Forum of Mathematics, Sigma (2016), Vol. 4, e2, 13 pages
doi:10.1017/fms.2015.34

A NOTE ON RICH LINES IN TRULY HIGH DIMENSIONAL SETS

JOSHUA ZAHL
Massachusetts Institute of Technology, Cambridge, MA, USA;
email: jzahl@mit.edu

Received 3 March 2015; accepted 26 December 2015

Abstract

We modify an argument of Hablicsek and Scherr to show that if a collection of points in $\mathbb{C}^d$ spans many r-rich lines, then many of these lines must lie in a common $(d-1)$-flat. This is closely related to a previous result of Dvir and Gopi.

2010 Mathematics Subject Classification: 52C35 (primary); 52C10 (secondary)

1. Introduction

This note shows that the techniques of Hablicsek and Scherr from [6] can be extended from $\mathbb{R}^d$ to $\mathbb{C}^d$, with an ϵ loss in the exponent. In [3], Dvir and Gopi proved a new upper bound on the number of r-rich lines in 'truly' d-dimensional configurations of points in $\mathbb{C}^d$. Given a collection of points in $\mathbb{C}^d$, Dvir and Gopi proved that either many of these points lie on a $(d-1)$-flat (that is, a $(d-1)$-dimensional affine subspace) or the points span few r-rich lines. Specifically, they established the following result.

THEOREM 1.1 (Dvir and Gopi). *For all $d \geqslant 1$, there exist constants c_d, C_d such that the following holds. Let $\mathcal{P} \subset \mathbb{C}^d$ be a set of n points, let $r \geqslant 2$ be an integer, and let $\mathcal{L}_r(\mathcal{P})$ be the set of lines that are incident to at least r points from $\mathcal{P}$. Suppose that for some $\alpha \geqslant 1$,*

$$|\mathcal{L}_r(\mathcal{P})| \geqslant C_d \cdot \alpha \cdot \frac{n^2}{r^d}.$$

Then there exists a subset $\mathcal{P}' \subset \mathcal{P}$ of size at least $c_d \cdot \alpha \cdot n/r^{d-2}$ contained in a $(d-1)$-flat.

The bounds in Theorem 1.1 are not believed to be tight. Dvir and Gopi conjectured the following bound, which (if correct) would be tight.

CONJECTURE 1.1 (Dvir and Gopi). *For $r \geqslant 2$, suppose $\mathcal{P} \subset \mathbb{C}^d$ is a set of n points with*

$$|\mathcal{L}_r(\mathcal{P})| \gg_d \frac{n^2}{r^{d+1}} + \frac{n}{r}.$$

Then there exists an integer $1 < t < d$ and a subset $\mathcal{P}' \subset \mathcal{P}$ of size $\gtrsim_d n/r^{d-t}$ contained in a t-flat.

This bound is a d-dimensional generalization of the (complex) Szemerédi–Trotter theorem [10–12]. In [6], Hablicsek and Scherr proved a stronger version of Theorem 1.1, except that they needed to replace complex lines in $\mathbb{C}^d$ with real lines in $\mathbb{R}^d$. This allowed them to use the discrete polynomial partitioning theorem [5, Theorem 4.1] and the joints theorem [4] of Guth and Katz.

THEOREM 1.2 (Hablicsek and Scherr). *For all $d \geqslant 1$, there exist constants c_d, C_d such that the following holds. Let $\mathcal{P} \subset \mathbb{R}^d$ be a set of n points, let $r \geqslant 2$ be an integer, and let $\mathcal{L}_r(\mathcal{P})$ be the set of lines that are incident to at least r points from $\mathcal{P}$. Suppose that*

$$|\mathcal{L}_r(\mathcal{P})| \geqslant C_d \cdot \frac{n^2}{r^{d+1}}.$$

Then there exists a subset $\mathcal{P}' \subset \mathcal{P}$ of size at least $c_d \cdot n/r^{d-1}$ contained in a $(d-1)$-flat.

In this paper, we will prove a theorem similar to Theorem 1.2 in the original setting of Dvir and Gopi (that is, for complex lines in $\mathbb{C}^d$). As in the Hablicsek and Scherr proof, the discrete polynomial partitioning theorem will play a major role. The joints theorem will not be used directly, but similar types of arguments will be employed.

1.1. Incidence theorems over the complex numbers. Since its introduction in 2010, the discrete polynomial partitioning theorem has been used to prove many incidence bounds in $\mathbb{R}^d$. The theorem makes crucial use of the fact that removing a point from $\mathbb{R}$ disconnects the real line into two components. This property is not true for $\mathbb{C}$, which means that the discrete polynomial partitioning theorem cannot be employed directly to prove incidence theorems in $\mathbb{C}^d$. Of course, one can identify $\mathbb{C}$ with $\mathbb{R}^2$ and then the discrete polynomial partitioning

https://doi.org/10.1017/fms.2015.34 Published online by Cambridge University Press

theorem becomes available again. Unfortunately, moving from $\mathbb{C}^d$ to $\mathbb{R}^{2d}$ often makes the problem appear more complicated, since the dimension of the objects being studied has now doubled. However, since the problem originally arose from a configuration in $\mathbb{C}^d$, the new configuration in $\mathbb{R}^{2d}$ may have some special properties that can aid in the analysis of the problem. This strategy was employed by Tóth in [11] to prove the Szemerédi–Trotter theorem in $\mathbb{C}^2$ and by Sheffer and the author in [8] to obtain an incidence theorem for points and curves in $\mathbb{C}^2$. In the present work, we will note several elementary relationships between collections of lines in $\mathbb{C}^d$ and the corresponding two-flats in $\mathbb{R}^{2d}$, and these observations will allow us to transfer the Hablicsek–Scherr argument from $\mathbb{R}^d$ to $\mathbb{C}^d$.

1.2. Statement of the theorem

THEOREM 1.3. *For all $d \geqslant 1$ and $\epsilon > 0$, there exist constants $c_{d,\epsilon}, C_{d,\epsilon}$ such that the following holds. Let $\mathcal{P} \subset \mathbb{C}^d$ be a set of n points and let $r \geqslant 2$ be an integer. Suppose that for some $\alpha \geqslant 1$,*

$$|\mathcal{L}_r(\mathcal{P})| > C_{d,\epsilon} \cdot \alpha \cdot \frac{n^{2+\epsilon}}{r^{d+1}}. \tag{1}$$

Then there exists a subset $\mathcal{P}' \subset \mathcal{P}$ of size at least $c_{d,\epsilon} \cdot \alpha \cdot n^{1+\epsilon}/r^{d-1}$ contained in a $(d-1)$-flat.

Theorem 1.3 is not strictly stronger than Theorem 1.1 because Theorem 1.3 contains the term $n^{2+\epsilon}$ rather than n^2. However, if r is not too small compared to n, then one can 'trade' the term n^ϵ for a term of the form r^{ϵ_1}. More precisely, we have the following.

COROLLARY 1.1 (Cheap Dvir–Gopi). *For all $d \geqslant 1$ and $\epsilon_0 > 0$, there exist constants $c_{d,\epsilon_0}, C_{d,\epsilon_0}$ such that the following holds. Let $\mathcal{P} \subset \mathbb{C}^d$ be a set of n points and let $r \geqslant n^{\epsilon_0}$ be an integer. Suppose that for some $\alpha' \geqslant 1$,*

$$|\mathcal{L}_r(\mathcal{P})| > C_{d,\epsilon_0} \cdot \alpha' \cdot \frac{n^2}{r^d}. \tag{2}$$

Then there exists a subset $\mathcal{P}' \subset \mathcal{P}$ of size at least $c_{d,\epsilon_0} \cdot \alpha' \cdot n/r^{d-2}$ contained in a $(d-1)$-flat.

Proof. Let $\rho = \log r/\log n$; by assumption $\rho \geqslant \epsilon_0$, and for each $\epsilon > 0$, $n^\epsilon = r^{\epsilon/\rho}$. Select $\epsilon = \rho/2$, and let $\alpha = \alpha' r^{1/2}$. Applying Theorem 1.3, we conclude that either

$$\begin{aligned} |\mathcal{L}_r(\mathcal{P})| &> C_{d,\epsilon} \cdot \alpha \cdot \frac{n^{2+\epsilon}}{r^{d+1}} \\ &> C_{d,\epsilon_0} \cdot \alpha' \cdot \frac{n^2}{r^d}, \end{aligned} \tag{3}$$

https://doi.org/10.1017/fms.2015.34 Published online by Cambridge University Press

or there exists a subset $\mathcal{P}' \subset \mathcal{P}$ of size at least $c_{d,\epsilon} \cdot \alpha \cdot n^{1+\epsilon}/r^{d-1} = c_{d,\epsilon} \cdot \alpha' \cdot n/r^{d-2}$ contained in a $(d-1)$-flat. □

1.3. Initial reductions. In this section, we will show that in order to prove Theorem 1.3 it suffices to prove the following lemma.

LEMMA 1.1. *For all $d \geqslant 1$ and $\epsilon > 0$, there exist constants $c'_{d,\epsilon}, C'_{d,\epsilon}$ such that the following holds. Let $\mathcal{P}_1 \subset \mathbb{C}^d$ be a set of n_1 points and let $\mathcal{L}_1$ be a set of ℓ_1 lines in $\mathbb{C}^d$. Then at least one of the following statements must hold:*

(A) *There are a point $p \in \mathcal{P}_1$ and a complex $(d-1)$-flat $\Pi \subset \mathbb{C}^d$ so that*

$$|\{L \in \mathcal{L}_1 : p \in L,\ L \subset \Pi\}| \geqslant c'_{d,\epsilon} |\{L \in \mathcal{L}_1 : p \in L\}|. \tag{4}$$

(B)

$$I(\mathcal{P}_1, \mathcal{L}_1) \leqslant C'_{d,\epsilon}\big(n_1^{(2+\epsilon)/(d+1)} \ell_1^{d/(d+1)} + n_1 + \ell_1\big). \tag{5}$$

To obtain Theorem 1.3 from Lemma 1.1, fix $\epsilon > 0$. Let $\mathcal{P} \subset \mathbb{C}^d$ be a set of n points and suppose (1) holds. If $r \geqslant \alpha^{1/d} n^{(1+\epsilon)/d}$, then

$$\alpha \frac{n^{1+\epsilon}}{r^{d-1}} \leqslant r,$$

so Theorem 1.3 holds with $c_{d,\epsilon} = 1$. Just select any line from $\mathcal{L}_r(\mathcal{P})$; let $\mathcal{P}'$ be the set of points incident to the line, and select any $(d-1)$-flat containing the line.

Henceforth we will assume that $r < \alpha^{1/d} n^{(1+\epsilon)/d}$. Repeating an argument of Hablicsek and Scherr [6], we can find a set $\mathcal{L}_1 \subset \mathcal{L}_r(\mathcal{P})$ and a set $\mathcal{P}_1 \subset \mathcal{P}$ so that each line in $\mathcal{L}_1$ is incident to at least $r/4$ points from $\mathcal{P}_1$, each point in $\mathcal{P}_1$ is incident to at least $\frac{1}{4} C'_{d,\epsilon} \cdot \alpha \cdot n^{1+\epsilon}/r^d$ lines from $\mathcal{L}_1$, and $I(\mathcal{P}_1, \mathcal{L}_1) \geqslant \frac{1}{2} I(\mathcal{P}, \mathcal{L}_r(\mathcal{P}))$. In brief, the argument is as follows. If $G = (A \sqcup B, E)$ is a bipartite graph, then we can find subsets $A' \subset A,\ B' \subset B$ so that in the induced subgraph, each vertex from A has degree at least $|E|/4|A|$, each vertex from B has degree at least $|E|/4|B|$, and $|E'| \geqslant |E|/2$. We apply this lemma to the bipartite graph with edge set $A = \mathcal{P},\ B = \mathcal{L}_r(\mathcal{P})$, and where $a \sim b$ if the point corresponding to a lies on the line corresponding to b. See [6] for details.

Let $p \in \mathcal{P}_1$ and let $\rho > 0$. Observe that if a ρ-fraction of the lines from $\mathcal{L}_1$ passing through p lie in a common $(d-1)$-flat, then at least $(\rho/16) C_{d,\epsilon} \cdot \alpha \cdot n^{1+\epsilon}/r^{d-1}$ points from $\mathcal{P}_1$ lie in a common $(d-1)$-flat.

Apply Lemma 1.1 to the arrangement $(\mathcal{P}_1, \mathcal{L}_1)$ (with the same value of ϵ). If condition (A) from Lemma 1.1 holds, then there exists a subset of $\mathcal{P}' \subset \mathcal{P}$ of cardinality at least $\frac{1}{16} c'_{d,\epsilon} \cdot \alpha \cdot n^{1+\epsilon}/r^{d-1}$ contained in a $(d-1)$-flat.

https://doi.org/10.1017/fms.2015.34 Published online by Cambridge University Press

If condition (A) does not hold, then condition (B) must hold, and this implies that

$$I(\mathcal{P}, \mathcal{L}_r(\mathcal{P})) \leqslant C'_{d,\epsilon}(n_1^{(2+\epsilon)/(d+1)}\ell_1^{d/(d+1)} + n_1 + \ell_1).$$

But since each line from $\mathcal{L}_r(\mathcal{P})$ is r-rich, we conclude that

$$\begin{aligned} r|\mathcal{L}_r(\mathcal{P})| &\leqslant 2C'_{d,\epsilon}(n_1^{(2+\epsilon)/(d+1)}\ell_1^{d/(d+1)} + n_1 + \ell_1) \\ &\leqslant 2C'_{d,\epsilon}(n^{(2+\epsilon)/(d+1)}|\mathcal{L}_r|^{d/(d+1)} + n + |\mathcal{L}_r|), \end{aligned}$$

and thus

$$\begin{aligned} |\mathcal{L}_r| &\leqslant 2^{d+1}C'_{d,\epsilon}\left(\frac{n^{2+\epsilon}}{r^{d+1}} + \frac{n}{r}\right) \\ &\leqslant 2^{d+1}C'_{d,\epsilon}\alpha\frac{n^{2+\epsilon}}{r^{d+1}}, \end{aligned} \tag{6}$$

where on the second line we used the fact that $\alpha \geqslant 1$ and $r < \alpha^{1/d}n^{(1+\epsilon)/d}$. This contradicts the assumption that (1) holds. Thus we obtain Theorem 1.3 with $c_{d,\epsilon} = \frac{1}{16}c'_{d,\epsilon}$ and $C_{d,\epsilon} = 2^{d+1}C'_{d,\epsilon}$.

1.4. Main proof ideas. The rest of this paper will be devoted to proving Lemma 1.1. The basic idea is to use a bounded-degree partitioning polynomial and prove the lemma by induction on $n_1 + \ell_1$. Theorem 1.3 does not survive the process of induction, but Lemma 1.1 does—this is why we prove Lemma 1.1 first rather than proving Theorem 1.3 directly.

Here are the main steps. We regard complex lines in $\mathbb{C}^d$ as two-flats in $\mathbb{R}^{2d}$. We will then prove Lemma 1.1 by induction on the number of points and flats. We partition $\mathbb{R}^{2d}$ into cells using a bounded-degree partitioning polynomial; inside each cell we can apply the induction hypothesis. Either there is a point inside a cell satisfying condition (A) from the lemma, or the total number of incidences inside the cells is controlled by (5).

We must now deal with incidences occurring on the boundary of the partition. If p is a point lying on the boundary of the partition and X is a two-flat (arising from a complex line) that is incident to p, then X is either contained in the boundary of the partition or intersects the boundary in a bounded-degree algebraic set of dimension at most 1.

If the former option occurs most of the time (for a given point p), then there must exist a $(2d-1)$-flat in $\mathbb{R}^{2d}$ containing many two-flats, each of which contains p, and these two-flats in turn arise from complex lines. This implies that many complex lines are contained in a complex $(d-1)$-flat in $\mathbb{C}^d$.

If the latter option occurs most of the time, then after applying a generic linear transformation we are reduced to a problem that is very similar to our original

https://doi.org/10.1017/fms.2015.34 Published online by Cambridge University Press

one, except now we are dealing with points and bounded-degree curves in $\mathbb{R}^d$, rather than points and two-flats in $\mathbb{R}^{2d}$. A similar argument shows that either the number of incidences is controlled by (5), or there are a point p and a $(d-1)$-flat in $\mathbb{R}^{2d}$ containing p so that many curves passing through p are tangent to this $(d-1)$-flat at the point p. This implies that in the original configuration of complex lines there are many lines passing through a common point that are contained in a complex $(d-1)$-flat.

2. Preliminaries

2.1. Real and complex vectors. Let $\iota : \mathbb{C}^d \to \mathbb{R}^{2d}$ be the map $\iota(x_1 + iy_1, \ldots, x_d + iy_d) \mapsto (x_1, y_1, \ldots, x_d, y_d)$. If $v = (a_1, b_1, \ldots, a_d, b_d) \in \mathbb{R}^{2d}$, then $\iota^{-1}(v) = (a_1 + ib_1, \ldots, a_d + ib_d) \in \mathbb{C}^d$.

If $X_1, \ldots, X_k$ are vector spaces in $\mathbb{R}^d$ (respectively $\mathbb{C}^d$), let $\operatorname{span}_{\mathbb{R}}\{X_1, \ldots, X_k\}$ (respectively $\operatorname{span}_{\mathbb{C}}\{X_1, \ldots, X_k\}$) be the linear span of these vector spaces, regarded as subspaces of $\mathbb{R}^d$ or $\mathbb{C}^d$. Abusing notation, we will identify the nonzero vector $v \in \mathbb{R}^d$ with the one-dimensional vector space $\mathbb{R}v$, and similarly for vectors $v \in \mathbb{C}^d$.

If $v = (a_1, b_1, \ldots, a_d, b_d) \in \mathbb{R}^{2d}$, define

$$v^\dagger = \operatorname{span}_{\mathbb{R}}\{(a_1, b_1, \ldots, a_d, b_d), (-b_1, a_1, \ldots, -b_d, a_d)\}.$$

By definition, if $v \in \mathbb{R}^{2d}$ then

$$v^\dagger = \iota(\operatorname{span}_{\mathbb{C}}(\iota^{-1}(v))).$$

That is, if $v \in \mathbb{R}^{2d}$ is a vector, let $\iota^{-1}(v) \in \mathbb{C}^d$ be its preimage under ι. Then the span of $\iota^{-1}(v)$ is a one-dimensional complex subspace of $\mathbb{C}^d$. The image of this subspace under ι is a two-dimensional real subspace of $\mathbb{R}^{2d}$. This is precisely $v^\dagger$.

Finally, if $\Pi \subset \mathbb{R}^{2d}$ is a vector space, define $\Pi^\dagger = \operatorname{span}_{\mathbb{R}}\{v^\dagger : v \in \Pi\}$. Note that $\dim(\Pi^\dagger) \leqslant 2\dim(\Pi)$.

2.2. Linear algebra.

LEMMA 2.1. *Let $L_1, \ldots, L_k$ be lines in $\mathbb{C}^d$ passing through the origin. If there is a $(2k-1)$-flat $\Pi_0 \subset \mathbb{R}^{2d}$ containing $\iota(L_1), \ldots, \iota(L_k)$, then there is a $(d-1)$-flat $\Pi \subset \mathbb{C}^d$ containing $L_1, \ldots, L_d$.*

Proof. Since $L_1, \ldots, L_k$ pass through the origin, Π_0 must also contain the origin. Let $v_0 \in \mathbb{R}^{2d}$ be a vector orthogonal to Π_0. Then v_0 is orthogonal to $\iota(L_j)$ for each $j = 1, \ldots, k$. This implies that $\iota^{-1}(v_0)$ is orthogonal to L_j for each index $j = 1, \ldots, k$. Let Π be the orthogonal complement of $\iota^{-1}(v_0)$. □

https://doi.org/10.1017/fms.2015.34 Published online by Cambridge University Press

LEMMA 2.2. *Let $L_1, \dots, L_k$ be lines in $\mathbb{C}^d$ passing through the origin. For each index i, let $v_i \in \iota(L_i)$ be a nonzero vector. Let $\Pi_0 \subset \mathbb{R}^{2d}$ be a $(d-1)$-dimensional vector space. If $v_1, \dots, v_k \in \Pi_0$, then there is a $(d-1)$-dimensional subspace $\Pi \subset \mathbb{C}^d$ containing $L_1, \dots, L_k$.*

Proof. Let $w_1, \dots, w_{d-1} \in \mathbb{R}^{2d}$ be vectors that span Π_0. For each index j we have $v_j = \sum_i a_{ij} w_i$, and thus $\iota^{-1}(v_j) = \sum_i \iota^{-1}(a_{ij} w_i)$. But this implies that $L = \mathbb{C}\iota^{-1}(v_j) \subset \mathrm{span}_{\mathbb{C}}\{\iota^{-1}(w_1), \dots, \iota^{-1}(w_d)\}$. Define $\Pi = \mathrm{span}_{\mathbb{C}}\{\iota^{-1}(w_1), \dots, \iota^{-1}(w_d)\}$. □

LEMMA 2.3. *Let $v_1, \dots, v_k \in \mathbb{R}^{2d}$. Let $\Pi \subset \mathbb{R}^{2d}$ be a $(d-1)$-dimensional vector space, and suppose $v_1, \dots, v_k \in \Pi$. Then $v_1^\dagger, \dots, v_k^\dagger \subset \Pi^\dagger$.*

2.3. Real and complex varieties. We will need several properties of real and complex affine varieties. For complex varieties, a good introduction can be found in [7], while for real varieties [2] is a good source.

Let $K = \mathbb{R}$ or $\mathbb{C}$. Let $V \subset K^d$ be an irreducible variety, and let $I(K)$ be the ideal of polynomials in $K[x_1, \dots, x_d]$ that vanish on V. For $x \in K$, we define the Zariski tangent space of V at x to be

$$T_x V = \{v \in K^d : v \cdot \nabla f(x) = 0 \text{ for all } f \in I(V)\}.$$

We always have $\dim(T_x V) \geqslant \dim(V)$ (in the case $K = \mathbb{R}$, the dimension of a real variety is slightly subtle; see [2] for details). If $\dim(T_x V) \geqslant \dim(V)$, then x is a singular point of V. Let V_{sing} be the set of singular points of V and let $V_{\mathrm{reg}} = V \backslash V_{\mathrm{sing}}$ be the set of regular points of V. See [2, 7] for further background and details.

If $V \subset \mathbb{C}^d$ is a complex variety, let $V(\mathbb{R}) \subset \mathbb{R}^d$ be its real locus. In particular, if $V \subset \mathbb{C}^d$ is a one-dimensional variety, then $V(\mathbb{R})$ is a real variety of dimension at most 1. If $V \subset \mathbb{R}^d$ is a real variety, let $V^* \subset \mathbb{C}^d$ be the complexification of V—this is the smallest complex variety whose real locus contains V. If $p \in \mathbb{R}^d$ is a point, then $p^* \subset \mathbb{C}^d$ is the image of p under the usual embedding from $\mathbb{R}^d \to \mathbb{C}^d$.

Most of our arguments will occur over the reals. However, we will sometimes need to work over $\mathbb{C}$ in order to make use of the following result of Solymosi and Tao.

LEMMA 2.4 (Solymosi and Tao [9, Corollary 4.5]). *Let $V \subset \mathbb{C}^d$ be a k-dimensional algebraic variety of degree at most D. Then one can cover V by V_{reg} and $O_{D,d}(1)$ sets of the form W_{reg}, where each W is an algebraic variety contained in V of dimension at most $k-1$ and degree $O_{D,d}(1)$.*

Finally, if $P \in \mathbb{R}[x_1, \dots, x_d]$ is a polynomial, define $\mathbf{Z}(P) = \{x \in \mathbb{R}^d : P(x) = 0\}$.

https://doi.org/10.1017/fms.2015.34 Published online by Cambridge University Press

2.4. The discrete polynomial partitioning theorem. We will make crucial use of the discrete polynomial partitioning theorem of Guth and Katz.

THEOREM 2.1 (Guth and Katz [5, Theorem 4.1]). *Let $\mathcal{P}$ be a set of n points in $\mathbb{R}^d$. For each $D \geqslant 1$, there exists a polynomial P of degree $\leqslant D$ so that $\mathbb{R}^d \backslash \mathbf{Z}(P)$ is a union of $O_d(D^d)$ connected components (cells), and each cell contains $O_d(nD^{-d})$ points of $\mathcal{P}$.*

2.5. Connected components of real varieties. The discrete polynomial partitioning theorem will be used to partition the point-line arrangement (technically, the point-two-flat arrangement) into connected components, which are called cells. Each point lies in at most one cell. While each two-flat can enter several cells, the following theorem controls how many cells a given two-flat can enter.

THEOREM 2.2 (Barone and Basu [1], special case). *Let $Y \subset \mathbb{R}^d$ be a real variety of dimension e that can be defined by polynomials of degree at most D_1, and let $Z \subset \mathbb{R}^d$ be a real variety that can be defined by polynomials of degree at most D_2. Then $Y \backslash Z$ contains $O_{d,D_1}(D_2^e)$ connected components.*

REMARK 1. In the special case where $Y \subset \mathbb{R}^d$ is a two-flat, the above result also follows from the Milnor–Thom theorem. However, we will also be interested in the case where Y is defined by polynomials of larger (though still bounded) degree.

3. Proof of Lemma 1.1, Step 1: Lines in $\mathbb{C}^d$

Proof of Lemma 1.1. We prove the lemma by induction on $n_1 + \ell_1$. First, note that $I(\mathcal{P}_1, \mathcal{L}_1) \leqslant n_1^2 + \ell_1$. Thus if we select $C'_{d,\epsilon}$ sufficiently large, we can assume that $\ell_1 < \rho n_1^{2/(d+1)} \ell_1^{d/(d+1)}$; we can make $\rho > 0$ arbitrarily small by making the constant $C'_{d,\epsilon}$ larger.

Let P be a partitioning polynomial in $\mathbb{R}^{2d}$ of degree D adapted to the set $\iota(\mathcal{P}_1) \subset \mathbb{R}^{2d}$, as given by Theorem 2.1; there are $O_d(D^{2d})$ cells. We will choose D later, and it will depend only on d and ϵ. For each cell Ω, let ℓ_Ω be the number of lines $L \in \mathcal{L}_1$ so that $\iota(L) \cap \Omega \neq \emptyset$. By Theorem 2.2, we have

$$\sum_{\Omega} \ell_\Omega = O_d(D^2 \ell_1). \tag{7}$$

Apply the induction hypothesis inside each cell. Either property (A) holds inside some cell, or we have

https://doi.org/10.1017/fms.2015.34 Published online by Cambridge University Press

$$
\begin{aligned}
&I(\iota(\mathcal{P}_1)\backslash\mathbf{Z}(P), \iota(\mathcal{L}_1)) \\
&\quad = \sum_{\Omega} I(\iota(\mathcal{P}) \cap \Omega,\ \iota(\mathcal{L}_1)) \\
&\quad \leqslant \sum_{\Omega} C'_{d,\epsilon}(O_d(n_1 D^{-2d})^{(2+\epsilon)/(d+1)} \ell_\Omega^{d/(d+1)} + |\iota(\mathcal{P}_1) \cap \Omega| + \ell_\Omega) \\
&\quad \leqslant C'_{d,\epsilon}(O_d(1) D^{-(2d\epsilon)/(d+1)} n_1^{(2+\epsilon)/(d+1)} \ell_1^{d/(d+1)} + |\iota(\mathcal{P}_1)\backslash Z(P)| + O_d(D^2 \ell_1)) \\
&\quad \leqslant C'_{d,\epsilon}((O_d(1) D^{-(2d\epsilon)/(d+1)}) n_1^{(2+\epsilon)/(d+1)} \ell_1^{d/(d+1)} + |\iota(\mathcal{P}_1)\backslash Z(P)| \\
&\qquad + \tfrac{1}{4} n_1^{2/(d+1)} \ell_1^{d/(d+1)}) \\
&\quad \leqslant C'_{d,\epsilon}(\tfrac{1}{2} n_1^{(2+\epsilon)/(d+1)} \ell_1^{d/(d+1)} + |\iota(\mathcal{P}_1)\backslash Z(P)|). \qquad (8)
\end{aligned}
$$

Here we selected D sufficiently large (depending on d and ϵ) so that the term $O_d(1) D^{-2d\epsilon/(d+1)}$ is at most $\frac{1}{4}$, and we selected $C'_{d,\epsilon}$ sufficiently large (depending on D, d, and ϵ, which in turn depend only on d and ϵ) to guarantee that the term $O_d(D^2 \ell_1)$ is at most $\frac{1}{4} n_1^{2/(d+1)} \ell_1^{d/(d+1)}$.

Applying Lemma 2.4 to $\mathbf{Z}(P)^*$, we can find a collection $\mathcal{V}$ of $O_{d,D}(1)$ complex varieties in $\mathbb{C}^{2d}$ and sets $\{\mathcal{P}_V\}_{V\in\mathcal{V}}$ so that the following properties hold:

- Each variety in $\mathcal{V}$ is of degree $O_{d,D}(1)$.
- For each $V \in \mathcal{V}$, $\mathcal{P}_V \subset \iota(\mathcal{P}) \cap V(\mathbb{R})$.
- The sets $\{\mathcal{P}_V\}_{V\in\mathcal{V}}$ are disjoint, and $\bigcup_{V\in\mathcal{V}} \mathcal{P}_V = \iota(\mathcal{P}) \cap \mathbf{Z}(P)$.
- For each $p \in \mathcal{P}_V$, p^* is a regular point of V.

For each $V \in \mathcal{V}$, let

$$\mathcal{L}_V = \{L \in \mathcal{L}_1 : (\iota(L))^* \subset V\}.$$

One of the following must hold:

(A.1) There are a variety $V \in \mathcal{V}$ and a point $p \in \mathcal{P}_V$ so that

$$|\{L \in \mathcal{L}_V : p \in L\}| \geqslant \tfrac{1}{2}|\{L \in \mathcal{L}_1 : p \in L\}|. \qquad (9)$$

(B.1) Equation (9) fails for every $V \in \mathcal{V}$ and every $p \in \mathcal{P}_V$.

Suppose (A.1) holds for some variety V and some point $p \in \mathcal{P}_V$. Let $\Pi_0 \subset \mathbb{R}^{2d}$ be a $(2d-1)$-flat containing $(T_{p^*}V)(\mathbb{R})$. If $\iota(L)^* \subset V$ and $p \in L$, then $L \subset \Pi_0$. Thus $\iota(L) \subset \Pi_0$ for at least half of the complex lines in $\{L \in \mathcal{L}_1 : p \in L\}$. By Lemma 2.1, there exists a complex $(d-1)$-flat Π so that at least half of the complex lines in $\{L \in \mathcal{L}_1 : p \in L\}$ are contained in Π. We conclude that (A) holds, provided $c'_{d,\epsilon} \geqslant \frac{1}{2}$.

https://doi.org/10.1017/fms.2015.34 Published online by Cambridge University Press

Now suppose (B.1) holds. We will consider each $V \in \mathcal{V}$ in turn. If (A) holds for some point $p \in \mathcal{P}_V$, then we are done. If (A) fails for every point $p \in \mathcal{P}_V$, then we will establish the bound

$$I(\mathcal{P}_V, \iota(\mathcal{L}_1)) \leqslant C''_{d,\epsilon}(n_1^{(2+\epsilon)/(d+1)}\ell_1^{d/(d+1)} + |\mathcal{P}_V| + \ell_1). \tag{10}$$

Since there are $O_{d,\epsilon}(1)$ elements of $\mathcal{V}$, if we establish (10) for each $V \in \mathcal{V}$ then we will have proved the bound

$$I(\mathcal{P} \cap Z(P), \iota(\mathcal{L}_1)) \leqslant C'_{d,\epsilon}(\tfrac{1}{2}n_1^{(2+\epsilon)/(d+1)}\ell_1^{d/(d+1)} + |\mathcal{P} \cap Z(P)| + \ell_1), \tag{11}$$

provided $C'_{d,\epsilon}$ is chosen sufficiently large (depending only on d and ϵ), and this will conclude the proof of Lemma 1.1. It remains to prove (10) for each $V \in \mathcal{V}$; recall that by assumption, (B.1) holds and (A) fails.

Since (A) fails, by Lemma 2.3, for each $p \in \mathcal{P}_V$ and each $(d-1)$-flat $\Pi \subset \mathbb{R}^{2d}$ containing v, we have

$$|\{L \in \mathcal{L}_1 : p \in L, \dim(\Pi \cap \iota(L)) \geqslant 1\}| \leqslant c'_{d,\epsilon}|\{L \in \mathcal{L}_1 : p \in L\}|.$$

Since $\dim(\Pi \cap \iota(L)) = \dim(\Pi^* \cap \iota(L)^*)$ (where the left dim is the real dimension of the real algebraic set and the right dim is the complex dimension of the complex algebraic set), we have

$$|\{L \in \mathcal{L}_1 : p \in L, \dim(\Pi^* \cap \iota(L)^*) \geqslant 1\}| \leqslant c'_{d,\epsilon}|\{L \in \mathcal{L}_1 : p \in L\}|,$$

and thus

$$|\{L \in \mathcal{L}_1 : p \in L, \dim(\Pi^* \cap T_p(\iota(L)^* \cap V)) \geqslant 1\}| \leqslant c'_{d,\epsilon}|\{L \in \mathcal{L}_1 : p \in L\}|, \tag{12}$$

where $T_p(\iota(L)^* \cap V) \subset \mathbb{C}^{2d}$ is the Zariski tangent space of the (complex) variety $\iota(L)^* \cap V$ at p.

Let $\pi : \mathbb{R}^{2d} \to \mathbb{R}^d$ be a generic (with respect to $\mathcal{P}$, $\mathcal{L}$, and V) linear transformation. The transformation π also extends to a map from $\mathbb{C}^{2d} \to \mathbb{C}^d$. Since the projection of a linear algebraic variety cannot increase its dimension (this is actually true for all algebraic varieties, though we do not need this fact), we have

$$\begin{aligned}|\{L \in \mathcal{L}_1 : p \in L, \dim(\pi(\Pi^*) \cap \pi(T_p(\iota(L)^* \cap V))) \geqslant 1\}| \\ \leqslant c'_{d,\epsilon}|\{L \in \mathcal{L}_1 : p \in L\}|\end{aligned}$$

for every $(d-1)$-flat $\Pi \subset \mathbb{R}^{2d}$ containing p. This implies that

$$\begin{aligned}|\{L \in \mathcal{L}_1 : p \in L, \dim(\Pi^* \cap \pi(T_p(\iota(L)^* \cap V))) \geqslant 1\}| \\ \leqslant c'_{d,\epsilon}|\{L \in \mathcal{L}_1 : p \in L\}|\end{aligned} \tag{13}$$

for every $(d-1)$-flat $\Pi \subset \mathbb{R}^d$ containing p.

https://doi.org/10.1017/fms.2015.34 Published online by Cambridge University Press

Let Γ_0 be the set of irreducible components of $\pi(\iota(L)^* \cap V)$ as L ranges over the set $\mathcal{L}_1 \backslash \mathcal{L}_V$. We have

$$|\Gamma_0| = O_{d,D}(\ell_1) = O_{d,\epsilon}(1). \tag{14}$$

Furthermore, if $L \in \mathcal{L}_1$, then $O_{d,\epsilon}(1)$ curves from Γ_0 can be contained in $\pi(\iota(L)^*)$ (this is because the projection π was chosen generically with respect to $\mathcal{L}$). Since there is at most one complex line from $\mathcal{L}_1$ passing through any two points in $\mathbb{C}^d$, there are at most $O_{d,\epsilon}(1)$ curves from Γ_0 passing through any two points in $\mathbb{C}^{2d}$.

Let $\mathcal{P}' = \pi(\iota(\mathcal{P}_V))$; technically $\mathcal{P}' \subset \mathbb{C}^d$, but the points in $\mathcal{P}'$ are real, that is, all coordinates are real. As noted above, the arrangement $(\mathcal{P}', \Gamma_0)$ has two degrees of freedom and multiplicity type $O_{d,\epsilon}(1)$—this means that any two curves intersect in at most $O_{d,\epsilon}(1)$ points, and if we fix two points from $\mathcal{P}'$ then at most $O_{d,\epsilon}(1)$ curves pass through both of them.

We will need the following lemma.

LEMMA 3.1. *Let $\mathcal{P}_2 \subset \mathbb{R}^d$ be a collection of n_2 points. Let Γ_0 be a collection of ℓ_2 irreducible complex curves in $\mathbb{C}^d$, each of degree at most C. Let $\Gamma = \{\alpha(\mathbb{R}) : \alpha \in \Gamma_0\}$. Suppose that $(\mathcal{P}_2, \Gamma)$ has two degrees of freedom and multiplicity type s. Then at least one of the following must hold.*

(A.2) *There are a point $p \in \mathcal{P}_2$ and a $(d-1)$-flat $\Pi \subset \mathbb{R}^d$ containing p so that*

$$|\{\gamma \in \Gamma : p \in \gamma, \dim(T_p\gamma \cap \Pi) \geqslant 1\}| \geqslant \tfrac{1}{2}|\{\gamma \in \Gamma : p \in \gamma\}|, \tag{15}$$

where $T_p\gamma$ is the (real) Zariski tangent space of γ at p (see Section 2.3 for a definition of the Zariski tangent space).

(B.2)

$$I(\mathcal{P}_2, \Gamma) \leqslant C''_\epsilon(n_2^{(2+\epsilon)/(d+1)}\ell_2^{d/(d+1)} + n_2 + \ell_2), \tag{16}$$

where the constant C''_ϵ depends on ϵ, d, C, and s.

To avoid breaking the flow of the argument, we will defer the proof of Lemma 3.1 to Section 4.

For each $V \in \mathcal{V}$, apply Lemma 3.1 to the collection $(\mathcal{P}_V, \Gamma_0)$. If (B.2) holds, then

$$I(\mathcal{P}_V, \iota(\Gamma_0)) \leqslant C''_\epsilon(|\mathcal{P}_V|^{(2+\epsilon)/(d+1)}|\Gamma_0|^{d/(d+1)} + |\mathcal{P}_V| + |\Gamma_0|).$$

Recall that the constant C''_ϵ depends only on d, ϵ, C, and s. The constants C and s in turn depend on D, which depends only on ϵ and d. Thus C''_ϵ depends only on ϵ and d. By (14), $|\Gamma_0| = O_{d,\epsilon}(\ell_1)$. We conclude that

$$I(\mathcal{P}_V, \iota(\mathcal{L}_1 \backslash \mathcal{L}_V)) \leqslant C''_{d,\epsilon}(n_1^{(2+\epsilon)/(d+1)}\ell_1^{d/(d+1)} + |\mathcal{P}_V| + \ell_1).$$

https://doi.org/10.1017/fms.2015.34 Published online by Cambridge University Press

Finally, since (A.1) fails, we have

$$I(\mathcal{P}_V, \iota(\mathcal{L}_1)) \leqslant 2I(\mathcal{P}_V, \iota(\mathcal{L}_1 \backslash \mathcal{L}_V)),$$

so (10) holds. Thus if (B.2) holds for each $V \in \mathcal{V}$ then we are done.

Now suppose (A.2) holds for some $V \in \mathcal{V}$ and some point $p_1 \in \mathcal{P}'$, and let $\Pi \subset \mathbb{R}^d$ be the corresponding $(d-1)$-flat. Then

$$|\{\gamma \in \Gamma : p \in \gamma, \dim(T_p\gamma^* \cap \Pi^*) \geqslant 1\}| \geqslant \tfrac{1}{2}|\{\gamma \in \Gamma : p \in \gamma\}|.$$

Since each $L \in \mathcal{L}$ contributes $O_{d,\epsilon}(1)$ curves to Γ, if we choose the constant $c'_{d,\epsilon}$ sufficiently small (depending only on d and ϵ), then

$$|\{L \in \mathcal{L} : \dim(T_p\pi(L^* \cap V) \cap \Pi^*) \geqslant 1\}| \geqslant c'_{d,\epsilon}|\{L \in \mathcal{L} : p \in L\}|,$$

which violates (13). This concludes the proof of Lemma 1.1, modulo the proof of Lemma 3.1. □

4. Proof of Lemma 1.1, Step 2: Proof of Lemma 3.1

Proof of Lemma 3.1. First, note that $I(\mathcal{P}, \Gamma) \leqslant sn_2^2 + \ell_2$. Arguing as above, we can assume $\ell_2 < \rho n_2^{(2+\epsilon)/(d+1)} \ell_2^{d/(d+1)}$ or (16) holds immediately; we can make $\rho > 0$ smaller by making the constant C''_ϵ larger.

We will prove the statement by induction on $n_2 + \ell_2$. Let P be a partitioning polynomial of degree D, as given by Theorem 2.1; there are $O_d(D^d)$ cells. By Theorem 2.2, we have

$$|\{(\gamma, \Omega) : \gamma \in \Gamma, \Omega \text{ a cell}, \gamma \cap \Omega \neq \emptyset\}| = O_{d,C}(D\ell_2). \tag{17}$$

Apply the induction hypothesis inside each cell. Either there is a point $p \in \mathcal{P}_2 \backslash \mathbf{Z}(P)$ satisfying property (A.2), or by the same calculation as in (8) we have

$$I(\mathcal{P}_2 \backslash \mathbf{Z}(P), \Gamma) \leqslant C''_\epsilon(n_2^{(2+\epsilon)/(d+1)} \ell_2^{d/(d+1)} + n_2). \tag{18}$$

It remains to count incidences involving points on $\mathbf{Z}(P) \cap \mathcal{P}_2$. Applying Lemma 2.4 to $\mathbf{Z}_{\mathbb{R}}(P)^*$, we can find a collection $\mathcal{W}$ of $O_D(1)$ complex algebraic varieties, and a partition $\{\mathcal{P}_W\}_{W \in \mathcal{W}}$ of $\mathcal{P}$ so that for each $W \in \mathcal{W}$, $\mathcal{P}_W \subset W_{\text{reg}}$.

By Bézout's theorem (recall, we are currently working with complex curves and varieties), for each $W \in \mathcal{W}$ we have

$$|\{(p, \alpha) \in \mathcal{P}_W \times \Gamma_0 : p^* \in \alpha, \alpha \not\subset W\}| = O_{C,D}(\ell_2). \tag{19}$$

The total contribution from terms of the form (19) summed over all $W \in \mathcal{W}$ is $O_{C,D}(\ell_2)$. If we select C''_ϵ large enough, we conclude that either

$$I(\mathcal{P}_2 \cap \mathbf{Z}(P), \Gamma) \leqslant C''_\epsilon \ell_2, \tag{20}$$

https://doi.org/10.1017/fms.2015.34 Published online by Cambridge University Press

or there are a variety $W \in \mathcal{W}$ and a point $p \in \mathcal{P}_W$ so that at least half of the curves $\{\alpha \in \Gamma_0 : p^* \in \alpha\}$ are contained in W. If the latter happens then

$$|\{\alpha \in \Gamma_0 : p^* \in \alpha, \dim(T_{p^*}\alpha \cap T_{p^*}(W)) \geqslant 1\}| \geqslant \tfrac{1}{2}|\{\alpha \in \Gamma^* : p^* \in \alpha\}|,$$

where $T_{p^*}(W))$ is the (complex) Zariski tangent space of W.

Let $\Pi \subset \mathbb{R}^d$ be a $(d-1)$-flat containing $(T_{p^*}(W))(\mathbb{R})$. Then (15) holds with this choice of p and Π. Thus, either (A.2) holds, or combining (18) and (20) we obtain (16); that is, (B.2) holds. □

Acknowledgments

The author would like to thank Márton Hablicsek for pointing out errors in an earlier version of this manuscript and to the two anonymous referees for their helpful suggestions. The author was supported by an NSF Postdoctoral Fellowship.

References

[1] S. Barone and S. Basu, 'Refined bounds on the number of connected components of sign conditions on a variety', *Discrete Comput. Geom.* **47**(3) (2012), 577–597.
[2] J. Bochnak, M. Coste and M.-F. Roy., *Real Algebraic Geometry* (Springer, Berlin, 1998).
[3] Z. Dvir and S. Gopi, 'On the number of rich lines in truly high dimensional sets', *31st International Symposium on Computational Geometry (SoCG 2015)*, *Leibniz International Proceedings in Informatics (LIPIcs), 34*, 584–598. (2015).
[4] L. Guth and N. Katz., 'Algebraic methods in discrete analogues of the Kakeya problem.', *Adv. Math.* **225** (2010), 2828–2839.
[5] L. Guth and N. Katz., 'On the Erdős distinct distance problem in the plane', *Ann. of Math. (2)* **181** (2015), 155–190.
[6] M. Hablicsek and Z. Scherr, 'On the number of rich lines in high dimensional real vector spaces', 2014, preprint arXiv:1412.7025.
[7] J. Harris, *Algebraic Geometry: A First Course*, Graduate Texts in Mathematics, 133 (Springer, New York, 1995).
[8] A. Sheffer and J. Zahl, 'Point-curve incidences in the complex plane', 2015, preprint arXiv:1502.07003.
[9] J. Solymosi and T. Tao, 'An incidence theorem in higher dimensions', *Discrete Comput. Geom.* **48**(2) (2012), 255–280.
[10] E. Szemerédi and W. T. Trotter Jr., 'Extremal problems in discrete geometry', *Combinatorica* **3**(3–4) (1983), 381–392.
[11] C. Tóth, 'The Szemerédi–Trotter theorem in the complex plane', *Combinatorica* **35**(1) (2015), 95–126.
[12] J. Zahl., 'A Szeméredi–Trotter type theorem in $\mathbb{R}^4$', *Discrete Comput. Geom.* **54**(3) (2015), 513–572.

https://doi.org/10.1017/fms.2015.34 Published online by Cambridge University Press