Annals of Mathematics **181** (2015), 155–190
http://dx.doi.org/10.4007/annals.2015.181.1.2

On the Erdős distinct distances problem in the plane

By Larry Guth and Nets Hawk Katz

Abstract

In this paper, we prove that a set of N points in $\mathbf{R}^2$ has at least $c\frac{N}{\log N}$ distinct distances, thus obtaining the sharp exponent in a problem of Erdős. We follow the setup of Elekes and Sharir which, in the spirit of the Erlangen program, allows us to study the problem in the group of rigid motions of the plane. This converts the problem to one of point-line incidences in space. We introduce two new ideas in our proof. In order to control points where many lines are incident, we create a cell decomposition using the polynomial ham sandwich theorem. This creates a dichotomy: either most of the points are in the interiors of the cells, in which case we immediately get sharp results or, alternatively, the points lie on the walls of the cells, in which case they are in the zero-set of a polynomial of suprisingly low degree, and we may apply the algebraic method. In order to control points incident to only two lines, we use the flecnode polynomial of the Rev. George Salmon to conclude that most of the lines lie on a ruled surface. Then we use the geometry of ruled surfaces to complete the proof.

1. Introduction

In [Erd46], Paul Erdős posed the question "how few distinct distances are determined by N points in the plane?" Erdős checked that if the points are arranged in a square grid, then the number of distinct distances is $\sim \frac{N}{\sqrt{\log N}}$. He conjectured that for any arrangement of N points, the number of distinct distances is $\gtrsim \frac{N}{\sqrt{\log N}}$. (Throughout this paper, we use the notation $A \gtrsim B$ to mean that there is a universal constant $C > 0$ with $A > CB$.)

In the present paper, we prove

Theorem 1.1. *A set of N points in the plane determines $\gtrsim \frac{N}{\log N}$ distinct distances.*

Various authors have proved lower bounds for the number of distinct distances. These include, but are not limited to, [Mos52], [CST92], [Szé97],

[ST01], and [Tar03]. The most recent lower bound, in [KT04], says that the number of distances is $\gtrsim N^{.8641}$. For a more thorough presentation of the history of the subject, see the recent book [GIS11].

In [ES11], Elekes and Sharir introduced a completely new approach to the distinct distances problem, which uses the symmetries of the problem in a novel way. They laid out a plan to prove Theorem 1.1, which we follow in this paper. Their approach connects the distinct distances problem to three-dimensional incidence geometry. Using their arguments, Theorem 1.1 follows from the following estimate about the incidences of lines in $\mathbb{R}^3$.

THEOREM 1.2. *Let $\mathfrak{L}$ be a set of N^2 lines in $\mathbb{R}^3$. Suppose that $\mathfrak{L}$ contains $\lesssim N$ lines in any plane or any regulus. Suppose that $2 \leq k \leq N$. Then the number of points that lie in at least k lines is $\lesssim N^3 k^{-2}$.*

A regulus is a quadratic surface in $\mathbb{R}^3$ which is doubly ruled. Doubly ruled means that each point in the surface lies in two lines in the surface. An example is the surface defined by the equation $z = xy$. Taking finitely many of the lines in a regulus one can build a configuration of lines with many intersection points. The term regulus is used in the incidence geometry community (cf. the first paper on the joints problem [CEGea92]) and in the harmonic analysis community (cf. [Sch98]).

Recently, there has been a lot of progress in incidence geometry coming from the polynomial method. In [Dvi09], Dvir used the polynomial method to prove the finite field Kakeya conjecture, which can be considered as a problem in incidence geometry over finite fields. In [GK10], the polynomial method was applied to incidence geometry problems in $\mathbb{R}^3$, solving the joints problem. The method was simplified and generalized in [EKS11], [KSS10], and [Qui10]. Kaplan, Sharir, and Shustin [KSS10] and Quilodrán [Qui10] solved the joints problem in higher dimensions. For context, we mention here the joints theorem in n dimensions.

THEOREM 1.3 ([KSS10], [Qui10]). *Let $n \geq 3$. Let $\mathfrak{L}$ be a set of L lines in $\mathbb{R}^n$. A joint of $\mathfrak{L}$ is a point that lies in n lines of $\mathfrak{L}$ with linearly independent directions. The number of joints of $\mathfrak{L}$ is $\leq C_n L^{\frac{n}{n-1}}$.*

One of the remarkable things about the polynomial method is how short the proofs are. The finite field Kakeya problem and the joints problem were considered to be very difficult, and many ideas were tried in both cases. The proof of the finite field Kakeya result ([Dvi09]) and the simplified proof of the joints theorem ([KSS10] or [Qui10]) are each about one page long. This simplicity gives the feeling that these are the "right" proofs for these theorems.

In [EKS11], Elekes, Kaplan, and Sharir used the polynomial method to prove the case $k = 3$ of Theorem 1.2. (It is a special case of Theorem 9 of that paper.) It remains to prove Theorem 1.2 when $k = 2$ and when k is large. This requires two new ideas. When $k = 2$, the key new idea is an application of ruled surfaces. When k is large, the key new idea is an application of a ham sandwich theorem from topology. Let us discuss these new ideas in more detail.

First we explain the extra difficulty that occurs when $k = 2$, as opposed to $k = 3$. The fundamental idea of the polynomial method is to find a polynomial p of controlled degree whose zero-set Z contains the set of lines $\mathfrak{L}$. Then one uses the geometry of Z to study $\mathfrak{L}$. A point where three lines of $\mathfrak{L}$ intersect is an unusual point of the surface Z: it is either a critical point of Z or else a 'flat' point of Z. One can use algebraic geometry to control the critical points and flat points of Z in terms of the degree of p. But a point of Z where two lines of $\mathfrak{L}$ intersect does not have to be either critical or flat, and we do not know of any special property of such a point.

Reguli play an important role in the case $k = 2$. If l_1, l_2, and l_3 are three pairwise-skew lines in $\mathbf{R}^3$, there is a one-parameter family of lines that intersects all three lines. The union of the lines in the one-parameter family is a surface called a regulus. A regulus is a degree 2 algebraic surface. An example is the surface defined by the equation $z = xy$. A set of N^2 lines in a regulus can have $\sim N^4$ points of intersection. A regulus is an example of a ruled surface. In this paper, a ruled surface means an algebraic surface that contains a line through each point.

We apply the theory of ruled surfaces to prove our estimate in the case $k = 2$. We first observe that if a surface Z of controlled degree contains too many lines, then some component of the surface Z must be ruled. In this way, we can reduce to the case of a set of lines contained in a ruled surface of controlled degree. Ruled surfaces have some special structure, and we use that structure to bound the intersections between the lines. A ruled surface is called singly-ruled if a generic point in the surface lies in only one line in the surface. It is well known that planes and reguli are not singly-ruled, but every other irreducible ruled surface is. The reason is that if a surface is not singly-ruled, it is easy to find three lines l_1, l_2, and l_3 in the surface which meet infinitely many lines not at one of the possibly three points of intersection of l_1, l_2, and l_3. This implies that the surface has a factor which is a plane, if any two of l_1, l_2, and l_3 are coplanar and a regulus if they are pairwise skew. A point where two lines intersect inside of a singly-ruled surface must be critical — except for points lying in the union of a controlled number of lines and a finite set of additional exceptions. By using this type of result, the structure of ruled surfaces helps us to prove our estimate.

Next we try to explain the extra difficulty that occurs for large k. An indication of the difficulty is that for large k, Theorem 1.2 does not hold over finite fields. (When $k = 2$ or 3, it is an open question whether Theorem 1.2 holds over finite fields, but we suspect that it does.) The counterexample occurs when one considers $\mathfrak{L}$ to be all of the lines in $\mathbb{F}^3$. (Here, $\mathbb{F}$ denotes a finite field.) This situation is reminiscent of the situation for the Szemerédi-Trotter theorem.

The Szemerédi-Trotter incidence theorem ([ST83]) is the most fundamental and important result in extremal incidence geometry. It was partly inspired by Erdős's distances problem, and it has played a role in all the recent work on the subject.

Theorem 1.4 (Szemerédi-Trotter). *Let $\mathfrak{L}$ be a set of L lines in $\mathbb{R}^2$. Then the number of points that lie in at least k lines is* $\le C(L^2k^{-3} + Lk^{-1})$.

The Szemerédi-Trotter theorem is also false over finite fields: the counterexample occurs when one considers $\mathfrak{L}$ to be all of the lines in $\mathbb{F}^2$. All of the proofs of the theorem involve in some way the topology of $\mathbb{R}^2$. One approach, which is important in our paper, is the cellular method introduced in the seminal paper [CEG$^+$90] by Clarkson, Edelsbrunner, Guibas, Sharir, and Welzl. The cellular method is a kind of divide-and-conquer argument. One carefully picks some lines, which divide the plane into cells, and then one studies $\mathfrak{L}$ inside of each cell.

The cellular method has been very successful for problems in the plane, but only partly successful in higher dimensions. For example, in [FS05], Feldman and Sharir attacked the (three-dimensional) joints problem using the cellular method (among other tools). They were able to prove that the number of joints determined by L lines is $\lesssim L^{1.62}$. (For contrast, the algebraic method gives $\lesssim L^{3/2}$.)

It seems to us that there are strong analogies between Theorem 1.2 and the Szemerédi-Trotter theorem, and also between Theorem 1.2 and the joints theorem. As in the Szemerédi-Trotter theorem, topology must play some role. As in the joints theorem, it is natural for polynomials to play some role.

To prove Theorem 1.2 when k is large, we construct a cell decomposition where the walls of the cells form an algebraic surface Z defined by a polynomial p. The polynomial is found by a topological argument, using the general ham sandwich theorem of Stone and Tukey [ST42]. At this point, our argument involves a dichotomy. Let $\mathfrak{S}$ denote the points that lie in at least k lines of $\mathfrak{L}$. In one extreme case, the points of $\mathfrak{S}$ are evenly distributed among the open cells of our decomposition. In this case, we prove our estimate by the cellular method, similar to arguments from [CEG$^+$90]. In another extreme case, the points of $\mathfrak{S}$ all lie in Z. In this case, the main contribution comes from lines of $\mathfrak{L}$ that lie in Z. In this case, we prove our estimate by the polynomial method, studying the critical and flat points of Z as in [GK10] or [EKS11].

In Section 2, we explain the plan laid out by Elekes and Sharir. In particular, we explain how Theorem 1.1 follows from Theorem 1.2. In Section 3, we prove Theorem 1.2 in the case $k = 2$ using the ruled surfaces method. We begin with the necessary background on ruled surfaces. In Section 4, we prove Theorem 1.2 for $k \geq 3$ using the polynomial cell method. We begin with background on the polynomial ham sandwich theorem. In the appendix, we show how our argument plays out when the set of points is a square grid. This example shows that several of our estimates are sharp up to constant factors, including Theorem 1.2.

Acknowledgements. The first author is partially supported by NSERC, by NSF grant DMS-0635607, and by the Monell Foundation. The second author is partially supported by NSF grant DMS-1001607. He would like to thank Michael Larsen for some very helpful discussions about algebraic geometry. He would also like to thank the Institute of Advanced Study for the use of its magnificent duck pond during a visit that resulted in this paper. Both authors would like to thank the helpful referee because of whom the exposition in the paper is significantly improved.

2. **Elekes-Sharir framework**

Elekes and Sharir [ES11] developed a completely new approach to the distinct distances problem, connecting it to incidence geometry in three-dimensional space. In this section, we present (a small variation of) their work.

Let $P \subset \mathbf{R}^2$ be a set of N points. We let $d(P)$ denote the set of nonzero distances among points of P:

$$d(P) := \{d(p, q)\}_{p,q \in P, p \neq q}.$$

To obtain a lower bound on the size of $d(P)$, we will prove an upper bound on a set of quadruples. We let $Q(P)$ be the set of quadruples, $(p_1, p_2, p_3, p_4) \in P^4$ satisfying

$$d(p_1, p_2) = d(p_3, p_4) \neq 0. \tag{2.1}$$

We refer to the elements of $Q(P)$ as distance quadruples. If $d(P)$ is small, then $Q(P)$ needs to be large. By applying the Cauchy-Schwarz inequality, we easily obtain the following inequality.

Lemma 2.1. *For any set $P \subset \mathbf{R}^2$ with N points, the following inequality holds:*

$$|d(P)| \geq \frac{N^4 - 2N^3}{|Q(P)|}.$$

Proof. Consider the distances in $d(P)$, which we denote by $d_1, \ldots, d_m$ with $m = |d(P)|$. There are $N^2 - N$ ordered pairs $(p_i, p_j) \in P^2$ with $p_i \neq p_j$. Let n_i be the number of these pairs at distance d_i. So $\sum_{i=1}^m n_i = N^2 - N$.

The cardinality $|Q(P)|$ is equal to $\sum_{i=1}^m n_i^2$. But by Cauchy-Schwarz,

$$(N^2 - N)^2 = \left(\sum_{i=1}^m n_i\right)^2 \leq \left(\sum_{i=1}^m n_i^2\right) m = |Q(P)||d(P)|.$$

Rearranging, we see that $|d(P)| \geq (N^2 - N)^2 |Q(P)|^{-1}$. $\square$

To prove Theorem 1.1, it suffices to prove the following upper bound on $|Q(P)|$.

Proposition 2.2. *For any set $P \subset \mathbf{R}^2$ of N points, the number of quadruples in $Q(P)$ is bounded by $|Q(P)| \lesssim N^3 \log N$.*

This proposition is sharp up to constant factors when P is a square grid (see the appendix).

Elekes and Sharir study $Q(P)$ from a novel point of view related to the symmetries of the plane. We let G denote the group of *positively oriented* rigid motions of the plane. The first connection between $Q(P)$ and G comes from the following simple proposition.

Proposition 2.3. *Let (p_1, p_2, p_3, p_4) be a distance quadruple in $Q(P)$. Then there is a unique $g \in G$ so that $g(p_1) = p_3$ and $g(p_2) = p_4$.*

Proof. All positively oriented rigid motions taking p_1 to p_3 can be obtained from the translation from p_1 to p_3 by applying a rotation R about the point p_3. Since $d(p_3, p_4) = d(p_1, p_2) > 0$, there is a unique such rotation sending $p_2 + p_3 - p_1$ into p_4. $\square$

Using Proposition 2.3, we get a map E from $Q(P)$ to G, which associates to each distance quadruple $(p_1, p_2, p_3, p_4) \in Q(P)$, the unique $g \in G$ with $g(p_1) = p_3$ and $g(p_2) = p_4$. The letter E here stands for Elekes, who introduced this idea.

Our goal is to use the map E to help us estimate $|Q(P)|$ by counting appropriate rigid motions. It is important to note that the map E is not necessarily injective. The number of quadruples in $E^{-1}(g)$ depends on the size of $P \cap gP$. We make this precise in the following lemma.

Lemma 2.4. *Suppose that $g \in G$ is a rigid motion and that $|P \cap gP| = k$. Then the number of quadruples in $E^{-1}(g)$ is $2\binom{k}{2}$.*

Proof. Suppose that $P \cap gP$ is $\{q_1, \ldots, q_k\}$. Let $p_i = g^{-1}(q_i)$. Since q_i lies in gP, each point p_i lies in P. For any ordered pair (q_i, q_j) with $q_i \neq q_j$, the set (p_i, p_j, q_i, q_j) is a distance quadruple. This assertion is easy to check. We have seen that p_i, p_j, q_i, q_j all lie in P. Since g preserves distances, $d(p_i, p_j) = d(q_i, q_j)$. Since $q_i \neq q_j$, the distance $d(q_i, q_j) \neq 0$.

Now we check that every distance quadruple in $E^{-1}(g)$ is of this form. Let (p_1, p_2, p_3, p_4) be a distance quadruple in $E^{-1}(g)$. We know that $g(p_1) = p_3$ and $g(p_2) = p_4$. So p_3, p_4 lie in $P \cap gP$. Say $p_3 = q_i$ and $p_4 = q_j$. Now $p_1 = g^{-1}(p_3) = p_i$ and $p_2 = g^{-1}(p_4) = p_j$. □

Let $G_{=k}(P) \subset G$ be the set of $g \in G$ with $|P \cap gP| = k$. Notice that $G_{=N}(P)$ is a subgroup of G. It is the group of orientation-preserving symmetries of the set P. For other k, $G_{=k}(P)$ is not a group, but these sets can still be regarded as sets of "partial symmetries" of P. Since P has N elements, $G_{=k}(P)$ is empty for $k > N$.

By Lemma 2.4, we can count $|Q(P)|$ in terms of $|G_{=k}(P)|$:

$$|Q(P)| = \sum_{k=2}^{N} 2\binom{k}{2} |G_{=k}(P)|.$$

Let $G_k(P) \subset G$ be the set of $g \in G$ so that $|P \cap gP| \geq k$. We see that $|G_{=k}(P)| = |G_k(P)| - |G_{k+1}(P)|$. Plugging this into the last equation and rearranging, we get the following:

$$|Q(P)| = \sum_{k=2}^{N} 2\binom{k}{2} \left(|G_k(P)| - |G_{k+1}(P)|\right) = \sum_{k=2}^{N} (2k-2)|G_k(P)|. \tag{2.2}$$

We will bound the number of partial symmetries as follows.

Proposition 2.5. *For any set $P \subset \mathbf{R}^2$ of N points, and any $2 \leq k \leq N$, the size of $G_k(P)$ is bounded as follows*:

$$|G_k(P)| \lesssim N^3 k^{-2}.$$

When P is a square grid, this estimate is sharp up to constant factors for all $2 \leq k \leq N/2$ (see the appendix). Plugging this bound into equation (2.2), we get $|Q(P)| \lesssim N^3 \log N$, proving Proposition 2.2. This in turn implies our main theorem, Theorem 1.1. So it suffices to prove Proposition 2.5.

Next Elekes and Sharir related the sets $G_k(P)$ to an incidence problem involving certain curves in G. For any points $p, q \in \mathbf{R}^2$, define the set $S_{pq} \subset G$ given by

$$S_{pq} = \{g \in G : g(p) = q\}.$$

Each S_{pq} is a smooth one-dimensional curve in the three-dimensional Lie group G. The sets $G_k(P)$ are closely related to the curves S_{pq}.

Lemma 2.6. *A rigid motion g lies in $G_k(P)$ if and only if it lies in at least k of the curves $\{S_{pq}\}_{p,q \in P}$.*

Proof. First suppose that g lies in $G_k(P)$. By definition, $|P \cap gP| \geq k$. Let $q_1, \dots, q_k$ be distinct points in $P \cap gP$. Let $p_i = g^{-1}(q_i)$. Since $q_i \in gP$, we see that p_i lies in P. Since $g(p_i) = q_i$, we can say that g lies in $S_{p_i q_i}$ for $i = 1, \dots, k$. Since the q_i are all distinct, these are k distinct curves.

On the other hand, suppose that g lies in the curves $S_{p_1 q_1}, \dots, S_{p_k q_k}$, where we assume that the pairs $(p_1, q_1), \dots, (p_k, q_k)$ are all distinct. We claim that $q_1, \dots, q_k$ are distinct points. To see this, suppose that $q_i = q_j$. Since g is a bijection, we see that $p_i = g^{-1}(q_i) = g^{-1}(q_j) = p_j$, and this gives a contradiction. But the points $q_1, \dots, q_k$ all lie in $P \cap gP$. $\square$

Bounding $G_k(P)$ is a problem of incidence geometry about the curves $\{S_{pq}\}_{p,q \in P}$ in the group G. By making a careful change of coordinates, we can reduce this problem to an incidence problem for lines in $\mathbf{R}^3$. (Our change of coordinates is slightly nicer than the one in [ES11]. In the coordinates of [ES11], the curves $\{S_{pq}\}$ become parabolas.)

Let G' denote the open subset of the orientable rigid motion group G given by rigid motions that are not translations. We can write G as a disjoint union $G' \cup G^{\text{trans}}$, where G^{trans} denotes the translations. We then divide $G_k(P) = G'_k \cup G^{\text{trans}}_k$. Translations are a very special class of rigid motions, and it is fairly easy to bound $|G^{\text{trans}}_k(P)| \lesssim N^3 k^{-2}$. We carry out this minor step at the end of this section. The main point is to bound $|G'_k(P)|$. To do this, we pick a nice set of coordinates $\rho : G' \to \mathbf{R}^3$.

Each element of G' has a unique fixed point (x, y) and an angle θ of rotation about the fixed point with $0 < \theta < 2\pi$. We define the map

$$\rho : G' \longrightarrow \mathbf{R}^3$$

by

$$\rho(x, y, \theta) = \left(x, y, \cot \frac{\theta}{2}\right).$$

PROPOSITION 2.7. *Let $p = (p_x, p_y)$ and $q = (q_x, q_y)$ be points in $\mathbf{R}^2$. Then with ρ as above, the set $\rho(S_{pq} \cap G')$ is a line in $\mathbf{R}^3$.*

Proof. Noting that the fixed point of any transformation taking p to q must lie on the perpindicular bisector of p and q, the reader will easily verify that the set $\rho(S_{pq} \cap G')$ can be parametrized as

$$\left(\frac{p_x + q_x}{2}, \frac{p_y + q_y}{2}, 0\right) + t\left(\frac{q_y - p_y}{2}, \frac{p_x - q_x}{2}, 1\right). \tag{2.3}$$ $\square$

For any $p, q \in \mathbf{R}^2$, let L_{pq} denote the line $\rho(S_{pq} \cap G')$. The line L_{pq} is parametrized by equation (2.3). Let $\mathfrak{L}$ be the set of lines $\{L_{pq}\}_{p,q \in P}$. By examining the parametrization in equation (2.3), it is easy to check that these are N^2 distinct lines. If g lies in $G'_k(P)$, then $\rho(g)$ lies in at least k lines of $\mathfrak{L}$. In the remainder of the paper, we will study the set of lines $\mathfrak{L}$ and estimate the number of points lying in k lines.

We would like to prove that there are $\lesssim N^3 k^{-2}$ points that lie in at least k lines of $\mathfrak{L}$. Such an estimate does not hold for an arbitrary set of N^2 lines. For example, if all the lines of $\mathfrak{L}$ lie in a plane, then one may expect $\sim N^4$ points that lie in at least two lines. This number of intersection points is far too high. There is another important example, which occurs when all the lines lie in a regulus. Recall that a regulus is a doubly-ruled surface, and each line from one ruling intersects all the lines from the other ruling. If $\mathfrak{L}$ contained $N^2/2$ lines in each of the rulings, then we would have $\sim N^4$ points that lie in at least two lines. Because of this example, we have to show that not too many lines of $\mathfrak{L}$ lie in a plane or a regulus.

PROPOSITION 2.8. *No more than N lines of $\mathfrak{L}$ lie in a single plane. No more than $O(N)$ lines of $\mathfrak{L}$ lie in a single regulus.*

Proof. For each $p \in P$, we consider the subset $\mathfrak{L}_p \subset \mathfrak{L}$ given by

$$\mathfrak{L}_p = \{L_{pq}\}_{q \in P}.$$

Notice that if $q \neq q'$, then for each p, L_{pq} and $L_{pq'}$ are disjoint. So the lines of $\mathfrak{L}_p$ are disjoint. From equation (2.3), it follows that the lines of $\mathfrak{L}_p$ all have different directions. So the lines of $\mathfrak{L}_p$ are pairwise skew, and no two of them lie in the same plane. Therefore, any plane contains at most N lines of $\mathfrak{L}$.

The situation for reguli is more complicated because all N lines of $\mathfrak{L}_p$ may lie in a single regulus. But we will prove that this can only occur for at most two values of p. To formulate this argument, we define $\mathfrak{L}'_p := \{L_{pq}\}_{q \in \mathbf{R}^2}$, so that $\mathfrak{L}_p \subset \mathfrak{L}'_p$.

LEMMA 2.9. *Suppose that a regulus R contains at least seven lines of $\mathfrak{L}'_p$. Then all the lines in one ruling of R lie in $\mathfrak{L}'_p$.*

Given this lemma, the rest of the proof of Proposition 2.8 is straightforward. If a regulus R contains at least seven lines of $\mathfrak{L}_p$, then all the lines in one ruling of R lie in $\mathfrak{L}'_p$. If $p_1 \neq p_2$, then the two sets of lines $\mathfrak{L}'_{p_1}$ and $\mathfrak{L}'_{p_2}$ are disjoint. This follows from the explicit formula in equation (2.3). Since a regulus has only two rulings, there are at most two values of p such that R contains ≥ 7 lines of $\mathfrak{L}_p$. These two values of p contribute $\leq 2N$ lines of $\mathfrak{L}$ in the surface R. The other $N-2$ values of p contribute at most $6(N-2)$ lines of $\mathfrak{L}$ in the surface R. Therefore, the surface R contains at most $2N+6(N-2) \lesssim N$ lines of $\mathfrak{L}$.

Proof of Lemma 2.9. We fix the value of p. We will check below that each point of $\mathbf{R}^3$ lies in exactly one line of $\mathfrak{L}'_p$. We will construct a nonvanishing vector field $V = (V_1, V_2, V_3)$ on $\mathbf{R}^3$ tangent to the lines of $\mathfrak{L}'_p$. Moreover, the coefficients V_1, V_2 and V_3 are all polynomials in (x, y, z) of degree ≤ 2. This construction is slightly tedious but straightforward. We postpone it to the end of the proof.

The regulus R is defined by an irreducible polynomial f of degree 2. Now suppose that a line L_{pq} lies in R. At each point $x \in L_{pq}$, the vector $V(x)$ points tangent to the line L_{pq}, and so the directional derivative of f in the direction $V(x)$ vanishes at the point x. In other words, the dot product $V \cdot \nabla f$ vanishes on the line L_{pq}. Since f has degree 2, the dot product $V \cdot \nabla f$ is a polynomial of degree at most 3.

Suppose that R contains seven lines of $\mathfrak{L}'_p$. We know that f vanishes on each line, and the previous paragraph shows that $V \cdot \nabla f$ vanishes on each line. By Bezout's theorem (see Lemma 3.1), f and $V \cdot \nabla f$ must have a common factor. Since f is irreducible, we must have that f divides $V \cdot \nabla f$. In other words, $V \cdot \nabla f$ vanishes on the surface R, and so V is tangent to R at every point of R. If x denotes any point in R, and we let L be the line of $\mathfrak{L}'_p$ containing x, then we see that this line lies in R. In this way, we get a ruling of R consisting of lines from $\mathfrak{L}'_p$.

It remains to define the vector field V. We begin by checking that each point (x, y, z) lies in exactly one line of $\mathfrak{L}'_p$. By equation (2.3), (x, y, z) lies in L_{pq} if and only if the following equation holds for some t:

$$\left(\frac{p_x + q_x}{2}, \frac{p_y + q_y}{2}, 0\right) + t\left(\frac{q_y - p_y}{2}, \frac{p_x - q_x}{2}, 1\right) = (x, y, z).$$

Given p and (x, y, z), we can solve uniquely for t and (q_x, q_y). First of all, we see that $t = z$. Next we get a matrix equation of the following form:

$$\begin{pmatrix} 1 & z \\ -z & 1 \end{pmatrix} \begin{pmatrix} q_x \\ q_y \end{pmatrix} = a(x, y, z).$$

In this equation, $a(x, y, z)$ is a vector whose entries are polynomials in x, y, z of degree ≤ 1. (The polynomials also depend on p, but since p is fixed, we suppress the dependence.) Since the determinant of the matrix on the left-hand side is $1 + z^2 > 0$, we can solve this equation for q_x and q_y. The solution has the form

$$\begin{pmatrix} q_x \\ q_y \end{pmatrix} = (z^2 + 1)^{-1} b(x, y, z). \tag{2.4}$$

In this equation, $b(x, y, z)$ is a vector whose entries are polynomials in x, y, z of degree ≤ 2.

The vector field $V(x, y, z)$ is $(z^2 + 1)(\frac{q_y - p_y}{2}, \frac{p_x - q_x}{2}, 1)$. Recall that p is fixed, and q_x and q_y can be expressed in terms of (x, y, z) by the equation above. By equation (2.3), this vector field is tangent to the line L_{pq}. After multiplying out, the third entry of V is $z^2 + 1$, so V is nonvanishing. Plugging in equation (2.4) for q_x and q_y and multiplying out, we see that the entries of $V(x, y, z)$ are polynomials of degree ≤ 2. $\square$

This concludes the proof of Proposition 2.8. $\square$

We have now connected the distinct distances problem to the incidence geometry problem we mentioned in the introduction. We know that $\mathfrak{L}$ consists of N^2 lines with $\lesssim N$ lines in any plane or regulus. We now state our two results on incidence geometry.

THEOREM 2.10. *Let $\mathfrak{L}$ be any set of N^2 lines in $\mathbf{R}^3$ for which no more than N lie in a common plane and no more than $O(N)$ lie in a common regulus. Then the number of points of intersection of two lines in $\mathfrak{L}$ is $O(N^3)$.*

THEOREM 2.11. *Let $\mathfrak{L}$ be any set of N^2 lines in $\mathbf{R}^3$ for which no more than N lie in a common plane, and let k be a number $3 \leq k \leq N$. Let $\mathfrak{S}_k$ be the set of points where at least k lines meet. Then*

$$|\mathfrak{S}_k| \lesssim N^3 k^{-2}.$$

Elekes and Sharir essentially conjectured these two theorems (Conjecture 1 in [ES11]). (The difference is that they used different coordinates, so their conjectures are about certain types of parabolas.) In the case $k = 3$, Theorem 2.11 was proven in [EKS11].

Combining these theorems with the coordinates ρ and Proposition 2.8, we get bounds for $|G'_k(P)|$. Theorem 2.10 shows that $|G'_2(P)| \lesssim N^3$. Theorem 2.11 shows that $|G'_k(P)| \lesssim N^3 k^{-2}$ for $3 \leq k \leq N$.

We now prove similar bounds for $|G_k^{\text{trans}}(P)|$. These bounds are completely elementary

LEMMA 2.12. *Let P be any set of N points in $\mathbf{R}^2$. The number of quadruples in $E^{-1}(G^{\text{trans}})$ is $\leq N^3$. Moreover, $|G_k^{\text{trans}}(P)| \lesssim N^3 k^{-2}$ for all $2 \leq k \leq N$.*

Proof. Suppose that (p_1, p_2, p_3, p_4) is a distance quadruple in $E^{-1}(G^{\text{trans}})$. By definition, there is a *translation* g so that $g(p_1) = p_3$ and $g(p_2) = p_4$. Therefore, $p_3 - p_1 = p_4 - p_2$. This equation allows us to determine p_4 from p_1, p_2, p_3. Hence there are $\leq N^3$ quadruples in $E^{-1}(G^{\text{trans}})$.

By Proposition 2.4 we see that

$$|E^{-1}(G^{\text{trans}})| = \sum_{k=2}^{N} 2\binom{k}{2} |G_{=k}^{\text{trans}}(P)|.$$

Noting that $|G_k^{\text{trans}}(P)| = \sum_{l \geq k} |G_{=l}^{\text{trans}}(P)|$, we see that

$$N^3 \geq |E^{-1}(G^{\text{trans}})| \geq 2\binom{k}{2} |G_k^{\text{trans}}(P)|.$$

This inequality shows that $|G_k^{\text{trans}}(P)| \lesssim N^3 k^{-2}$ for all $2 \leq k \leq N$. □

This substantially ends Section 2. To conclude, we give a summary and make some comments.

The new ingredients in this paper are Theorems 2.10 and 2.11, which we prove in Sections 3 and 4. These theorems allow us to bound the partial symmetries of P in G': they imply that $|G'_k(P)| \lesssim N^3 k^{-2}$ for all $2 \le k \le N$. An elementary argument in Lemma 2.12 shows the same estimates for $|G_k^{\text{trans}}(P)|$. Combining these, we see that $|G_k(P)| \lesssim N^3 k^{-2}$ for $2 \le k \le N$, proving Proposition 2.5. Now the number of quadruples in $Q(P)$ is expressed in terms of $|G_k(P)|$ in equation (2.2). Plugging in our bound for $|G_k(P)|$, we get that $|Q(P)| \lesssim N^3 \log N$, proving Proposition 2.2. Finally, the number of distinct distances is related to $|Q(P)|$ by Lemma 2.1. Plugging in our bound for $|Q(P)|$, we see that $|d(P)| \gtrsim N(\log N)^{-1}$, proving our main theorem.

The group G acts as a bridge connecting the original problem on distinct distances to the incidence geometry of lines in $\mathbf{R}^3$. The distance set $d(P)$ is related to the set of quadruples $Q(P)$, which is related to the partial symmetries $G_k(P)$, which correspond to k-fold intersections of the lines in $\mathfrak{L}$. The group G is a natural symmetry group for the problem of distinct distances, but this way of using the symmetry group is new and rather surprising.

Our estimates show that sets with few distinct distances must have many partial symmetries. For example, if $G_3(P)$ is empty, then our results show that $|Q(P)| \lesssim N^3$ and $|d(P)| \gtrsim N$. Also, any set with $|d(P)| \lesssim N(\log N)^{-1/2}$ must have a partial symmetry with $k \ge \exp(c \log^{1/2} N)$ for a universal constant $c > 0$. Any set with $|d(P)| \lesssim N(\log N)^{-1}$ must have a partial symmetry with $k \ge N^c$ for a universal $c > 0$.

3. Flecnodes

Our goal in this section is to prove Theorem 2.10. We will do this by purely algebraic methods following essentially the proof strategy of [GK10]. That is, we will show that an important subset of our lines lies in the zero-set of a fairly low degree polynomial p. What requires a new idea is the next step. We need a polynomial q derived from p with similar degree on which the lines also vanish. With that information we will apply a variant of Bezout's lemma.

Lemma 3.1. *Let $p(x, y, z)$ and $q(x, y, z)$ be polynomials on $\mathbf{R}^3$ of degrees m and n respectively. If there is a set of $mn + 1$ distinct lines simultaneously contained in the zero-set of p and the zero-set of q, then p and q have a common factor.*

Thus we will conclude that p and the derived polynomial q must have a common factor, and we will arrive at some geometrical conclusion from this based on the way that q was derived. In the paper [GK10], the derived polynomials that we used were the gradient of p and the algebraic version of the second fundamental form of the surface given by $p = 0$. These were good choices because when three or more lines were incident at each point, we knew on geometric grounds that one or the other would vanish at each point, because the point would be either critical or flat. However, here we are faced with points at which only two lines intersect, and so we must make a more clever choice of the derived polynomial.

We begin with the definition of a flecnode. Given an algebraic surface in $\mathbf{R}^3$ given by the equation $p(x, y, z) = 0$ where p is a polynomial of degree d at least 3, a flecnode is a point (x, y, z) where a line agrees with the surface to order three. To find all such points, we might solve the system of equations:

$$p(x,y,z)=0; \quad \nabla_v p(x,y,z)=0; \quad \nabla_v^2 p(x,y,z)=0; \quad \nabla_v^3 p(x,y,z)=0.$$

These are four equations for six unknowns, (x, y, z) and the components for the direction v. However, the last three equations are homogeneous in v and may be viewed as three equations in five unknowns (and the whole system as four equations in five unknowns.) We may reduce the last three equations to a single equation in three unknowns (x, y, z). We write the reduced equation as

$$\mathrm{Fl}(p)(x,y,z)=0.$$

The polynomial $\mathrm{Fl}(p)$ is of degree $11d-24$. It is called the flecnode polynomial of p and vanishes at any flecnode of any level set of p. (See [Sal58, Art. 588, pp. 277–78].)

The term flecnode was apparently first coined by Cayley. The polynomial $\mathrm{Fl}(p)$ was discovered by the Rev. George Salmon, but its most important property to us was communicated to him by Cayley.

Proposition 3.2. *The surface $p = 0$ is ruled if and only if* $\mathrm{Fl}(p)$ *is everywhere vanishing on it.*

Proposition 3.2 was used in a famous paper of Segre [Seg43]. For a generalization to manifolds in higher dimensions, see [Lan99]. One direction of Proposition 3.2 is obvious. If the surface is ruled, there is a line contained in the surface at every point. If the line is contained in the surface, it certainly agrees to order three.

We take a moment to briefly review the argument given in Salmon for Proposition 3.2. It is written in language a little uncomfortable for the modern reader and broken into two parts in Salmon's book. First one observes that the property that an algebraic surface is ruled (that is, has a line going through every point) is equivalent to a differential equation satisfied by the polynomial that defines the surface. This is done in [Sal58, Art. 437, pp. 19–20]. Basically, we write down the parametric equation of the one-parameter family of lines contained in the surface. We differentiate three times until we can eliminate all parameters having to do with the direction of the line. We then obtain a third-order differential equation satisfied by the surface and observe that given a surface satisfying this equation, one can recover the one-parameter family of lines. The remainder of the proof is in the footnote in [Sal58, p. 278]. It is merely the observation that this differential equation is the same as the statement that everywhere a line vanishes to third order.

An algebraic surface (in $\mathbf{R}^3$) is ruled if it contains a line passing through every point. The set of all lines contained in an algebraic surface (of some degree d) is an algebraic set of lines. (This is because a line is contained in the surface if and only if it is contained to order $d+1$ at one of its points. So a line is contained in the surface if and only if $d+1$ polynomial equations in the parameters of the line are satisfied.) The set of lines contained in a surface may have two-dimensional components, one-dimensional components and zero-dimensional components. It is easy to see that an algebraic surface in $\mathbf{R}^3$ contains a two-dimensional set of lines only if it has a plane as a factor. (The way to see this is to find a regular point of the surface with an infinite number of lines going through it. Then the surface must contain the tangent plane to this point.) Thus an algebraic surface that is ruled and plane-free will contain both a one-dimensional set of lines (the generators) and possibly a zero-dimensional set of lines. A detailed classical treatment of ruled surfaces is given in [Sal58, Chap. XIII, Part 3].

One important example of a ruled surface is a regulus. A regulus is actually doubly-ruled: every point in the regulus lies in two lines in the regulus. A ruled surface is called singly-ruled if a generic point in the surface lies in only one line in the surface. (Some points in a singly-ruled surface may lie in two lines.) Except for reguli and planes, every irreducible ruled surface (in $\mathbf{R}^3$) is singly-ruled. (See the explanation in Section 1.)

An immediate corollary of the proposition is

COROLLARY 3.3. *Let $p=0$ be a degree d hypersurface in $\mathbf{R}^3$. Suppose that the surface contains more than $11d^2-24d$ lines. Then p has a ruled factor.*

Proof. By Lemma 3.1, since both p and $\mathrm{Fl}(p)$ vanish on the same set of more than $11d^2-24d$ lines, they must have a common factor q. Since q is a factor of p and $\mathrm{Fl}(p)$ vanishes on the surface $q=0$, it must be that at every regular point of the surface $q=0$, there is a line that meets the surface to order three. Thus $\mathrm{Fl}(q)=0$, which implies by Proposition 3.2 that q is ruled. □

Now we would like to consider ruled surfaces of degree less than N. Thus our surfaces are the sets

$$p(x,y,z)=0$$

for a polynomial p (which we may choose square free) of degree less than N. We may uniquely factorize the polynomial into irreducibles:

$$p=p_1p_2\cdots p_m.$$

We say that p is plane-free and regulus-free if none of the zero-sets of the factors is a plane or a regulus. Thus if p is plane-free and regulus-free, the zero-set of each of the factors is an irreducible algebraic singly-ruled surface. We now state the main geometrical lemma for proving Theorem 2.10.

LEMMA 3.4. *Let p be a polynomial of degree less than N so that $p = 0$ is ruled and so that p is plane-free and regulus-free. Let $\mathfrak{L}_1$ be a set of lines contained in the surface $p = 0$ with $|\mathfrak{L}_1| \lesssim N^2$. Let Q_1 be the set of points of intersection of lines in $\mathfrak{L}_1$. Then*

$$|Q_1| \lesssim N^3.$$

Before we begin in earnest the proof of Lemma 3.4, we will nail down a few delicate points of the geometry of irreducible singly-ruled surfaces.

We let $p(x, y, z)$ be an irreducible polynomial so that $p(x, y, z) = 0$ is a ruled surface that is not a plane or a regulus. In other words, the surface $S = \{(x, y, z) : p(x, y, z) = 0\}$ is irreducible and singly-ruled. We say that a point $(x_0, y_0, z_0) \in S$ is an *exceptional point* of the surface if it lies on infinitely many lines contained in the surface. We say that a line l contained in S is an *exceptional line* of the surface if there are infinitely many lines in S that intersect l at nonexceptional points. We prove a structural lemma about exceptional points and exceptional lines of irreducible singly-ruled surfaces.

LEMMA 3.5. *Let $p(x, y, z)$ be an irreducible polynomial. Let $S = \{(x, y, z) : p(x, y, z) = 0\}$ be an irreducible ruled surface which is neither a plane nor a regulus.*

(1) *Let (x_0, y_0, z_0) be an exceptional point of S. Then every other point (x, y, z) of S is on a line l which is contained in S and which contains the point (x_0, y_0, z_0).*
(2) *Let l be an exceptional line of S. Then there is an algebraic curve C so that every point of S not lying on C is contained in a line contained in S and intersecting l.*

We proceed to give an elementary proof of Lemma 3.5. (We have been advised by a helpful referee that the second part of the lemma is true without the exceptional curve C, but perhaps the proof is a little less elementary.)

Proof. To prove the first part, we observe that by a change of coordinates we can move (x_0, y_0, z_0) to the origin. We let Q be the set of points q different from the origin so that the line from q to the origin is contained in S. We observe that Q is the intersection of an algebraic set with the complement of the origin. That is, there is a finite set of polynomials E so that a point q different from the origin lies in Q if and only if each polynomial in E vanishes at q. This is because if d is the degree of p, to test whether $q \in Q$, we need only check that the line containing q and the origin is tangent to S to degree $d+1$ at q. Now by assumption, the zero-set of each polynomial in E contains the union of infinitely many lines contained in S. Thus by Lemma 3.1 and by the irreducibility of p, it must be that each polynomial in E has p as a factor. Therefore, Q is all of S except the origin. We have proved the first part.

Now to prove the second part, we observe that by a change of coordinates, we may choose l to be the coordinate line $y = 0; z = 0$. We let Q be the set of points q not on l so that there is a line from q to a nonexceptional point of l that is contained in S. We will prove that Q contains all the points $(x, y, z) \in S$ where $\frac{\partial p}{\partial x}(x, y, z) \neq 0$.

Consider a point (x, y, z) on S for which $\frac{\partial p}{\partial x}(x, y, z) \neq 0$. In particular, the point (x, y, z) is a regular point of S. Since $\frac{\partial p}{\partial x}(x, y, z) \neq 0$, there is a unique point $(x', 0, 0)$ of l that lies in the tangent plane to S at the point (x, y, z). In fact, we can solve for x' as a rational function of (x, y, z) with only the polynomial $\frac{\partial p}{\partial x}$ in the denominator. Thus we can find a set E of rational functions having only powers of $\frac{\partial p}{\partial x}$ in their denominators, so that for any (x, y, z) at which $\frac{\partial p}{\partial x}$ does not vanish, we have that $(x, y, z) \in Q$ if and only if every function in E vanishes on (x, y, z).

In order for the previous paragraph to be useful to us, we need to know that $\frac{\partial p}{\partial x}$ does not vanish identically on S. Suppose that it did. Since $\frac{\partial p}{\partial x}$ is of lower degree than p and p is irreducible, it must be that $\frac{\partial p}{\partial x}$ vanishes identically as a polynomial so that p depends only on y and z. In this case, since S contains l and it contains a line l' intersecting l, it must contain all translates of l' in the x-direction. Thus it contains a plane, which is a contradiction.

Thus, we let C be the algebraic curve where both p and $\frac{\partial p}{\partial x}$ vanish. Away from C, there is a finite set of polynomials F (which we obtain from E by multiplying by a large enough power of $\frac{\partial p}{\partial x}$) so that a point (x, y, z) of S outside of C is in Q if and only if each polynomial in F vanishes at (x, y, z). Since we know that p is irreducible and Q contains an infinite number of lines, it must be that each polynomial in F has p as a factor. Thus every point of S that is outside of C lies in Q, which was to be shown. $\square$

Now that we have established our structural result, Lemma 3.5, we may use it to obtain a corollary giving quantitative bounds on the number of exceptional points and lines.

COROLLARY 3.6. *Let $p(x, y, z)$ be an irreducible polynomial. Let*

$$S = \{(x, y, z) : p(x, y, z) = 0\}$$

be an irreducible surface that is neither a plane nor a regulus. Then S has at most one exceptional point and at most two exceptional lines.

We now prove Corollary 3.6

Proof. Let (x_0, y_0, z_0) and (x_1, y_1, z_1) be distinct exceptional points of S. Since S is singly-ruled, the generic point of S is contained in only a single line l contained in S. Thus by Lemma 3.5, if the point is different from (x_0, y_0, z_0) and (x_1, y_1, z_1), this line l must contain both (x_0, y_0, z_0) and (x_1, y_1, z_1) . But there is only one such line, and that is a contradiction.

Now let l_1, l_2, l_3 be exceptional lines of S. There are curves C_1, C_2, and C_3 so that the generic point in the complement of C_1, C_2, and C_3 lies on only one line contained in S and this line must intersect each of l_1, l_2, and l_3. Thus there are infinitely many lines contained in S that intersect each of l_1, l_2, and l_3. (Moreover, since the lines are exceptional, there must be an infinite set of lines that intersect the three away from the possible three points of intersection of any two of l_1, l_2, and l_3.) If any two of the three lines are coplanar, this means there is an infinite set of lines contained in S that lie in one plane. This contradicts the irreducibility and nonplanarity of S. If contrariwise, the three lines l_1, l_2, and l_3 are pairwise skew, then the set of all lines that intersect all three are one ruling of a regulus. In this case, S contains infinitely many lines of a regulus, which contradicts the fact that S is irreducible and not a regulus. □

For context, we remark that an irreducible singly-ruled surface with an exceptional point is often referred to as a cone and the exceptional point is referred to as the cone point. Irreducible ruled surfaces with two exceptional lines do exist: one way of constructing a ruled surface with two exceptional lines is to choose a curve in the two-dimensional set of lines that intersect a pair of skew lines. At last, we may begin the proof of Lemma 3.4.

Proof. We say that a point (x, y, z) is exceptional for the surface $p = 0$ if it is exceptional for $p_j = 0$ where p_j is one of the irreducible factors of p. We say that a line l is exceptional for the surface $p = 0$ if it is exceptional for $p_j = 0$ where p_j is one of the irreducible factors of p. Thus, in light of Corollary 3.6, there are no more than N exceptional points and $2N$ exceptional lines for $p = 0$. Thus there are $\lesssim N^3$ intersections between exceptional lines and lines of $\mathfrak{L}_1$. Thus to prove the lemma, we need only consider intersections between nonexceptional lines of $\mathfrak{L}_1$ at nonexceptional points.

We recall that the set of lines in $Z(p)$ can be divided into generators and nongenerators. The set of lines in $Z(P)$ forms an algebraic variety of dimension one in the affine Grassmannian of all lines in $\mathbb{R}^3$. This algebraic variety may have several irreducible components, and the components may have dimension zero or one. The lines in the components of dimension one are called generators. There are only finitely many nongenerators.

We note that any line contained in a ruled surface that is not a generator must be an exceptional line since each point of the line will have a generator going through it. (The definition of a ruled surface is that every point lies in a line in the surface. Since there are only finitely many nongenerators, almost every point must lie in a generator. But in fact every point lies in a generator by a limiting argument. Let q be a point in the ruled surface, and let q_i be a sequence of points that converge to q with q_i lying in a generator l_i. By taking a subsequence, we can arrange that the directions of the l_i converge, and so the lines l_i converge to a limit line l that contains q and lies in the surface. This line is a limit of generators, and so it is a generator.)

Let l be a nonexceptional line in the ruled surface. In particular, l is a generator. We claim that there are at most $N-1$ nonexceptional points in l where l intersects another nonexceptional line in the ruled surface. This claim implies that there are at most $(N-1)N^2$ nonexceptional points where two nonexceptional lines intersect, proving the bound we want.

To prove the claim, we repeat an argument found in [Sal58, Art 485, pp. 88–89]. Choose a plane π through the generator l. The plane intersects the surface in a curve of degree N. One component is the generator itself. The other component is an algebraic curve c of degree $N-1$. There are at most $N-1$ points of intersection between l and c. Suppose that l' is another nonexceptional line and that l' intersects l at a nonexceptional point q. It suffices to prove that q lies in the curve c. If l' lies in π, then $l' \subset c$ and $q \in c$. So we can assume that l' is transverse to π. Since l' is a generator, it lies in a continuous one-parameter family of other generators. Consider a small open set of generators around l'. These generators intersect the plane π. So each of them intersects either l or c. Since q is nonexceptional, only finitely many of them intersect q. Since there are only finitely many exceptional points, we can arrange that each generator in our small open set intersects π in a nonexceptional point. Since l is nonexceptional, only finitely many of our generators can intersect l. Therefore, almost all of our generators must intersect c. This is only possible if q lies in c. $\square$

Now we are ready to begin the proof of Theorem 2.10. We assume we have a set $\mathfrak{L}$ of at most N^2 lines for which no more than N lie in a plane and no more than N lie in a regulus. We suppose, by way of contradiction, that for Q, a positive real number sufficiently large, there are QN^3 points of intersection of lines of $\mathfrak{L}$, and we assume that this is an optimal example, so that for no $M < N$ do we have a set of M^2 lines so that no more than M lie in a plane and no more than M lie in a regulus is it the case that there are more than QM^3 intersections. (N need not be an integer.)

We now apply a degree reduction argument similar to the one in [GK10]. We let $\mathfrak{L}'$ be the subset of $\mathfrak{L}$ consisting of lines that intersect other lines of $\mathfrak{L}$ in at least $\frac{QN}{10}$ different points. The lines not in $\mathfrak{L}'$ participate in at most

$\frac{QN^3}{10}$ points of intersection. Thus there are at least $\frac{9QN^3}{10}$ points of intersection between lines of $\mathfrak{L}'$. We define a number α with $0<\alpha\leq 1$ so that $\mathfrak{L}'$ has αN^2 lines.

Now we select a random subset $\mathfrak{L}''$ of the lines of $\mathfrak{L}'$ choosing lines independently with probability $\frac{100}{Q}$. With positive probability, there will be no more than $\frac{200\alpha N^2}{Q}$ lines in $\mathfrak{L}''$ and each line of $\mathfrak{L}'$ will intersect lines of $\mathfrak{L}''$ in at least N different points. Now pick $\frac{R\sqrt{\alpha}N}{\sqrt{Q}}$ points on each line of $\mathfrak{L}''$. (R is a constant that is sufficiently large but universal.) Call the set of all of the points $\mathfrak{S}$. There are $O\left(\frac{R\alpha^{\frac{3}{2}}N^3}{Q^{\frac{3}{2}}}\right)$ points in $\mathfrak{S}$, so we may find a polynomial p of degree $O\left(\frac{R^{\frac{1}{3}}\alpha^{\frac{1}{2}}N}{Q^{\frac{1}{2}}}\right)$ that vanishes on every point of $\mathfrak{S}$. With R sufficiently large, p must vanish identically on every line of $\mathfrak{L}''$. Since each line of $\mathfrak{L}'$ meets $\mathfrak{L}''$ at N different points, it must be that p vanishes identically on each line of $\mathfrak{L}'$. Thus ends the degree reduction argument, and we will now study the relatively low degree polynomial p.

We may factor $p = p_1p_2$, where p_1 is the product of the ruled irreducible factors of p and p_2 is the product of unruled irreducible factors of p. Each of p_1 and p_2 is of degree $O\left(\frac{\alpha^{\frac{1}{2}}N}{Q^{\frac{1}{2}}}\right)$. (We have suppressed the R dependence since R is universal.) We break up the set of lines of $\mathfrak{L}'$ into the disjoint subsets $\mathfrak{L}_1$ consisting of those lines in the zero-set of p_1 and $\mathfrak{L}_2$ consisting of all the other lines in $\mathfrak{L}'$.

There are no more than $O(N^3)$ points of intersection between lines of $\mathfrak{L}_1$ and $\mathfrak{L}_2$ since each line of $\mathfrak{L}_2$ contains no more than $O\left(\frac{\alpha^{\frac{1}{2}}N}{Q^{\frac{1}{2}}}\right)$ points where p_1 is zero. Thus we are left with two (not mutually exclusive) cases which cover all possibilities. There are either $\frac{3QN^3}{10}$ points of intersection between lines of $\mathfrak{L}_1$ or there are $\frac{3QN^3}{10}$ points of intersection between lines of $\mathfrak{L}_2$. We will handle these separately.

Suppose there are $\frac{3QN^3}{10}$ intersections between lines of $\mathfrak{L}_1$. We factor $p_1 = p_3p_4$, where p_3 is plane-free and regulus-free and p_4 is a product of planes and reguli. We break $\mathfrak{L}_1$ into disjoint sets $\mathfrak{L}_3$ and $\mathfrak{L}_4$, with $\mathfrak{L}_3$ consisting of lines in the zero-set of p_3 and $\mathfrak{L}_4$ consisting of all other lines of $\mathfrak{L}_1$. As before there $O(N^3)$ points of intersection between lines of $\mathfrak{L}_3$ and $\mathfrak{L}_4$ since lines of $\mathfrak{L}_4$ are not in the zero-set of p_3. Moreover, there are at most $O(N^3)$ points of intersection between lines of $\mathfrak{L}_4$ because they lie in at most N planes and reguli each containing at most N lines. (We just see that each line has at most $O(N)$ intersections with planes and reguli it is not contained in and there are at most $O(N^2)$ points of intersection between lines internal to each plane and regulus.) However, there cannot be more than $O(N^3)$ points of intersection between lines of $\mathfrak{L}_3$ by applying the key Lemma 3.4. (Here we used that p_3 is plane-free and regulus-free.)

Thus we must be in the second case, where many of the points of intersection are between lines of $\mathfrak{L}_2$, all of which lie in the zero-set of p_2 which is totally unruled. Recall that p_2 is of degree $O\big(\frac{\alpha^{\frac{1}{2}}N}{Q^{\frac{1}{2}}}\big)$. Thus by Corollary 3.3, its zero-set contains no more than $O\big(\frac{\alpha N^2}{Q}\big)$ lines. We would like to now invoke the fact that the example we started with was optimal and reach a contradiction. But we cannot quite do that. Our set $\mathfrak{L}_2$ has βN^2 lines with $\beta = O\big(\frac{\alpha}{Q}\big)$, and we only know that there are no more than N lines in any plane or regulus, whereas we need to know that there are no more than $\sqrt{\beta}N$ lines. If this is the case, we are done. If not, we construct a subset $\mathfrak{L}_5$ as follows. If there is a plane or regulus containing more than $\sqrt{\beta}N$ lines of $\mathfrak{L}_2$, we put those lines in $\mathfrak{L}_5$ and remove them from $\mathfrak{L}_2$. We repeat as needed, labelling the remaining lines $\mathfrak{L}_6$. Since we removed $O(N)$ planes and reguli, there are $O(N^3)$ points of intersection between lines of $\mathfrak{L}_5$. Since no lines of $\mathfrak{L}_6$ belong to any plane or regulus of $\mathfrak{L}_5$, there are fewer than $O(N^3)$ points of intersection between lines of $\mathfrak{L}_5$ and $\mathfrak{L}_6$. Now we apply optimality of our original example to rule out more than $O\big(\frac{N^3}{Q^{\frac{1}{2}}}\big)$ points of intersection between lines of $\mathfrak{L}_6$. Thus we have reached a contradiction.

4. **Cell decompositions**

In this section, we construct a new type of cell decomposition of $\mathbf{R^n}$, where the walls of the cells are the zero-set of a polynomial. We use this type of cell decomposition to prove an incidence theorem for lines in $\mathbf{R^3}$ when not too many lines lie in a plane. The cell decomposition is described in the following theorem.

THEOREM 4.1. *If $\mathfrak{S}$ is a set of S points in $\mathbf{R^n}$ and $J \geq 1$ is an integer, then there is a polynomial surface Z of degree $d \lesssim 2^{J/n}$ with the following property. The complement $\mathbf{R^n} \setminus Z$ is the union of 2^J open cells O_i, and each cell contains $\leq 2^{-J}S$ points of $\mathfrak{S}$.*

Remark. Some or all of the points of $\mathfrak{S}$ may lie inside the surface Z. Recall that Z is not part of any of the open sets O_i. So there are two extreme cases in Theorem 4.1. In one extreme, all the points of $\mathfrak{S}$ lie in the open cells O_i, and there are exactly $2^{-J}S$ points in each cell. In the other extreme, all the points of $\mathfrak{S}$ lie in the surface Z. When the points all lie in Z, the theorem does not give any information about where in Z they lie.

The proof of Theorem 4.1 is based on the polynomial ham sandwich theorem of Stone and Tukey [ST42]. For context, we first recall the original ham sandwich theorem.

THEOREM 4.2 (Ham sandwich theorem). *If $U_1, \ldots, U_n \subset \mathbf{R}^\mathbf{n}$ are finite volume open sets, then there is a hyperplane that bisects each set U_i.*

The ham sandwich theorem was proven in the case $n = 3$ by Banach in the late 30's, using the Borsuk-Ulam theorem. In 1942, Stone and Tukey generalized Banach's proof to all dimensions. They also observed that the same argument applies to many other situations. In particular, they proved the following polynomial version of the ham sandwich theorem.

We say that an algebraic hypersurface $p(x_1, \ldots, x_n) = 0$ bisects a finite volume open set U if

$$\mathrm{Vol}(U \cap \{p < 0\}) = \mathrm{Vol}(U \cap \{p > 0\}) = (1/2)\mathrm{Vol}(U).$$

THEOREM 4.3 (Stone-Tukey, [ST42]). *For any degree $d \geq 1$, the following holds. Let $U_1, \ldots, U_M$ be any finite volume open sets in $\mathbf{R}^n$, with $M = \binom{n+d}{n} - 1$. Then there is a real algebraic hypersurface of degree at most d that bisects each U_i.*

(For a recent exposition of the proof, see [Gut10].)

We now adapt Theorem 4.3 to finite sets of points. Instead of open sets U_i, we will have finite sets S_i. We say that a polynomial p bisects a finite set S if at most half the points in S are in $\{p > 0\}$ and at most half the points in S are in $\{p < 0\}$. Note that p may vanish on some or all of the points of S.

COROLLARY 4.4. *Let $S_1, \ldots, S_M$ be finite sets of points in $\mathbf{R}^n$ with $M = \binom{n+d}{n} - 1$. Then there is a real algebraic hypersurface of degree at most d that bisects each S_i.*

Proof. For each $\delta > 0$, define $U_{i,\delta}$ to be the union of δ-balls centered at the points of S_i. By the polynomial ham sandwich theorem, Theorem 4.3, we can find a nonzero polynomial p_δ of degree $\leq d$ that bisects each set $U_{i,\delta}$.

We want to take a limit of the polynomials p_δ as $\delta \to 0$. To help make this work, we pick a norm $\|\|$ on the space of polynomials of degree $\leq d$. Any norm will do — to be definite, let $\|p\|$ denote the maximal absolute value of the coefficients of p. By scaling p_δ, we can assume that $\|p_\delta\| = 1$ for all δ. Now we can find a sequence $\delta_m \to 0$ so that p_{δ_m} converges in the space of degree $\leq d$ polynomials. We let p be the limit polynomial and observe that $\|p\| = 1$. In particular, p is not 0. Since the coefficients of p_{δ_m} converge to the coefficients of p, it is easy to check that p_δ converges to p uniformly on compact sets.

We claim that p bisects each set S_i. We prove the claim by contradiction. Suppose instead that $p > 0$ on more than half of the points of S_i. (The case $p < 0$ is similar.) Let $S_i^+ \subset S_i$ denote the set of points of S_i where $p > 0$. By choosing ε sufficiently small, we can assume that $p > \varepsilon$ on the ε-ball around each point of S_i^+. Also, we can choose ε small enough that the ε-balls around the points of S_i are disjoint. Since p_{δ_m} converges to p uniformly on compact sets, we can find m large enough that $p_{\delta_m} > 0$ on the ε-ball around each point of S_i^+. By making m large, we can also arrange that $\delta_m < \varepsilon$. Therefore, $p_{\delta_m} > 0$ on the δ_m-ball around each point of S_i^+. But then $p_{\delta_m} > 0$ on more than half of U_{i,δ_m}. This contradiction proves that p bisects S_i. □

Using this finite polynomial ham sandwich theorem, we can quickly prove Theorem 4.1.

Proof of Theorem 4.1. We do the construction in J steps. In the first step, we pick a linear polynomial p_1 that bisects $\mathfrak{S}$. We let $\mathfrak{S}^+$ and $\mathfrak{S}^-$ be the sets where p_1 is positive and negative, respectively. In the second step, we find a polynomial p_2 that bisects $\mathfrak{S}^+$ and $\mathfrak{S}^-$. And so on. At each new step, we use Corollary 4.4 to bisect the sets from the previous step.

We now describe the inductive procedure a little more precisely. At the end of step j, we have defined j polynomials $p_1, \dots, p_j$. We define 2^j subsets of $\mathfrak{S}$ by looking at the points where the polynomials $p_1, \dots, p_j$ have specified signs. Then we use Corollary 4.4 to bisect each of these 2^j sets. It follows by induction that each subset contains $\le 2^{-j} S$ points.

Finally, we let p be the product $p_1 \cdots p_J$, and we let Z denote the zero-set of p.

First we estimate the degree of p. By Corollary 4.4, the degree of p_j is $\lesssim 2^{j/n}$. Hence the degree of p is $d \lesssim \sum_{j=1}^{J} 2^{j/n} \lesssim 2^{J/n}$.

Now we define the 2^J open sets O_i as the sets where the polynomials $p_1, \dots, p_J$ have specified signs. For example, one of the sets O_i is defined by the inequalities $p_1(x) > 0, p_2(x) < 0, p_3(x) > 0, \dots, p_J(x) > 0$. The sets O_i are open and disjoint. Their union is exactly the complement of Z. As we saw above, the number of points in $\mathfrak{S} \cap O_i$ is at most $2^{-J} S$. □

Using this type of cell decomposition, we will prove an estimate for incidences of lines when not too many lines lie in a plane.

THEOREM 4.5. *Let $k \ge 3$. Let $\mathfrak{L}$ be a set of L lines in $\mathbf{R}^3$ with at most B lines in any plane. Let $\mathfrak{S}$ be the set of points in $\mathbf{R}^3$ intersecting at least k lines of $\mathfrak{L}$. Then the following inequality holds*:

$$|\mathfrak{S}| \le C[L^{3/2}k^{-2} + LBk^{-3} + Lk^{-1}].$$

Theorem 4.5 implies Theorem 2.11 by setting $L = N^2$ and $B = N$.

This theorem is sharp up to constant factors in a number of cases. These examples help to give a sense of the right-hand side.

Example 1. Choose L/k points. Let $\mathfrak{L}$ consist of k lines through each point. The set $\mathfrak{L}$ has a k-fold incidence at each of the L/k points. (We can also arrange that no three lines lie in a plane.)

Example 2. Choose L/B planes. Put B lines in each of the planes. The B lines in each plane can be arranged to create B^2k^{-3} k-fold incidences. (See the examples in [ST83].) This set of lines has a total of LBk^{-3} k-fold incidences.

Example 3. Let G_0 denote the integer lattice $\{(a,b,0)\}$ with $1 \le a, b \le L^{1/4}$. Let G_1 denote the integer lattice $\{(a,b,1)\}$ with $1 \le a, b \le L^{1/4}$. Let $\mathfrak{L}$ denote all the lines from a point of G_0 to a point of G_1. The horizontal planes $z = 0$ and $z = 1$ do not contain any lines of $\mathfrak{L}$. Any other plane contains at most $L^{1/4}$ points of each G_i, and so at most $L^{1/2}$ lines of $\mathfrak{L}$. We will prove in the appendix that there are $\sim L^{3/2}k^{-2}$ points that lie in $\ge k$ lines of $\mathfrak{L}$ for each k in the range $2 \le k \le L^{1/2}/400$.

For context, we should compare Theorem 4.5 to the Szemerédi-Trotter theorem, which holds in all dimensions as we now recall.

THEOREM 4.6. *If $\mathfrak{L}$ is a set of L lines in $\mathbf{R^n}$, and $\mathfrak{S}$ denotes the set of points lying in at least k lines of $\mathfrak{L}$, then*

$$|\mathfrak{S}| \lesssim L^2k^{-3} + Lk^{-1}.$$

The higher-dimensional case follows easily from the two-dimensional case by taking a generic projection from $\mathbf{R^n}$ to $\mathbf{R^2}$. The set of lines $\mathfrak{L}$ will project to L distinct lines in $\mathbf{R^2}$, and the points of $\mathfrak{S}$ project to distinct points in $\mathbf{R^2}$.

Theorem 4.5 is a refinement of Theorem 4.6. When $B = L$, Theorem 4.5 is Theorem 4.6. Theorem 4.5 tells us how much we can improve the Szemerédi-Trotter theorem if we know in addition that not too many lines lie in a plane.

We will use Theorem 4.6 in our proof. Recently, in [KMS12], Kaplan, Matoušek, and Sharir gave a new proof of the Szemerédi-Trotter theorem using polynomial cell decompositions.

Now we turn to the proof of Theorem 4.5. An important special case is the uniform case where each point has $\sim k$ lines through it and each line contains about the same number of points. We will first prove the theorem under some uniformity hypotheses.

PROPOSITION 4.7. *Let $k \ge 3$. Let $\mathfrak{L}$ be a set of L lines in $\mathbf{R^3}$ with at most B lines in any plane. Let $\mathfrak{S}$ be a set of S points in $\mathbf{R^3}$ so that each point intersects between k and $2k$ lines of $\mathfrak{L}$.*

Also, we assume that there are $\ge \frac{1}{100}L$ lines in $\mathfrak{L}$ that each contain $\ge \frac{1}{100}SkL^{-1}$ points of $\mathfrak{S}$. Then $S \le C[L^{3/2}k^{-2} + LBk^{-3} + Lk^{-1}]$.

The second paragraph of Proposition 4.7 is a uniformity assumption about the lines. Note that there are $\sim Sk$ total incidences between lines of $\mathfrak{L}$ and points of $\mathfrak{S}$. Therefore, an average line of $\mathfrak{L}$ contains $\sim SkL^{-1}$ points of $\mathfrak{S}$. We assume here that there are many lines that are about average. Proposition 4.7 is the main part of the proof of Theorem 4.5. The general case reduces to this special case by easy inductive arguments.

Proof. We begin by outlining our strategy. We suppose that

$$S \geq AL^{3/2}k^{-2} + Lk^{-1}. \tag{4.1}$$

In this equation, A represents a large constant that we will choose below. Assuming 4.1, we need to show that many lines of $\mathfrak{L}$ lie in a plane. In particular, we will find a plane that contains $\gtrsim SL^{-1}k^3$ lines of $\mathfrak{L}$. This means that $B \gtrsim SL^{-1}k^3$, and hence $S \lesssim BLk^{-3}$, and we will be done.

Let us outline how we find the plane. First we prove that a definite fraction of the lines of $\mathfrak{L}$ lie in an algebraic surface Z of degree $\lesssim L^2S^{-1}k^{-3}$. Second we prove that this variety Z contains some plannes and that a definite fraction of the lines of $\mathfrak{L}$ lie in the planes. Since there are at most d planes, one plane must contain $\gtrsim L/d$ lines. Because $d \lesssim L^2S^{-1}k^{-3}$, this plane contains $\gtrsim SL^{-1}k^3$ lines, which is what we wanted to prove.

Our bound for the degree d is sharp up to a constant factor because of Example 2 above. In this example, the lines $\mathfrak{L}$ lie in $\sim L^2S^{-1}k^{-3}$ planes. Since the planes can be taken in general position, the lines $\mathfrak{L}$ do not lie in an algebraic surface of lower degree.

(Our bound for the degree d is the new ingredient in this section. We will find the algebraic surface Z by using the polynomial cell decomposition of Theorem 4.1. We initially tried to find Z by using the purely algebraic degree reduction argument from [GK10], as in Section 3. With this method, we proved that a definite fraction of the lines of Z lie in an algebraic surface of degree $L^2S^{-1}k^{-2}$. But this degree is too large to make our argument work.)

Now we begin the detailed proof of Proposition 4.7. First we prove that almost all points of $\mathfrak{S}$ lie in a surface Z with controlled degree. This lemma is the most important step in the proof of Theorem 4.5.

Lemma 4.8. *If the constant A in inequality 4.1 is sufficiently large, then there is an algebraic surface Z of degree $\lesssim L^2S^{-1}k^{-3}$ that contains at least $(1-10^{-8})S$ points of $\mathfrak{S}$.*

Proof. We let θ denote a large constant that we will choose later, and we let d be the greatest integer less than $\theta L^2S^{-1}k^{-3}$. This d will be the degree of our surface Z. First we check that $d \geq 1$. By the Szemerédi-Trotter theorem, $S \lesssim L^2k^{-3} + Lk^{-1}$. But by inequality (4.1), $S \geq Lk^{-1}$. Therefore, $S \lesssim L^2k^{-3}$. Hence we can choose θ so that $d \geq 1$.

Now we apply Theorem 4.1 to construct a degree d surface Z such that $\mathbf{R^3} \setminus Z$ is a union of $\sim d^3$ open cells O_i, each containing $\lesssim Sd^{-3}$ points of $\mathfrak{S}$.

Let us suppose that Z contains $< (1-10^{-8})S$ points of $\mathfrak{S}$. So the open cells O_i all together contain $\geq 10^{-8}S$ points of $\mathfrak{S}$. Since each cell contains $\lesssim Sd^{-3}$ points of $\mathfrak{S}$, there must be $\gtrsim d^3$ cells that each contain $\gtrsim Sd^{-3}$ points of $\mathfrak{S}$. We call these full cells.

We now prove an upper bound for S using the cellular method from [CEG$^+$90].

We let $\mathfrak{L}(O_i)$ denote the subset of lines of $\mathfrak{L}$ that intersect O_i. We let L_{cell} be the minimum of $|\mathfrak{L}(O_i)|$ among all the full cells O_i. We apply the Szemerédi-Trotter inequality to the full cell with the fewest lines. Since this full cell still contains $\gtrsim Sd^{-3}$ points, we get the following inequality:

$$Sd^{-3} \lesssim L_{\text{cell}}^2 k^{-3} + L_{\text{cell}} k^{-1}.$$

Next we estimate L_{cell} in terms of the degree of Z. A line either lies in Z or else it intersects Z at most d times. Every time a line moves from one open cell O_i to another, it needs to pass through Z. So each line of $\mathfrak{L}$ intersects at most $d+1$ cells O_i. So there are $\leq L(d+1)$ pairs (l, O_i) where $l \in \mathfrak{L}(O_i)$. But there are $\sim d^3$ full cells O_i. Hence $L_{\text{cell}} \lesssim Ld^{-2}$. Plugging in this estimate for L_{cell}, we get the following inequality:

$$Sd^{-3} \lesssim L^2 d^{-4} k^{-3} + Ld^{-2} k^{-1}.$$

Recalling that $d \sim \theta L^2 S^{-1} k^{-3}$ and rearranging, we get the following inequality:

$$S \leq C(\theta^{-1} S + \theta L^3 S^{-1} k^{-4}).$$

Note that the constant C does not depend on θ. (We could work it out explicitly using an explicit constant in Theorem 4.1 and in the Szemerédi-Trotter theorem.) At this point, we choose θ sufficiently large so that $C\theta^{-1} < 1/2$. We can then move the term $C\theta^{-1} S$ to the left-hand side and rearrange to get the inequality

$$S \lesssim \theta^{1/2} L^{3/2} k^{-2}.$$

If the constant A is sufficiently large, this inequality contradicts (4.1). We conclude that there are less than $10^{-8} S$ points of $\mathfrak{S}$ outside of Z.

Finally, the degree of Z is $d \leq \theta L^2 S^{-1} k^{-3}$. The constant θ is a particular number that we chose above. In particular, θ does not depend on A. And so $d \lesssim L^2 S^{-1} k^{-3}$ as desired. □

We let $\mathfrak{S}_Z$ denote the points of $\mathfrak{S}$ that lie in Z. By Lemma 4.8, $|\mathfrak{S} \setminus \mathfrak{S}_Z| \leq 10^{-8} S$. Our next goal is to prove that many lines of $\mathfrak{L}$ lie in the surface Z. This result depends on a quick calculation about the degree d. Recall that an average line of $\mathfrak{L}$ contains SkL^{-1} points of $\mathfrak{S}$. We prove that the degree d is much smaller than SkL^{-1}.

Lemma 4.9. *If the constant A is sufficiently large, then*

$$d < 10^{-8} SkL^{-1}.$$

Proof. Inequality (4.1) can be rewritten as

$$1 \leq A^{-1} S L^{-3/2} k^2.$$

Squaring this, we see that

$$d \leq dA^{-2}S^2L^{-3}k^4 \lesssim A^{-2}SkL^{-1}.$$

Now choosing A sufficiently large finishes the proof. $\square$

As an immediate corollary, we get the following lemma.

LEMMA 4.10. *If l is a line of $\mathfrak{L}$ that contains at least $10^{-8}SkL^{-1}$ points of $\mathfrak{S}_Z$, then l is contained in Z.*

Proof. The line l contains at least $10^{-8}SkL^{-1}$ points of Z. Since $d > 10^{-8}SkL^{-1}$, the line l must lie in the surface Z. $\square$

Let $\mathfrak{L}_Z$ denote the set of lines in $\mathfrak{L}$ that are contained in Z.

LEMMA 4.11. *The set $\mathfrak{L}_Z$ contains at least $(1/200)L$ lines.*

Proof. We assumed that there are $\geq (1/100)L$ lines of $\mathfrak{L}$ that each contain $\geq (1/100)SkL^{-1}$ points of $\mathfrak{S}$. Let $\mathfrak{L}_0 \subset \mathfrak{L}$ be the set of these lines. We claim that most of these lines lie in $\mathfrak{L}_Z$. Suppose that a line l lies in $\mathfrak{L}_0 \setminus \mathfrak{L}_Z$. It must contain at least $(1/100)SkL^{-1}$ points of $\mathfrak{S}$. But by Lemma 4.10, it contains $< 10^{-8}SkL^{-1}$ points of $\mathfrak{S}_Z$. Therefore, it must contain at least $(1/200)SkL^{-1}$ points of $\mathfrak{S} \setminus \mathfrak{S}_Z$. This gives us the following inequality:

$$(1/200)SkL^{-1}|\mathfrak{L}_0 \setminus \mathfrak{L}_Z| \leq I(\mathfrak{S} \setminus \mathfrak{S}_Z, \mathfrak{L}_0 \setminus \mathfrak{L}_Z).$$

Here we write I to abbreviate the number of incidences between a set of points and a set of lines.

On the other hand, each point of $\mathfrak{S}$ lies in at most $2k$ lines of $\mathfrak{L}$, giving us an upper bound on incidences:

$$I(\mathfrak{S} \setminus \mathfrak{S}_Z, \mathfrak{L}_0 \setminus \mathfrak{L}_Z) \leq 2k|\mathfrak{S} \setminus \mathfrak{S}_Z| \leq 2 \cdot 10^{-8}Sk.$$

Comparing these two inequalities, we see that $|\mathfrak{L}_0 \setminus \mathfrak{L}_Z| \leq 4 \cdot 10^{-6}L$, which implies that $|\mathfrak{L}_Z| \geq (1/200)L$. $\square$

We have now carried out the first step of our outline: we found a surface Z of degree $\lesssim L^2S^{-1}k^{-3}$ that contains a definite fraction of the lines from $\mathfrak{L}$.

We now turn to the second step of our outline. We will prove that Z contains some planes and that these planes contain many lines of $\mathfrak{L}$. This step is closely based on the techniques in [GK10] and [EKS11]. The paper [EKS11] contains a clear introduction to the techniques. In particular, Section 2 of [EKS11] proves all of the fundamental lemmas from algebraic geometry that we need.

Each point of $\mathfrak{S}_Z$ lies in at least k lines of $\mathfrak{L}$. But such a point does not necessarily lie in any lines of $\mathfrak{L}_Z$. Therefore, we make the following definition.

We define $\mathfrak{S}'_Z$ to be the set of points in $\mathfrak{S}_Z$ that lie in at least three lines of $\mathfrak{L}_Z$.

This subset is important because each point of $\mathfrak{S}'_Z$ is a special point of the surface Z: either a critical point or a flat point. Let us recall the definitions of critical points and flat points.

The surface Z is the vanishing set of a polynomial p. The polynomial p can be factored into irreducible polynomials $p = p_1 p_2 \cdots$. We assume that each irreducible factor of p appears only once. Now a point $x \in Z$ is called critical if the gradient ∇p vanishes at x. If $x \in Z$ is not critical, we say that x is regular. In a small neighborhood of a regular point $x \in Z$, Z is a smooth submanifold. We say that a regular point $x \in Z$ is flat if the second fundamental form of Z vanishes at x.

LEMMA 4.12. *Each point of $\mathfrak{S}'_Z$ is either a critical point or a flat point of Z.*

Proof. Let $x \in \mathfrak{S}'_Z$. By definition, x lies in three lines that all lie in Z. If x is a critical point of Z, we are done. If x is a regular point of Z, then all three lines must lie in the tangent space of Z at x. In particular, the three lines are coplanar. Let v_1, v_2, v_3 be nonzero tangent vectors of the three lines at x. The second fundamental form of Z vanishes in each of these three directions. Since the second fundamental form is a symmetric bilinear form on the two-dimensional tangent space, it must vanish. Therefore, x is a flat point of Z. □

(See also [EKS11, Props. 4, 6] for a more detailed proof.) Lemma 4.12 shows that the points of $\mathfrak{S}'_Z$ are important. Next we show that almost every point of $\mathfrak{S}$ lies in $\mathfrak{S}'_Z$.

LEMMA 4.13. *The set $\mathfrak{S} \setminus \mathfrak{S}'_Z$ contains at most $10^{-7} S$ points.*

Proof. Lemma 4.8 tells us that $|\mathfrak{S} \setminus \mathfrak{S}_Z| < 10^{-8} S$. Suppose x is a point in $\mathfrak{S}_Z \setminus \mathfrak{S}'_Z$. The point x lies in at least k lines from $\mathfrak{L}$, but it lies in at most two lines from $\mathfrak{L}_Z$. So x lies in $\geq k-2$ lines of $\mathfrak{L} \setminus \mathfrak{L}_Z$:

$$(k-2)|\mathfrak{S}_Z \setminus \mathfrak{S}'_Z| \leq I(\mathfrak{S}_Z \setminus \mathfrak{S}'_Z, \mathfrak{L} \setminus \mathfrak{L}_Z).$$

On the other hand, we showed in Lemma 4.10 that each line of $\mathfrak{L} \setminus \mathfrak{L}_Z$ contains $\leq 10^{-8} S k L^{-1}$ points of $\mathfrak{S}_Z$. Therefore,

$$I(\mathfrak{S}_Z \setminus \mathfrak{S}'_Z, \mathfrak{L} \setminus \mathfrak{L}_Z) \leq I(\mathfrak{S}_Z, \mathfrak{L} \setminus \mathfrak{L}_Z) \leq (10^{-8} S k L^{-1}) L.$$

Combining these inequalities and recalling that $k \geq 3$, we see that

$$|\mathfrak{S}_Z \setminus \mathfrak{S}'_Z| \leq 10^{-8} \frac{k}{k-2} S \leq 3 \cdot 10^{-8} S. \qquad \square$$

We let $\mathfrak{S}_{\text{crit}} \subset \mathfrak{S}'_Z$ denote the critical points in $\mathfrak{S}'_Z$ and we let $\mathfrak{S}_{\text{flat}} \subset \mathfrak{S}'_Z$ denote the flat points of $\mathfrak{S}'_Z$. We call a line $l \subset Z$ a critical line of Z if every point of l is a critical point of Z. We call a line $l \subset Z$ a flat line if it is not a critical line and every regular point in l is flat.

Our next goal is to show that Z contains many flat lines, which is a step to showing that Z contains a plane. In order to do this, we show that the flat points of Z are defined by the vanishing of certain polynomials.

LEMMA 4.14. *Let x be a regular point of Z. Then x is flat if and only if the following three polynomial vectors vanish at x:*

$$\nabla_{e_j \times \nabla p} \nabla p \times \nabla p, j = 1, 2, 3.$$

Here, e_j are the coordinate vectors of $\mathbf{R^3}$, and $\times$ denotes the cross product of vectors. Each vector above has three components, so we have a total of nine polynomials. Each polynomial has degree $\leq 3d$. For more explanation, see Section 3 of [GK10] or Section 2 of [EKS11]. In [EKS11], they use a more efficient set of polynomials: only three polynomials.

To find critical or flat lines, we use the following simple lemmas.

LEMMA 4.15. *Suppose that a line l contains more than d critical points of Z. Then l is a critical line of Z.*

Proof. At each critical point of Z, the polynomial p and all the components of ∇p vanish. Since p has degree d, we conclude that p vanishes on every point of l. Since ∇p has degree $d - 1$, we conclude that ∇p vanishes on every point of l. Hence l is a critical line of Z. □

LEMMA 4.16. *Suppose that a line l contains more than $3d$ flat points of Z. Then l is a flat line of Z.*

Proof. Let $x_1, \ldots, x_{3d+1}$ be flat points of Z contained in l. By Lemma 4.14, each polynomial $\nabla_{e_j \times \nabla p} \nabla p \times \nabla p$ vanishes at x_i. Since the degree of these polynomials is $\leq 3d$, we conclude that each of these polynomials vanishes on l. Similarly, p vanishes on l. Therefore, the line l lies in Z and every regular point in l is a flat point. But by definition, x_i are regular points of Z. Therefore, l is not a critical line, and it must be a flat line. □

Using these lemmas, we will prove that a definite fraction of the lines of $\mathfrak{L}$ are either critical or flat. We define $\mathfrak{L}'_Z$ to be the set of lines of $\mathfrak{L}_Z$ that contain at least $(1/200)SkL^{-1}$ points of $\mathfrak{S}'_Z$.

LEMMA 4.17. *Each line in $\mathfrak{L}'_Z$ is either critical or flat.*

Proof. Since every point of $\mathfrak{S}'_Z$ is either critical or flat, each line in $\mathfrak{L}'_Z$ contains either $(1/400)SkL^{-1}$ critical points or $(1/400)SkL^{-1}$ flat points. But by Lemma 4.9, $d \leq 10^{-8}SkL^{-1}$. So by Lemmas 4.15 and 4.16, each line of $\mathfrak{L}'_Z$ is either critical or flat. □

Now we show that $\mathfrak{L}'_Z$ contains a definite fraction of the lines of $\mathfrak{L}$.

LEMMA 4.18. *The number of lines in $\mathfrak{L}'_Z$ is $\geq (1/200)L$.*

Proof. Recall that we assumed in the statement of Proposition 4.7 that there are at least $(1/100)L$ lines of $\mathfrak{L}$ that each contain $\geq (1/100)SkL^{-1}$ points of $\mathfrak{S}$. We denote these lines by $\mathfrak{L}_0 \subset \mathfrak{L}$.

Suppose a line l lies in $\mathfrak{L}_0 \setminus \mathfrak{L}'_Z$. Then l contains at least $(1/100)SkL^{-1}$ points of $\mathfrak{S}$. But it contains less than $(1/200)SkL^{-1}$ points of $\mathfrak{S}'_Z$. Therefore, it contains at least $(1/200)SkL^{-1}$ points of $\mathfrak{S} \setminus \mathfrak{S}'_Z$. So we get the following inequality:

$$(1/200)SkL^{-1}|\mathfrak{L}_0 \setminus \mathfrak{L}'_Z| \leq I(\mathfrak{S} \setminus \mathfrak{S}'_Z, \mathfrak{L}_0 \setminus \mathfrak{L}'_Z).$$

But since each point of $\mathfrak{S}$ lies in at most $2k$ lines of $\mathfrak{L}$,

$$I(\mathfrak{S} \setminus \mathfrak{S}'_Z, \mathfrak{L}_0 \setminus \mathfrak{L}'_Z) \leq I(\mathfrak{S} \setminus \mathfrak{S}'_Z, \mathfrak{L}) \leq 2k|\mathfrak{S} \setminus \mathfrak{S}'_Z|.$$

Lemma 4.13 says that $|\mathfrak{S} \setminus \mathfrak{S}'_Z| \leq 10^{-7}S$. Assembling all our inequalities, we see that

$$(1/200)SkL^{-1}|\mathfrak{L}_0 \setminus \mathfrak{L}'_Z| \leq 2k \cdot 10^{-7}S.$$

Simplifying this expression, we see that

$$|\mathfrak{L}_0 \setminus \mathfrak{L}'_Z| \leq 4 \cdot 10^{-5}L.$$

So almost all the lines of $\mathfrak{L}_0$ lie in $\mathfrak{L}'_Z$. In particular, $\mathfrak{L}'_Z$ contains $\geq (1/200)L$ lines. □

Next we bound the number of critical lines in Z.

LEMMA 4.19. *A surface Z of degree d contains $\leq d^2$ critical lines.*

This lemma follows from Bezout's theorem applied to p and ∇p. See Proposition 3 in [EKS11].

If the constant A from inequality (4.1) is sufficiently large, then d^2 will be much less than L. We record this calculation in the next lemma.

LEMMA 4.20. *If A is sufficiently large, then $d \leq 10^{-4}L^{1/2}$.*

Proof. Inequality (4.1) implies that $1 \leq A^{-1}SL^{-3/2}k^2$. Therefore,

$$d \leq dA^{-1}SL^{-3/2}k^2 \lesssim A^{-1}L^{1/2}k^{-1}.$$

Choosing A sufficiently large finishes the proof. □

In particular, we see that Z contains at most $d^2 < 10^{-8}L$ critical lines. Since $\mathfrak{L}'_Z$ contains at least $(1/200)L$ lines, we see that most of these lines must be flat. In particular, $\mathfrak{L}'_Z$ contains at least $(1/300)L$ flat lines of Z.

We are trying to prove that Z contains some planes. Let Z_{pl} denote the union of all planes contained in Z. We let $\tilde{Z}$ denote the rest of Z so that $Z = Z_{\mathrm{pl}} \cup \tilde{Z}$. In terms of polynomials, Z is the vanishing set of p. The polynomial p factors into irreducibles: $p = p_1 p_2 \cdots$. Some of these factors have degree 1, and some factors have degree more than 1. Each factor of degree 1 defines a plane, and Z_{pl} is the union of these planes. The product of the remaining factors is a polynomial $\tilde{p}$, and $\tilde{Z}$ is the zero-set of $\tilde{p}$. A line that lies in both Z_{pl} and $\tilde{Z}$ is actually a critical line of Z. So a flat line of Z lies either in Z_{pl} or in $\tilde{Z}$, but not both. A flat line of Z that lies in $\tilde{Z}$ is a flat line of $\tilde{Z}$. The number of flat lines in a surface of degree $\leq d$ is bounded by the following lemma from [EKS11].

Lemma 4.21 ([EKS11, Prop. 8]). *If Z is an algebraic surface of degree $\leq d$ with no planar component, then Z contains $\leq 3d^2$ flat lines.*

We have seen that $\mathfrak{L}$ contains at least $(1/300)L$ flat lines of Z. But $\tilde{Z}$ contains only $3 \cdot 10^{-8}L$ flat lines. The rest of the flat lines lie in Z_{pl}. In particular, $\mathfrak{L}$ contains at least $(1/400)L$ lines in Z_{pl}.

Finally, we observe that the number of planes in Z_{pl} is $\leq d \lesssim L^2S^{-1}k^{-3}$. So one of these planes must contain $\gtrsim Sk^3L^{-1}$ lines of $\mathfrak{L}$. In other words, $B \gtrsim Sk^3L^{-1}$.

At several points in the argument, we needed A to be sufficiently large. We now choose A large enough for those steps. We conclude that either $S \leq AL^2k^{-3/2}+Lk^{-1}$ or else $S \lesssim LBk^{-3}$. This finishes the proof of Proposition 4.7. $\square$

Proposition 4.7 is the heart of the proof of Theorem 4.5. We are going to reduce the general case to Proposition 4.7. First we remove the assumption that many lines have roughly the average number of points.

Proposition 4.22. *Let $k \geq 3$. Let $\mathfrak{L}$ be a set of L lines in $\mathbf{R}^3$ with $\leq B$ lines in any plane. Let $\mathfrak{S}$ be a set of S points so that each point meets between k and $2k$ lines of $\mathfrak{L}$.*

Then $S \leq C[L^{3/2}k^{-2} + LBk^{-3} + Lk^{-1}]$.

Proof. Let $\mathfrak{L}_1$ be the subset of lines in $\mathfrak{L}$ that contain $\geq (1/100)SkL^{-1}$ points of $\mathfrak{S}$. If $|\mathfrak{L}_1| \geq (1/100)L$, then we have all the hypotheses of Proposition 4.7, and we may conclude

$$S \leq C_0[L^{3/2}k^{-2} + LBk^{-3} + Lk^{-1}].$$

We are going to prove that S obeys this same estimate, with the same constant, regardless of the size of $\mathfrak{L}_1$. The proof will go by induction on the number of lines.

From now on we assume that $|\mathfrak{L}_1| \leq (1/100)L$. The lines in $\mathfrak{L}_1$ contribute most of the incidences. In particular, we have the following inequality:

$$I(\mathfrak{S}, \mathfrak{L} \setminus \mathfrak{L}_1) \leq (1/100)SkL^{-1} \cdot L = (1/100)Sk.$$

We define $\mathfrak{S}' \subset \mathfrak{S}$ to be the set of points with $\geq (9/10)k$ incidences with lines of $\mathfrak{L}_1$. If x is in $\mathfrak{S} \setminus \mathfrak{S}'$, then x lies in at least k lines of $\mathfrak{L}$, but less than $(9/10)k$ lines of $\mathfrak{L}_1$. So x lies in at least $(1/10)k$ lines of $\mathfrak{L} \setminus \mathfrak{L}_1$. Therefore,

$$(1/10)k|\mathfrak{S} \setminus \mathfrak{S}'| \leq I(\mathfrak{S} \setminus \mathfrak{S}', \mathfrak{L} \setminus \mathfrak{L}_1) \leq I(\mathfrak{S}, \mathfrak{L} \setminus \mathfrak{L}_1) \leq (1/100)Sk.$$

Rearranging, we see that $|\mathfrak{S} \setminus \mathfrak{S}'| \leq (1/10)S$, and so $|\mathfrak{S}'| \geq (9/10)S$.

A point of $\mathfrak{S}'$ has at least $(9/10)k$ incidences with $\mathfrak{L}_1$ and at most $2k$ incidences with $\mathfrak{L}_1$. This is a slightly larger range than we have considered before. In order to do induction, we need to reduce the range. We observe $\mathfrak{S}' = \mathfrak{S}'_+ \cup \mathfrak{S}'_-$, where $\mathfrak{S}'_+$ consists of points with $\geq k$ incidences to $\mathfrak{L}_1$ and $\mathfrak{S}'_-$ consists of points with $\leq k$ incidences with $\mathfrak{L}_1$. We define $\mathfrak{S}_1$ to be the larger of $\mathfrak{S}'_+$ and $\mathfrak{S}'_-$. It has $\geq (9/20)S$ points in it.

If we picked $\mathfrak{S}_1 = \mathfrak{S}'_+$, then we define $k_1 = k$. If we picked $\mathfrak{S}_1 = \mathfrak{S}'_-$, then we define k_1 to be the smallest integer $\geq (9/10)k$. Each point in $\mathfrak{S}_1$ has at least k_1 and at most $2k_1$ incidences with lines of $\mathfrak{L}_1$. Also, k_1 is an integer $\geq (9/10)k \geq 27/10$, so $k_1 \geq 3$.

The set of lines $\mathfrak{L}_1$ and the set of points $\mathfrak{S}_1$ obey all the hypotheses of Theorem 4.22 (using k_1 in place of k and using the same B). There are fewer lines in $\mathfrak{L}_1$ than in $\mathfrak{L}$. Doing induction on the number of lines, we may assume that our result holds for these sets. If we denote $|\mathfrak{L}_1| = L_1$ and $|\mathfrak{S}_1| = S_1$, we get

$$S_1 \leq C_0[L_1^{3/2}k_1^{-2} + BL_1k_1^{-3} + L_1k_1^{-1}].$$

Now $S \leq (20/9)S_1$. Also, $L_1 \leq (1/100)L$. And $k_1 \geq (9/10)k$. Therefore,

$$S \leq (20/9)S_1 \leq [(20/9)(1/100)(10/9)^3]C_0[L^{3/2}k^{-2} + LBk^{-3} + Lk^{-1}].$$

The bracketed product of fractions is < 1, and so S obeys the desired bound. $\square$

Finally, we can prove Theorem 4.5.

Proof of Theorem 4.5. Let $k \geq 3$. Suppose that $\mathfrak{L}$ is a set of L lines with $\leq B$ in any plane. Suppose that $\mathfrak{S}$ is a set of points, each intersecting at least k lines of $\mathfrak{L}$.

We subdivide the points $\mathfrak{S} = \cup_{j=0}^{\infty}\mathfrak{S}_j$, where $\mathfrak{S}_j$ consists of the points incident to at least $2^j k$ lines and at most $2^{j+1}k$ lines. We define k_j to be $2^j k$. Then Theorem 4.22 applies to $(\mathfrak{L}, \mathfrak{S}_j, k_j, B)$, and we conclude that

$$\begin{aligned}|\mathfrak{S}_j| &\leq C_0[L^{3/2}k_j^{-2} + LBk_j^{-3} + Lk_j^{-1}] \\ &\leq 2^{-j}C_0[L^{3/2}k^{-2} + LBk^{-3} + Lk^{-1}].\end{aligned}$$

Now $S \leq \sum_j |\mathfrak{S}_j| \leq 2C_0[L^{3/2}k^{-2} + LBk^{-3} + Lk^{-1}]$. $\square$

Appendix A. The example of a square grid

In this section, we return to Erdős's example of a square grid of points. When P is a square grid of N points, we show that $|Q(P)| \gtrsim N^3 \log N$ and $|G_k(P)| \gtrsim N^3 k^{-2}$ for all $2 \leq k \leq N/2000$. So the estimates in Propositions 2.2 and 2.5 are sharp up to constant factors. We also study the set of lines $\mathfrak{L}$ associated to a square grid P. This set of lines shows that many of our incidence estimates are sharp up to constant factors.

Let $S \geq 1$ be an integer. Let P be the grid of points (x, y) where x and y are integers with norm $\leq 2S$. Note that the number of points in $|P|$ is $N = (4S+1)^2$. Let $\mathfrak{L}$ be the set of lines in $\mathbf{R}^3$ associated to the set P, as in Section 2.

LEMMA A.1. *If a, b, c, and d are positive integers with norm $\leq S$, then the line from $(a, b, 0)$ to $(c, d, 1)$ is contained in $\mathfrak{L}$.*

Proof. Using the parametrization in (2.3), we see that the line from $(a, b, 0)$ to $(c, d, 1)$ is the line L_{pq}, where p and q are defined by the following equations:

$$\begin{aligned} (1/2)(p_x + q_x) &= a, & (1/2)(p_y + q_y) &= b, \\ (1/2)(q_y - p_y) &= c - a, & (1/2)(p_x - q_x) &= d - b. \end{aligned}$$

Solving these equations, we get $p = (a+d-b, b-c+a)$ and $q = (a-d+b, b+c-a)$. Since a, b, c, and d are positive integers of norm $\leq S$, it follows that p_x, p_y, q_x, and q_y are integers of norm $\leq 2S$, and so p and q lie in P. $\square$

Let $\mathfrak{L}_0 \subset \mathfrak{L}$ be the set of lines from $(a, b, 0)$ to $(c, d, 1)$ where a, b, c, and d are positive integers with norm $\leq S$. In the proposition below, we study the incidences of $\mathfrak{L}_0$. Note that $|\mathfrak{L}_0| = S^4$.

PROPOSITION A.2. *Let $\mathfrak{S}_k$ be the set of points in $\mathbf{R}^3$ that lie in at least k lines of $\mathfrak{L}_0$. For any k in the range $2 \leq k \leq (1/400)S^2$, $|\mathfrak{S}_k| \gtrsim S^6 k^{-2}$.*

Proof. Consider a point x in $\mathbf{R}^3$ contained in the slab $0 < x_3 < 1$. We define a map $F_x : \mathbf{R}^2 \to \mathbf{R}^2$ by saying that $F_x(a, b) = (c, d)$ if the line from $(a, b, 0)$ through x hits $(c, d, 1)$. We define G to be the integral grid in the plane given by $\{(a, b)\}$ with $1 \leq a, b \leq S$. The number of lines from $\mathfrak{L}_0$ that pass through x is exactly the cardinality of $F_x(G) \cap G$. Now any intersection of two lines from $\mathfrak{L}_0$ will have rational coordinates, so we can assume the coordinates of x are rational. Let us say that the x_3 coordinate of x is p/q, written in lowest terms.

By a similar triangles argument, $F_x(G)$ is a square grid with spacing $\frac{q-p}{p}$. Since p and q are in lowest terms, the intersection $F_x(G) \cap G$ will be a rectangular grid with spacing $q - p$. The edges of this rectangle will have length $< S$. So the number of points in $F_x(G) \cap G$ is at most $S^2(q-p)^{-2}$. On the other hand, the edges of this rectangle have length $< S\frac{q-p}{p}$. Therefore, the number of points of $F_x(G) \cap G$ is at most $S^2 p^{-2}$. Combining these estimates, we see that $|F_x(G) \cap G| \leq 4S^2 q^{-2}$.

Let us say that the middle half of G, written $G_{\text{middle}} \subset G$, is the integral grid $\{(a, b)\}$ with $(1/4)S \leq a, b \leq (3/4)S$. If F_x maps a vertex from G_{middle} into G, then the number of intersections between $F_x(G)$ and G is fairly close to this upper bound. Using the arguments from the last paragraph, it is straightforward to show that $|F_x(G) \cap G| \geq (1/100)S^2q^{-2}$ whenever $F_x(G_{\text{middle}})$ intersects G.

Let us define $X(p,q)$ to be the set of $x = (x_1, x_2, p/q)$ so that $F_x(G_{\text{middle}}) \cap G$ is nonempty. The set $X(p,q)$ lies in $\mathfrak{S}_k$ whenever $k \leq (1/100)S^2q^{-2}$. Equivalently, $X(p,q)$ lies in $\mathfrak{S}_k$ whenever $q \leq (1/10)Sk^{-1/2}$.

For any pair of points $(a_1,b_1) \in G_{\text{middle}}$ and $(a_2,b_2) \in G$, there is a unique $x \in X(p,q)$ so that $F_x(a_1,b_1) = (a_2,b_2)$. There are $\sim S^4$ such pairs of points. Each element of $X(p,q)$ corresponds to at least $(1/100)S^2q^{-2}$ pairs of points and at most $4S^2q^{-2}$ pairs of points. Therefore, $|X(p,q)| \sim S^2q^2$.

Now we fix $k \leq (1/400)S^2$. We pick q in the range $(1/20)Sk^{-1/2} \leq q \leq (1/10)Sk^{-1/2}$. Because k is not too big, this range of q contains some integers. For each p coprime to q, $X(p,q) \subset \mathfrak{S}_k$. The sets $X(p,q)$ are clearly disjoint, and so

$$|\mathfrak{S}_k| \gtrsim \sum_{q=(1/20)Sk^{-1/2}}^{(1/10)Sk^{-1/2}} \sum_{0<p<q, gcd(p,q)=1} |X(p,q)| \gtrsim \sum_{q=(1/20)Sk^{-1/2}}^{(1/10)Sk^{-1/2}} \phi(q)S^2q^2.$$

The sums of the Euler totient function $\phi(n)$ are well studied. Theorem 3.7 in [Apo76] gives the asymptotic $\sum_{q=1}^{x}\phi(q) = \frac{3}{\pi^2}x^2 + O(x\log x)$. Therefore, $\sum_{q=x}^{2x}\phi(q) \sim x^2$. Therefore,

$$|\mathfrak{S}_k| \gtrsim \left(Sk^{-1/2}\right)^2 S^2q^2 \sim S^6k^{-2}. \qquad \square$$

Recall that $|G_k(P)|$ is at least $|G'_k(P)|$, which is the number of points lying in at least k lines of $\mathfrak{L}$. So we see that $|G_k(P)| \gtrsim S^6k^{-2} \sim N^3k^{-2}$ for all $2 \leq k \leq (1/400)S^2 \leq N/2000$. Equation (2.2) gives

$$|Q(P)| \sim \sum_{k=2}^{N} k|G_k(P)| \gtrsim \sum_{k=2}^{N/2000} N^3k^{-1} \gtrsim N^3\log N.$$

Now we consider how sharp our incidence theorems are. The set of lines $\mathfrak{L}_0 \subset \mathfrak{L}$ has $\lesssim N \sim S^2$ lines in any plane or doubly ruled surface by Proposition 2.8. This example shows that Theorems 2.10 and 2.11 are sharp up to constant factors.

Next we consider Theorem 4.5. The lines $\mathfrak{L}_0$ correspond to Example 3 in Section 4. The three examples show that Theorem 4.5 is sharp up to constant factors as long as $B \gtrsim L^{1/2}$. The example $\mathfrak{L}_0$ has $B \sim L^{1/2}$. For much smaller values of B, we do not know what happens. For example, suppose that $\mathfrak{L}$ is a set of L lines in $\mathbf{R}^3$ with at most 100 lines in any plane. How many points can be incident to three lines of $\mathfrak{L}$? Or suppose that $\mathfrak{L}$ is a set of L lines in $\mathbf{R}^3$ with at most 100 lines in any plane or doubly ruled surface. How many points can be incident to two lines of $\mathfrak{L}$?

References

[Apo76] T. M. Apostol, *Introduction to Analytic Number Theory*, *Undergraduate Texts in Mathematics*, Springer-Verlag, New York, 1976. MR 0434929. Zbl 0335.10001. http://dx.doi.org/10.1007/978-1-4757-5579-4.

[CEGea92] B. Chazelle, H. Edelsbrunner, L. J. Guibas, and et al., Counting and cutting cycles of lines and rods in space, *Comput. Geom.* **1** (1992), 305–323. MR 1172840. Zbl 0748.68082. http://dx.doi.org/10.1016/0925-7721(92)90009-H.

[CST92] F. R. K. Chung, E. Szemerédi, and W. T. Trotter, The number of different distances determined by a set of points in the Euclidean plane, *Discrete Comput. Geom.* **7** (1992), 1–11. MR 1134448. Zbl 0755.52005. http://dx.doi.org/10.1007/BF02187820.

[CEG+90] K. L. Clarkson, H. Edelsbrunner, L. J. Guibas, M. Sharir, and E. Welzl, Combinatorial complexity bounds for arrangements of curves and spheres, *Discrete Comput. Geom.* **5** (1990), 99–160. MR 1032370. Zbl 0704.51003. http://dx.doi.org/10.1007/BF02187783.

[Dvi09] Z. Dvir, On the size of Kakeya sets in finite fields, *J. Amer. Math. Soc.* **22** (2009), 1093–1097. MR 2525780. Zbl 1202.52021. http://dx.doi.org/10.1090/S0894-0347-08-00607-3.

[EKS11] G. Elekes, H. Kaplan, and M. Sharir, On lines, joints, and incidences in three dimensions, *J. Combin. Theory Ser. A* **118** (2011), 962–977. MR 2763049. Zbl 1232.05041. http://dx.doi.org/10.1016/j.jcta.2010.11.008.

[ES11] G. Elekes and M. Sharir, Incidences in three dimensions and distinct distances in the plane, *Combin. Probab. Comput.* **20** (2011), 571–608. MR 2805398. Zbl 1222.52016. http://dx.doi.org/10.1017/S0963548311000137.

[Erd46] P. Erdös, On sets of distances of n points, *Amer. Math. Monthly* **53** (1946), 248–250. MR 0015796. Zbl 0060.34805. http://dx.doi.org/10.2307/2305092.

[FS05] S. Feldman and M. Sharir, An improved bound for joints in arrangements of lines in space, *Discrete Comput. Geom.* **33** (2005), 307–320. MR 2121298. Zbl 1068.52032. http://dx.doi.org/10.1007/s00454-004-1093-7.

[GIS11] J. Garibaldi, A. Iosevich, and S. Senger, *The Erdős Distance Problem*, *Student Math. Library* **56**, Amer. Math. Soc., 2011. MR 2721878. Zbl 1234.00002.

[Gut10] L. Guth, The endpoint case of the Bennett-Carbery-Tao multilinear Kakeya conjecture, *Acta Math.* **205** (2010), 263–286. MR 2746348. Zbl 1210.52004. http://dx.doi.org/10.1007/s11511-010-0055-6.

[GK10] L. Guth and N. H. Katz, Algebraic methods in discrete analogs of the Kakeya problem, *Adv. Math.* **225** (2010), 2828–2839. MR 2680185. Zbl 1202.52022. http://dx.doi.org/10.1016/j.aim.2010.05.015.

[KMS12] H. Kaplan, J. Matoušek, and M. Sharir, Simple proofs of classical theorems in discrete geometry via the Guth-Katz polynomial partitioning technique, *Discrete Comput. Geom.* **48** (2012), 499–517. MR 2957631. Zbl 1259.52008. http://dx.doi.org/10.1007/s00454-012-9443-3.

[KSS10] H. Kaplan, M. Sharir, and E. Shustin, On lines and joints, *Discrete Comput. Geom.* **44** (2010), 838–843. MR 2728035. Zbl 1250.52012. http://dx.doi.org/10.1007/s00454-010-9246-3.

[KT04] N. H. Katz and G. Tardos, A new entropy inequality for the Erdős distance problem, in *Towards a Theory of Geometric Graphs*, *Contemp. Math.* **342**, Amer. Math. Soc., Providence, RI, 2004, pp. 119–126. MR 2065258. Zbl 1069.52017. http://dx.doi.org/10.1090/conm/342/06136.

[Lan99] J. M. Landsberg, Is a linear space contained in a submanifold? On the number of derivatives needed to tell, *J. Reine Angew. Math.* **508** (1999), 53–60. MR 1676869. Zbl 1067.14514.

[Mos52] L. Moser, On the different distances determined by n points, *Amer. Math. Monthly* **59** (1952), 85–91. MR 0046663. Zbl 0046.14101. http://dx.doi.org/10.2307/2307105.

[Qui10] R. Quilodrán, The joints problem in $\mathbb{R}^n$, *SIAM J. Discrete Math.* **23** (2009/10), 2211–2213. MR 2594983. Zbl 1228.05106. http://dx.doi.org/10.1137/090763160.

[Sal58] G. Salmon, *A Treatise on the Analytic Geometry of Three Dimensions*, edited by C. H. Rowe, Chelsea Publishing Company, New York, 1958, revised by R. A. P. Rogers., 7th ed., Vol. 1. MR 0094753. JFM 55.0996.05.

[Sch98] W. Schlag, A geometric inequality with applications to the Kakeya problem in three dimensions, *Geom. Funct. Anal.* **8** (1998), 606–625. MR 1631271. Zbl 0939.42012. http://dx.doi.org/10.1007/s000390050068.

[Seg43] B. Segre, The maximum number of lines lying on a quartic surface, *Quart. J. Math., Oxford Ser.* **14** (1943), 86–96. MR 0010431. Zbl 0063.06860.

[ST01] J. Solymosi and C. D. Tóth, Distinct distances in the plane, *Discrete Comput. Geom.* **25** (2001), 629–634. MR 1838423. Zbl 0988.52027. http://dx.doi.org/10.1007/s00454-001-0009-z.

[ST42] A. H. Stone and J. W. Tukey, Generalized "sandwich" theorems, *Duke Math. J.* **9** (1942), 356–359. MR 0007036. Zbl 0061.38405. http://dx.doi.org/10.1215/S0012-7094-42-00925-6.

[Szé97] L. A. Székely, Crossing numbers and hard Erdős problems in discrete geometry, *Combin. Probab. Comput.* **6** (1997), 353–358. MR 1464571. Zbl 0890.05022. http://dx.doi.org/10.1017/S0963548397002976.

[ST83] E. Szemerédi and W. T. Trotter, Jr., Extremal problems in discrete geometry, *Combinatorica* **3** (1983), 381–392. MR 0729791. Zbl 0541.05012. http://dx.doi.org/10.1007/BF02579194.

[Tar03] G. Tardos, On distinct sums and distinct distances, *Adv. Math.* **180** (2003), 275–289. MR 2019225. Zbl 1039.52014. http://dx.doi.org/10.1016/S0001-8708(03)00004-5.

(Received: November 18, 2010)
(Revised: July 15, 2014)

Massachusetts Institute of Technology, Cambridge, MA
E-mail: larry.guth.work@gmail.com

California Institute of Technology, Pasadena, CA
E-mail: nets@caltech.edu