Contributions to Discrete Mathematics

Volume 8, Number 1, Pages 100–121
ISSN 1715-0868

AN IMPROVED BOUND ON THE NUMBER OF POINT-SURFACE INCIDENCES IN THREE DIMENSIONS

JOSHUA ZAHL

Abstract. We show that m points and n smooth algebraic surfaces of bounded degree in $\mathbb{R}^3$ satisfying suitable non-degeneracy conditions can have at most $O(m^{2k/(3k-1)}n^{(3k-3)/(3k-1)}+m+n)$ incidences, provided that any collection of k points have at most $O(1)$ surfaces passing through all of them, for some $k \geq 3$. In the case where the surfaces are spheres and no three spheres meet in a common circle, this implies there are $O((mn)^{3/4}+m+n)$ point-sphere incidences. This is a slight improvement over the previous bound of $O((mn)^{3/4}\beta(m,n)+m+n)$ for $\beta(m,n)$ an (explicit) very slowly growing function. We obtain this bound by using the discrete polynomial ham sandwich theorem to cut $\mathbb{R}^3$ into open cells adapted to the set of points, and within each cell of the decomposition we apply a Turan-type theorem to obtain crude control on the number of point-surface incidences. We then perform a second polynomial ham sandwich decomposition on the irreducible components of the variety defined by the first decomposition. As an application, we obtain a new bound on the maximum number of unit distances amongst m points in $\mathbb{R}^3$.

1. Introduction

In [6], Clarkson, Edelsbrunner, Guibas, Sharir, and Welzl obtained the following bound on the number of incidences between points and spheres in $\mathbb{R}^3$.

Theorem 1.1 (Clarkson et al.). *The number of incidences between m points and n spheres in $\mathbb{R}^3$ with no three spheres meeting at a common circle is*

$$O\left((mn)^{3/4}\beta(m,n)+m+n\right), \tag{1.1}$$

where $\beta(m,n)$ is a very slowly growing function of m and n. In particular, $\beta(m,n)\leq 2^{C\alpha(m^3/n)^2}$, where $\alpha(s)$ is the inverse Ackerman function and C is a large constant.

We obtain the following slight sharpening.

Received by the editors May 2, 2011, and in revised form September 3, 2012.
2010 *Mathematics Subject Classification.* 52C10.
Key words and phrases. incidence geometry, polynomial ham sandwich theorem, polynomial method, Erdős unit distance problem.

Theorem 1.2. *Let $k \geq 3$, and let $\mathcal{P} \subset \mathbb{R}^3$ be a collection of m points and $\mathcal{S}$ a collection of n smooth algebraic surfaces of bounded degree (the degree is allowed to depend on k) such that for some constant C we have $|S \cap S' \cap S''| \leq C$ for all $S, S', S'' \in \mathcal{S}$, and for any collection of k points in $\mathbb{R}^3$, there are at most C surfaces that contain all k points. Then the number of incidences between points in $\mathcal{P}$ and surfaces in $\mathcal{S}$ is*

$$O\left(m^{\frac{2k}{3k-1}} n^{\frac{3k-3}{3k-1}} + m + n\right), \tag{1.2}$$

where the implicit constant depends only on k, C, and the degree of the algebraic surfaces.

In particular, the number of incidences between m points and n spheres in $\mathbb{R}^3$ with no three spheres meeting at a common circle is

$$O\left((mn)^{\frac{3}{4}} + m + n\right). \tag{1.3}$$

Remark 1.3. The requirement that every three surfaces meet in a complete intersection, or some variant thereof, is necessary to prevent the situation in which all of the surfaces meet in a common curve and all of the points lie on that curve, yielding mn incidences (i.e. if we don't place any restrictions on how the surfaces can intersect, then the trivial bound of mn incidences is sharp).

Remark 1.4. When $k = 2$ and $n = m$, the following example shows that Theorem 1.2 is sharp. Let $\mathcal{P}$ be the set $[-2k, 2k]^2 \times [0, 2k^2]$, and let

$$\mathcal{S} = \{z = (x - x_0)^2 + (y - y_0)^2 + z_0 \mid x_0, y_0 = -k, \dots, k;\ z_0 = 0, \dots, k^2\}.$$

Then $|\mathcal{P}| = 32k^4$, $|\mathcal{S}| = k^4$, and we can verify that for every triple S, S', S'' of surfaces in $\mathcal{S}$, we have $|S \cap S' \cap S''| \leq 8$, and for every three points of $\mathcal{P}$, there are at most four surfaces from $\mathcal{S}$ that contain all three. Since each $S \in \mathcal{S}$ hits at least k^2 points from $\mathcal{P}$, there are at least k^6 incidences total.

Remark 1.5. The requirements that every three surfaces meet in C points and that every k points have at most C surfaces passing through them are analogous to the definition of "curves with k degrees of freedom" from [18], though in [18] the curves do not need to be algebraic.

Remark 1.6. Theorem 1.1 can be extended to the more general case of bounded degree algebraic surfaces using the decomposition techniques described in [1, §8.3] to obtain an analogue of (1.2). Doing so yields a bound of

$$O\left(m^{\frac{2k}{3k-1}} n^{\frac{3k-3}{3k-1}} \beta(m, n) + m + n\right),$$

where β is a slowly growing function.

1.1. **Previous results.** Similar results to Theorem 1.1 and 1.2 have been obtained by Łaba and Solymosi in [16] and by Iosevich, Jorati, and Łaba in [12]. In [16] and [12], however, the authors consider a more general class of surfaces (they need not be algebraic), but they require that the point set be "homogeneous" in a suitable sense.

Our techniques do not work well when $k = 2$, i.e. for obtaining bounds on point-hyperplane incidences, but this case has been studied by other authors (see e.g. [7], where the authors obtain sharp bounds on point-hyperplane incidences under a slightly different set of non-degeneracy conditions).

1.2. **Update 7/4/2011.** The author has recently become aware that concurrently with this paper, Kaplan, Matoušek, Safernová, and Sharir in [13] obtained results similar to the bound (1.3) using similar methods. Kaplan et. al. are able to avoid some of the technical difficulties present in this paper by using an explicit parameterization of the sphere by rational functions.

1.3. **Proof sketch.** Clarkson et al. obtain Theorem 1.1 through their "Canham threshold plus divide and conquer" technique: the arrangement of spheres in $\mathbb{R}^3$ is subdivided into smaller collections through a careful partitioning of $\mathbb{R}^3$, and the number of incidences between these smaller collections of spheres and points is controlled by a Turan-type bound on the number of edges in a bipartite graph with certain forbidden subgraphs.

In this paper, we employ similar ideas, except instead of dividing the problem into smaller subproblems by partitioning $\mathbb{R}^3$ into cells using a decomposition adapted to the collection of spheres (or more general nonsingular algebraic surfaces), we employ a partition adapted to the collection of points. This partition is obtained from the discrete polynomial ham sandwich theorem recently used to great effect by Guth and Katz in [11] and more recently by Solymosi and Tao in [19] and by Kaplan, Matoušek, and Sharir in [14]. Specifically, we find a polynomial P such that the complement of the zero set of P consists of open "cells," none of which contain too many points. We can then apply a Turan-type bound to the points and surfaces inside each cell. However, some points may lie on the zero set of P, and thus do not lie in any of the cells. To deal with these points, we perform a second polynomial ham sandwich decomposition to find a polynomial Q whose zero set partitions the zero set of P into cell-like objects, and we apply the Turan-type bound to each of these "cells." While it is possible that a point could lie in the zero set of both P and Q, we can use Bézout-type theorems to control how often this can occur.

1.4. **Some difficulties with real algebraic sets.** There are several technical difficulties that have to be dealt with while executing the above strategy. In contrast to the situation over $\mathbb{C}$, there exist polynomials $P_1, \ldots, P_d \in \mathbb{R}[x_1, \ldots, x_d]$ of degrees $D_1, \ldots, D_d$ such that $\{x \in \mathbb{R}^d | P_1(x) = 0\} \cap \cdots \cap \{x \in \mathbb{R}^d | P_d(x) = 0\}$ contains more than $D_1, \ldots, D_d$ isolated points, i.e. the naïve analogue of Bézout's theorem fails over $\mathbb{R}$. To deal with this problem, we will sometimes be forced to embed our varieties into $\mathbb{C}$ and use the (usual) Bézout's theorem (though we have to be careful that the intersection of the embedded varieties does not contain new, unexpected components of positive dimension).

A second difficulty concerns the failure of the Nullstellensatz for varieties defined over $\mathbb{R}$. In contrast to the complex case, if (P) is a principal prime ideal and Q is a real polynomial, it need not be the case that if Q vanishes identically on $\{x \in \mathbb{R}^d \mid P(x) = 0\}$ then $Q \in (P)$. Luckily, there is a special type of ideal known as a "real ideal" for which an analogue of the Nullstellensatz does hold. Frequently we will be required to replace our polynomials with new polynomials that generate real ideals.

Finally, if $P \in \mathbb{R}[x_1, \dots, x_d]$, then the dimension of $\{x \in \mathbb{R}^d \mid P(x) = 0\}$ may be less than $d - 1$, and even if P is square-free, ∇P may vanish on $\{x \in \mathbb{R}^d | P(x) = 0\}$. Again, we can remedy this problem by working with (irreducible) polynomials that generate real ideals.

2. Main Result

2.1. **Notation.** Throughout the paper, c and C will denote sufficiently small and large constants, respectively, which are allowed to vary from line to line. We will write $A \lesssim B$ to mean $A < CB$, and we say that a quantity A is $O(B)$ if $A \lesssim B$.

Let $\mathcal{S}$ be a collection of smooth (real) surfaces and $\mathcal{P}$ a collection of points. Then $\mathcal{I}(\mathcal{P}, \mathcal{S})$ is the number of incidences between the surfaces in $\mathcal{S}$ and the points in $\mathcal{P}$. If $S \in \mathcal{S}$ is a surface, then f_S is the polynomial whose zero set is S.

All ideals and varieties will be assumed to be affine. Unless otherwise specified, all ideals are subsets of $\mathbb{R}[x_1, \dots, x_d]$, and all varieties are defined over $\mathbb{R}$ and thus are subsets of $\mathbb{R}^d$, though sometimes we will specialize to the case $d = 3$. If P is a polynomial, $(P) \subset \mathbb{R}[x_1, \dots, x_d]$ is the ideal generated by P.

Special emphasis will be placed on "real ideals." These are described in Definition A.2 of Appendix A, and they should not be confused with ideals that are merely subsets of $\mathbb{R}[x_1, \dots, x_d]$. On the other hand, a "real variety" is merely a variety defined over $\mathbb{R}$ (as opposed to $\mathbb{C}$).

If I is an ideal, we use

$$\mathbf{Z}(I) = \{x \in \mathbb{R}^d \mid P(x) = 0 \text{ for all } P \in I\}$$

to denote the zero set of I. If P is a polynomial we shall abuse notation and use $\mathbf{Z}(P)$ to denote $\mathbf{Z}((P)) = \{x \in \mathbb{R}^d \mid P(x) = 0\}$. If $Z \subset \mathbb{R}^d$ is a real variety, then we define

$$\mathbf{I}(Z) = \{P \in \mathbb{R}[x_1, \dots, x_d] \mid P(x) = 0 \text{ for all } x \in Z\}$$

to be the ideal of polynomials that vanish on Z.

If $Z \subset \mathbb{R}^d$ is a real variety, then $Z^* \subset \mathbb{C}^d$ denotes the smallest complex variety containing Z. Conversely, if $Z \subset \mathbb{C}^d$ is a complex variety, then $\mathfrak{R}(Z) \subset \mathbb{R}^d$ is its set of real points.

If $\mathcal{Q} \subset \mathbb{R}[x_1, \dots, x_d]$ is a collection of real polynomials, then we can partition $\mathbb{R}^d \backslash \bigcup_{Q \in \mathcal{Q}} \mathbf{Z}(Q)$ into a collection of open sets such that on each open set, the polynomials from $\mathcal{Q}$ do not change sign. These sets will be called the *realizations of realizable strict sign conditions of* $\mathcal{Q}$. Similarly, if $Z \subset \mathbb{R}^d$ is a variety, then we can consider the restriction of the above open sets to Z, and these are called the *realizations of realizable strict sign conditions of* $\mathcal{Q}$ *on* Z. These notions are defined more precisely in Appendix A.

2.2. **Preliminaries.** Following [6], we shall need the following Turan-type bound:

Theorem 2.1 (Kővari, Sós, Turan [15])**.** *Let s, t be fixed, and let $G = G_1 \sqcup G_2$ be a bipartite graph with $|G_1| = m$, $|G_2| = n$ that contains no copy of $K_{s,t}$. Then G has at most $O(nm^{1-1/s} + m)$ edges. Symmetrically, G has at most $O(mn^{1-1/t} + n)$ edges. All implicit constants depend only on s and t.*

In our case, we have that $|S \cap S' \cap S''| \leq C$ for every three surfaces S, S', S'', and any k points have at most C surfaces passing through all of them. Thus we have the bounds

$$\mathcal{I}(\mathcal{P}, \mathcal{S}) \lesssim |\mathcal{P}||\mathcal{S}|^{1-1/k} + |\mathcal{S}|, \tag{2.1}$$

$$\mathcal{I}(\mathcal{P}, \mathcal{S}) \lesssim |\mathcal{P}|^{2/3}|\mathcal{S}| + |\mathcal{P}|. \tag{2.2}$$

Recall the discrete polynomial partitioning theorem from [11]:

Theorem 2.2. *Let $\mathcal{P}$ be a collection of points in $\mathbb{R}^d$, and let $D > 0$. Then there exists a non-zero polynomial P of degree at most D such that each connected component of $\mathbb{R}^d \backslash \mathbf{Z}(P)$ contains $O(|\mathcal{P}|/D^d)$ points of $\mathcal{P}$.*

Remark 2.3. Without loss of generality, we can assume that P is square-free. Indeed if P is not square-free then we can replace P by its square-free part, and the new polynomial still has all of the desired properties.

Example 2.4. Consider the set of 24 points

$$\mathcal{P}_1 = \{(0, \pm 1, \pm 1), (0, \pm 2, \pm 2), (\pm 1, \pm 1, \pm 1), (\pm 2, \pm 2, \pm 2)\} \subset \mathbb{R}^3,$$

and let $D = 3$. Then the polynomial $P_1(x_1, x_2, x_3) = x_1x_2x_3$ partitions $\mathbb{R}^3$ into 8 octants, each of which contains 2 points from $\mathcal{P}_1$.

Remark 2.5. Note as well that in the above example, the 8 points $\{0, \pm 1, \pm 1\}$, $\{0, \pm 2, \pm 2\}$ lie on the set $\mathbf{Z}(P_1)$ and thus they do not lie inside any of the open components of $\mathbb{R}^3 \backslash \mathbf{Z}(P_1)$. This is not merely a consequence of us choosing P_1 poorly; it is an unavoidable phenomena that occurs when performing the discrete polynomial partitioning decomposition. In order to control the number of incidences between points lying on $\mathbf{Z}(P_1)$ and surfaces in $\mathcal{S}$, we shall have to perform a second polynomial partitioning decomposition "on" the surface $\mathbf{Z}(P_1)$. For technical reasons, we cannot simply consider the complement of the zero set of our second partitioning polynomial as a union of relatively open subsets of $\mathbf{Z}(P_1)$. Instead, we need to perform a somewhat more detailed decomposition that partitions $\mathbf{Z}(P_1)$ into sets that are realizations of realizable strict sign conditions of a certain family of polynomials. This is made precise in the theorem below. See Appendix A for the definition of a real ideal, a strict sign condition, and the realization of a strict sign condition.

Theorem 2.6 (Discrete polynomial partitioning on a hypersurface)**.** *Let $P \in \mathbb{R}[x_1, \ldots, x_d]$ be an irreducible polynomial of degree D such that (P) is a real ideal, and let $\mathcal{P}$ be a collection of points contained in $Z = \mathbf{Z}(P) \subset \mathbb{R}^d$. Let $E \geq D$. Then there exists a collection of polynomials $\mathcal{Q} \subset \mathbb{R}[x_1, \ldots, x_d]$ with the following properties:*

(1) $|\mathcal{Q}| \leq \log_2(DE^{d-1}) + O(1)$,

(2) $\sum_{\mathcal{Q}} \deg Q \lesssim E$,

(3) *None of the polynomials in $\mathcal{Q}$ vanish identically on Z,*

(4) *The realization of each of the $O(DE^{d-1})$ strict sign conditions of $\mathcal{Q}$ on Z contains $O\left(\frac{|\mathcal{P}|}{DE^{d-1}}\right)$ points of $\mathcal{P}$.*

All implicit constants depend only on d.

We shall defer the proof of Theorem 2.6 to Appendix B. In our applications, we will always have $d = 3$.

Example 2.7. Let us continue Example 2.4. The polynomial P_1 from Example 2.4 was not irreducible, but we can factor it into the three irreducible factors x_1, x_2, x_3. All of the points lying on $\mathbf{Z}(P_1)$ actually lie on the irreducible component $\mathbf{Z}(x_1)$, so we let $P_2(x_1, x_2, x_3) = x_1$. Note that $(P_2) = (x_1)$ is a real ideal, and $D = \deg(P_2) = 1$. Select $E = 2$ (which is larger than D). Then the collection of polynomials $\mathcal{Q} = \{x_2, x_3\}$ satisfies the requirements of Theorem 2.6. The realizations of realizable strict sign conditions of $\mathcal{Q}$ on Z are the 4 sets of the form

$$\{(x_1, x_2, x_3) \,|\, x_1 = 0,\ \pm x_2 > 0,\ \pm x_3 > 0\}. \tag{2.3}$$

Note that each of these sets contains 2 points of $\mathcal{P}_1 \cap \mathbf{Z}(P_2)$. Two coincidences occur in this example that are not present in general. First, in this example the realizations of the four strict sign conditions of $\mathcal{Q}$ on $\mathbf{Z}(P_2)$ correspond to the four connected components of $\mathbf{Z}(P_2) \backslash \bigcup_{\mathcal{Q}} \mathbf{Z}(Q)$. In general, each realization of a strict sign condition may be a union of multiple connected components of $\mathbf{Z}(P_2) \backslash \bigcup_{\mathcal{Q}} \mathbf{Z}(Q)$. Second, each of the polynomials in $\mathcal{Q}$ were irreducible factors of P_1. In general this does not occur.

2.3. **Proof of Theorem 1.2.** We are now ready to prove Theorem 1.2. For the reader's convenience, we will restate the theorem below.

Theorem 1.2. *Let $k \geq 3$, and let $\mathcal{P} \subset \mathbb{R}^3$ be a collection of m points and $\mathcal{S}$ a collection of n smooth algebraic surfaces of bounded degree (the* *degree is allowed to depend on k) such that for some constant C we have $|S \cap S' \cap S''| \leq C$ for all $S, S', S'' \in \mathcal{S}$, and for any collection of k points in $\mathbb{R}^3$, there are at most C surfaces that contain all k points. Then the number of incidences between points in $\mathcal{P}$ and surfaces in $\mathcal{S}$ is*

$$O\left(m^{\frac{2k}{3k-1}} n^{\frac{3k-3}{3k-1}} + m + n\right), \tag{1.2}$$

where the implicit constant depends only on k, C, and the degree of the algebraic surfaces.

Proof. From (2.1) and (2.2), we have that if $n > cm^k$ or $m > cn^3$ for some fixed small constant $c > 0$ to be specified later, then Theorem 1.2 immediately holds. Thus we may assume

$$n < cm^k \text{ and } m < cn^3. \tag{2.4}$$

We may also assume that the surfaces in $\mathcal{S}$ are irreducible varieties. Indeed, if this were not the case, then we could let $\mathcal{S}'$ be the set of all irreducible components of surfaces in $\mathcal{S}$. We would have $|\mathcal{S}'| \lesssim |\mathcal{S}|$, and the surfaces in $\mathcal{S}'$ would satisfy the same bounds as the surfaces in $\mathcal{S}$. We could then run our arguments below with $\mathcal{S}'$ in place of $\mathcal{S}$.

Let P be a square-free polynomial of degree at most D (D will be determined later, but the impatient reader can jump to (2.23)) that cuts $\mathbb{R}^3$ into $O(D^3)$ cells with $O(m/D^3)$ points in each cell, and let $Z = \mathbf{Z}(P)$. Let m_i be the number of points lying in the i–th cell of the above decomposition, and let n_i be the number of surfaces that meet the interior of the i–th cell.

Lemma 2.8.

$$\sum n_i \lesssim D^2 n. \tag{2.5}$$

Proof. Let S be a surface that is not contained in Z. Since there are finitely many cells, we can select a large closed ball $B \subset \mathbb{R}^3$ so that the number of cells that meet S is equal to the number of cells that meet $S \cap B$. We can apply a small generic translation to S, and doing so can only increase the number of cells that meet $S \cap B$ (and thus can only increase the number of cells that meet S). Select a generic vector $v \in R^3$ and let $T(x) = v \wedge \nabla f_S(x) \wedge \nabla P(x)$, so if $x \in S \cap Z$ and $\nabla f_S(x)$ and $\nabla P(x)$ are both non-zero and non-collinear, then $T(x) = 0$ if the curve $S \cap Z$ is tangent at x to a plane with normal vector v.

For every cell Ω that meets S, either Ω contains an entire connected component of S (since S has bounded degree, at most $O(1)$ cells can contain an entire connected component of S), or there is a point $x \in \partial\Omega \cap S$ satisfying the following properties.

(1) x is a smooth point of the space curve $Z \cap S$,
(2) x is a non-singular intersection point of $\mathbf{Z}(T) \cap Z \cap S$,
(3) x is a smooth point of $\partial\Omega$.

These three properties follow from the fact that v is generic and we picked a generic translation of S. From Item 3, each point x satisfying the above properties can be associated to at most two distinct cells. By Item 2 and the real Bézout inequality (see e.g. [3, §4.7]), there can be at most $\deg(P)\deg(T)\deg(f_s) = O(D^2)$ such points, and thus S can enter at most $O(D^2)$ such cells. Since there are n surfaces in $\mathcal{S}$, the result follows. $\square$

Using Lemma 2.8 and the bound from (2.1) we can control the number of incidences between points not lying in Z and surfaces in $\mathcal{S}$:

$$\begin{aligned}\mathcal{I}(\mathcal{P}\backslash Z, \mathcal{S}) &= \sum_i \mathcal{I}(\mathcal{P}\cap\Omega_i, \mathcal{S})\\ &\lesssim \sum_i m_i n_i^{1-\frac{1}{k}} + n_i\\ &\lesssim \left(\sum_i m_i^k\right)^{\frac{1}{k}}\left(\sum_i n_i\right)^{1-\frac{1}{k}} + D^2 n \qquad (2.6)\\ &\lesssim \left(D^3\frac{m^k}{D^{3k}}\right)^{\frac{1}{k}}(D^2n)^{1-\frac{1}{k}} + D^2n\\ &\lesssim \frac{mn^{1-1/k}}{D^{1-1/k}} + D^2 n.\end{aligned}$$

We must now control $\mathcal{I}(\mathcal{P}\cap Z, \mathcal{S})$. We have

$$\mathcal{I}(\mathcal{P}\cap Z, \mathcal{S}) = \mathcal{I}(\mathcal{P}\cap Z, \mathcal{S}_1) + \mathcal{I}(\mathcal{P}\cap Z, \mathcal{S}_2), \qquad (2.7)$$

where $\mathcal{S}_1$ is the set of surfaces contained in Z, and $\mathcal{S}_2$ are the remaining surfaces. Since Z has degree D, Z can contain at most D surfaces from $\mathcal{S}$, i.e. $|\mathcal{S}_1| \leq D$. By (2.2),

$$\begin{aligned}\mathcal{I}(\mathcal{P}\cap Z, \mathcal{S}_1) &\lesssim |\mathcal{S}_1|\ |\mathcal{P}|^{\frac{2}{3}} + |\mathcal{P}|\\ &\lesssim Dm^{\frac{2}{3}} + m.\end{aligned} \qquad (2.8)$$

Thus it remains to control $\mathcal{I}(\mathcal{P}\cap Z, \mathcal{S}_2)$. Write $P = P_1\cdots P_\ell$, where each P_j is irreducible of degree D_j, and let $Z_j = \mathbf{Z}(P_j)$. Thus we have $D_1 + \cdots + D_\ell \leq D$, and $Z = \bigcup Z_j$. We would like to use Lemma 2.6 to perform a second discrete polynomial ham sandwich decomposition on each variety Z_j, but if (P_j) is not a real ideal then we cannot apply the lemma. Luckily, the following lemma lets us remedy this situation.

Lemma 2.9. *Let $\mathcal{A}\subset\mathbb{R}[x_1,\ldots,x_d]$ be a collection of irreducible polynomials. Then we can find a new collection $\mathcal{A}'$ of irreducible polynomials such that:*

(1) $\bigcup_{P\in\mathcal{A}}\mathbf{Z}(P)\subset\bigcup_{P\in\mathcal{A}'}\mathbf{Z}(P)$,

(2) $\sum_{P\in\mathcal{A}}\deg P \leq \sum_{P\in\mathcal{A}'}\deg P$,

(3) (P) *is a real ideal for each* $P\in\mathcal{A}'$.

Proof. We shall proceed by induction on $\sum_{P\in\mathcal{A}}\deg P$. If the sum is 1, then the result is trivial since in that case $\mathcal{A}$ consists of a single linear polynomial, so we can let $\mathcal{A}'=\mathcal{A}$. Suppose the lemma has been established for all families $\widetilde{\mathcal{A}}$ with $\sum_{P\in\widetilde{\mathcal{A}}}\deg P<w$, and let $\sum_{P\in\mathcal{A}}\deg P=w$. If (P) is a real ideal for every $P\in\mathcal{A}$ then the result is immediate. If not, select $P_0\in\mathcal{A}$ such that (P_0) is not a real ideal. By Proposition A.3 in Appendix A, ∇P_0 vanishes on $\mathbf{Z}(P_0)$. Let $v\in\mathbb{R}^d$ be a generic unit vector. Then $\mathbf{Z}(P_0)\subset\mathbf{Z}(\nabla_v P_0)$ and $\deg(\nabla_v P_0)<\deg P_0$. Write $\nabla_v P_0=Q_1\cdots Q_a$ as a product of irreducible components, and let $\widetilde{\mathcal{A}}=\mathcal{A}\backslash\{P_0\}\cup\{Q_1,\ldots,Q_a\}$. We have $\sum_{P\in\widetilde{\mathcal{A}}}\deg P<\sum_{P\in\mathcal{A}}\deg P=w$, and $\bigcup_{P\in\mathcal{A}}\mathbf{Z}(P)\subset\bigcup_{P\in\widetilde{\mathcal{A}}}\mathbf{Z}(P)$. Apply the induction hypothesis to $\widetilde{\mathcal{A}}$ to obtain a family $\widetilde{\mathcal{A}}'$ satisfying Properties 1–3 with $\widetilde{\mathcal{A}}$ in place of $\mathcal{A}$. We can verify that $\widetilde{\mathcal{A}}'$ has the desired properties. □

After applying Lemma 2.9, we can assume that each irreducible polynomial P_j in the factorization of P generates a real ideal. Write $\mathcal{P}\cap Z=\bigsqcup\mathcal{P}_j$, where $\mathcal{P}_j$ consists of those points lying in Z_j. If a point lies on two or more such varieties, place it into only one of the sets. We need to distinguish between several cases. Let

$$
\begin{aligned}
\mathcal{A}_1&=\{j\mid |\mathcal{P}_j|^k<D_j^k n\},\\
\mathcal{A}_2&=\{j\mid D_j^k n\le|\mathcal{P}_j|^k<D_j^{3k-1}n\},\\
\mathcal{A}_3&=\{j\mid |\mathcal{P}_j|^k\ge D_j^{3k-1}n\}.
\end{aligned}
\tag{2.9}
$$

For each $j\in\mathcal{A}_1$ we have

$$
\begin{aligned}
\mathcal{I}(\mathcal{P}_j,\mathcal{S}_2)&\lesssim|\mathcal{P}_j|n^{1-\frac{1}{k}}+n\\
&\lesssim D_j n,
\end{aligned}
\tag{2.10}
$$

where the second inequality uses the assumption $|\mathcal{P}_j|<D_j n^{1/k}$. Summing (2.10) over all $j\in\mathcal{A}_1$, we obtain

$$
\begin{aligned}
\mathcal{I}\left(\bigcup_{j\in A_1}\mathcal{P}_j,\mathcal{S}_2\right)&\lesssim\sum_{\mathcal{A}_1}D_j n\\
&\le Dn.
\end{aligned}
\tag{2.11}
$$

Now we must control the incidences between surfaces and points lying on varieties Z_j, $j\in\mathcal{A}_2$ or $j\in\mathcal{A}_3$. If $j\in\mathcal{A}_2$, use Theorem 2.2 to select a square-free polynomial Q_j of degree at most E_j,

$$
E_j=\left(\frac{|\mathcal{P}_j|^k}{nD_j^k}\right)^{\frac{1}{2k-1}},
\tag{2.12}
$$

that cuts $\mathbb{R}^3$ into $O(E_j^3)$ cells, each of which contains $O(|\mathcal{P}_j|/E_j^3)$ points of $\mathcal{P}_j$. Recall that P_j is irreducible, (P_j) is real, and $j \in \mathcal{A}_2$ implies $\deg(Q_j) \leq E_j < \deg(P_j)$. Thus Q_j does not vanish identically on Z_j. Let $\mathcal{Q}_j = \{Q_j\}$ and let $W_j = \mathbf{Z}(Q_j)$.

If $j \in \mathcal{A}_3$, let E_j be as in (2.12) and use Theorem 2.6 (with $E = E_j$) to find a family $\mathcal{Q}_j$ of polynomials satisfying properties 1–4 of the theorem. In particular, the realizations of the realizable strict sign conditions of $\mathcal{Q}_j$ on Z_j partition Z_j into $O(D_j E_j^2)$ (not necessarily connected) sets, each of which contains $O\big(|\mathcal{P}_j|/D_j E_j^2\big)$ points, plus the "boundary" $Z_j \cap \bigcup_{\mathcal{Q}_j} \mathbf{Z}(Q)$. Define $W_j = \bigcup_{\mathcal{Q}_j} \mathbf{Z}(Q)$ (thus the definition of W_j depends on whether $j \in \mathcal{A}_2$ or $j \in \mathcal{A}_3$).

Regardless of whether $j \in \mathcal{A}_2$ or $\mathcal{A}_3$, we have

$$\mathcal{I}(\mathcal{P}_j, \mathcal{S}_2) = \mathcal{I}(\mathcal{P}_j \backslash W_j, \mathcal{S}_2) + \mathcal{I}(\mathcal{P}_j \cap W_j, \mathcal{S}_2). \tag{2.13}$$

We shall begin by bounding the first term of (2.13). If $j \in \mathcal{A}_2$, then through the same computation performed in (2.6) we have

$$\begin{aligned} \mathcal{I}(\mathcal{P}_j \backslash W_j, \mathcal{S}_2) &\lesssim \frac{|\mathcal{P}_j| n^{1-1/k}}{E_j^{1-1/k}} + nE_j^2 \\ &\leq \frac{|\mathcal{P}_j| n^{1-1/k}}{E_j^{1-1/k}} + nD_jE_j. \end{aligned} \tag{2.14}$$

If $j \in \mathcal{A}_3$, then let Ω_{ij} be the realization of the i–th realizable strict sign condition of $\mathcal{Q}_j$ on Z_j. Recall that there are $O(D_j E_j^2)$ such realizable strict sign conditions. Let $m_{ij} = |\mathcal{P}_j \cap \Omega_{ij}|$, and let n_{ij} be the number of surfaces in $\mathcal{S}_2$ that intersect Ω_{ij}.

Lemma 2.10.

$$\sum_i n_{ij} \lesssim nD_jE_j. \tag{2.15}$$

Proof. If a surface $S \in \mathcal{S}_2$ lies in W_j then it does not contribute to the above sum, so we need only consider those surfaces that do not lie in Z_j or W_j. First, we can replace each $Q \in \mathcal{Q}$ by the polynomial $Q + \epsilon$ for $\epsilon > 0$ a sufficiently small constant. If $S \cap \{x \in \mathbb{R}^3 \,|\, Q(x) > 0\} \cap Z_j \neq \emptyset$, then there must be a point on $S \cap Z_j$ where Q is positive, so $S \cap \{x \in \mathbb{R}^3 \,|\, Q(x) + \epsilon > 0\} \cap Z_j \neq \emptyset$ where ϵ is sufficiently small, and similarly for $S \cap \{x \in \mathbb{R}^3 \,|\, Q(x) < 0\} \cap Z_j$. Thus replacing each $Q \in \mathcal{Q}$ by $Q + \epsilon$ does not increase the number of realizations of realizable strict sign conditions that meet S. We shall select a small generic (with respect to S and Z_j) choice of ϵ.

By Corollary A.7 in Appendix A, we can assume that each irreducible component of each polynomial in $\mathcal{Q}_j$ generates a real ideal. Instead of counting $\sum_i n_{ij}$ directly, we shall bound the number of times a surface S enters a connected component of $Z_j \backslash W_j$, as this quantity controls $\sum_i n_{ij}$ (i.e. if the same surface enters multiple connected components of the same realization of a realizable strict sign condition, then we will over-count, but this is acceptable). The proof is essentially topological.

Let $S \in \mathcal{S}_2$ with S not contained in W_j. As in Lemma 2.8, we can select a large closed ball B so that the number of connected components of $Z_j \backslash W_j$ that S enters is equal to the number of connected components that $S \cap B$ enters. Now, replace S by $S' = \mathbf{Z}((f_S + \epsilon)(f_S - \epsilon))$ where $\epsilon > 0$ is a sufficiently small generic number. Provided ϵ is sufficiently small, if S meets a connected component Δ of $Z \backslash W_j$ then S' also meets Δ, since f_S is a continuous function on the (relatively) open set Δ, so f_S vanishes somewhere on Δ but does not vanish identically on Δ. Thus it suffices to count the number of times S' meets a connected component of $Z_j \backslash W_j$. After replacing S by S' (and recalling that we applied a small generic perturbation to each $Q \in \mathcal{Q}$), every point in $Z_j \cap W_j \cap S'$ is a point of non-singular intersection.

Now, if S meets a connected component Δ of $Z_j \backslash W_j$, then one of the following two things must occur:

(1) Δ contains (all of) a connected component of $S' \cap Z_j$,
(2) $S' \cap \Delta$ contains a (topological) curve that meets the boundary of Δ at a point $x \in S' \cap Z_j \cap W_j$. Furthermore, there is at most one other connected component Δ' for which Item 2 holds for the same point x.

We will first bound the number of times Item 1 can occur by showing that $S' \cap Z_j$ contains $O(D_j^2) = O(D_j E_j)$ connected components. Apply a generic rotation to the coordinate axes, and consider the plane curve $\gamma = \mathbf{Z}(\mathrm{res}_{x_3}(f_{S'}, P_j))$, where res_{x_3} is the resultant of $f_{S'}$ and P_j in the x_3 variable. Since neither of the (two) irreducible components of S' are contained in Z_j, γ is indeed a plane curve, and γ contains the image of the projection of $S' \cap Z_j$ in the x_3 direction. Thus, the number of connected components of $S' \cap Z_j$ is bounded by the number of connected components of γ plus the number of singular points of γ. Since γ has degree $O(D_j)$, both these quantities are $O(D_j^2)$ (this follows from Bézout's theorem in the plane and the Harnack curve theorem).

We will now bound the number of times Item 2 can occur. By the real Bézout's inequality, $S' \cap Z_j \cap W_j$ contains $O(D_j E_j)$ points of non-singular intersection, and thus Item 2 can occur at most $O(D_j E_j)$ times.

Thus S' can enter at most $O(D_j E_j)$ connected components of $Z_j \backslash W_j$. Since there are at most n surfaces, the result follows. □

Remark 2.11. A similar result to Lemma 2.10 can be obtained from the recent work of Barone and Basu in [2] and Solymosi and Tao in [19].

Using Lemma 2.10, we have

$$
\begin{aligned}
\mathcal{I}(\mathcal{P}_j \backslash W_j, \mathcal{S}_2) &= \sum_i \mathcal{I}(\mathcal{P}_j \cap \Omega_{ij}, \mathcal{S}_2) \\
&\lesssim \sum_i m_{ij} n_{ij}^{1-\frac{1}{k}} + n_{ij} \\
&\leq \left(\sum_i m_{ij}^k\right)^{\frac{1}{k}} \left(\sum_i n_{ij}\right)^{1-\frac{1}{k}} + n_{ij} \qquad (2.16) \\
&\lesssim \left(D_j E_j^2 \frac{|\mathcal{P}_j|^k}{(D_j E_j^2)^k}\right)^{\frac{1}{k}} (nD_jE_j)^{1-\frac{1}{k}} + nD_jE_j \\
&= \frac{|\mathcal{P}_j| n^{1-1/k}}{E_j^{1-1/k}} + nD_jE_j.
\end{aligned}
$$

Our analysis of the second term of (2.13) will be the same regardless of whether $j \in \mathcal{A}_2$ or $\mathcal{A}_3$. We shall express this bound as a lemma.

Lemma 2.12. *For $j \in \mathcal{A}_2 \cup \mathcal{A}_3$, let Z_j, W_j, $\mathcal{P}_j$, and $\mathcal{S}_2$ be as above. Then*

$$\mathcal{I}(\mathcal{P}_j \cap W_j, \mathcal{S}_2) \lesssim nD_jE_j + |\mathcal{P}_j|. \qquad (2.17)$$

Proof. We shall write

$$\mathcal{I}(\mathcal{P}_j \cap W_j, \mathcal{S}_2) = \mathcal{I}_1(\mathcal{P}_j \cap W_j, \mathcal{S}_2) + \mathcal{I}_2(\mathcal{P}_j \cap W_j, \mathcal{S}_2), \qquad (2.18)$$

where $\mathcal{I}_1$ counts those incidences between points $p \in \mathcal{P}_j \cap W_j$ and surfaces $S \in \mathcal{S}_2$ such that p^* lies on a 1 (complex) dimensional component of $S^* \cap Z_j^* \cap W_j^*$, and $\mathcal{I}_2$ counts the remaining incidences. To control $\mathcal{I}_2$, note that by Bézout's inequality (over $\mathbb{C}$), for each $S \in \mathcal{S}_2$, $S^* \cap Z_j^* \cap W_j^*$ contains $O(D_jE_j)$ isolated points. Since $|\mathcal{S}_2| \leq n$ we obtain

$$\mathcal{I}_2(\mathcal{P}_j \cap W_j, \mathcal{S}_2) \lesssim nD_jE_j. \qquad (2.19)$$

Thus it remains to control $\mathcal{I}_1$. First, we shall replace $\mathcal{Q}_j$ with a new family of polynomials $\widetilde{\mathcal{Q}}_j$ with the following properties:

(1) $Z_j \cap W_j \subset Z_j \cap \bigcup_{Q \in \widetilde{\mathcal{Q}}_j} \mathbf{Z}(Q)$,
(2) $\sum_{Q \in \widetilde{\mathcal{Q}}_j} \deg Q \leq E_j$,
(3) Each $Q \in \widetilde{\mathcal{Q}}_j$ is irreducible,
(4) For each $Q \in \widetilde{\mathcal{Q}}_j$, every irreducible component of $Z_j^* \cap \mathbf{Z}(Q)^*$ that contains a real point has (complex) dimension 1.

The procedure will be similar to that in the proof of Lemma 2.9: For each $Q \in \mathcal{Q}_j$, write $Q = Q_1, \ldots, Q_a$ as a product of irreducible factors. Discard those factors Q_b with $\mathbf{Z}(Q_b) \cap Z_j = \emptyset$. Of the remaining factors, place each irreducible factor that generates a real ideal in $\widetilde{\mathcal{Q}}_j$. If Q_b is a factor that does not generate a real ideal then consider $\nabla_v Q_b$ for v a generic vector. By assumption, Q_b does not vanish identically on Z_j, but it does vanish on at least one point of Z_j. Thus Q_b is not constant on Z_j, so ∇Q_j does not vanish identically on Z_j and hence if v is a generic vector then $\nabla_v Q_b$ does not vanish identically on Z_j. Thus we can repeat the above procedure with $\nabla_v Q_b$ in place of Q_b. This process will eventually terminate, and the resulting collection of polynomials $\widetilde{\mathcal{Q}}_j$ has the desired properties; Properties 1–3 are immediate. To obtain Property 4, suppose that for some $Q \in \widetilde{\mathcal{Q}}_j$, $Z_j^* \cap \mathbf{Z}(Q)^*$ fails to be a complete intersection. Then there exists some variety Y that is an irreducible component of both Z_j^* and $\mathbf{Z}(Q)^*$. by Proposition A.8 in Appendix A, $\mathcal{R}(Y)$ is an irreducible component of Z_j and $\mathbf{Z}(Q)$, and thus either $\mathcal{R}(Y) = \emptyset$ or $\mathcal{R}(Y) = Z_j = \mathbf{Z}(Q)$. The latter is impossible since Z_j and $\mathbf{Z}(Q)$ have dimension 2, while $Z_j \cap \mathbf{Z}(Q)$ has dimension at most 1.

Let

$$\widetilde{W}_j = \bigcup_{Q \in \widetilde{\mathcal{Q}}_j} \mathbf{Z}(Q).$$

We can write

$$Z_j^* \cap \widetilde{W}_j^* = \bigcup Y_j \tag{2.20}$$

as a union of irreducible (complex) varieties. By Property 4 above, we need only consider those components with (complex) dimension 1. We shall discard all components that have dimension 2. Let

$$\widetilde{\mathcal{P}}_j = \{p \in \mathcal{P}_j \mid \text{there exists a (Euclidean) neighborhood } U \subset \mathbb{C}^3 \text{ of } p^* \text{ such that } Z_j^* \cap \widetilde{W}_j^* \cap U \text{ is a (topological) 1–complex-dimensional curve}\}.$$

We shall establish several claims.

(1) $Z_j^* \cap \widetilde{W}_j^*$ is a union of $O(D_j E_j)$ irreducible varieties,

(2) If $p \in \widetilde{\mathcal{P}}_j$ then p^* lies on at most one of the irreducible components from (2.20),

(3) Let Y be a variety from the above decomposition. If there exist three surfaces $S_1, S_2, S_3 \in \mathcal{S}_2$ such that $Y \subset S_i^*$, $i = 1, 2, 3$, then $|\mathcal{P}_j \cap \mathfrak{R}(Y)| \leq C$,

(4) If $S \in \mathcal{S}_2$, then there are $O(D_j E_j)$ points $p \notin \widetilde{\mathcal{P}}_j$ such that p^* is contained in a 1–dimensional component of $S^* \cap Z_j^* \cap W_j^*$.

For Item 1, see e.g. [10]. Item 2 follows from the assumption that every variety in the decomposition (2.20) has dimension 1. Item 3 follows from the requirement that any three surfaces intersect in at most C points. To obtain Item 4, suppose that $D_j \leq E_j$ (if not, we can interchange the roles of Z_j and W_j). Note that if p satisfies the requirements of Item 4, then $S^* \cap Z_j^* \cap W_j^*$ fails to be a complex (C^0) curve in a small neighborhood of p^* (i.e. in a small neighborhood of p^*, $S^* \cap Z_j^* \cap W_j^*$ is a union of several complex curves all passing though p^*), and thus $S^* \cap Z_j^*$ fails to be a complex (C^0) curve in a small neighborhood of p^*. Thus after a generic rotation of the coordinate axis, the image of p^* under the projection $(x_1, x_2, x_3) \mapsto$

(x_1, x_2) is a singular point of the (complex) plane curve $\mathbf{Z}(\mathrm{res}_{x_3}(f_S, P_j))^*$, where res_{x_3} is the bivariate polynomial obtained by taking the resultant of f_S and P_j in the x_3 variable. This curve has degree $O(D_j)$ and thus has $O(D_j^2) = O(D_j E_j)$ singular points.

Now, for each $S \in \mathcal{S}_2$, at most $O(D_j E_j)$ points $p \in \mathcal{P}_j \backslash \widetilde{\mathcal{P}}_j$ can contribute to $\mathcal{I}_1(\mathcal{P}_j, \{S\})$, so the total contribution from all surfaces in $\mathcal{S}_2$ is $O(nD_jE_j)$. To control the remaining incidences, use Item 3 to write $\{Y_j\} = \{Y_j'\} \sqcup \{Y_j''\}$, where the first set consists of varieties that are contained in at most 2 surfaces $S \in \mathcal{S}_2$, and the second consists of varieties that contain at most C points. Each point $p \in \widetilde{\mathcal{P}}_j$ with $p^* \in \bigcup Y_j'$ can be incident to at most two surfaces, so the total contribution from such points is $O(|\mathcal{P}_j|)$. On the other hand, by Item 1 at most $O(D_jE_j)$ points can be contained in $\Re(\bigcup Y_j'')$, so these points can contribute at most $O(nD_jE_j)$ incidences. □

Combining (2.14), (2.16), and (2.17) and optimizing in E_j, we see that our choice of E_j from (2.12) yields the bound

$$\mathcal{I}(\mathcal{P}_j, \mathcal{S}_2) \lesssim |\mathcal{P}_j|^{\frac{k}{2k-1}} n^{\frac{2k-2}{2k-1}} D_j^{\frac{k-1}{2k-1}} + m_j. \tag{2.21}$$

Summing (2.21) over all $j \in \mathcal{A}_2 \cup \mathcal{A}_3$ and noting that $(2k-1)/k$ and $(2k-1)/(k-1)$ are conjugate exponents, we obtain

$$\begin{aligned} \mathcal{I}\Bigg(\bigcup_{j \in \mathcal{A}_2 \cup \mathcal{A}_3} \mathcal{P}_j, \mathcal{S}_2\Bigg) &\lesssim \sum_{\mathcal{A}_2 \cup \mathcal{A}_3} |\mathcal{P}_j|^{\frac{k}{2k-1}} n^{\frac{2k-2}{2k-1}} D_j^{\frac{k-1}{2k-1}} + |\mathcal{P}_j| \\ &\lesssim n^{\frac{2k-2}{2k-1}} \Big(\sum_j |\mathcal{P}_j|\Big)^{\frac{k}{2k-1}} \Big(\sum_j D_j\Big)^{\frac{k-1}{2k-1}} + m \\ &\lesssim m^{\frac{k}{2k-1}} n^{\frac{2k-2}{2k-1}} D^{\frac{k-1}{2k-1}} + m. \end{aligned} \tag{2.22}$$

Finally, selecting

$$D = m^{\frac{k}{3k-1}} n^{\frac{-1}{3k-1}}, \tag{2.23}$$

which by (2.4) satisfies $D > C$, and combining (2.4), (2.6), (2.8), (2.11), and (2.22), we obtain

$$\begin{aligned} \mathcal{I}(\mathcal{P}, \mathcal{S}) &\lesssim D^2 n + m + \frac{mn^{1-1/k}}{D^{1-1/k}} + Dm^{\frac{2}{3}} \\ &\quad + nD + m + m^{\frac{k}{2k-1}} n^{\frac{2k-2}{2k-1}} D^{\frac{k-1}{2k-1}} \\ &\lesssim m^{\frac{2k}{3k-1}} n^{\frac{3k-3}{3k-1}} + m^{\frac{2}{3} + \frac{k}{3k-1}} n^{\frac{-1}{3k-1}} + m + n \\ &\lesssim m^{\frac{2k}{3k-1}} n^{\frac{3k-3}{3k-1}} + m + n. \end{aligned} \tag{2.24}$$

□

3. Applications

In [8, 9], Erdős asked how many unit distances could there be amongst m points in the plane or in $\mathbb{R}^3$. Theorem 1.2 yields new bounds for the $\mathbb{R}^3$ version of this question. Let $\mathcal{P}$ be a collection of m points in $\mathbb{R}^3$, and let $\mathcal{S}$ be a collection of unit spheres centered about the points in $\mathcal{P}$. We can immediately verify that any three spheres have at most eight points in common, so Theorem 1.2 tells us that there are $O(m^{3/2})$ point-sphere incidences.

Theorem 3.1. *The maximum number of unit-distance pairs in a set of m points in $\mathbb{R}^3$ is $O(m^{3/2})$.*

This is a slight improvement over the previous bound of $O(m^{3/2}\beta(m))$ from [6], where β is a very slowly growing function.

As observed in [6], Theorem 1.2, combined with the method outlined in [5] can be used to establish bounds on the number of incidences between points and spheres in $\mathbb{R}^d$. Specifically, we have the following theorem:

Theorem 3.2. *The maximum number of incidences between m points and n spheres in $\mathbb{R}^d$ is*

$$O\left(m^{\frac{d}{d+1}}n^{\frac{d}{d+1}}+m+n\right), \tag{3.1}$$

provided no d of the spheres intersect in a common circle.

Again, this is a slight improvement (by a $\beta(m,n)$ factor) from the analogous bounds established in [6]. See [6, §6.5] for additional applications of Theorem 1.2. In each case, we are able to slightly sharpen the bound from [6] by removing the $\beta(m)$ factor.

4. Generalizations to higher dimensions

It is reasonable to ask whether Theorem 1.2 can be generalized to incidences between points and hypersurfaces in higher dimensions. This task appears to be quite involved, as the necessary algebraic geometry becomes more difficult. In particular, it appears that in order to generalize the proof of Theorem 1.2 to (say) spheres in $\mathbb{R}^d$, we need to perform $d-1$ polynomial ham sandwich decompositions, with each successive decomposition performed on the variety defined by the previous decompositions. As d increases, the number of cases to be considered increases dramatically, and certain difficulties such as the failure of the connected components of a complete intersection to themselves be a complete intersection, and the failure of an arbitrary complete intersection to be a nonsingular complete intersection, etc. become increasingly problematic.

One could also consider two dimensional surfaces in $\mathbb{R}^d$, $d>3$, and this appears to be more promising. However, the analogues of (2.8) and Lemma 2.12 become more difficult: an algebraic variety of dimension $d-1$ can contain many 2–dimensional surfaces without obvious constraints being imposed on its structure, and in higher dimensions there are more (and more complicated) ways in which varieties can fail to intersect completely. Nevertheless, this is certainly a promising area for future work.

A. Real algebraic geometry

A.1. Sign conditions.

Definition A.1. Let $\mathcal{Q} \subset \mathbb{R}[x_1, \ldots, x_d]$ be a collection of non-zero real polynomials. A *strict sign condition* on $\mathcal{Q}$ is a map $\sigma \colon \mathcal{Q} \to \{\pm 1\}$. If $Q \in \mathcal{Q}$, we will denote the evaluation of σ at Q either by σ_Q or $\sigma(Q)$, depending on context. If σ is a strict sign condition on $\mathcal{Q}$ we define its *realization* by

$$\mathrm{Reali}(\sigma, \mathcal{Q}) = \{x \in \mathbb{R}^d \,|\, Q(x)\sigma_Q > 0 \text{ for all } Q \in \mathcal{Q}\}. \tag{A.1}$$

If $\mathrm{Reali}(\sigma, \mathcal{Q}) \neq \emptyset$ then we say that σ is *realizable.* We define

$$\Sigma_{\mathcal{Q}} = \{\sigma \mid \mathrm{Reali}(\sigma, \mathcal{Q}) \neq \emptyset\}, \tag{A.2}$$

and

$$\mathrm{Reali}(\mathcal{Q}) = \{\mathrm{Reali}(\sigma, \mathcal{Q}) \,|\, \sigma \in \Sigma_{\mathcal{Q}}\}. \tag{A.3}$$

We call $\mathrm{Reali}(\mathcal{Q})$ the collection of "realizations of realizable strict sign conditions of $\mathcal{Q}$."

If $Z \subset \mathbb{R}^d$ is a variety, and σ is a strict sign condition on $\mathcal{Q}$, then we can define the *realization of σ on Z* by

$$\mathrm{Reali}(\sigma, \mathcal{Q}, Z) = \{x \in Z \,|\, Q(x)\sigma_Q > 0 \text{ for all } Q \in \mathcal{Q}\}, \tag{A.4}$$

and we can define analogous sets

$$\Sigma_{\mathcal{Q},Z} = \{\sigma \mid \mathrm{Reali}(\sigma, \mathcal{Q}, Z) \neq \emptyset\}, \tag{A.5}$$

and

$$\mathrm{Reali}(\mathcal{Q}, Z) = \{\mathrm{Reali}(\sigma, \mathcal{Q}, Z) \,|\, \sigma \in \Sigma_{\mathcal{Q},Z}\}. \tag{A.6}$$

We call $\mathrm{Reali}(\mathcal{Q}, Z)$ the collection of "realizations of realizable strict sign conditions of $\mathcal{Q}$ on Z." Note that if some $Q \in \mathcal{Q}$ vanishes identically on Z then $\Sigma_{\mathcal{Q},Z} = \emptyset$ and thus $\mathrm{Reali}(\mathcal{Q}, Z) = \emptyset$.

A.2. Real ideals.

Definition A.2. An ideal $I \subset \mathbb{R}[x_1, \ldots, x_d]$ is *real* if for every sequence $a_1, \ldots, a_\ell \in \mathbb{R}[x_1, \ldots, x_d]$, $a_1^2 + \cdots + a_\ell^2 \in I$ implies $a_j \in I$ for each $j = 1, \ldots, \ell$.

The following proposition shows that real principal prime ideals and their corresponding real varieties have some of the "nice" properties of ideals and varieties defined over $\mathbb{C}$.

Proposition A.3 (see [4, §4.5])**.** Let $P \in \mathbb{R}[x_1, \ldots, x_d]$ be irreducible. Then the following are equivalent:

(1) (P) is real.
(2) $(P) = \mathbf{I}(\mathbf{Z}(P))$.
(3) $\dim(\mathbf{Z}(P)) = d - 1$.
(4) ∇P does not vanish identically on $\mathbf{Z}(P)$.
(5) The sign of P changes somewhere on $\mathbb{R}^d$ (i.e. from strictly positive to strictly negative).

A.3. **Removing non-real components from a polynomial.**

Definition A.4. If $P \subset \mathbb{R}[x_1, \ldots, x_d]$ is a polynomial and $P = P_1, \ldots, P_\ell$ is its factorization, we define $\widehat{P}$ to be the polynomial obtained by removing those irreducible components that generate ideals that aren't real. If every irreducible component of P generates an ideal that is not real, then we define $\widehat{P} = 1$.

Example A.5. Let $P = (x_1^2 + x_2^2 + x_3^2 - 1)(x_1^2 + x_2^2)$. Then $\widehat{P} = x_1^2 + x_2^2 + x_3^2 - 1$. Geometrically, if $\widehat{P} \neq 1$, then $\mathbf{Z}(P)$ is a $(d-1)$–dimensional (real) variety, but some of the components of $\mathbf{Z}(P)$ may have dimension less than $d-1$. $\widehat{P}$ keeps only those factors that generate components that have dimension $d-1$, and discards the rest. Note that $\mathbf{Z}(\widehat{P})$ may still contain points whose local dimension is less than $d-1$.

The existence of polynomials that do not generate real ideals complicates our analysis, but since the zero sets of such polynomials have codimension at least 2, we can ignore them when we are computing the number of times a surface meets the realization of a realizable strict sign condition of a family of polynomials. The following theorem helps make this statement precise.

Theorem A.6. *Let $\mathcal{Q} \subset \mathbb{R}[x_1, \ldots, x_d]$, $d \geq 3$ be a collection of real polynomials and let $\widehat{\mathcal{Q}} = \{\widehat{Q} \mid Q \in \mathcal{Q}\}$. Then there exists a bijection*

$$\tau\colon \operatorname{Reali}(\mathcal{Q}) \to \operatorname{Reali}(\widehat{\mathcal{Q}})$$

such that

$$X \subset \tau(X) \text{ for every } X \in \operatorname{Reali}(\mathcal{Q}). \tag{A.7}$$

Similarly, if $Z = \mathbf{Z}(P)$ where $P \in \mathbb{R}[x_1, \ldots, x_d]$ generates a real ideal and no polynomial $Q \in \mathcal{Q}$ vanishes identically on Z, then there exists a bijection

$$\tau\colon \operatorname{Reali}(\mathcal{Q}, Z) \to \operatorname{Reali}(\widehat{\mathcal{Q}}, Z)$$

such that

$$X \subset \tau(X) \text{ for every } X \in \operatorname{Reali}(\mathcal{Q}, Z). \tag{A.8}$$

Proof. First, by Item 5 of Proposition A.3, for each $Q \in \mathcal{Q}$ we have that $Q/\widehat{Q} \geq 0$ or $Q/\widehat{Q} \leq 0$ on all of $\mathbb{R}^d$. Choose $\varepsilon_Q \in \{\pm 1\}$ so that $\varepsilon_Q Q/\widehat{Q} \geq 0$. Now, note that if there exist $Q_1, Q_2 \in \mathcal{Q}$ with $\widehat{Q}_1 = \widehat{Q}_2$ and if σ is a strict sign condition on $\mathcal{Q}$, then either $\varepsilon_{Q_1}\sigma(Q_1) = \varepsilon_{Q_2}\sigma(Q_2)$ or $\mathrm{Reali}(\sigma, \mathcal{Q}) = \emptyset$. Thus if σ is a realizable strict sign condition on $\mathcal{Q}$, then we can define $\widehat{\sigma}\colon \widehat{\mathcal{Q}} \to \{\pm 1\}$ by $\widehat{\sigma}(T) = \varepsilon_Q \sigma(Q)$, where $Q \in \mathcal{Q}$ satisfies $T = \widehat{Q}$, and $\widehat{\sigma}$ is well-defined.

We shall show that the map $\Sigma_{\mathcal{Q}} \to \Sigma_{\widehat{\mathcal{Q}}}$, $\sigma \mapsto \widehat{\sigma}$ is a bijection. To prove injectivity, note that if distinct σ_1, σ_2 both map to the same element $\widehat{\sigma}$, then $\varepsilon_Q\sigma_1(Q) = \varepsilon_Q\sigma_2(Q)$ for all $Q \in \mathcal{Q}$, so clearly $\sigma_1 = \sigma_2$. To establish surjectivity, note that each $\sigma_1 \in \Sigma_{\widehat{Q}}$ has a pre-image under the map $\sigma \mapsto \widehat{\sigma}$. Thus every element of $\Sigma_{\widehat{Q}}$ may be written as $\widehat{\sigma}$ for some strict sign condition σ on $\mathcal{Q}$. All that we must establish is that σ is realizable. For each $Q \in \mathcal{Q}$, we have

$$\dim\Big(\{x \in \mathbb{R}^d | \widehat{Q}(x) > 0\} \backslash \{x \in \mathbb{R}^d | \varepsilon_Q Q(x) > 0\}\Big) \leq d - 2, \tag{A.9}$$

(see [4] for the dimension of a semi-algebraic set). On the other hand, the realization of each realizable strict sign condition of $\widehat{\mathcal{Q}}$ has dimension d. Thus, if $\mathrm{Reali}(\widehat{\sigma}, \widehat{\mathcal{Q}}) \neq \emptyset$, then $\mathrm{Reali}(\sigma, \mathcal{Q})$ can be written as a (non-empty) dimension d semi-algebraic set minus a dimension $d-2$ semi-algebraic set, and in particular, $\mathrm{Reali}(\sigma, \mathcal{Q}) \neq \emptyset$.

Thus the map $\mathrm{Reali}(\mathcal{Q}) \to \mathrm{Reali}(\widehat{\mathcal{Q}})$ given by $\mathrm{Reali}(\sigma, \mathcal{Q}) \mapsto \mathrm{Reali}(\widehat{\sigma}, \widehat{\mathcal{Q}})$ is well-defined and is a bijection. Now, note that by Items 3 and 5 of Proposition A.3, $\{x \in \mathbb{R}^d \,|\, \varepsilon_Q Q(x) > 0\} \subset \{x \in \mathbb{R}^d \,|\, \widehat{Q}(x) > 0\}$, and similarly with "$>$" replaced by "$<$"). Thus

$$\mathrm{Reali}(\sigma, \mathcal{Q}) \subset \mathrm{Reali}(\widehat{\sigma}, \widehat{\mathcal{Q}}),$$

so (A.7) holds.

The same arguments establish the second part of the theorem. The only new thing that must be verified is that the map $\Sigma_{\mathcal{Q},Z} \to \Sigma_{\widehat{\mathcal{Q}},Z}$, $\sigma \mapsto \widehat{\sigma}$ is onto. However, this is established by (A.9) plus the fact that the realization of each realizable strict sign condition of $\mathcal{Q}$ on Z has dimension $d-1$. $\square$

Corollary A.7. Let $S \subset \mathbb{R}^3$ be a smooth surface, let $\mathcal{Q}$ be a collection of polynomials, and let $\widehat{\mathcal{Q}}$ be as in Theorem A.6. Then

$$|\{X \in \mathrm{Reali}(\mathcal{Q}) \,|\, X \cap S \neq \emptyset\}| \leq |\{X \in \mathrm{Reali}(\widehat{\mathcal{Q}}) \,|\, X \cap S \neq \emptyset\}|. \tag{A.10}$$

Similarly, let $S \subset \mathbb{R}^3$ be a smooth surface, let $Z = \mathbf{Z}(P)$ where $P \in \mathbb{R}[x_1, x_2, x_3]$ generates a real ideal, let $\mathcal{Q}$ be a collection of polynomials, none of which vanish identically on Z, and let $\widehat{\mathcal{Q}}$ be as in Theorem A.6. Then

$$|\{X \in \mathrm{Reali}(\mathcal{Q}, Z) \,|\, X \cap S \neq \emptyset\}| \leq |\{X \in \mathrm{Reali}(\widehat{\mathcal{Q}}, Z) \,|\, X \cap S \neq \emptyset\}|. \tag{A.11}$$

A.4. **Real and complex varieties.** As noted in Section 1.4, the number of intersection points of a collection of real polynomials may exceed the product of their degrees, even if those polynomials intersect completely. Over $\mathbb{C}$ things are much better behaved, so there will be times when we will wish to embed everything into $\mathbb{C}$. The following proposition relates the properties of a variety defined over $\mathbb{R}$ and the corresponding variety defined over $\mathbb{C}$:

Proposition A.8 (see [21, §10])**.** Let $Z \subset \mathbb{R}^d$ be a real variety and let $Z_1^*, \ldots, Z_\ell^*$ be the irreducible components of Z^*. Then $\mathfrak{R}(Z_1^*), \ldots, \mathfrak{R}(Z_\ell^*)$ are the irreducible components of Z.

B. Proof of Theorem 2.6

For the reader's convenience, we will restate Theorem 2.6 below:

Theorem 2.6. *Let $P \in \mathbb{R}[x_1, \ldots, x_d]$ be an irreducible polynomial of degree D such that (P) is a real ideal, and let $\mathcal{P}$ be a collection of points contained in $Z = \mathbf{Z}(P) \subset \mathbb{R}^d$. Let $E \geq D$. Then there exists a collection of polynomials $\mathcal{Q} \subset \mathbb{R}[x_1, \ldots, x_d]$ with the following properties:*

(1) $|\mathcal{Q}| \leq \log_2(DE^{d-1}) + O(1)$,

(2) $\sum_{\mathcal{Q}} \deg Q \lesssim E$,

(3) *None of the polynomials in $\mathcal{Q}$ vanish identically on Z,*

(4) *The realization of each of the $O(DE^{d-1})$ strict sign conditions of $\mathcal{Q}$ on Z contains $O\left(\frac{|\mathcal{P}|}{DE^{d-1}}\right)$ points of $\mathcal{P}$.*

All implicit constants depend only on d.

Our proof of Theorem 2.6 will be similar to the original proof of the discrete polynomial ham sandwich theorem in [11, §4], which can be stated as follows:

Proposition B.1 (Discrete polynomial ham sandwich theorem)**.** Let $V \subset \mathbb{R}[x_1, \ldots, x_d]$ be a vector space of dimension ℓ, and let $F_1, \ldots, F_\ell \subset \mathbb{R}^d$ be finite families of points. Then there exists a polynomial $P \in V$ such that

$$|F_j \cap \{x \in \mathbb{R}^d \mid P(x) > 0\}| \leq \frac{|F_j|}{2}, \text{ and}$$
$$|F_j \cap \{x \in \mathbb{R}^d \mid P(x) < 0\}| \leq \frac{|F_j|}{2}, \ j = 1, \ldots, \ell.$$

Proposition B.1 is proved in [11] only in the special case where V is the vector space of all polynomials of degree at most e (where e is chosen large enough to ensure that V has the required dimension). However, the proof carries over verbatim to the general case where V is arbitrary. To prove Theorem 2.6, we will iterate the following lemma:

Lemma B.2. *Let $Z = \mathbf{Z}(P) \subset \mathbb{R}^d$ for P an irreducible polynomial of degree D such that (P) is a real ideal. Let $E > 0$, and let $F_1, \ldots, F_\ell$, $\ell = c\min\{E^d, DE^{d-1}\}$ be finite families of points in $\mathbb{R}^d$, with $F_j \subset Z$ for each j. Then provided c is sufficiently small (depending only on d), there* *exists a polynomial Q of degree at most E that does not vanish identically on $\mathbf{Z}(P)$ such that*

$$\begin{aligned} &|F_j \cap \{x \in \mathbb{R}^d \,|\, Q(x) > 0\}| \leq \frac{|F_j|}{2}, \ \textit{and} \\ &|F_j \cap \{x \in \mathbb{R}^d \,|\, Q(x) < 0\}| \leq \frac{|F_j|}{2}, \ j = 1, \ldots, \ell. \end{aligned} \tag{B.1}$$

Proof. Let $\mathbb{R}[x_1, \ldots, x_d]_{\leq E}$ be the vector space of all polynomials in d variables of degree at most E, and let $(P)_{\leq E}$ be the vector space of all polynomials in the ideal (P) that have degree at most E (of course, if $E < \deg P$ then $(P)_{\leq E} = 0$). We have

$$\dim\left(\mathbb{R}[x_1, \ldots, x_d]_{\leq E}\right) - \dim\left((P)_{\leq E}\right) > c \min\left\{E^d, DE^{d-1}\right\}$$

for some (explicit) constant c depending only on the dimension d. Thus, we can find a vector space $V \subset \mathbb{R}[x_1, \ldots, x_d]_{\leq E}$ with $\dim(V) > c \min\{E^d,\ DE^{d-1}\}$ such that $V \cap (P)_{\leq E} = 0$. By Proposition B.1, we can find a polynomial $Q \in V$ satisfying (B.1). Since $Q \in \mathbb{R}[x_1, \ldots, x_d]_{\leq E}$ but $Q \notin (P)_{\leq E}$, we have $Q \notin (P)$. Since P is irreducible and generates a real ideal, by Item 2 of Proposition A.3, Q does not vanish identically on $\mathbf{Z}(P)$. ☐

Proof of Theorem 2.6. Let $\mathcal{Q}_0 = \{1\}$. For each $i = 1, \ldots, t$, with

$$t = \left\lceil \log_2(DE^{d-1}) \right\rceil, \tag{B.2}$$

use Lemma B.2 to find a polynomial Q_i with

$$\deg(Q_i) \lesssim \max\left\{ \left(\frac{2^i}{D}\right)^{\frac{1}{d-1}}, 2^{\frac{i}{d}} \right\}$$

(the implicit constant depends only on d) such that for each $\sigma \in \Sigma_{\mathcal{Q}_{i-1}}$ we have

(B.3)

$$\begin{aligned} &\left|\{x \in \mathbb{R}^d \,|\, Q_i(x) > 0\} \cap \left(\mathcal{P} \cap \operatorname{Reali}(\sigma, \mathcal{Q}_{i-1})\right)\right| \leq \frac{1}{2}|\mathcal{P} \cap \operatorname{Reali}(\sigma, \mathcal{Q}_{i-1})|, \\ &\left|\{x \in \mathbb{R}^d \,|\, Q_i(x) < 0\} \cap \left(\mathcal{P} \cap \operatorname{Reali}(\sigma, \mathcal{Q}_{i-1})\right)\right| \leq \frac{1}{2}|\mathcal{P} \cap \operatorname{Reali}(\sigma, \mathcal{Q}_{i-1})|. \end{aligned}$$

Some of the above sets may be empty, but this does not pose a problem. Let $\mathcal{Q}_i = \mathcal{Q}_{i-1} \cup \{Q_i\}$. None of the polynomials in $\mathcal{Q} = \mathcal{Q}_t$ vanish on P, so

Item 3 of the theorem is satisfied. Since $E \geq D$ we have

$$\begin{aligned}\sum_{\mathcal{Q}} \deg Q &\lesssim \sum_{i=1}^{t} \left(\frac{2^i}{D}\right)^{\frac{1}{d-1}} + \sum_{i=1}^{t} 2^{\frac{i}{d}} \\ &\lesssim \left(\frac{DE^{d-1}}{D}\right)^{\frac{1}{d-1}} + \left(DE^{d-1}\right)^{\frac{1}{d}} \\ &\lesssim E,\end{aligned}$$

which satisfies Item 2. By (B.3), for each $\sigma \in \Sigma_{\mathcal{Q}}$,

$$\begin{aligned}|\mathcal{P} \cap \mathrm{Reali}(\sigma, \mathcal{Q})| &\lesssim 2^{-t}|\mathcal{P}| \\ &\lesssim \frac{|\mathcal{P}|}{DE^{d-1}},\end{aligned} \tag{B.4}$$

which satisfies Item 4. Finally, Item 1 follows from (B.2). □

Acknowledgments

The author is very grateful to Jordan Ellenberg, Larry Guth, Netz Katz, Jozsef Solymosi, and Terence Tao for helpful discussions, and to Haim Kaplan, Jiří Matoušek, Zuzana Safernová, and Micha Sharir for pointing out errors in an earlier version of this paper. The author would also like to thank Poh-Shen Loh for bringing the example in Remark 1.4 to the author's attention. Finally, the author would like to thank the anonymous referee for his or her careful proofreading and corrections. The author was supported in part by the Department of Defense through the National Defense Science & Engineering Graduate Fellowship (NDSEG) Program.

References

1. P. Agarwal and M. Sharir, *Davenport–Schinzel sequences and their geometric applications*, Cambridge University Press, NY, 1995.
2. S. Barone and S. Basu, *Refined bounds on the number of connected components of sign conditions on a variety*, Discrete Comput. Geom. **47** (2012), no. 3, 577–597.
3. S. Basu, R. Pollack, and M. Roy, *Algorithms in real algebraic geometry*, Springer-Verlag, Berlin, 2006.
4. J. Bochnak, M. Coste, and M. Roy, *Real algebraic geometry*, Springer-Verlag, 1998.
5. F. Chung, *Sphere-and-point incidence relations in high dimensions with applications to unit distances and furthest-neighbor pairs*, Discrete Comput. Geom. **4** (1989), no. 1, 183–190.
6. K. Clarkson, H. Edelsbrunner, L. Guibas, M. Sharir, and E. Welzl, *Combinatorial complexity bounds for arrangements of curves and surfaces*, Discrete Comput. Geom. **5** (1990), 99–160.
7. G. Elekes and C. Tóth, *Incidences of not-too-degenerate hyperplanes*, Proc. 21st Annu. Sympos. Comput. Geom., 2005, pp. 16–21.
8. P. Erdős, *On sets of distances of n points*, Amer. Math. Monthly **53** (1946), 248–250.
9. ______, *On sets of distances of n points in Euclidean space*, Magyar Tud. Akad. Mat. Kutaló Int. Kozl. **5** (1960), 165–169.
10. W. Fulton, *Intersection theory*, 2 ed., Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge, Springer-Verlag, Berlin, 1998.

11. L. Guth and N. Katz, *On the Erdős distinct distance problem in the plane*, arXive:1011.4105v3, 2011.
12. A. Iosevich, H. Jorati, and I. Łaba, *Geometric incidence theorems via Fourier analysis*, Trans. Amer. Math. Soc. **361** (2009), no. 12, 6595–6611.
13. H. Kaplan, J. Matoušek, Z. Safernová, and M. Sharir, *Unit distances in three dimensions*, Comb. Probab. Comput. **21** (2012), no. 4, 597–610.
14. 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), no. 3, 499–517.
15. T. Kővari, V. Sós, and P. Turan, *On a problem of K. Zarankiewicz*, Colloquium Mathematicum **3** (1954), 50–57.
16. I. Łaba and J. Solymosi, *Incidence theorems for pseudoflats*, Discrete Comput. Geom. **37** (2007), no. 2, 163–174.
17. J. Milnor, *On the Betti numbers of real varieties*, Proc. AMS, vol. 15, 1964, pp. 275–280.
18. J. Pach and M. Sharir, *On the number of incidences between points and curves*, Combin. Probab. Comput. **7** (1998), no. 1, 121–127.
19. J. Solymosi and T. Tao, *An incidence theorem in higher dimensions*, Discrete Comput. Geom. **48** (2012), no. 2, 255–280.
20. R. Thom, *Sur l'homologie des variétés algébriques réelles*, Differential and Combinatorial Topology (S.S. Cairns, ed.), Princeton Univ. Press, 1965, (Symposium in Honor of Marston Morse), pp. 255–265.
21. H. Whitney, *Elementary structure of real algebraic varieties*, Annals of Math **66** (1957), 545–556.

Department of Mathematics, UCLA,
Los Angeles CA 90095-1555, USA
E-mail address: `jzahl@zahl.ca`