Discrete Comput Geom (2015) 53:428–444
DOI 10.1007/s00454-014-9648-8

Distinct Distance Estimates and Low Degree Polynomial Partitioning

Larry Guth

Received: 17 April 2014 / Revised: 6 November 2014 / Accepted: 11 November 2014 /
Published online: 24 December 2014

Abstract We give a shorter proof of a slightly weaker version of a theorem from Guth and Katz (Ann Math 181:155–190, 2015): we prove that if $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$ with at most $L^{1/2}$ lines in any low degree algebraic surface, then the number of r-rich points of $\mathfrak{L}$ is $\lesssim L^{(3/2)+\varepsilon} r^{-2}$. This result is one of the main ingredients in the proof of the distinct distance estimate in Guth and Katz (2015). With our slightly weaker theorem, we get a slightly weaker distinct distance estimate: any set of N points in $\mathbb{R}^2$ determines at least $c_\varepsilon N^{1-\varepsilon}$ distinct distances.

Keywords Incidence geometry · Distinct distances · Polynomial method · Combinatorics

In [4], Erdős asked how few distinct distances may be determined by a set of N points in the plane. He conjectured that a square grid of points is near-optimal, giving a conjectural lower bound of $cN(\log N)^{-1/2}$. Quite recently, in [2], Elekes and Sharir suggested a new approach to this problem, connecting it to the incidence geometry of curves in 3-dimensional space. This approach was carried out by Katz and the author in [6], proving that any set of N points determines $\geq cN(\log N)^{-1}$ distinct distances. In this paper, we give a variation of the most difficult step of the proof. We will prove a slightly weaker result, but using a shorter argument.

The main work in [6] is an estimate about lines in $\mathbb{R}^3$. If $\mathfrak{L}$ is a set of lines in $\mathbb{R}^3$, then a point x is called r-rich if it lies in at least r lines of $\mathfrak{L}$. We write $P_r(\mathfrak{L})$ for the set of r-rich points of $\mathfrak{L}$.

Theorem 0.1 [6, Theorem 1.2] *If $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$ with at most $L^{1/2}$ lines in any plane or regulus, and if $2 \leq r \leq L^{1/2}$, then $|P_r(\mathfrak{L})| \leq CL^{3/2}r^{-2}$.*

L. Guth
MIT, Cambridge, MA, USA
e-mail: lguth@math.mit.edu

The distinct distance estimate follows from combining the approach of Elekes and Sharir with this bound. The proof of Theorem 0.1 is somewhat involved. There are different arguments for the cases $r = 2$ and $r \geq 3$ and each argument is pretty long. The case $r \geq 3$ uses the idea of polynomial partitioning, which will also be central to this paper. The case $r = 2$ uses the theory of ruled surfaces. We will prove a slightly weaker result using only polynomial partitioning.

Theorem 0.2 *For any $\varepsilon > 0$, there are $D(\varepsilon)$, $K(\varepsilon)$ so that the following holds.*

If $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$, and there are less than $L^{(1/2)+\varepsilon}$ lines of $\mathfrak{L}$ in any irreducible algebraic surface of degree at most D, and if $2 \leq r \leq 2L^{1/2}$, then

$$|P_r(\mathfrak{L})| \leq K L^{(3/2)+\varepsilon} r^{-2}.$$

Using Theorem 0.2 in place of Theorem 0.1 in the arguments of [6], one gets the following slightly weaker distinct distance estimate.

Theorem 0.3 *For any $\varepsilon > 0$, there is a constant $c_\varepsilon > 0$ so that any set of N points in the plane determines at least $c_\varepsilon N^{1-\varepsilon}$ distinct distances.*

Polynomial partitioning is one of the main new ideas in [6], and it will also be the key tool in our proof. We recall the statement of the partitioning theorem.

Theorem 0.4 [6, Theorem 4.1] *For each dimension n and each degree $D \geq 1$, the following holds. For any finite set $S \subset \mathbb{R}^n$, we can find a non-zero polynomial P of degree at most D so that $\mathbb{R}^n \setminus Z(P)$ is a union of disjoint open sets O_i, and for each of these sets,*

$$|S \cap O_i| \leq C_n D^{-n} |S|.$$

This polynomial partitioning result is a corollary of the Stone–Tukey ham sandwich theorem [12]. Polynomial partitioning is useful in divide and conquer arguments. The set S is divided into a part in each cell O_i plus a part in a lower-dimensional surface $Z(P)$. In a divide and conquer argument, we estimate each of these contributions separately and then add up the results.

Kaplan et al. wrote a paper [7] on the polynomial partitioning technique. They give a good exposition of the topic. They show how to use polynomial partitioning to give new proofs of some classical results in incidence geometry, such as the Szemerédi–Trotter theorem. They also discuss how polynomial partitioning compares with other partitioning methods, such as the cutting method (see [7, Sect. 2.3]).

The arguments of [6] use polynomial partitioning with degree D equal to a power of L. This gives good bounds on what happens in the cells O_i, but it also makes $Z(P)$ rather complicated. In [11], Solymosi and Tao gave a modification of this argument using partitioning with degree D equal to a large constant, and using induction to control what happens in each cell. In [10], Sharir and Solomon further developed this method, proving estimates for lines in $\mathbb{R}^4$. We will use this low degree partitioning method to prove Theorem 0.2.

Here is the main new issue that arises in the proof of Theorem 0.2. Recall that we use a low degree polynomial to partition $\mathbb{R}^3$ into cells O_i, and we plan to use induction to study the behavior of the lines entering each cell. Let $\mathfrak{L}_i$ denote the lines of $\mathfrak{L}$ that intersect the cell O_i. By hypothesis, we know that $\mathfrak{L}$ contains less than $|\mathfrak{L}|^{(1/2)+\varepsilon}$ lines in any low degree surface. Since $\mathfrak{L}_i \subset \mathfrak{L}$, $\mathfrak{L}_i$ contains less than $|\mathfrak{L}|^{(1/2)+\varepsilon}$ lines in any low degree surface. But that doesn't mean that $\mathfrak{L}_i$ contains less than $|\mathfrak{L}_i|^{(1/2)+\varepsilon}$ lines in any low degree surface. Therefore, we cannot immediately apply induction to $\mathfrak{L}_i$. At first sight, the inductive argument doesn't look like it will close. The main new ingredient in this paper is a way to organize the low degree surfaces containing many lines. By keeping track of their contribution, we can make the induction close.

1 Background and Notation

Our proof is based on polynomial partitioning. Here we restate the partitioning theorem with an extra condition bounding the number of cells O_i.

Theorem 1.1 *For each dimension n and each degree $D \geq 1$, the following holds. For any finite set $S \subset \mathbb{R}^n$, we can find a non-zero polynomial P of degree at most D so that $\mathbb{R}^n \setminus Z(P)$ is a union of disjoint open sets O_i obeying the following:*

- *For each i, $|S \cap O_i| \leq C_n D^{-n} |S|$.*
- *The number of open sets O_i is at most $C_n D^n$.*

Proof The first claim is [6, Theorem 4.1]. So we just need to prove the second claim.

The number of connected components of the complement $\mathbb{R}^n \setminus Z(P)$ is at most $C_n D^n$, by estimates proven independently by Oleinik–Petrovsky [9], Milnor [8], and Thom [15]. A short proof was also given by Solymosi and Tao, as [11, Theorem A.1]. This implies the second claim.

However, we don't need to appeal to these results. The statement of the theorem does not require that each open set O_i is connected. By Theorem 4.1 of [6], we can write $\mathbb{R}^n \setminus Z(P)$ as a union of open sets U_j with $|S \cap U_j| \leq C_n D^{-n} |S|$. We can then define each O_i to be a union of some of the U_j so that each O_i contains $\leq C_n D^{-n} |S|$ points of S and the number of sets O_i is at most $C_n D^n$. □

We will also need a version of the Bézout theorem. The simplest version of the Bézout theorem is the following.

Theorem 1.2 *If P, Q are non-zero polynomials in $\mathbb{R}[x_1, x_2]$ with no common factor, then $Z(P) \cap Z(Q) \subset \mathbb{R}^2$ contains at most* (Deg P)(Deg Q) *points.*

We need a version of this theorem for polynomials in three variables where we count the number of lines in $Z(P) \cap Z(Q)$.

Theorem 1.3 *If P, Q are non-zero polynomials in $\mathbb{R}[x_1, x_2, x_3]$ with no common factor, then $Z(P) \cap Z(Q)$ contains at most* (Deg P)(Deg Q) *lines.*

Proofs of these classical results appear in [5]. They are Corollaries 2.3 and 2.4. See also Sect. 2 of [3] for a proof of Theorem 1.3 and a review of related material. A more general version of Theorem 1.2 can be found in van der Waerden's book *Modern Algebra* [16, Vol. 2, p. 16].

We will also use the Szemerédi–Trotter theorem, which we record here in the following form:

Theorem 1.4 ([14]) *If $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^n$, then*

$$|P_r(\mathfrak{L})| \leq C\big(L^2 r^{-3} + L r^{-1}\big).$$

There are several nice proofs of the Szemerédi–Trotter theorem that have appeared since the original article. In [1], Clarkson et al. gave a proof using the method of cuttings. In [13], Székely gave a proof using the crossing number lemma. In [7], Kaplan et al. gave a proof using the polynomial partitioning theorem. Their proof is closely related to the ideas in this paper.

We end with a note on constants. We will use C to denote a constant that may change from line to line. If we want to label a particular constant to refer to later, we will call it C_1, C_2, etc.

2 A Stronger Result for Inductive Purposes

We will prove Theorem 0.2 by induction. To make the induction work, we prove a slightly stronger result. The stronger result says that for any set of lines $\mathfrak{L}$ in $\mathbb{R}^3$, there is a small set of low degree surfaces that account for all but $\sim L^{(3/2)+\varepsilon} r^{-2}$ of the r-rich points of $\mathfrak{L}$.

To state our theorem we need a piece of notation. If $\mathfrak{L}$ is a set of lines and Z is an algebraic surface, we define $\mathfrak{L}_Z \subset \mathfrak{L}$ to be the set of lines of $\mathfrak{L}$ that lie in Z.

Theorem 2.1 *For any $\varepsilon > 0$, there are $D(\varepsilon)$, and $K(\varepsilon)$ so that the following holds. For any $r \geq 2$, let $r' = \lceil (9/10) r \rceil$, the least integer which is at least $(9/10)r$.*

If $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$, and if $2 \leq r \leq 2L^{1/2}$, then there is a set $\mathcal{Z}$ of algebraic surfaces so that

- *Each surface $Z \in \mathcal{Z}$ is an irreducible surface of degree at most D.*
- *Each surface $Z \in \mathcal{Z}$ contains at least $L^{(1/2)+\varepsilon}$ lines of $\mathfrak{L}$.*
- $|\mathcal{Z}| \leq 2L^{(1/2)-\varepsilon}$.
- $|P_r(\mathfrak{L}) \setminus \bigcup_{Z \in \mathcal{Z}} P_{r'}(\mathfrak{L}_Z)| \leq K L^{(3/2)+\varepsilon} r^{-2}$.

Theorem 2.1 implies Theorem 0.2. If there are less than $L^{(1/2)+\varepsilon}$ lines of $\mathfrak{L}$ in any irreducible algebraic surface of degree at most D, then the set $\mathcal{Z}$ must be empty, and so Theorem 2.1 implies that $|P_r(\mathfrak{L})| \leq K L^{(3/2)+\varepsilon} r^{-2}$.

In our theorems above, we always assumed that $r \leq 2L^{1/2}$. Studying r-rich points for $r > 2L^{1/2}$ is much simpler. We recall the following elementary estimate, which will also be useful in our proof.

Proposition 2.2 *If $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^d$ for $d \geq 2$, and if $r > 2L^{1/2}$, then $|P_r(\mathfrak{L})| \leq 2Lr^{-1}$.*

We include the well-known proof here, because it is a model for a different proof below.

Proof Let $P_r(\mathfrak{L})$ be $\{x_1, x_2, \ldots, x_M\}$, with $M = |P_r(\mathfrak{L})|$. Now x_1 lies in at least r lines of $\mathfrak{L}$. The point x_2 lies in at least $(r-1)$ lines of $\mathfrak{L}$ that did not contain x_1. More generally, the point x_j lies in at least $r-(j-1)$ lines of $\mathfrak{L}$ that did not contain any of the previous points $x_1,\ldots,x_{j-1}$. Therefore, we have the following inequality for the total number of lines:

$$L \geq \sum_{j=1}^{M} \max(r-j,0).$$

If $M \geq r/2$, then we would get $L \geq (r/2)(r/2) = r^2/4$. But by hypothesis, $r > 2L^{1/2}$, giving a contradiction. Therefore, $M < r/2$, and we get $L \geq M(r/2)$ which proves the proposition. □

3 Proof of Theorem 2.1

Here is an outline of our proof. We will use induction on the number of lines in $\mathfrak{L}$.

First, we use a low degree polynomial partitioning argument to cut $\mathbb{R}^3$ into cells O_i. For each cell, we use induction to study the lines of $\mathfrak{L}$ that enter that cell. For each cell, we get a set of surfaces $\mathcal{Z}_i$ that accounts for all but a few of the r-rich points in O_i. Combining these surfaces with the polynomial partitioning surface, we will get a large set of surfaces $\tilde{\mathcal{Z}}$ with the following properties:

- Each surface $Z \in \tilde{\mathcal{Z}}$ is an irreducible algebraic surface of degree at most D.
- $|\tilde{\mathcal{Z}}| \leq \mathrm{Poly}(D) L^{(1/2)-\varepsilon} \log L$.
- $|P_r(\mathfrak{L}) \setminus \bigcup_{Z\in\tilde{\mathcal{Z}}} P_{r'}(\mathfrak{L}_Z)| \leq (1/100) K L^{(3/2)+\varepsilon} r^{-2}$.

(We write $A \leq \mathrm{Poly}(D) B$ to mean that is an exponent p and a constant C so that $A \leq C D^p B$.)

This set of surfaces $\tilde{\mathcal{Z}}$ does not close the induction. There are too many surfaces in $\tilde{\mathcal{Z}}$, and we don't know that each surface contains $L^{(1/2)+\varepsilon}$ lines of $\mathfrak{L}$. The second step is to prune $\tilde{\mathcal{Z}}$. We will define

$$\mathcal{Z} := \{Z \in \tilde{\mathcal{Z}} \mid Z \text{ contains at least } L^{(1/2)+\varepsilon} \text{ lines of } \mathfrak{L}\}.$$

Then we will check that $\mathcal{Z}$ satisfies the conclusions of the theorem. First, we will prove that $|\mathcal{Z}| \leq 2L^{(1/2)-\varepsilon}$. This follows from a simple counting argument, similar to the proof of Proposition 2.2 above. Second, we will check that the surfaces in $\tilde{\mathcal{Z}} \setminus \mathcal{Z}$ did not contribute too much to controlling the r-rich points of $\mathfrak{L}$. More precisely we will prove that

$$\sum_{Z\in\tilde{\mathcal{Z}}\setminus\mathcal{Z}} |P_{r'}(\mathfrak{L}_Z)| \leq (1/100) K L^{(3/2)+\varepsilon} r^{-2}.$$

To prove this bound, we use Szemerédi–Trotter to bound the size of $P_{r'}(\mathfrak{L}_Z)$ in terms of $|\mathfrak{L}_Z|$ for each surface $Z \in \tilde{\mathcal{Z}} \setminus \mathcal{Z}$, and we use a simple counting argument to control how many surfaces Z have large $|\mathfrak{L}_Z|$. This finishes our outline. Now we begin the proof of Theorem 2.1.

We remark that if $\varepsilon \geq 1/2$ then the theorem is trivial: we can take $\mathcal{Z}$ to be empty, and it is easy to check that $|P_r(\mathfrak{L})| \leq 2L^2r^{-2}$. (This follows from Szemerédi–Trotter, which gives a stronger estimate. But it also follows from a simple double-counting argument.) So we can assume that $\varepsilon \leq 1/2$.

We start by discussing how to choose $D = D(\varepsilon)$ and $K = K(\varepsilon)$. We will choose D a large constant depending on ε and then we will choose K a large constant depending on ε and D. As long as these are large enough at certain points in the proof, the argument goes through. For example, we will choose K large enough that

$$K \geq 10(2D)^{2/\varepsilon}. \tag{1}$$

The proof is by induction on L. We start by checking the base of the induction. Because of (1), we claim the theorem holds when $L^\varepsilon \leq 2D$. Suppose that $\mathfrak{L}$ is a set of L lines with $L^\varepsilon \leq 2D$, and that $2 \leq r \leq 2L^{1/2}$. We choose $\mathcal{Z}$ to be the empty set. Using (1), we see that

$$|P_r(\mathfrak{L})| \leq L^2 \leq (2D)^{2/\varepsilon} \leq K/10 \leq KL^{(3/2)+\varepsilon}r^{-2}.$$

We have now established the base of the induction. By the inductive hypothesis, we can assume that the theorem holds for sets of at most $L/2$ lines.

3.1 Building $\tilde{\mathcal{Z}}$.

Let S be any subset of $P_r(\mathfrak{L})$. An important case is $S = P_r(\mathfrak{L})$, but we will have to consider other sets as well. We use Theorem 1.1 to do a polynomial partitioning of the set S with a polynomial of degree at most D. The polynomial partitioning theorem, Theorem 1.1, says that there is a non-zero polynomial P of degree at most D so that

- $\mathbb{R}^3 \setminus Z(P)$ is the union of at most CD^3 disjoint open cells O_i, and
- for each cell O_i, $|S \cap O_i| \leq CD^{-3}|S|$.

We define $\mathfrak{L}_i \subset \mathfrak{L}$ to be the set of lines from $\mathfrak{L}$ that intersect the open cell O_i. We note that $S \cap O_i \subset P_r(\mathfrak{L}_i)$. If a line does not lie in $Z(P)$, then it can have at most D intersection points with $Z(P)$, which means that it can enter at most $D+1$ cells O_i. So each line of $\mathfrak{L}$ intersects at most $D+1$ cells O_i. Therefore, we get the following inequality:

$$\sum_i |\mathfrak{L}_i| \leq (D+1)L \leq 2DL. \tag{2}$$

Let $\beta > 0$ be a large parameter that we will choose below. We say that a cell O_i is β-good if

$$|\mathfrak{L}_i| \leq \beta D^{-2}L. \tag{3}$$

The number of β-bad cells is at most $2\beta^{-1}D^3$. Each cell contains at most $CD^{-3}|S|$ points of S. Therefore, the bad cells all together contain at most $C\beta^{-1}|S|$ points of S. We now choose β so that $C\beta^{-1} \leq (1/100)$. β is an absolute constant, independent of ε. We now have the following estimate:

$$\text{The union of the bad cells contains at most } (1/100)|S| \text{ points of } S. \tag{4}$$

For each good cell O_i, we apply induction to understand $\mathfrak{L}_i$. By choosing D sufficiently large, we can guarantee that for each good cell, $|\mathfrak{L}_i| \leq (1/2)L$. Now there are two cases, depending on whether $r \leq 2|\mathfrak{L}_i|^{1/2}$.

If $r \leq 2|\mathfrak{L}_i|^{1/2}$, then we can apply the inductive hypothesis. In this case, we see that there is a set $\mathcal{Z}_i$ of irreducible algebraic surfaces of degree at most D with the following two properties:

$$|\mathcal{Z}_i| \leq 2|\mathfrak{L}_i|^{(1/2)-\varepsilon} \leq 2(\beta D^{-2}L)^{(1/2)-\varepsilon}. \tag{5}$$

Because $S \cap O_i \subset P_r(\mathfrak{L}_i)$, we also get

$$\begin{aligned}|(S \cap O_i)\backslash \bigcup_{Z \in \mathcal{Z}_i} P_{r'}(\mathfrak{L}_Z)| &\leq K|\mathfrak{L}_i|^{(3/2)+\varepsilon}r^{-2} \leq K(\beta D^{-2}L)^{(3/2)+\varepsilon}r^{-2}\\ &\leq C_1 K D^{-3-2\varepsilon}L^{(3/2)+\varepsilon}r^{-2}. \end{aligned}\tag{6}$$

On the other hand, if $r > 2|\mathfrak{L}_i|^{1/2}$, then we define $\mathcal{Z}_i$ to be empty, and Proposition 2.2 gives the bound

$$|S \cap O_i| \leq |P_r(\mathfrak{L}_i)| \leq 2|\mathfrak{L}_i|r^{-1} \leq 2Lr^{-1} \leq 4L^{3/2}r^{-2}. \tag{7}$$

By choosing K sufficiently large compared to D, we can arrange that $4L^{3/2}r^{-2} \leq C_1 K D^{-3-2\varepsilon}L^{(3/2)+\varepsilon}r^{-2}$. Therefore, inequality 6 holds for the good cells: with $r > 2|\mathfrak{L}_i|^{1/2}$ as well as the good cells with $r \leq 2|\mathfrak{L}_i|^{1/2}$. We sum this inequality over all the good cells:

$$\begin{aligned}\sum_{O_i \text{ good}} |(S \cap O_i)\backslash \bigcup_{Z \in \mathcal{Z}_i} P_{r'}(\mathfrak{L}_Z)| &\leq CD^3 \cdot C_1 K D^{-3-2\varepsilon}L^{(3/2)+\varepsilon}r^{-2}\\ &\leq C_2 D^{-2\varepsilon} K L^{(3/2)+\varepsilon}r^{-2}.\end{aligned}$$

We choose $D(\varepsilon)$ large enough so that $C_2 D^{-2\varepsilon} \leq (1/400)$. Therefore, we get the following:

$$\sum_{O_i \text{ good}} |(S \cap O_i)\backslash \bigcup_{Z \in \mathcal{Z}_i} P_{r'}(\mathfrak{L}_Z)| \leq (1/400)KL^{(3/2)+\varepsilon}r^{-2}. \tag{8}$$

We have studied the points of S in the good cells. Next we study the points of S in the zero set of the partioning polynomial $Z(P)$. Let Z_j be an irreducible component of $Z(P)$. If $x \in S \cap Z_j$, but $x \notin P_{r'}(\mathfrak{L}_{Z_j})$, then x must be contained in at least $r/10$ lines of $\mathfrak{L}\backslash\mathfrak{L}_{Z_j}$. Each line of $\mathfrak{L}$ that is not contained in Z_j has at most $\text{Deg}(Z_j)$ intersection points with Z_j. Therefore

$$|(S \cap Z_j)\backslash P_{r'}(\mathfrak{L}_{Z_j})| \leq 10r^{-1}(\text{Deg } Z_j)L.$$

If $\{Z_j\}$ are all the irreducible components of $Z(P)$, then we see that

$$|(S \cap Z(P))\backslash \bigcup_j P_{r'}(\mathfrak{L}_{Z_j})| \leq 10r^{-1}DL.$$

We choose $K = K(\varepsilon, D)$ sufficiently large so that $10D \leq (1/800)K$. Since $r \leq 2L^{1/2}$, we have

$$|(S \cap Z(P)) \setminus \bigcup_j P_{r'}(\mathfrak{L}_{Z_j})| \leq (1/800)KLr^{-1} \leq (1/400)KL^{3/2}r^{-2}. \tag{9}$$

Now we define $\tilde{\mathcal{Z}}_S$ to be the union of $\mathcal{Z}_i$ over all the good cells O_i together with all the irreducible components Z_j of $Z(P)$. Each surface in $\tilde{\mathcal{Z}}_S$ is an algebraic surface of degree at most D. By (5), we have the following estimate for $|\tilde{\mathcal{Z}}_S|$:

$$|\tilde{\mathcal{Z}}_S| \leq CD^3(\beta D^{-2}L)^{(1/2)-\varepsilon} + D \leq \text{Poly}(D)L^{(1/2)-\varepsilon}. \tag{10}$$

Summing the contribution of the bad cells in (4), the contribution of the good cells in (8), and the contribution of the cell walls in (9), we get

$$|S \setminus \bigcup_{Z \in \tilde{\mathcal{Z}}_S} P_{r'}(\mathfrak{L}_Z)| \leq (1/100)|S| + (1/200)KL^{(3/2)+\varepsilon}r^{-2}. \tag{11}$$

If we didn't have the $(1/100)|S|$ term coming from the bad cells, we could simply take $S = P_r(\mathfrak{L})$ and $\tilde{\mathcal{Z}} = \tilde{\mathcal{Z}}_S$. Because of this term, we need to run the above construction repeatedly.

Let $S_1 = P_r(\mathfrak{L})$, and let $\tilde{\mathcal{Z}}_{S_1}$ be the set of surfaces constructed above. Now we define $S_2 = S_1 \setminus \bigcup_{Z \in \tilde{\mathcal{Z}}_{S_1}} P_{r'}(\mathfrak{L}_Z)$. We iterate this procedure, defining

$$S_{j+1} := S_j \setminus \bigcup_{Z \in \tilde{\mathcal{Z}}_{S_j}} P_{r'}(\mathfrak{L}_Z).$$

Each set S_j is a subset of $P_r(\mathfrak{L})$. Each set of surfaces $\tilde{\mathcal{Z}}_{S_j}$ has cardinality at most $\text{Poly}(D)L^{(1/2)-\varepsilon}$. Iterating (11) we see

$$|S_{j+1}| \leq (1/100)|S_j| + (1/200)KL^{(3/2)+\varepsilon}r^{-2}. \tag{12}$$

We define $J = C \log L$ for a large constant C. Because of the iterative formula in (12), we get

$$|S_J| \leq (1/100)KL^{(3/2)+\varepsilon}r^{-2}. \tag{13}$$

We define $\tilde{\mathcal{Z}} = \bigcup_{j=1}^{J-1} \tilde{\mathcal{Z}}_{S_j}$. This set of surfaces has the following properties. Since each set $\tilde{\mathcal{Z}}_{S_j}$ has at most $\text{Poly}(D)L^{(1/2)-\varepsilon}$ surfaces, we get

$$|\tilde{\mathcal{Z}}| \leq \text{Poly}(D)L^{(1/2)-\varepsilon} \log L. \tag{14}$$

Also, $P_r(\mathfrak{L}) \setminus \bigcup_{Z \in \tilde{\mathcal{Z}}} P_{r'}(\mathfrak{L}_Z) = S_J$, and so (13) gives

$$|P_r(\mathfrak{L}) \setminus \bigcup_{Z \in \tilde{\mathcal{Z}}} P_{r'}(\mathfrak{L}_Z)| \leq (1/100)KL^{(3/2)+\varepsilon}r^{-2}. \tag{15}$$

This finishes our construction of $\tilde{\mathcal{Z}}$. Next we prune $\tilde{\mathcal{Z}}$ down to our desired set of surfaces $\mathcal{Z}$.

3.2 Pruning $\tilde{\mathcal{Z}}$

We define

$$\mathcal{Z} := \{Z \in \tilde{\mathcal{Z}} \mid Z \text{ contains at least } L^{(1/2)+\varepsilon} \text{ lines of } \mathfrak{L}\}.$$

To close our induction, we have to check two properties of $\mathcal{Z}$.

(1) $|\mathcal{Z}| \leq 2L^{(1/2)-\varepsilon}$.
(2) $|P_r(\mathfrak{L}) \setminus \bigcup_{Z\in\mathcal{Z}} P_{r'}(\mathfrak{L}_Z)| \leq KL^{(3/2)+\varepsilon}r^{-2}$.

We begin with a simple lemma about surfaces that each contain many lines.

Lemma 3.1 *Suppose $\mathfrak{L}$ is a set of lines in $\mathbb{R}^3$, and $\mathcal{Y}$ is a set of irreducible algebraic surfaces of degree at most D, and suppose that each surface $Z \in \mathcal{Y}$ contains at least A lines of $\mathfrak{L}$.*
If $A > 2D|\mathfrak{L}|^{1/2}$, then $|\mathcal{Y}| \leq 2|\mathfrak{L}|A^{-1}$.

Proof The proof of this lemma follows the same idea as the proof of Proposition 2.2. By the Bézout theorem for lines, Theorem 1.3, the intersection of any two surfaces $Z_1, Z_2 \in \mathcal{Y}$ contains at most D^2 lines of $\mathfrak{L}$.

We choose an ordering of the surfaces of $\mathcal{Y}$. We consider the surfaces one at a time in order and count the number of new lines.

Z_1 contains at least A lines of $\mathfrak{L}$. Z_2 contains at least $A - D^2$ lines of $\mathfrak{L}$ that are not in Z_1. Z_{j+1} contains at least $A - jD^2$ lines of $\mathfrak{L}$ that are not in the previous surfaces $Z_1, \ldots, Z_j$. Therefore, we get the following inequality:

$$|\mathfrak{L}| \geq \sum_{j=1}^{|\mathcal{Y}|} \max(A - jD^2, 0).$$

If $j \leq (1/2)AD^{-2}$, then $A - jD^2 \geq A/2$. Therefore, if $|\mathcal{Y}| \geq (1/2)AD^{-2}$, then we see that $|\mathfrak{L}| \geq (1/2)AD^{-2}(A/2)$. By hypothesis, we know $A > 2D|\mathfrak{L}|^{1/2}$, which gives the contradiction $|\mathfrak{L}| > |\mathfrak{L}|$. Therefore, $|\mathcal{Y}| \leq (1/2)AD^{-2}$. Now we see that $|\mathfrak{L}| \geq |\mathcal{Y}|(A/2)$, and this completes the proof of the lemma. □

We apply this lemma with $\mathcal{Y} = \mathcal{Z}$ and $A = L^{(1/2)+\varepsilon}$. We can assume that $L^\varepsilon > 2D$, because the case of $L^\varepsilon \leq 2D$ was the base of our induction, and we handled it by choosing K sufficiently large. Therefore, $A = L^{(1/2)+\varepsilon} > 2DL^{1/2}$, and the hypotheses of Lemma 3.1 are satisfied. The lemma tells us that $|\mathcal{Z}| \leq 2L^{(1/2)-\varepsilon}$, which proves item (1) above. Now we turn to item (2). We recall (15):

$$|P_r(\mathfrak{L}) \setminus \bigcup_{Z\in\tilde{\mathcal{Z}}} P_{r'}(\mathfrak{L}_Z)| \leq (1/100)KL^{(3/2)+\varepsilon}r^{-2}.$$

Therefore, it suffices to check that

$$\sum_{Z\in\tilde{\mathcal{Z}}\setminus\mathcal{Z}} |P_{r'}(\mathfrak{L}_Z)| \leq (1/100)KL^{(3/2)+\varepsilon}r^{-2}.$$

We sort $\tilde{\mathcal{Z}}\setminus\mathcal{Z}$ according to the number of lines in each surface. For each integer $s \geq 0$, we define

$$\tilde{\mathcal{Z}}_s := \{Z \in \tilde{\mathcal{Z}} \text{ so that } |\mathfrak{L}_Z| \in [2^s, 2^{s+1})\}.$$

Since each surface of $\tilde{\mathcal{Z}}$ with at least $L^{(1/2)+\varepsilon}$ lines of $\mathfrak{L}$ lies in $\mathcal{Z}$, we see that

$$\tilde{\mathcal{Z}}\setminus\mathcal{Z} \subset \bigcup_{2^s\leq L^{(1/2)+\varepsilon}} \tilde{\mathcal{Z}}_s. \tag{16}$$

For each $Z \in \tilde{\mathcal{Z}}_s$, $|\mathfrak{L}_Z| \leq 2^{s+1}$. We use the Szemerédi–Trotter theorem, Theorem 1.4, to bound $P_{r'}(\mathfrak{L}_Z)$. Since $r' \geq (9/10)r$, Szemerédi–Trotter gives

$$P_{r'}(\mathfrak{L}_Z) \leq C\big(2^{2s}r^{-3} + 2^s r^{-1}\big). \tag{17}$$

Using Lemma 3.1 with $A = 2^s$, we get the following estimate for $|\tilde{\mathcal{Z}}_s|$:

$$\text{If } 2^s > 2DL^{1/2}, \text{ then } |\tilde{\mathcal{Z}}_s| \leq 2L2^{-s}. \tag{18}$$

We can now estimate $\sum_{Z\in\tilde{\mathcal{Z}}\setminus\mathcal{Z}} |P_{r'}(\mathfrak{L}_Z)|$.

$$\begin{aligned}\sum_{Z\in\tilde{\mathcal{Z}}\setminus\mathcal{Z}} \big|P_{r'}(\mathfrak{L}_Z)\big| &\leq \sum_{2^s\leq L^{(1/2)+\varepsilon}} \Big(\sum_{Z\in\tilde{\mathcal{Z}}_s} |P_{r'}(\mathfrak{L}_Z)|\Big) \\ &\leq C \sum_{2^s\leq L^{(1/2)+\varepsilon}} \Big|\tilde{\mathcal{Z}}_s\Big| \big(2^{2s}r^{-3} + 2^s r^{-1}\big). \end{aligned} \tag{19}$$

We consider the contribution to the last sum from s in the range $2DL^{1/2} < 2^s \leq L^{(1/2)+\varepsilon}$. Using (18) to estimate $|\tilde{\mathcal{Z}}_s|$ gives:

$$\begin{aligned} &\sum_{2DL^{1/2}<2^s\leq L^{(1/2)+\varepsilon}} |\tilde{\mathcal{Z}}_s|\big(2^{2s}r^{-3} + 2^s r^{-1}\big) \\ &\quad\leq \sum_{2^s\leq L^{(1/2)+\varepsilon}} (2L2^{-s})\big(2^{2s}r^{-3} + 2^s r^{-1}\big) \\ &\quad\leq C \sum_{2^s\leq L^{(1/2)+\varepsilon}} (L2^s r^{-3} + Lr^{-1}) \\ &\quad\leq C(L^{(3/2)+\varepsilon}r^{-3} + L(\log L)r^{-1}) \\ &\quad\leq CL^{(3/2)+\varepsilon}r^{-2}. \end{aligned}$$

Next we consider the contribution to the last sum in (19) from s in the range $2^s \le 2DL^{1/2}$. In this range of s, we use (14) to bound $|\tilde{\mathcal{Z}}_s|$: $|\tilde{\mathcal{Z}}_s| \le |\tilde{\mathcal{Z}}| \le \mathrm{Poly}(D)L^{(1/2)-\varepsilon}\log L$.

$$\sum_{2^s \le 2DL^{1/2}} |\tilde{\mathcal{Z}}_s|\big(2^{2s}r^{-3}+2^s r^{-1}\big) \le \mathrm{Poly}(D)\big(L^{(1/2)-\varepsilon}\log L\big)\big(2^{2s}r^{-3}+2^s r^{-1}\big). \quad (20)$$

Since $2^s \le 2DL^{1/2}$ we see that $2^{2s}r^{-3} \le \mathrm{Poly}(D)Lr^{-3}$ and $2^s r^{-1} \le \mathrm{Poly}(D)L^{1/2}r^{-1} \le \mathrm{Poly}(D)Lr^{-2}$. Plugging these into the right-hand side of (20), we get

$$\sum_{2^s \le 2DL^{1/2}} |\tilde{\mathcal{Z}}_s|\big(2^{2s}r^{-3} + 2^s r^{-1}\big) \le \mathrm{Poly}(D)L^{3/2}r^{-2}.$$

All together, we see

$$\sum_{Z\in\tilde{\mathcal{Z}}\setminus\mathcal{Z}} |P_{r'}(\mathfrak{L}_Z)| \le \mathrm{Poly}(D)L^{(3/2)+\varepsilon}r^{-2}.$$

Choosing $K = K(\varepsilon, D)$ sufficiently large, we see that

$$\sum_{Z\in\tilde{\mathcal{Z}}\setminus\mathcal{Z}} |P_{r'}(\mathfrak{L}_Z)| \le (1/100)KL^{(3/2)+\varepsilon}r^{-2}.$$

This proves item (2), closing the induction, and finishing the proof of Theorem 2.1.

4 Distinct Distances

In [2], Elekes and Sharir proposed a new approach to the distinct distance problem, connecting it to incidence estimates about curves in $\mathbb{R}^3$. A tiny modification of these ideas is explained in Sect. 2 of [6], connecting the distinct distance problem to an estimate about incidences of lines in $\mathbb{R}^3$. The paper [6] then uses Theorem 0.1 to control these incidences. We can also use our slightly weaker Theorem 0.2 to prove a slightly weaker bound on the number of distinct distances.

In this section, we give a concise review of the Elekes–Sharir approach to the distinct distance problem. Using our incidence bound, Theorem 0.2, we prove the following distinct distance bound.

Theorem 4.1 *For any $\varepsilon > 0$, there is a constant $c_\varepsilon > 0$ so that the following holds. If P is a set of N points in $\mathbb{R}^2$, then P determines at least $c_\varepsilon N^{1-\varepsilon}$ distinct distances.*

If $P \subset \mathbb{R}^2$ is a set of points, we let $d(P)$ be the set of distinct distances:

$$d(P) := \{|p_1 - p_2|\}_{p_1,p_2\in P}.$$

The approach of Elekes and Sharir involves the set of distance quadruples $Q(P)$:

$$Q(P) := \{(p_1, p_2, p_3, p_4) \in P^4 \text{ so that } |p_1 - p_2| = |p_3 - p_4| \neq 0\}.$$

A simple Cauchy–Schwarz inequality proves the following estimate (Lemma 2.1 in [6]):

$$|d(P)| \geq \frac{N^4 - 2N^3}{|Q(P)|}. \tag{21}$$

The heart of the matter is to prove an upper bound for $|Q(P)|$. The next step is to introduce a family of lines in $\mathbb{R}^3$, $\mathfrak{L}(P)$, associated to the set $P \subset \mathbb{R}^2$. The incidence geometry of this family of lines encodes the distance quadruples.

For any two points $p_1, p_2 \in \mathbb{R}^2$, we define a line $l_{p_1,p_2} \subset \mathbb{R}^3$ as follows. Suppose that $p_1 = (x_1, y_1)$ and $p_2 = (x_2, y_2)$. We use x, y, z for the coordinates of $\mathbb{R}^3$. Then l_{p_1,p_2} is the line defined by the following equations:

$$2x = (x_1 + x_2) + (y_1 - y_2)z, \tag{22}$$

$$2y = (y_1 + y_2) + (x_2 - x_1)z. \tag{23}$$

The set $\mathfrak{L}(P)$ is defined to be $\{l_{p_1,p_2}\}_{p_1,p_2 \in P}$. If P is a set of N points, then $\mathfrak{L}(P)$ is a set of N^2 lines. The connection between $Q(P)$ and $\mathfrak{L}(P)$ appears in the following lemma.

Lemma 4.2 *A quadruple $(p_1, p_2, p_3, p_4) \in P^4$ is a distance quadruple if and only if the line l_{p_1,p_3} and the line l_{p_2,p_4} are intersecting or parallel.*

Remark The condition of being intersecting or parallel is natural from the projective point of view. Two lines $l, \bar{l}$ are intersecting or parallel in $\mathbb{R}^n$ if and only if they intersect in $\mathbb{RP}^n$.

We now give a proof by direct computation. The paper [2] gives a nice motivation for introducing these lines. The motivation comes from the group of rigid motions of the plane, which is a symmetry group of the distinct distance problem. This point of view is also explained in [6, Sect. 2]. Lemma 4.2 is proven in [6, Sect. 2] using the point of view of rigid motions.

Proof First we describe the projective completion of the line l_{p_1,p_2} in $\mathbb{RP}^3$. A point in $\mathbb{RP}^3$ is an equivalence class of non-zero vectors $(w, x, y, z) \in \mathbb{R}^4$, where two vectors are equivalent if one is a scalar multiple of the other. In these coordinates, the equations for the line $l_{p_1,p_2} \subset \mathbb{RP}^3$ are as follows:

$$2x = (x_1 + x_2)w + (y_1 - y_2)z, \tag{24}$$

$$2y = (y_1 + y_2)w + (x_2 - x_1)z. \tag{25}$$

Next we investigate when two lines in $\mathbb{RP}^3$ intersect. Suppose that l is defined by the equations

$$2x = a_x w + b_x z; \; 2y = a_y w + b_y z \tag{26}$$

and $\bar{l}$ is defined by the equations

$$2x = \bar{a}_x w + \bar{b}_x z; 2y = \bar{a}_y w + \bar{b}_y z. \tag{27}$$

The lines l and $\bar{l}$ intersect in $\mathbb{RP}^3$ if and only if the following system of two equations in w, z has a non-zero solution:

$$a_x w + b_x z = \bar{a}_x w + \bar{b}_x z; \qquad a_y w + b_y z = \bar{a}_y w + \bar{b}_y z. \tag{28}$$

By standard linear algebra, this system of equations has a non-zero solution if and only if an appropriate determinant vanishes, which we can rewrite as the following equation:

$$(a_x - \bar{a}_x)(b_y - \bar{b}_y) = (a_y - \bar{a}_y)(b_x - \bar{b}_x). \tag{29}$$

Now we take $l = l_{p_1,p_3}$ and $\bar{l} = l_{p_2,p_4}$. Using (24) and (25), we can find the values of a_x etc. In particular, we see that $a_x = x_1 + x_3$, $a_y = y_1 + y_3$, $b_x = y_1 - y_3$ and $b_y = x_3 - x_1$, and similarly $\bar{a}_x = x_2 + x_4$, $\bar{a}_y = y_2 + y_4$, $\bar{b}_x = y_2 - y_4$, and $\bar{b}_y = x_4 - x_2$. When we plug these values into (29), we get a homogeneous quadratic equation in x_i and y_i. We claim that this equation is equivalent to $(x_1 - x_2)^2 + (y_1 - y_2)^2 = (x_3 - x_4)^2 - (y_3 - y_4)^2$. Here is the computation. Plugging the values of a_x etc. into (29), we immediately get:

$$\begin{aligned}&[(x_1 + x_3) - (x_2 + x_4)]\,[(x_3 - x_1) - (x_4 - x_2)]\\ &\quad = [(y_1 + y_3) - (y_2 + y_4)]\,[(y_1 - y_3) - (y_2 - y_4)]\,.\end{aligned}$$

Rearranging the terms inside of each large parentheses, this is equivalent to

$$\begin{aligned}&[(x_3 - x_4) + (x_1 - x_2)]\,[(x_3 - x_4) - (x_1 - x_2)]\\ &\quad = [(y_1 - y_2) + (y_3 - y_4)]\,[(y_1 - y_2) - (y_3 - y_4)]\end{aligned}$$

Expanding both sides, this is equivalent to

$$(x_3 - x_4)^2 - (x_1 - x_2)^2 = (y_1 - y_2)^2 - (y_3 - y_4)^2.$$

Moving the negative terms to the other sides, this is equivalent to

$$(x_3 - x_4)^2 + (y_3 - y_4)^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2.$$

This is equivalent to $|p_3 - p_4| = |p_1 - p_2|$. □

Because of Lemma 4.2, each distance quadruple $(p_1, p_2, p_3, p_4) \in Q(P)$ can be labelled as an intersecting quadruple or a parallel quadruple, depending on whether l_{p_1,p_3} and l_{p_2,p_4} are intersecting or parallel.

The number of parallel quadruples is straightforward to bound. If l_{p_1,p_3} and l_{p_2,p_4} are parallel, then (22) and (23) imply that $y_1 - y_3 = y_2 - y_4$ and $x_3 - x_1 = x_4 - x_2$. In other words, l_{p_1,p_3} and l_{p_2,p_4} are parallel if and only if $p_1 - p_2 = p_3 - p_4$. For any p_1, p_2, p_3, there is at most one $p_4 \in P$ so that $p_1 - p_2 = p_3 - p_4$, and so there are at most N^3 parallel distance quadruples.

From now on, we sometimes abbreviate $\mathfrak{L}(P)$ by $\mathfrak{L}$.

The number of intersecting distance quadruples can be counted as follows. We let $P_{=r}(\mathfrak{L})$ denote the set of points that lie in exactly r lines of $\mathfrak{L}$. At each point of $P_{=r}(\mathfrak{L})$ there are $r^2 - r$ intersecting pairs $(l_1, l_2) \in \mathfrak{L}^2$. Therefore, the number of intersecting distance quadruples is

$$|Q(P)_{\text{inter}}| = \sum_{r \geq 2} (r^2 - r)|P_{=r}(\mathfrak{L})|.$$

Since $|P_{=r}(\mathfrak{L})| = |P_r(\mathfrak{L})| - |P_{r+1}(\mathfrak{L})|$, we can rewrite this formula as

$$|Q(P)_{\text{inter}}| = \sum_{r \geq 2} (2r - 2)|P_r(\mathfrak{L})|. \tag{30}$$

Therefore, a bound on $|P_r(\mathfrak{L})|$ gives a bound on $|Q(P)|$.

To bound $|P_r(\mathfrak{L})|$ the paper [6] proves the following result (Proposition 2.8 in [6]):

Lemma 4.3 *If $P \subset \mathbb{R}^2$ is a set of N points, then $\mathfrak{L}(P)$ contains at most CN lines in any plane or regulus, and at most N lines of $\mathfrak{L}(P)$ contain any point.*

With this lemma in hand, we [6] can apply Theorem 0.1, giving the bound $|P_r(\mathfrak{L})| \leq CN^3 r^{-2}$ for all $2 \leq r \leq N$. (And for $r > N + 1$, Lemma 4.3 says that $|P_r(\mathfrak{L})| = 0$.) Plugging these bounds into (30) shows that $|Q(P)| \leq N^3 + \sum_{r=2}^{N} CN^3 r^{-1} \leq CN^3 \log N$.

We will use Theorem 0.2 in place of Theorem 0.1 to give a slightly weaker bound on the number of distance quadruples. In order to apply Theorem 0.2 we need a slightly stronger lemma.

Lemma 4.4 *For any degree $D \geq 1$ there is a constant C_D so that the following holds. If $P \subset \mathbb{R}^2$ is a set of N points, then $\mathfrak{L}(P)$ contains at most $C_D N$ lines in any algebraic surface of degree at most D. Also $\mathfrak{L}(P)$ contains at most N lines that pass through any point.*

We will give the proof of Lemma 4.4 below. Using Lemma 4.4, we can apply Theorem 0.2, giving the following bound: for any $\varepsilon > 0$, there is a constant C_ε so that

$$|P_r(\mathfrak{L})| \leq C_\varepsilon N^{3+\varepsilon} r^{-2}.$$

Plugging this bound into (30), we see that

$$|Q(P)| \leq N^3 + \sum_{r=2}^{N} (2r - 2)|P_r(\mathfrak{L})| \leq N^3 + \sum_{r=2}^{N} C_\varepsilon N^{3+\varepsilon} r^{-1} \leq C_\varepsilon N^{3+\varepsilon}.$$

Plugging this bound into (21), we see that $|d(P)| \geq c_\varepsilon N^{1-\varepsilon}$ for any $\varepsilon > 0$. This proves Theorem 4.1.

4.1 The Proof of the Non-clustering Lemma

It only remains to prove Lemma 4.4. Suppose that $P \subset \mathbb{R}^2$ is a set of N points.

We first observe that if $p \in \mathbb{R}^2$ and $q_1 \neq q_2 \in \mathbb{R}^2$ then the lines l_{p,q_1} and l_{p,q_2} are skew. By Lemma 4.2, l_{p,q_1} and l_{p,q_2} are non-skew if and only if $|p - p| = |q_1 - q_2|$. But $|p - p| = 0$ and $|q_1 - q_2| \neq 0$.

From this observation, we can quickly establish two parts of Lemma 4.4. First, for any plane in $\mathbb{R}^3$, at most one of the lines $\{l_{p,q}\}_{q \in P}$ can lie in the plane. Therefore, any plane contains at most N lines of $\mathfrak{L}(P)$. Second, for any point $\mathbb{R}^3$, at most one of the lines $\{l_{p,q}\}_{q \in P}$ can contain the point. Therefore, for any point in $\mathbb{R}^3$, at most N lines of $\mathfrak{L}(P)$ contain the point.

Now consider an irreducible polynomial Q with $1 < \mathrm{Deg}\, Q \leq D$. We will prove that $Z(Q)$ contains $\leq 3D^2N$ lines of $\mathfrak{L}(P)$, and this will finish the proof of Lemma 4.4.

We let $\mathfrak{L}_p := \{l_{p,q}\}_{q \in \mathbb{R}^2}$. We would like to understand how many lines of $\mathfrak{L}_p$ may lie in $Z(Q)$.

Lemma 4.5 *If Q is an irreducible polynomial with $1 < \mathrm{Deg}\, Q \leq D$, then there is at most one point $p \in \mathbb{R}^2$ so that $Z(Q)$ contains at least $2D^2$ lines of $\mathfrak{L}_p$.*

Given Lemma 4.5, we now check that $Z(Q)$ contains at most $3D^2N$ lines of $\mathfrak{L}(P)$. For $N - 1$ of the points $p \in P$, $Z(Q)$ contains at most $2D^2$ of the lines $\{l_{p,p'}\}_{p' \in P}$. For the last point $p \in P$, $Z(Q)$ contains at most all N of the lines $\{l_{p,p'}\}_{p' \in P}$. In total, $Z(Q)$ contains at most $(2D^2 + 1)N$ lines of $\mathfrak{L}(P)$.

The proof of Lemma 4.5 is based on a more technical lemma which describes the algebraic structure of the set of lines $\{l_{p,q}\}$ in $\mathbb{R}^3$.

Lemma 4.6 *For each p, each point of $\mathbb{R}^3$ lies in a unique line from the set $\{l_{p,q}\}_{q \in \mathbb{R}^2}$. Moreover, for each p, there is a non-vanishing vector field $V_p(x_1, x_2, x_3)$, so that at each point, $V_p(x)$ is tangent to the unique line $l_{p,q}$ through x. Moreover, $V_p(x)$ is a polynomial in p and x, with degree at most 1 in the p variables and degree at most 2 in the x variables.*

Let us assume this technical lemma for the moment and use it to prove Lemma 4.5.

Fix a point $p \in \mathbb{R}^2$. Suppose $Z(Q)$ contains at least $2D^2$ lines from the set $\mathfrak{L}_p := \{l_{p,q}\}_{p,q \in \mathbb{R}^2}$. On each of these lines, Q vanishes identically, and V_p is tangent to the line. Therefore, $V_p \cdot \nabla Q$ vanishes on all these lines. But $V_p \cdot \nabla Q$ is a polynomial in x of degree at most $2D - 2$. If $V_p \cdot \nabla Q$ and Q have no common factor, then the Bezout theorem for lines, Theorem 1.3, implies that there are at most $2D^2 - 2D$ lines where the two polynomials vanish. Therefore, $V_p \cdot \nabla Q$ and Q have a common factor. Since Q is irreducible, Q must divide $V_p \cdot \nabla Q$, and we see that $V_p \cdot \nabla Q$ vanishes identically on $Z(Q)$.

Now suppose that $Z(P)$ contains at least $2D^2$ lines from $\mathfrak{L}_{p_1}$ and from $\mathfrak{L}_{p_2}$. We see that $V_{p_1} \cdot \nabla Q$ and $V_{p_2} \cdot \nabla Q$ vanish on $Z(Q)$. For each fixed x, the expression $V_p \cdot \nabla Q$ is a degree 1 polynomial in p. Therefore, for any point p in the affine span of p_1 and p_2, $V_p \cdot \nabla Q$ vanishes on $Z(Q)$.

Suppose that $Z(Q)$ has a non-singular point x, which means that $\nabla Q(x) \neq 0$. In this case, x has a smooth neighborhood $U_x \subset Z(Q)$ where ∇Q is non-zero. If $V_p \cdot \nabla Q$

vanishes on $Z(Q)$, then the vector field V_p is a vector field on U_x, and so its integral curves lie in U_x. But the integral curves of V_p are exactly the lines of $\mathfrak{L}_p$. Therefore, for each p on the line connecting p_1 and p_2, the line of $\mathfrak{L}_p$ through x lies in $Z(Q)$. Since x is a smooth point, all of these lines must lie in the tangent plane $T_x Z(Q)$, and we see that $Z(Q)$ contains infinitely many lines in a plane. Using Bezout's theorem, Theorem 1.3, again, we see that $Z(Q)$ is a plane, and that Q is a degree 1 polynomial. This contradicts our assumption that Deg $Q > 1$.

We have now proven Lemma 4.5 in the case that $Z(Q)$ contains a non-singular point. But if every point of $Z(Q)$ is singular, then we get an even stronger estimate on the lines in $Z(Q)$:

Lemma 4.7 *Suppose that Q is a non-zero irreducible polynomial of degree D on $\mathbb{R}^3$. If $Z(Q)$ has no non-singular point, then $Z(Q)$ contains at most D^2 lines.*

Proof Since every point of $Z(Q)$ is singular, ∇Q vanishes on $Z(Q)$. In particular, each partial derivative $\partial_i Q$ vanishes on $Z(Q)$. We suppose that $Z(Q)$ contains more than D^2 lines and derive a contradiction. Since $\partial_i Q = 0$ on $Z(Q)$ and $Z(Q)$ contains more than D^2 lines, then Bezout's theorem, Theorem 1.3, implies that Q and $\partial_i Q$ have a common factor. Since Q is irreducible, Q must divide $\partial_i Q$. Since Deg $\partial_i Q <$ Deg Q, it follows that $\partial_i Q$ is identically zero for each i. This implies that Q is constant. By assumption, Q is not the zero polynomial and so $Z(Q)$ is empty. But we assumed that $Z(Q)$ contains at least $D^2 + 1$ lines, giving a contradiction. □

This finishes the proof of Lemma 4.5 assuming Lemma 4.6. It only remains to prove Lemma 4.6.

First we check that each point $x \in \mathbb{R}^3$ lies in exactly one of the lines $\{l_{p,q}\}_{q \in \mathbb{R}^2}$. Suppose $p = (p_1, p_2)$ and $q = (q_1, q_2)$ are points in $\mathbb{R}^2$. Using (22) and (23), we see that $(x_1, x_2, x_3) \in l_{p,q}$ if and only if

$$2x_1 = (p_1 + q_1) + (p_2 - q_2)x_3, \tag{31}$$

$$2x_2 = (p_2 + q_2) + (q_1 - p_1)x_3. \tag{32}$$

We can rewrite these equations as a matrix equation for q as follows:

$$\begin{pmatrix} 1 & -x_3 \\ x_3 & 1 \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = (2x_1 - p_1 - x_3 p_2, 2x_2 - p_2 + p_1 x_3) =: a_p(x),$$

Note that $a_p(x)$ is a vector whose entries are polynomials in x, p and of degree ≤ 1 in x and of degree ≤ 1 in p. Since the determinant of the matrix on the left-hand side is $1 + x_3^2 > 0$, we can uniquely solve this equation for q_1 and q_2. The solution has the form

$$q_1 = (x_3^2 + 1)^{-1} b_{1,p}(x); \quad q_2 = (x_3^2 + 1)^{-1} b_{2,p}(x), \tag{33}$$

where b_1, b_2 are polynomials in x, p of degree ≤ 2 in x and degree ≤ 1 in p.

We have now proven that each point of $\mathbb{R}^3$ lies in a unique line from the set $\{l_{p,q}\}_{q \in \mathbb{R}^2}$. Now we can construct the vector field V_p. From (31) and (32), we see that the vector $(p_2 - q_2, q_1 - p_1, 2)$ is tangent to $l_{p,q}$. If $x \in l_{p,q}$, then we can use

(33) to expand q in terms of x, p, and we see that the following vector field is tangent to $l_{p,q}$ at x:

$$v_p(x) := \left(p_2 - \left(x_3^2+1\right)^{-1} b_{2,p}(x), \left(x_3^2+1\right)^{-1} b_{1,p}(x) - p_1, 2\right).$$

The coefficients of $v_p(x)$ are not polynomials because of the $(x_3^2+1)^{-1}$. We define $V_p(x) = (x_3^2+1)v_p(x)$, so

$$V_p(x) = \left(p_2\left(x_3^2+1\right) - b_{2,p}(x), b_{1,p}(x) - p_1\left(x_3^2+1\right), 2x_3^2+2\right).$$

The vector field $V_p(x)$ is tangent to the family of lines $\{l_{p,q}\}_{q\in\mathbb{R}^2}$. Moreover, V_p never vanishes because its last component is $2x_3^2+2$. Therefore, the integral curves of V_p are exactly the lines $\{l_{p,q}\}_{q\in\mathbb{R}^2}$. Moreover, each component of V_p is a polynomial of degree ≤ 2 in x and degree ≤ 1 in p.

This finishes the proof of Lemma 4.6 and hence the proof of Lemma 4.4.

Acknowledgments I would like to thank Nets Katz for many interesting discussions about these ideas over the last several years. I would also like to thank the referee for helpful suggestions about the exposition.

References

1. Clarkson, K.L., Edelsbrunner, H., Guibas, L., Sharir, M., Welzl, E.: Combinatorial complexity bounds for arrangements of curves and spheres. Discrete Comput. Geom. **5**, 99–160 (1990)
2. Elekes, G., Sharir, M.: Incidences in three dimensions and distinct distances in the plane. Comb. Probab. Comput. **20**, 571–608 (2011)
3. Elekes, G., Kaplan, H., Sharir, M.: On lines, joints, and incidences in three dimensions. J. Comb. Theory Ser. A **118**, 962–977 (2011)
4. Erdős, P.: On sets of distances of n points. Amer. Math. Monthly **53**, 248–250 (1946)
5. Guth, L., Katz, N.H.: Algebraic methods in discrete analogues of the Kakeya problem. Adv. Math. **225**, 2828–2839 (2010)
6. Guth, L., Katz, N.H.: On the Erdős distinct distances problem in the plane. Ann. Math. **181**, 155–190 (2015). arXiv:1011.4105
7. Kaplan, H., Matoušek, J., Sharir, M.: Simple proofs of classical theorems in discrete geometry via the Guth–Katz polynomial partitioning technique. Discrete Comput. Geom. **48**, 499–517 (2012)
8. Milnor, J.: On the Betti numbers of real varieties. Proc. AMS **15**, 275–280 (1964)
9. Oleinik, O.A., Petrovskii, I.B.: On the topology of real algebraic surfaces. Izv. Akad. Nauk SSSR **13**, 389–402 (1949)
10. Sharir, M., Solomon, N.: Incidences between points and lines in four dimensions. In: Proceedings of 30th ACM Symposium on Computational Geometry (2014)
11. Solymosi, J., Tao, T.: An incidence theorem in higher dimensions. Discrete Comput. Geom. **48**(2), 255–280 (2012)
12. Stone, A.H., Tukey, J.W.: Generalized sandwich theorems. Duke Math. J. **9**, 356–359 (1942)
13. Székely, L.: Crossing numbers and hard Erdős problems in discrete geometry. Comb. Probab. Comput. **6**(3), 353–358 (1997)
14. Szemerédi, E., Trotter Jr, W.T.: Extremal problems in discrete geometry. Combinatorica **3**, 381–392 (1983)
15. Thom, R.: Sur l'homologie des variétés algébriques réelles. In: Cairns, S.S. (ed.) Differential and Combinatorial Topology (Symposium in Honor of Marston Morse), pp. 255–265. Princeton University Press, Princeton (1965)
16. van der Waerden, B.L.: Modern Algebra. Frederick Ungar Publishing Co., New York (1950)