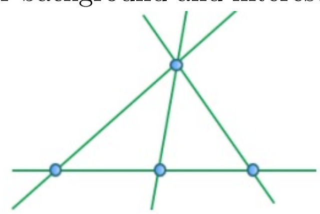
Project Example: Structural Szemerédi-Trotter Results

Adam Sheffer

This is an example 2024 REU project of Adam, which is only one option out of many. Adam finds a project that is a good fit for each participant, depending on their background and interests.

Given a set $\mathcal{P}$ of points and a set $\mathcal{L}$ of lines, both in $\mathbb{R}^2$, an *incidence* is a pair $(p,\ell) \in \mathcal{P} \times \mathcal{L}$ such that the point p is on the line ℓ. We denote by $I(\mathcal{P},\mathcal{L})$ the number of incidences in $\mathcal{P} \times \mathcal{L}$. For example, the figure to the right depicts a configuration with four points, four lines, and nine incidences.



For any n, the famous mathematician Erdős constructed a set $\mathcal{P}$ of n points and a set $\mathcal{L}$ of n lines with $c \cdot n^{4/3}$ incidences. Here, c is a constant close to 1 that we do not care about. In 1983, Szemerédi and Trotter [8] proved that this number of incidences is maximal, up to the constant c.

Theorem 1 (The Szemerédi-Trotter theorem)**.** *Let $\mathcal{P}$ be a set of n points and let $\mathcal{L}$ be a set of n lines, both in $\mathbb{R}^2$. Then $I(\mathcal{P},\mathcal{L}) \le c' \cdot n^{4/3}$ for a constant c'.*

To recap, there are configurations of n points, n lines, and about $n^{4/3}$ incidences. The is no such configuration with a larger number of incidences. This simple result about points and lines turned out to be surprisingly useful. It is used to obtain results in combinatorics, number theory, harmonic analysis, theoretical computer science, and more (for examples, see [5, 6]). Some of these results are considered as major breakthroughs in their field.

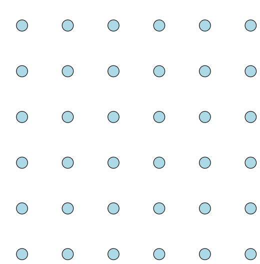
While the Szemerédi-Trotter theorem is a central result that has been known for over 40 years, not much is known about point–line configurations with about $n^{4/3}$ incidences. That is, *what is the structure of sets of points and lines that have many incidences?* One can ask many questions about such configurations. For example, is the point set always a lattice? Must there always be many parallel lines? Must there be a line that contains $\sqrt{n}$ points? Finding such structure may affect the many problems that rely on the Szemerédi-Trotter theorem. It is known as the *structural Szemerédi-Trotter problem.*

The problem has two main aspects:

- Proving that every point–line configuration with about $n^{4/3}$ incidences must have property X. See below for examples of potential properties.
- Finding point–line configurations with about $n^{4/3}$ incidences. These help us to make new conjectures and to disprove existing ones.

For past REU projects of Adam on this topic, see [1, 7].

Configurations. In Erdős's construction, the point set is the $\sqrt{n} \times \sqrt{n}$ lattice $\{1,2,3,\ldots,\sqrt{n}\} \times \{1,2,3,\ldots,\sqrt{n}\}$. For example, the figure on the right contains the lattice $\{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\}$. Erdős used Euler's Totient function to count lines that contain many points. We do no repeat his analysis here. For more information, see [4].

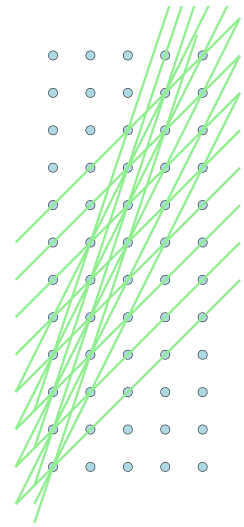


Decades later, Elekes [2] discovered another configuration that is easier to explain. We now describe the details of Elekes's construction. However, you might prefer to only skim it during a first read. One important point is that, in this configuration, the point set is an $n^{1/3} \times n^{2/3}$ lattice.

Let $r = (n/4)^{1/3}$ and $s = (2n)^{1/3}$ (for simplicity, assume that these are integers). The point set and line set are

$$\mathcal{P} = \{\ (i,j)\ :\ 1 \le i \le r \quad \text{and} \quad 1 \le j \le 2rs\ \},$$

$$\mathcal{L} = \{\ y = ax + b\ :\ 1 \le a \le s \quad \text{and} \quad 1 \le b \le rs\ \}.$$



For an example, see the figure on the right. We have that

$$|\mathcal{P}| = 2r^2 s = 2 \cdot \frac{n^{4/3}}{(4n)^{2/3}} \cdot \frac{(2n^2)^{1/3}}{n^{1/3}} = n,$$

$$|\mathcal{L}| = rs^2 = \frac{n^{2/3}}{(4n)^{1/3}} \cdot \frac{(2n^2)^{2/3}}{n^{2/3}} = n.$$

Consider a line $\ell \in \mathcal{L}$ that is defined by the equation $y = ax + b$. For each $x \in \{1, \ldots, r\}$, there exists a unique $y \in \{1, \ldots, 2rs\}$ such that the point (x, y) is incident to ℓ. That is, every line of $\mathcal{L}$ is incident to exactly r points of $\mathcal{P}$, so

$$I(\mathcal{P}, \mathcal{L}) = r \cdot |\mathcal{L}| = \frac{n^{2/3}}{(4n)^{1/3}} \cdot n = 2^{-2/3} n^{4/3}.$$

In the REU project of Olivine Silier with Adam [7], they discovered an infinite family of configurations. In particular, the point set can be a lattice of size $n^{\alpha} \times n^{1-\alpha}$ for any $1/3 \le \alpha \le 1/2$. The configurations of Erdős and Elekes are part of this infinite family. This discovery seemed to strengthen the possibility that the point set is a lattice in all optimal configurations. However, this possibility was disproved by a follow-up project of Olivine Silier with Larry Guth [3]. We may think of a lattice as a Cartesian product $A \times B$ with A and B being arithmetic progressions. Guth and Silier discovered a configuration where the point set is a Cartesian product $A \times A$ with A being a generalized arithmetic progression.[1]

The preceding paragraph illustrates how tricky this problem is. It is difficult to make any reasonable conjectures. A few example properties that are common to all known configurations:

- The point set is a Cartesian product. Is that always the case?
- There are about $n^{1/3}$ slopes, each with about $n^{2/3}$ lines. Is this always the case?
- The set of slopes is a generalized geometric progression. Is that always the case?

One can come up with many more potential properties. It is difficult to know which exist for all configurations with many incidences and which are red herrings.

Proofs. Silier's REU project [7] also proved the following structural result. It finds structure under the assumption that the point set is a Cartesian product. While Cartesian products are a central case, there might exist undiscovered configurations where the point set is not a Cartesian product. The statement ofthe following theorem is rather technical. You might prefer to skip it and read the intuition that follows.

[1] For a definition of this concept, see `https://en.wikipedia.org/wiki/Generalized_arithmetic_progression`.

Theorem 2.
(a) For $1/3 < \alpha < 1/2$*, let* $A, B \subset \mathbb{R}$ *satisfy* $|A| = n^{\alpha}$ *and* $|B| = n^{1-\alpha}$*. Let* $\mathcal{L}$ *be a set of* n *lines in* $\mathbb{R}^2$*, such that* $I(A \times B, \mathcal{L}) = \Theta(n^{4/3})$*. Then at least one of the following holds:*

- *There exists* $1 - 2\alpha \leq \beta \leq 2/3$ *such that* $\mathcal{L}$ *contains* $\Omega(n^{1-\beta}/\log n)$ *families of* $\Theta(n^{\beta})$ *parallel lines, each with a different slope.*
- *There exists* $1 - \alpha \leq \gamma \leq 2/3$ *such that* $\mathcal{L}$ *contains* $\Omega(n^{1-\gamma}/\log n)$ *disjoint families of* $\Theta(n^{\gamma})$ *concurrent lines, each with a different center.*

(b) Assume that we are in the case of $\Omega(n^{1-\beta}/\log n)$ *families of* $\Theta(n^{\beta})$ *parallel lines. There exists* $n^{2\beta} \leq t \leq n^{3\beta}$ *such that, for* $\Omega(n^{1-\beta}/\log^2 n)$ *of these families, the additive energy of the* y*-intercepts is* $\Theta(t)$*. Let* S *be the set of slopes of these families. Then* $E^{\times}(S) \cdot t = \Omega(n^{3-\alpha}/\log^{12} n)$*.*

Intuitively, part (a) of Theorem 2 states that, either most lines belong to large families of parallel lines, or most lines belong to large families of concurrent lines. This could be a first step towards proving that there must exist $n^{1/3}$ slopes, each with $n^{2/3}$ lines.

Intuitively and not rigorously, part (b) of Theorem 2 states that, either the set of slopes behaves similarly to a generalized geometric progression, or the y-intercepts behave similarly to a generalized arithmetic progression.

Shen's REU project wiht Adam [1] completely characterized the configurations when the point set is a lattice. It also provided a partial characterization when the point set is a Cartesian product of an arithmetic progression and an arbitrary set (half-lattice).

The 2024 project. In this project we will continue to study the Structural Szemerédi-Trotter problem, by attempting to prove new properties or finding new configurations. For example, one can continue where the previous project stopped [1]. Adam has more conrete ideas, which he prefers not to share here.

References

[1] S. Dasu, A. Sheffer, and J. Shen, Structural Szemerédi–Trotter for Lattices and their Generalizations, submitted.

[2] G. Elekes, On the number of sums and products, *ACTA ARITHMETICA — WARSZAWA* **81** (1997), 365–367.

[3] L. Guth and O. Silier, Sharp Szemerédi–Trotter Constructions in the Plane, submitted.

[4] A. Sheffer, Incidences: Lower Bounds (part 1) `https://adamsheffer.wordpress.com/2014/06/04/incidences-lower-bounds-part-1/`

[5] A. Sheffer, Incidences Outside of Discrete Geometry, `https://adamsheffer.wordpress.com/2016/01/22/incidences-outside-of-discrete-geometry/`

[6] A. Sheffer, Incidences Outside of Discrete Geometry part 2, `https://adamsheffer.wordpress.com/2017/10/06/incidences-outside-of-discrete-geometry-part-2/`

[7] A. Sheffer, and O. Silier, A Structural Szemerédi–Trotter Theorem for Cartesian Products, *Discrete & Computational Geometry*, accepted.

[8] E. Szemerédi and W. T. Trotter, Extremal problems in discrete geometry, *Combinatorica* **3** (1983), 381–392.