathrm{y}_0) \in X_2 \times Y_0$ to (147). By (134) there are $\lesssim M^{C\sqrt{\epsilon}} M^{1/2}$ possible values of $f(\mathrm{y}_2)$ which contribute. Since x_0 is determined by (133), we see that the contribution of each $(\mathrm{x}_2, \mathrm{y}_0)$ to the cardinality of (147) is at most (146) as before. From (121) and (120) we thus have

$$\#(147) \lesssim M^{C\sqrt{\epsilon}} M^{-C_7\sqrt{\epsilon}} M^2.$$

Define Ω_2 to be the set Ω_1 with (145) and (147) removed. By (131) we see that (143) holds, if we choose C_7 large enough. By construction we have

$$\mu_1(a) \gtrsim M^{-C\sqrt{\epsilon}} M^{1/2} \text{ for all } a \in A$$

and

$$\mu_2(b) \gtrsim M^{-C\sqrt{\epsilon}} M^{1/2} \text{ for all } b \in B$$

On the other hand, from (121), (120), and (122) we have

$$\|\mu_1\|_1 = \#X_2 \# Y_0 \lesssim M^{C\sqrt{\epsilon}} M, \quad \|\mu_2\|_1 = \#Y_2 \lesssim M^{C\sqrt{\epsilon}} M,$$

and (144) follows from Chebyshev's inequality. ∎

Now we can finally apply Bourgain's lemma. Let G denote the subset of $A \times B$ given by

$$G = \{(\mathrm{x}_2 + \mathrm{y}_0, f(\mathrm{y}_2)) : (\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2) \in \Omega_2\}.$$

From (143) and (135) we have

$$\#\{a - b : (a,b) \in G\} \gtrsim M^{-C\sqrt{\epsilon}} M.$$

On the other hand, from (134) we have

$$\#\{a + b : (a,b) \in G\} \lesssim M^{C\sqrt{\epsilon}} M^{1/2}.$$

Finally, the cardinality of A and B is controlled by (144). If we now apply Lemma 2.54 of [4] (or Lemma 14.1 in the Appendix of this paper) with $N = M^{C\sqrt{\epsilon}} M^{1/2}$ we obtain a contradiction, if ε is sufficiently small. Thus there are no Besicovitch sets of Minkowski dimension $\leq 5/2 + \varepsilon$, if ε is a sufficiently small absolute constant. This (finally!) completes the proof of Theorem 0.4. ∎

13. Remarks

• A large part of the above argument works for any dimension d, providing that one has an x-ray estimate at dimension d to begin with. This raises the possibility that one can obtain a much better result than Theorem 0.4 by a bootstrap argument. At present, however, there are two obstructions to this idea. Firstly, the conclusion of Theorem 0.4 is a Minkowski dimension statement, which is considerably weaker than the x-ray estimate hypothesis. Secondly, the additive combinatorial portion of the argument is only effective at dimension close to 5/2.

• We now consider the question of what happens when we relax the direction-separated assumption on the collection of tubes $\mathbf{T}$. Our discussion shall be fairly informal.

Definition 13.1. A family of δ-tubes $\mathbf{T}$ is said to obey the *Wolff axioms* if

- For any $\delta \leq \sigma \leq 1$ and any σ-tube T_σ, there are at most $\lessapprox \sigma^2 \delta^{-2}$ tubes in $\mathbf{T}$ which are contained in T_σ.

- For any $\delta \leq \sigma \leq 1$, any σ-tube T_σ, and any plane π there are at most $\lessapprox \sigma\delta^{-1}$ tubes in $\mathbf{T}$ which are contained in $T_\sigma \cap N_{C\delta}(\pi)$.

Clearly, any direction-separated collection of tubes obeys the Wolff axioms. The Kakeya results in [14] extend easily to these types of collections; for instance, we have

$$|\bigcup_{T\in\mathbf{T}} T| \gtrapprox \delta^{1/2} \tag{148}$$

whenever $\mathbf{T}$ is a collection of tubes with cardinality $\approx \delta^{-2}$ that satisfies the Wolff axioms. In contrast, the arguments in Bourgain [4] do not say anything concerning this generalized setting in any dimension, because these arguments crucially rely on the direction-separated nature of $\mathbf{T}$. For similar reasons the arguments in our paper also do not adapt to this setting even if one can somehow attain the analogue of stickiness. The estimate (126) in particular is critically dependent on the direction-separation of $\mathbf{T}$.

One may then ask if there is any improvement to (148) if one only assumes that $\mathbf{T}$ satisfies the Wolff axioms, or a similar condition which is weaker than direction-separation. It turns out in fact that there is no such improvement in three dimensions if one replaces the underlying field $\mathbf{R}$ by $\mathbf{C}$, and multiplies all relevant exponents by a factor of 2. Indeed, if we take E to be the unit ball of the Heisenberg group:

$$\{(z_1, z_2, z_3) \in \mathbf{C}^3 : Im(z_3) = Im(z_1\overline{z_2}), |z_1|, |z_2|, |z_3| \lesssim 1\},$$

then we see that E has Minkowski dimension 5, and for any $a, b \in \mathbf{R}$ and $w \in \mathbf{C}$ with $|a|, |b|, |w| \lesssim 1$ the (complex) line segment

$$\{(z, w + az, z\overline{w} + b) : z \in \mathbf{C}, |z| \lesssim 1\}$$

is contained in E. We can then find a collection $\mathbf{T}$ of complex δ-tubes in $N_{C\delta}(E)$ with cardinality $\approx \delta^{-4}$ which satisfy the complex version of the Wolff axioms, and such that

$$|\bigcup_{T\in\mathbf{T}} T| \sim \delta.$$

Informally, the collection $\mathbf{T}$ is also sticky, plany, and grainy in the spirit of the above argument. By "stickiness" we mean that the tubes $\mathbf{T}$ can be covered by a collection $\mathbf{T}_\rho$ of ρ-tubes which are not direction-separated, but satisfy the Wolff axioms and have cardinality $\approx \rho^{-2}$. The planiness and graininess properties are best seen by first verifying them at the origin $(0, 0, 0)$ (in which case the relevant plane is the $z_3 = 0$ plane), and then using the transitive action of the Heisenberg group. One can find a set which obeys all the properties of Proposition 11.1 (with the trivial distinction that the lattice $\mathbf{Z}$ is replaced by the Gaussian integer lattice $\mathbf{Z}[i]$) except for the crucial (126), which is lacking because $\mathbf{T}$ is not direction-separated. Namely, we may take

$$\begin{aligned} X_0 = Y_2 &= \{a + bi \in \mathbf{Z}[i] : |a + bi| \lesssim M^{1/2}\} \\ X_2 = Y_0 &= \{a \in \mathbf{Z} : |a| \lesssim M^{1/2}\} \\ \Omega &= \{(\mathrm{x}_0, \mathrm{y}_0, \mathrm{x}_2, \mathrm{y}_2) \in X_0 \times Y_0 \times X_2 \times Y_2 : \mathrm{x}_0 = \overline{\mathrm{y}_2}\} \\ g(\mathrm{x}_0, \mathrm{y}_0) &= \overline{\mathrm{x}_0} \\ h(\mathrm{x}_2, \mathrm{y}_2) &= \overline{\mathrm{y}_2}. \end{aligned}$$

We remark that the set (130) in this example resembles Figure 9 rather than the set in Figure 7. In particular, (126) fails dramatically.

Suppose that we had a subfield F of $\mathbf{R}$ which had dimension 1/2. Then one could (heuristically at least) modify the Heisenberg group example given above to show that (148) was sharp in $\mathbf{R}^3$ when one does not assume that the collection $\mathbf{T}$ is direction-separated, and only obeys the Wolff axioms. The point is that if F has dimension 1/2, then $\mathbf{R}$ must be a quadratic extension of F, and so there is a non-trivial conjugation $x \mapsto \overline{x}$ on $\mathbf{R}$ which plays the same role as complex conjugation does in the construction of the Heisenberg group.

On the other hand, it is well-known that such a field F does not exist in $\mathbf{R}$ because of the ordering properties of $\mathbf{R}$. Indeed, there are no non-trivial conjugations on $\mathbf{R}$ because such conjugations have to map square numbers to square numbers, hence are positivity preserving and therefore monotone. The claim then follows from observing that any conjugation must be the identity on the rationals $\mathbf{Q}$. This raises the possibility that one could exploit the ordering properties of $\mathbf{R}$ to improve upon (148) even when one does not assume direction-separation properties on $\mathbf{T}$. If one could find such an argument, then it is likely that this would also yield progress on the distance set conjecture of Falconer [5] or the Furstenburg problem mentioned in [16]. Two of us will pursue this connection in further detail in [8].

- Our main argument relied on only three slices of the Besicovitch set, and discovered that sets such as X_2 and Y_0 had very good additive and subtractive properties. One may naturally ask whether one could gain more information by looking at more slices. For instance, one can try to utilize planiness at the slice $\mathrm{z} = (\mathrm{z}_0 + \mathrm{z}_2)/2$; this property was not exploited in the arguments above. This property in fact leads to an interesting observation, that a large portion of the set $E[T_\rho]$ is approximately invariant under a δ-separated set of translations with cardinality $\rho^{-1/2}$ in the $(1, 1, 0)$ direction, and can be used to provide an alternate proof of the contradiction obtained in the arguments above.

If one exploits planiness and graininess on other horizontal slices, one can also discover *multiplicative* properties on X_2 and Y_0. Very roughly speaking, the sets X_2 and Y_0 are affinely equivalent to a ρ-separated set of diameter $O(1)$ which is an approximate ring, in the sense that it is often closed under addition and multiplication (up to an uncertainty of ρ). Thus there is (a very weak and heuristic) converse to the connections between 1/2-dimensional fields and 5/2-dimensional Kakeya sets mentioned earlier.

We hope to detail these results in a future paper.

• Define a *Nikodym set* is a subset E of $\mathbf{R}^3$ such that for every $x \in \mathbf{R}^3$ there exists a line l containing x such that $l \cap E$ contains a unit line segment. In analogy with the Kakeya conjecture, it is conjectured that all Nikodym sets have full dimension. At first glance, the above results do not appear to have any bearing on Nikodym sets as the associated collection of δ-tubes is not direction-separated. However this is easily resolved, as one can apply a projective transformation to turn a Nikodym set into a Kakeya set of the same dimension, and so Theorem 0.4 is also valid for Nikodym sets. See [10].

• It seems likely that these techniques could be extended to higher dimensions and give an improvement on (1) for all $n \geq 3$. Indeed, the fact that the codimension of a set with dimension $(n+1)/2$ increases rapidly with n is a favourable sign, especially when compared with the analogue of the graininess property in higher dimensions. Of course, for n sufficiently large ($n > 26$) the estimate (2) is superior to (1) anyway. The first step in this program would be to generalize the results in [15] to higher dimensions.

14. Appendix: A lemma of Bourgain

In this section we prove a weakened version of Lemma 2.54 of [4], namely:

Lemma 14.1. [4] *Let $N \gg 1$, $0 < \varepsilon \ll 1$ and let A, B be subsets of $\mathbf{Z}$ such that*

$$\#A, \#B \leq N. \tag{149}$$

Then, if G is a subset of $A \times B$ such that

$$\#\{a + b : (a, b) \in G\} \leq N, \tag{150}$$

then we have

$$\#\{a - b : (a, b) \in G\} \lesssim N^{2-\varepsilon} \tag{151}$$

for some sufficiently small absolute constant $\varepsilon > 0$.

In [4] this lemma was proven with the explicit constant $\varepsilon = 1/13$. Our argument is copied from [4] but is slightly simplified as we are not trying to optimize the constants. The argument in [4] is based in turn from the work of Gowers [6]. It is not known what the best value of ε is, but one must have

$$\varepsilon \leq 2 - \frac{\log(7/6)}{\log(7)} = 0.92078\ldots;$$

see [9].

Proof Let ε be a small constant to be chosen later. Suppose for contradiction that we had sets A, B, G for which (151) failed. Then we clearly have

$$\#G \approx N^2; \tag{152}$$

in this section the $\lessapprox$ notation will be used with δ replaced by N^{-1}.

Proposition 14.2. *Let A, B, G satisfy* (152) *and the hypotheses of Lemma* 14.1. *Then there exist $\approx N$ numbers d such that*

$$\#\{(a,b) \in A \times B : a - b = d\} \approx N. \tag{153}$$

A number d which satisfies (153) shall be called an (A,B)*-popular difference.*

Proof The upper bound in (153) is clear from (149), as is the upper bound on the number of popular differences. Thus it suffices to show the lower bound in both cases.

Let X denote the set in (150). From (152) we have

$$\|\chi_A * \chi_B\|_{L^1(X)} \geq \#G \gtrapprox N^2.$$

From (150) and (9) with $d = 2$ we thus have

$$\|\chi_A * \chi_B\|_2 \gtrapprox N^{3/2}.$$

From the identity

$$\|\chi_A * \chi_B\|_2 = \|\chi_A * \chi_{-B}\|_2$$

we thus obtain

$$\int (\chi_A * \chi_{-B})^2 \gtrapprox N^3.$$

On the other hand, from (149) we have

$$\int \chi_A * \chi_{-B} \lessapprox N^2.$$

Thus we have

$$\int_{\chi_A * \chi_{-B} \gtrapprox N} (\chi_A * \chi_{-B})^2 \gtrapprox N^3$$

for appropriate choices of constants in the $\gtrapprox$ symbols. On the other hand, from (149) again we have

$$\|\chi_A * \chi_{-B}\|_\infty \lessapprox N.$$

Combining the two estimates we obtain

$$\#\{\chi_A * \chi_{-B} \gtrapprox N\} N^2 \gtrapprox N^3,$$

and the claim follows. ■

Definition 14.3. If $a, a' \in A$, we say that a and a' (A,B)-*communicate* if

$$\#\{b \in B : a - b, a' - b \text{ are both } (A,B)\text{-popular differences}\} \gtrapprox N^{-\sqrt{\epsilon}} N.$$

It is quite plausible that under the conclusions of Proposition 14.2, that many pairs in A communicate. We shall go even further, and show that we can find a refinement of A for which *almost all* pairs communicate.

Definition 14.4. A *refinement* of A is any subset A' of A such that $\#A' \approx \#A$.

Proposition 14.5. *Let A, B be sets satisfying* (149) *and the conclusion of Proposition* 14.2. *Then there exists a refinement A' of A such that*

$$\tilde{\forall}_{a'} a' \in A' : [\tilde{\forall}_{a} a \in A' : a \text{ and } a' \ (A,B)\text{-communicate.}]$$

Here the notation $\tilde{\forall}$ is as before, but with δ replaced by N^{-1}.

Proof Let $I \subset A \times B$ be the incidence matrix

$$I = \{(a,b) \in A \times B : a - b \text{ is a } (A,B)\text{-popular difference}\};$$

by hypothesis we have

$$\#I \gtrapprox N^2. \tag{154}$$

We define the rows I_a and columns I^b of I by

$$I_a = \{b \in B : (a,b) \in I\}$$

and

$$I^b = \{a \in A : (a,b) \in I\}.$$

From (154) we have

$$\sum_{b \in B} \#I^b \gtrapprox N^2.$$

By (149) and Lemma 1.2 one can then find a refinement B' of B such that

$$\#I^b \approx N \text{ for all } b \in B'. \tag{155}$$

Fix B', and let $Y \subset A \times A$ be the set of non-communicative pairs, i.e.

$$Y = \{(a,a') \in A \times A : \#(I_a \cap I_{a'}) \leq N^{-\sqrt{\epsilon}} N.\}$$

Unfortunately, Y may be quite large. However, for many $b \in B'$ the set $(I^b \times I^b) \cap Y$ is quite small. Indeed, we have

$$\sum_{b \in B'} \#((I^b \times I^b) \cap Y) = \sum_{(a,a') \in Y} \#(I_a \cap I_{a'}) \lesssim N^{-\sqrt{\epsilon}} N (\#A)^2 \lessapprox N^{-\sqrt{\epsilon}} N^3$$

by (149). On the other hand, from (155) we have

$$\sum_{b \in B'} \#(I^b \times I^b) \approx N^2 \# B' \approx N^3. \tag{156}$$

Thus we have

$$\tilde{\forall}_{b,a,a'} \; b \in B', a \in I^b, a' \in I^b : (a, a') \notin Y.$$

By Corollary 5.9 we thus have

$$\tilde{\forall}_b b \in B' : [\, \tilde{\forall}_{a,a'} a \in I^b, a' \in I^b : (a, a') \notin Y].$$

Thus we may find a $b \in B'$ such that

$$\tilde{\forall}_{a,a'} a \in I^b, a' \in I^b : (a, a') \notin Y.$$

Fix this b. By Corollary 5.9 again we have

$$\tilde{\forall}_{a'} a' \in I^b : [\tilde{\forall}_a a \in I^b : (a, a') \notin Y].$$

The claim then follows from setting A' to be those elements of I^b for which

$$\tilde{\forall}_a a \in I^b : (a, a') \notin Y$$

holds. ∎

A variant of the above proposition is

Proposition 14.6. *Let A, B be sets satisfying* (149) *and the conclusion of Proposition* 14.2. *Then there exists a refinement B' of B such that for all $b' \in B'$ there are $\approx N$ numbers $a \in A$ for which $a - b'$ is a (A, B)-popular difference.*

Proof As in Proposition 14.2, the upper bound on the number of a is trivial from (149).

For each $b' \in B$ let $\mu(b')$ denote the quantity

$$\mu(b') = \#\{a \in A : a - b' \text{ is a } (A, B)\text{-popular difference}\}.$$

Then we have by hypothesis that

$$||\mu||_1 \gtrapprox N^2.$$

On the other hand, from (149) we clearly have

$$\mathrm{supp}(\mu), ||\mu||_\infty \lessapprox N.$$

The claim then follows from Lemma 1.2. ∎

We now are ready to give the key proposition.

Proposition 14.7. *Let A, B, G satisfy (152) and the hypotheses of Lemma 14.1. Then if ε is sufficiently small, there exists refinements A' and B' of A and B respectively such that*

$$|G \cap (A' \times B')| \gtrapprox N^2 \tag{157}$$

and

$$|\{a' - b' : (a', b') \in A' \times B'\}| \lessapprox N^{C\sqrt{\varepsilon}} N. \tag{158}$$

Proof Define the multiplicity $\mu_G(a)$ on A by

$$\mu_G(a) = \#\{b \in B : (a, b) \in G\}.$$

Then from (152) we have

$$\|\mu_G\|_1 \gtrapprox N^2$$

Thus by (149) and Lemma 1.2 we can find a refinement A_1 of A such that

$$\mu_G(a) \approx N \text{ for all } a \in A_1 \tag{159}$$

Fix this A_1. Note that $\#(G \cap (A_1 \times B)) \gtrapprox N^2$. Thus by Proposition 14.2 there are $\gtrapprox N$ (A_1, B)-popular differences. Thus by Proposition 14.5 there exists a refinement A' of A_1 such that

$$\tilde{\underset{a}{\forall}} a \in A' : a', a \ (A_1, B)\text{-communicate} \tag{160}$$

for all $a' \in A'$.

Fix A'. From (159) we have

$$\#(G \cap (A' \times B)) \approx N \# A' \approx N^2.$$

Thus if we define the multiplicity $\mu^G(b)$ on B by

$$\mu^G(b) = \#\{a \in A' : (a, b) \in G\}$$

then

$$\|\mu^G\|_1 \gtrapprox N^2.$$

Thus by another application of (149) and Lemma 1.2 we can find a refinement B_1 of B such that

$$\mu^G(b) \approx N \text{ for all } b \in B_1. \tag{161}$$

Fix this B_1. By Proposition 14.6 there exists a refinement B' of B_1 such that for every $b' \in B'$ there are $\approx N$ $a \in A'$ such that $a - b'$ is a (A', B_1)-popular difference.

Fix B'. Combining this with (160) we conclude (if ε is sufficiently small)

Fact 14.8. *For every $a' \in A'$ and $b' \in B'$, there are $\approx N$ elements $a \in A'$ such that a', a (A, B)-communicate, and $a - b'$ is a (A, B)-popular difference.*

We can now finish the proof of Proposition 14.7. Since (157) follows from (161), it suffices to show (158).

Fix a', b'. From the above Fact and the definition of communicativity, one can find $\gtrapprox N^{-C\sqrt{\epsilon}}N^2$ pairs $(a,b) \in A \times B$ such that $a'-b$, $a-b$, and $a-b'$ are (A,B)-popular differences. From the identity

$$a' - b' = (a' - b) - (a - b) + (a - b')$$

and the definition of a popular difference, we thus obtain

$$\begin{aligned}\#\{(a_1,a_2,a_3,b_1,b_2,b_3) \in A^3 \times B^3 : & a' - b' = (a_1 - b_1) - (a_2 - b_2) + (a_3 - b_3)\} \\ & \gtrapprox N^{-C\sqrt{\epsilon}}N^5.\end{aligned}$$

Since there are only $\approx N^6$ sextuplets $(a_1,a_2,a_3,b_1,b_2,b_3)$, (158) follows. ∎

We can now prove Lemma 14.1. We may refine G so that the map $(a,b) \mapsto a - b$ is injective on G without affecting the hypotheses of the Lemma. But this injectivity is clearly incompatible with the conclusions of Proposition 14.7 if ε is sufficiently small. This is our desired contradiction. ∎

References

1. J. Bourgain, *Besicovitch-type maximal operators and applications to Fourier analysis*, Geom. and Funct. Anal. **22** (1991), 147–187.
2. J. Bourgain, *Some new estimates on oscillatory integrals*, Essays in Fourier Analysis in honor of E. M. Stein, Princeton University Press (1995), 83–112.
3. J. Bourgain, *Estimates for cone multipliers*, Operator Theory: Advances and Applications, **77** (1995), 41–60.
4. J. Bourgain, *On the dimension of Kakeya sets and related maximal inequalities*, preprint.
5. K. J. Falconer, *The Hausdorff dimension of distance sets*, Mathematika, **32** (1985), 206–212.
6. T. Gowers, *A new proof of Szemerédi's theorem for arithmetic progressions of length four*, to appear in GAFA.
7. L.H. Loomis, H. Whitney, *An inequality related to the isoperimetric inequality*, Bull. Amer. Math. Soc **55**, (1949). 961–962.
8. N. Katz, T. Tao, *A connection between the distance set problem and a problem of Furstenburg*, in preparation.
9. I. Ruzsa, *Sums of finite sets*, Number Theory: New York Seminar; Springer-Verlag (1996), D.V. Chudnovsky, G.V. Chudnovsky and M.B. Nathanson editors.
10. T. Tao, *The Bochner-Riesz conjecture implies the Restriction conjecture*, to appear, Duke Math J.
11. T. Tao, A. Vargas, L. Vega, *A bilinear approach to the restriction and Kakeya conjectures*, J. Amer. Math. Soc. **11** (1998), pp. 967–1000.
12. T. Tao, A. Vargas, *A bilinear approach to cone multipliers I. Restriction Estimates*, to appear.
13. T. Tao, A. Vargas, *A bilinear approach to cone multipliers II. Applications*, to appear.
14. T. H. Wolff, *An improved bound for Kakeya type maximal functions*, Revista Mat. Iberoamericana. **11** (1995). 651–674.

15. T. H. Wolff, *A mixed norm estimate for the x-ray transform*, to appear in Revista Mat. Iberoamericana.
16. T. H. Wolff, *Recent work connected with the Kakeya problem*, to appear, Anniversary proceedings, Princeton 1996.

Department of Mathematics, University of Illinois at Chicago, Chicago IL 60607-7045

E-mail address: nets@math.uic.edu

Department of Mathematics, Princeton University, Princeton NJ 08544

E-mail address: laba@math.princeton.edu

Department of Mathematics, UCLA, Los Angeles CA 90095-1555

E-mail address: tao@math.ucla.edu