EPOCH AI

FrontierMath Open Problems

Epoch AI

New bounds for arithmetic Kakeya

A Kakeya set in $\mathbb{R}^d$ is a bounded set which contains a unit line segment in every direction. While such sets can have measure zero, the Kakeya conjecture posits that they always have the maximal possible (Minkowski and Hausdorff) dimension d. This conjecture was proved for $d = 2$ by Davies [2] and in a recent breakthrough for $d = 3$ by Wang and Zahl [10], but remains open for $d \geq 4$.

The arithmetic Kakeya conjecture was implicitly formulated by Katz and Tao [7], as a statement strong enough to imply (a form of) the Kakeya conjecture, but which can be formulated as a purely additive combinatorics statement, rather than harmonic analysis. The meta-theory of the arithmetic Kakeya conjecture was developed by Green and Ruzsa [4], who gave a number of equivalent formulations.

One way to state the arithmetic Kakeya problem is the following: we say that AK(α) holds if for any $\epsilon > 0$ there exists a constant $C_\epsilon > 0$ and a finite set $X \subset \mathbb{Z}^2$ such that $r_1 + r_2 \neq 0$ for all $\mathbf{x} \in X$, and for all finite $G \subset \mathbb{Z}^2$,

$$|(1, -1) \cdot G| \leq C_\epsilon \max_{\mathbf{x} \in X} |\mathbf{x} \cdot G|^{\alpha+\epsilon},$$

where

$$\mathbf{x} \cdot G = \{x_1 g_1 + x_2 g_2 : (g_1, g_2) \in G\}.$$

Note that we are not interested in the size or nature of X itself – for the applications to the Kakeya conjecture itself, all that matters is the exponent α.

It is trivial that AK(2) holds (for example by taking X to be any two linearly independent vectors in $\mathbb{Z}^2$, as then $(1, -1)$ can be expressed as a linear combination of these two). The arithmetic Kakeya conjecture itself is the following.

Conjecture 1 (Arithmetic Kakeya conjecture)**.** AK(1) *holds.*

The first non-trivial bound was obtained by Bourgain [1], who proved AK($\frac{25}{13}$). Katz and Tao improved this, first in [6] to AK($\frac{7}{4}$), and then later [7] to AK(γ) where $\gamma = 1.6751308\cdots$ is the largest root of $x^3 - 4x + 2 = 0$.

The relation to the Kakeya problem is the following.

Theorem 1. *If* AK(α) *holds then, for all $d \geq 2$, if $E \subseteq \mathbb{R}^d$ is a Kakeya set (a bounded set containing a unit line segment in every direction) then the Hausdorff dimension of E is at least*

$$\alpha^{-1} d + 1 - \alpha^{-1}.$$

This was implicitly established by Bourgain [1] using the 'method of slices', although Bourgain only established this for the Minkowski dimension. The stronger statement that this same bound holds for the Hausdorff dimension is a consequence of combining Bourgain's argument with recent progress on Szemerédi's theorem by Leng, Sah, and Sawhney [9].

The arguments of Katz and Tao [6] are very concrete, and proceed by finding explicit small graphs with certain properties. The goal of this problem is to find other explicit small graphs which lead, via the same argument, to better bounds.

Most optimistically, one could hope for finding a construction that establishes AK(α) for some $\alpha < 1.67\cdots$. This would immediately, via the above, improve the known lower bounds for the Kakeya conjecture in all sufficiently high dimensions.

Any new constructions would be of interest, however, as the argument of Katz and Tao [6] that produces the current limit of $\alpha = 1.6751308\cdots$ requires an extremely large set of dilates X (that increases without bound as

α approaches this value). If an exponent close to this could be found with a much simpler set then this would be interesting.

Katz [5] has shown that AK(3/2) is the limit of this kind of elementary argument, and hence establishing the full arithmetic Kakeya conjecture would be out of reach of this strategy.

The problem proposal

We will first explicitly define the kind of structure that is required, and how to 'score' such a structure (i.e. what value of α a structure would prove AK(α) for). We will then give, in this language, two examples of Katz and Tao. In the following subsection we sketch briefly how the proof goes, using the argument of Katz and Tao [6]. In this subsection we have presented things to be clear for the human reader; an equivalent, but more opaque, version which is more suitable for Python verification is given in Appendix **??**.

We first define a collection of constructible configurations iteratively.

Definition 1. *Let $X \subset \mathbb{Z}^2$. An X-constructible graph is a (simple undirected) finite graph, in which each edge is labelled with some $\mathbf{v} \in X$, defined in the following iterative fashion.*

1. *A single vertex is constructible.*
2. *If H is constructible then, for any integer $k \geq 1$, if $X_1, \ldots, X_k$ are sets of vertices of H, then the following graph is constructible: take $k+1$ disjoint copies of H and then, for each $1 \leq i \leq k$ and $x \in X_i$, either identify the vertices corresponding to x in both H_{i-1} and H_i, or put an edge between them labelled with some $\mathbf{v} \in X$.*

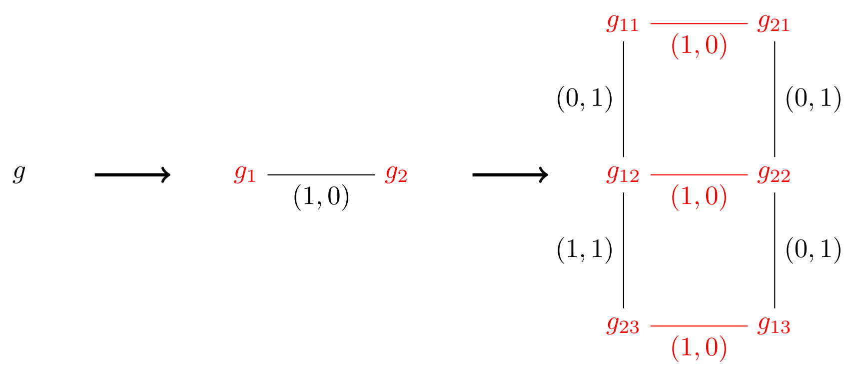

Figure 1: An example sequence of $\{(0,1),(1,0),(1,1)\}$-constructible graphs. The red colouring indicates the constructible components at each stage.

Given a finite set V, for any $g \in V$ and $\mathbf{x} \in \mathbb{Z}^2$ we write $\mathbf{x}g$ for the $2 \times |V|$ matrix which is zero everywhere except in the column corresponding to g, in which it is the column vector $\mathbf{x}$.

Definition 2. *Let G be a finite graph with vertex set V, such that each edge is labelled with some $\mathbf{x} \in \mathbb{Z}^2$. An X-*forcing pair *for G is a pair of sets (R, T) with $T \subseteq V$ and R a finite collection of $2 \times |V|$ integer matrices, with columns indexed by $g \in V$, such that*

$$R \subseteq \{\mathbf{x}g : \mathbf{x} \in X \text{ and } g \in V\},$$

and such that a finite sequence of the following operations on (R, T) eventually produces $T = V$:

1. *If $g_1 \sim g_2$ is an edge of G labelled $\mathbf{v}$ then add $\mathbf{v}g_1 - \mathbf{v}g_2$ to R.*
2. *If $(1,-1)g \in R$ then add g to T.*
3. *Any $\mathbb{Z}$-linear combination of matrices in R can be added to R.*

Informally speaking, the idea is that knowledge of $\mathbf{x} \cdot g$ for $\mathbf{x}g \in R$, together with knowledge of all $h \in T$, is enough, when combined with the identifications encoded in the edges of G and the injectivity of the map $g \mapsto (1,-1) \cdot g$, to know the values of all vertices of G.

Definition 3. *Let $X \subset \mathbb{Z}^2 \backslash \langle (1,-1) \rangle$ be a finite set. A constructible proof of*

$$\alpha(X) \leq \frac{m+r}{n-t}$$

is an X-constructible graph G with vertex set V, with n vertices and m edges, for which there is a forcing pair (R,T) with $|R| = r$ and $|T| = t$, where $T \subseteq V$ and

$$R \subseteq \{\mathbf{x}g : \mathbf{x} \in X \text{ and } g \in V\}.$$

In the following subsection we sketch how, as the terminology suggests, a constructible proof of $\alpha(X) \leq \alpha$ implies the arithmetic Kakeya statement $\mathrm{AK}(\alpha)$.

Goal 1 (New bounds for the arithmetic Kakeya conjecture)**.** *Find a finite $X \subset \mathbb{Z}^2 \backslash \langle (1,-1) \rangle$ together with a constructible proof of*

$$\alpha(X) < 1.67513 \cdots,$$

where the right-hand side is more precisely the largest root of $x^3 - 4x + 2 = 0$. More ambitiously, can this be done for some α close to $3/2$?

We reiterate that finding any X for which $\alpha(X) < 1.67513\cdots$ would already improve the current best bound known for the arithmetic Kakeya conjecture, and hence also improve the known lower bounds for the Kakeya conjecture in high dimensions.

A simple case of particular interest is when $X = \{(1,0),(0,1),(1,1)\}$ (although this has no direct applications, that we are aware of, to other problems such as Kakeya). This special case is sometimes referred as the Sums-Differences conjecture. The current best-known upper bound of

$$\alpha(\{(1,0),(0,1),(1,1)\}) \leq 11/6$$

is due to Katz and Tao [6].

Goal 2 (New bounds for the Sums-Differences conjecture)**.** *Find a constructible proof of*

$$\alpha(\{(1,0),(0,1),(1,1)\}) < 11/6.$$

The current best-known lower bound for such an α is $\alpha \geq 1.77898$, proved by Lemm [8] (with very slight further improvements found by AlphaEvolve [3]).

We will now give, in this language, two examples of constructible proofs, both adapted from Katz and Tao [6]. The first is a constructible proof of

$$\alpha(\{(1,0),(0,1),(1,1)\}) \leq 11/6.$$

The second achieves an improved upper bound of $\leq 7/4$, but at the cost of adding a new dilate $(1,2)$ to X.

The first Katz-Tao configuration

The configuration is depicted in Figure 2, which has $n = 6$ vertices and $m = 7$ edges. It has already been demonstrated in Figure 1 that this configuration is constructible. We claim that

$$R = \{(1,0)g_1, (1,1)g_3, (1,1)g_4, (0,1)g_6\}$$

and $T = \emptyset$ is a forcing pair, whence we achieve a score of

$$\frac{7+4}{6-0} = \frac{11}{6}$$

as claimed. To verify that this data fixes the configuration we first argue that, using the identifications built into the

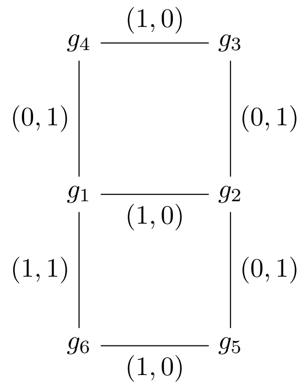

Figure 2: An achievable configuration with score 11/6.

configuration of Figure 2, the four elements of R fix the value of $(1,-1)g_5$. This follows from the following:

$$\begin{aligned}
(1,-1)g_5 &= (1,0)g_5 - (0,1)g_5 \\
&= (1,0)g_6 - (0,1)g_5 \\
&= (1,1)g_6 - (0,1)g_6 - (0,1)g_5 \\
&= (1,1)g_1 - (0,1)g_6 - (0,1)g_5 \\
&= (0,1)g_1 - (0,1)g_6 + (0,1)g_1 - (0,1)g_5 \\
&= (1,0)g_1 - (0,1)g_6 + (0,1)g_4 - (0,1)g_3 \\
&= (1,0)g_1 - (0,1)g_6 + ((0,1)g_4 + (1,0)g_3) - ((1,0)g_3 + (0,1)g_3) \\
&= (1,0)g_1 - (0,1)g_6 + (1,1)g_4 - (1,1)g_3.
\end{aligned}$$

By definition of a forcing pair we may now add g_5 to the 'known' values in T. We now observe that if we know $\mathbf{x}g$ and $\mathbf{y}g$ for any two linearly independent $\mathbf{x}$ and $\mathbf{y}$ then we may deduce the value of $(1,-1)g$ and hence add g itself to T. Proceeding in this fashion we can move all the remaining vertices of g_i into T: since we know g_5 we know $(0,1)g_3 = (0,1)g_2 = (0,1)g_5$, and $(1,1)g_3$ is known as an element of R, whence g_3 is known. Arguing similarly we can find the values of g_4, then g_1, then g_2, and finally g_6.

The second Katz-Tao configuration

We now give the second argument of Katz and Tao, that establishes

$$\{(1,0),(0,1),(1,2),(1,1)\} \to 7/4.$$

The required configuration is depicted in Figure 3. This can easily be checked as constructible as in the previous

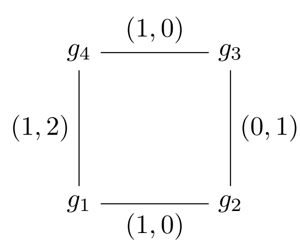

Figure 3: An achievable configuration with score 7/4.

case. THis has $n = 4$ vertices and $m = 4$ edges. We claim that

$$R = \{(1,1)g_1, (1,1)g_2, (0,1)g_4\}$$

and $T = \emptyset$ is a forcing pair, whence we achieve a score of

$$\frac{4+3}{4-0} = \frac{7}{4}$$

as claimed.

To verify that this data fixes the configuration we first argue that, similar to the above, that using the identifications built into the configuration of Figure 3, the three elements of R fix the value of $(1,-1)g_3$. This follows from:

$$\begin{aligned}
(1,-1)g_3 &= (1,0)g_3 - (0,1)g_3 \\
&= (1,0)g_4 - (0,1)g_3 \\
&= (1,2)g_4 - (0,2)g_4 - (0,1)g_3 \\
&= (1,2)g_1 - (0,2)g_4 - (0,1)g_2 \\
&= (2,2)g_1 - (0,2)g_4 - (1,0)g_1 - (0,1)g_2 \\
&= (2,2)g_1 - (0,2)g_4 - (1,0)g_2 - (0,1)g_2 \\
&= (2,2)g_1 - (1,1)g_2 - (0,2)g_4.
\end{aligned}$$

As before, this determines g_3, and by taking two known values of $\mathbf{x}g$ for linearly independent $\mathbf{x}$ we can then determine first g_4, then g_1, and finally g_2.

The link to arithmetic Kakeya

In this subsection we sketch how a 'constructible proof' as defined in the previous subsection actually implies a bound for the arithmetic Kakeya problem. We fix a finite set $G \subset \mathbb{Z}^2$ such that $|\mathbf{x} \cdot G| \leq N$ for all $\mathbf{x} \in X$. We seek an upper bound on $|(1,-1) \cdot G|$. Without loss of generality we can assume that $g \mapsto (1,-1) \cdot g$ is injective, whence it suffices to simply give an upper bound on $|G|$ itself.

An X-constructible graph H is to be interpreted as a subset $H \subseteq G$ such that if there is an edge between h_1 and h_2 labelled $\mathbf{x}$ then we have $\mathbf{x} \cdot h_1 = \mathbf{x} \cdot h_2$.

The goal is to find lower and upper bounds for the number of such H that appear as subsets of G, and then comparing these give the required upper bound on $|G|$. For the lower bound, we use the following lemma (which appears as [6, Lemma 2.1]).

Lemma 1. *If $X, Y_1, \ldots, Y_k$ are finite sets and $f_i : X \to Y_i$ for $1 \leq i \leq k$, with $|Y_i| \leq M_i$ for $1 \leq i \leq k$, then*

$$\left|(x_0, \ldots, x_k) \in X^{k+1} : f_i(x_{i-1}) = f_i(x_i) \text{ for } 1 \leq i \leq k\}\right| \geq \frac{|X|^{k+1}}{M_1 \cdots M_k}.$$

By the iterative nature of a constructible graph, one can use this lemma to give a lower bound of

$$\frac{|G|^n}{N^m} \leq \#\text{ copies of } H$$

for the number of copies of H that can be found in G, where n is the number of vertices of H and m is the number of edges.

On the other hand, if (R,T) is a forcing pair then fixing the values of all elements of R (which costs at most $N^{|R|}$) and fixing the values of all vertices in T (which costs $|G|^{|T|}$) determines everything about the copy of H, whence

$$\#\text{ copies of } H \leq N^{|R|} \, |G|^{|T|}.$$

Comparing these upper bound and lower bounds gives

$$|G| \leq N^{\frac{m+|R|}{n-|T|}},$$

and hence this proves $\mathrm{AK}(\alpha)$ with

$$\alpha = \frac{m+|R|}{n-|T|}$$

as claimed.

References

[1] Jean Bourgain. "On the dimension of Kakeya sets and related maximal inequalities". In: *Geom. Funct. Anal.* 9.2 (1999), pp. 256–282.

[2] Roy O. Davies. "Some remarks on the Kakeya problem". In: *Proc. Cambridge Philos. Soc.* 69 (1971), pp. 417–421.

[3] Bogdan Georgiev et al. *Mathematical exploration and discovery at scale*. 2025. arXiv: 2511.02864.

[4] Ben Green and Imre Z. Ruzsa. "On the arithmetic Kakeya conjecture of Katz and Tao". In: *Period. Math. Hungar.* 78.2 (2019), pp. 135–151.

[5] Nets Hawk Katz. "Elementary proofs and the sums differences problem". In: *Collectanea Mathematica* 57.Extra (2006), pp. 275–280.

[6] Nets Hawk Katz and Terence Tao. "Bounds on arithmetic projections, and applications to the Kakeya conjecture". In: *Math. Res. Lett.* 6.5–6 (1999), pp. 625–630.

[7] Nets Hawk Katz and Terence Tao. "New bounds for Kakeya problems". In: *J. Anal. Math.* 87 (2002), pp. 231–263.

[8] Marius Lemm. "New counterexamples for sums-differences". In: *Proc. Amer. Math. Soc.* 143.9 (2015), pp. 3863–3868.

[9] James Leng, Ashwin Sah, and Mehtaab Sawhney. *Improved bounds for Szemerédi's theorem*. 2024. arXiv: 2402.17995.

[10] Hong Wang and Joshua Zahl. *Volume estimates for unions of convex sets, and the Kakeya set conjecture in three dimensions*. 2025. arXiv: 2502.17655.