

Combinatorial Complexity Bounds for Arrangements of Curves and Spheres*

Kenneth L. Clarkson,¹ Herbert Edelsbrunner,² Leonidas J. Guibas,^{3,4} Micha Sharir,^{5,6} and Emo Welzl⁷

¹ AT&T Bell Laboratories, Murray Hill, NJ 07974, USA

² Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, Ill 61801, USA

³ DEC Systems Research Center, Palo Alto, Cal 94301, USA

⁴ Computer Science Department, Stanford University, Cal, 94305, USA

⁵ Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, USA

⁶ School of Mathematical Sciences, Tel Aviv University, 69978 Tel Aviv, Israel

⁷ Fachbereich Mathematik, Freie Universität Berlin, 1000 Berlin 33, Federal Republic of Germany

Abstract. We present upper and lower bounds for extremal problems defined for arrangements of lines, circles, spheres, and alike. For example, we prove that the maximum number of edges bounding m cells in an arrangement of n lines is $\Theta(m^{2/3}n^{2/3} + n)$, and that it is $O(m^{2/3}n^{2/3}\beta(n) + n)$ for n unit-circles, where $\beta(n)$ (and later $\beta(m, n)$) is a function that depends on the inverse of Ackermann's function and grows extremely slowly. If we replace unit-circles by circles of arbitrary radii the upper bound goes up to $O(m^{3/5}n^{4/5}\beta(n) + n)$. The same bounds (without the $\beta(n)$ -terms) hold for the maximum sum of degrees of m vertices. In the case of vertex degrees in arrangements of lines and of unit-circles our bounds match previous results, but our proofs are considerably simpler than the previous ones. The maximum sum of degrees of m vertices in an arrangement of n spheres in three dimensions is

* The research of the second author was supported by the National Science Foundation under Grant CCR-8714565. Work by the fourth author has been supported by Office of Naval Research Grant N00014-87-K-0129, by National Science Foundation Grant No. NSF-DCR-83-20085, by grants from the Digital Equipment Corporation and the IBM Corporation, and by a research grant from the NCRD, the Israeli National Council for Research and Development. A preliminary version of this paper has appeared in the *Proceedings of the 29th IEEE Symposium on Foundations of Computer Science*, 1988.

$O(m^{4/7}n^{9/7}\beta(m, n) + n^2)$, in general, and $O(m^{3/4}n^{3/4}\beta(m, n) + n)$ if no three spheres intersect in a common circle. The latter bound implies that the maximum number of unit-distances among m points in three dimensions is $O(m^{3/2}\beta(m))$ which improves the best previous upper bound on this problem. Applications of our results to other distance problems are also given.

1. Introduction

Combinatorial distance problems for finite point sets have a long history in the mathematical literature (see [49] and [51]). Some of this literature was originated by the following question asked by Paul Erdős in 1946 [24] (see also [25]):

What is the maximum number of pairs in a set of m points in two (or three) dimensions, that are exactly at distance 1 from each other?

By scaling we can make any distance the unit-distance, so this question is equivalent to asking how often the most popular distance can occur. This seemingly innocent problem turned out to be one of the most difficult problems in combinatorial geometry, and it is still far from being solved—in spite of some recent progress in its analysis to which this paper significantly contributes. In particular, we give a new and simpler proof of the $O(m^{4/3})$ upper bound in the plane [55] and improve the upper bound in three dimensions from $O(m^{8/5})$ [10] to $O(m^{3/2}\beta(m))$, where $\beta(m)$ is an extremely slowly growing function. The two upper bounds are corollaries of more general results about arrangements of curves in two dimensions and surfaces in three dimensions obtained in this paper. We change gears now and introduce these more general concepts and talk about the philosophy that motivates and directs our work.

Let us consider a collection $H = \{h_1, h_2, \dots, h_n\}$ of n bounded or unbounded $(d - 1)$ -dimensional manifolds in the Euclidean d -dimensional space. The *arrangement* $\mathcal{A}(H)$ is the subdivision of the space defined by these manifolds. Intuitively, we think of space as being filled with some solid material; the space is then “cut” along each of the manifolds in H , resulting in a collection of connected pieces, called *cells* or *d -faces*. The formal notion of an arrangement includes not only these “top-level” faces, but also the lower-dimensional ones that form the boundaries of these d -dimensional faces, as well as the incidence relationships between all these faces of various dimensions. Arrangements of lines and planes have been extensively considered in the literature; see [34] or [15] for further details. For additional material on the data structures needed to capture the topological information present in an arrangement, see also [35] and [14].

As an example, if the h_i are straight lines in the Euclidean plane, then the arrangement of these lines is a decomposition of the plane into open convex regions (the *2-faces* or *cells*), relatively open line segments bounding these convex regions (the *1-faces* or *edges*), and the intersection points of the lines h_i (the *0-faces* or *vertices*). Every point of the plane belongs to exactly one face of the arrangement $\mathcal{A}(H)$ when these additional lower-dimensional faces are considered.

Arrangements are ubiquitous structures in geometric computing: many geometric objects of interest can naturally be specified as a collection of cells or lower-dimensional faces in an appropriate arrangement. For instance, the convex hull of n points in the plane corresponds to the cell containing the origin in the arrangement of the lines dual to the points [15]. The Voronoi diagram of n points in the plane can be viewed as the planar projection of a cell's boundary complex in a three-dimensional arrangement of planes [50], [15] or cones [31]. In general, semi-algebraic sets (that is, geometric sets specified by unions and intersections of half-spaces bounded by algebraic surfaces) naturally correspond to collections of faces in the arrangement of the underlying surfaces. In fact, a whole branch of solid modeling has arisen, *constructive solid geometry* [47], which bases its solid shape descriptions on these boolean operations.

Although arrangements are commonly occurring objects, they are expensive to manipulate and to store in their entirety. An arrangement of n hyperplanes in E^d has $\Theta(n^d)$ faces [15], so even an arrangement of lines in E^2 is an unwieldy object. Of course, if our computation requires us to visit every cell of the arrangement, then we must pay at least this $\Theta(n^d)$ cost in time. In some situations [16], however, we can sweep over the arrangement by a topological wavefront while performing the computation, thus reducing the working storage required by our algorithm to that of storing only the cells intersected by the wavefront, plus some auxiliary structures. We omit further details here, because in this paper we focus on situations where not all but only some of the cells are of interest. We further assume that the “interesting” cells are given to us explicitly indicating a point contained in each cell we must consider.

In this and several companion papers [19], [20] we deal with the combinatorial problem of estimating the worst-case combinatorial size of m cells in an arrangement of n manifolds, and the algorithmic problem of computing (a boundary representation of) those cells—for various dimensions, and for various types of manifolds. We have a collection of powerful tools which in various combinations can be used to attack each specific problem. Section 2 spells out the specific results contained in this paper. For now we just wish to remark that our algorithmic and combinatorial tools are closely intertwined. We design efficient algorithms because we have access to detailed combinatorial knowledge about the underlying geometric objects. At the same time, we are able to prove purely combinatorial results by using algorithms as proof tools. We believe that this work fits well within the paradigm of the analysis of algorithms initiated by Knuth [42]–[44]. We should remark, however, that in contrast to [19] and [20] this paper is predominantly combinatorial and refrains from discussing algorithmic issues that do not also serve combinatorial ends. In another companion paper [17] some of the algorithmic problems suggested by the investigations of this paper are studied.

The Many-Faces Problem. In the *many-faces problem* our input consists of m points and n manifolds. We wish to count or construct the faces bounding the cells containing these points. Let $K(m, n)$ be used to denote the maximum number of faces of the m cells in a generic sense where the context defines the type of the n manifolds.

The techniques for studying this problem, both here and in [19] and [20], can be viewed as divide-and-conquer attacks on this problem. In the approach taken in this paper, which we call the *primal approach*, the points are partitioned into groups according to an underlying subdivision of space into cells, to be specified shortly. We call this subdivision the *funneling subdivision* and its cells *funnels*, to distinguish them from the cells of the arrangement that we are trying to compute. The funnels define the subproblems for the divide-and-conquer attack: the subproblem of a funnel f is the many-faces problem for the points lying in f and the manifolds intersecting f . Notice that this means that each point is passed to only one subproblem, while a manifold may be passed to many subproblems (all those corresponding to funnels it intersects). In fact, a key issue for the success of our method is the design of a strategy for choosing the funneling subdivision so that the average funnel is not cut by too many manifolds.

The technique we use for accomplishing this involves probabilistic methods: we choose a random sample R of our given manifolds and “triangulate” the cells of the arrangement $\mathcal{A}(R)$; the resulting decomposition is the funneling subdivision. The size r of the random sample R is a parameter that we use to optimize the effectivity of the method. The goal is that in an average funnel the number of points that lie in it is substantially smaller than the number of manifolds that cross it. In such cases we invoke a different kind of argument, specially tailored for each specific type of manifold, to obtain a bound on $K(m, n)$. We call these bounds *Canham thresholds*, after Raymond J. Canham who was the first to prove such a bound for the case of lines [6].

The “triangulation” involves breaking each cell of the arrangement $\mathcal{A}(R)$ into pieces such that each of them has constant descriptive complexity. We need this refining of the cells in order to be able to apply to funnels arguments related to the ε -net theory [38], [11], [12]. The triangulation method is specific to the manifolds that we have in each case. For instance, in the case of lines we can draw a vertical line segment through each vertex and thus cut the cells right above and right below the vertex. This results in a decomposition of each cell into trapezoids which have finite description since they have at most four sides each.

Unfortunately there is one more difficulty to overcome. It can happen that the cell surrounding a point p in a subproblem need not be the same as the cell surrounding p in the full problem. The two can be different if the cell surrounding p in the full problem “spills out” of the funnel in which p happens to fall. But then all these problematic cells are cells intersected by the triangulation manifolds, the manifolds introduced to cut down to constant size the complexity of the funnels forming the funneling subdivision. All cells intersecting a specific manifold define the *zone* of that manifold, and good combinatorial bounds for the complexity of zones are already known for many interesting cases. Thus we are able to deal with these difficult cells by a special argument.

In the approach taken [19] and [20], which we call the *dual approach*, we essentially use a funneling subdivision in the dual space. Our manifolds are partitioned into “bundles,” and a point is distributed to all the bundles that it splits, where points not splitting a bundle lie on the same side of all manifolds of the bundle. In this approach a new difficulty arises: since a point may be distributed to

many subproblems, we need efficient ways to merge the cells for that point returned by each of the relevant subproblems. We refer the reader to [19] and [20] for details.

The Incidence Problem. The *incidence problem* is similar to the many-faces problem in that its input consists of m points and n manifolds. An *incidence* is defined as a point-manifold pair so that the point lies on the manifold. The problem is to count or bound the number of incidences that are possible for certain classes of manifolds. We use $I(m, n)$ generically to denote the maximum number of such incidences.

The approach to proving bounds on $I(m, n)$ is essentially the same as to the many-faces problem, so we just note the differences that arise simply because the problems are different. Overall, the incidence problem is easier since we do not have to worry about cells of the arrangement that cross boundaries in the funneling subdivision. On the other hand, we have to worry about points that lie on the boundary between two funnels. In the case of the many-faces problem, we could assume that the points do not lie on any manifolds. But by the nature of the incidence problem configurations that maximize $I(m, n)$ will most likely be degenerate. For instance, we cannot assume that the points do not lie on any of the manifolds. This creates the need for upper bounds on the maximum number of incidences between the original manifolds and the points that lie on the sampled manifolds.

The results of this paper are summarized in the next section. There we also give pointers to the literature and to the places in this paper where the results can be found. Roughly, the structure of this paper is as follows. In Section 3 we describe the details of our method for the case of the many-faces problem in line arrangements. This problem has a relatively easy proof and is used to describe the general proof method without too much interference by technical difficulties that arise for particular types of manifolds. Section 4 gives a general technique for proving Canham thresholds. The major tools in that section are an extremal theorem for bipartite graphs and various elementary geometry and topology lemmas. All two-dimensional many-faces and incidence results, including applications to two-dimensional distance problems, are described in Section 5. In Section 6 we derive results on the incidence problem for points and spheres in three dimensions and on related distance problems in space. Finally, we discuss the contributions of this paper and mention a few problems that remain open in Section 7. Throughout the paper we ignore constants and concentrate on asymptotic results when we state and prove theorems. We do, however, remark on the constants that follow from our methods whenever this seems advisable.

2. Summary of Results

The results of this paper are in three categories: bounds for many-faces problems, bounds for incidence problems, and bounds for combinatorial distance problems.

Table 2.1. Summary of bounds for many-faces problems.

	$K(m, n)$	Section(s)
Lines	$\Theta(m^{2/3}n^{2/3} + n)$	3 and 5.6
Pseudolines	$\Theta(m^{2/3}n^{2/3} + n)$	5.6
Unit-circles	$O(m^{2/3}n^{2/3}\beta(n) + n)$	5.6
Circles	$O(m^{3/5}n^{4/5}\beta(n) + n)$	5.6
Pseudocircles	$O(m^{3/5}n^{4/5}\beta(n) + n)$	5.6

We summarize the results of each category using tables which give the bounds and provide pointers to places in this paper where the final results are obtained. Some of these bounds are given in terms of a generic function $\beta(n)$ or $\beta(m, n)$ which changes from case to case. However, in each case it is a function that depends on the inverse of Ackermann's function and grows extremely slowly. In addition we briefly discuss each bound and relate it with earlier results in the literature. Beyond the three categories, we also list here a few results of this paper which are needed to obtain the combinatorial bounds and are of independent interest.

Many-Faces Bounds. The general theme in this category is to derive upper and lower bounds on $K(m, n)$, the maximum number of edges bounding m distinct cells in an arrangement of n curves of a certain kind. We have results for lines, pseudolines,¹ unit-circles, circles, and pseudocircles² (see Table 2.1).

The upper bounds for lines and pseudolines improve previous bounds given in [6] and [23]; the matching lower bound can be found in [23] (see also [15]). The lower bound for lines extends to unit-circles (and therefore also to circles and pseudocircles) since they can be blown up so as to approximate any pattern of lines arbitrarily closely within a bounded region. This implies that our upper bound for unit-circles is off by at most a factor of $\beta(n)$ from the best-known lower bound—this factor is probably an artifact of the proof technique used. The most difficult part of the upper bound proof for unit-circles is the Canham threshold (Section 4.5) which states that any m cells in an arrangement of n unit-circles have at most $O(mn^{1/2} + n)$ edges.

The upper bounds for circles and pseudocircles are also new and follow readily from the general techniques of this paper. No matching lower bounds are known. In fact, the currently best lower bound is the same as for unit-circles except for a small range of values of m where a slightly better bound can be shown (see remark (2) after Corollary 5.6).

¹ For technical reasons we assume that a pseudoline intersects any vertical line in one point.

² A family of pseudocircles is a set of simple closed curves with the property that any two intersect in at most two points where they cross. For technical reasons we assume that any vertical line meets a pseudocircle in at most two points.

Table 2.2. Summary of bounds for incidence problems.

	$I(m, n)$	Section(s)
Lines	$\Theta(m^{2/3}n^{2/3} + m + n)$	3 and 5.3
Pseudolines	$\Theta(m^{2/3}n^{2/3} + m + n)$	5.3
Unit-circles	$O(m^{2/3}n^{2/3} + m + n)$	5.3
Circles	$O(m^{3/5}n^{4/5} + m + n)$	5.3
Pseudocircles	$O(m^{3/5}n^{4/5} + m + n)$	5.3
Spheres in general position	$O(m^{3/4}n^{3/4}\beta(m, n) + m + n)$	6.4
Spheres and vertices	$O(m^{4/7}n^{9/7}\beta(m, n) + n^2)$	6.4

Incidence Bounds. As in the introduction, we use $I(m, n)$ generically for the maximum number of incidences between m points and n curves or surfaces in two or three dimensions. Recall that we have an incidence if a point lies on a curve or surface. In two dimensions we have results for lines, pseudolines, unit-circles, circles, and pseudocircles. In three dimensions we consider spheres and distinguish two cases: (i) no three spheres intersect in a common circle (referred to as spheres in “general position”), and (ii) the points are restricted to vertices of the sphere arrangement. The bounds are listed in Table 2.2.

The bound for lines is not new and goes back to Szemerédi and Trotter [57]; however, our proof is much simpler than theirs. Furthermore, the constant of proportionality in our analysis is only $3\sqrt[3]{6}$ (see Section 3) whereas [57] establishes the upper bound with a constant equal to 10^{60} . Our proof also extends to pseudolines for which case our upper bound appears to be new.

The upper bound for unit-circles is also not new and was derived earlier in [55]. Again, our proof is simpler and the constant that follows from our proof is much smaller than theirs. It seems likely that the upper bound is not tight; the best lower bound known for the case $m = n$ is $n^{1+c/\log \log n}$ for some constant c (see [24]).

For circles (and pseudocircles) our upper bound improves the $O(m^{3/4}n^{3/4} + m + n)$ bound of [5] and [10]. Again, no matching lower bound is known, although $\Omega(m^{2/3}n^{2/3} + m + n)$ can be derived from the lower bound for lines (see remark (3) after Theorem 5.4) and a slight improvement of this lower bound for a small range of values of m follows from a result by Erdős [24] (see remark (2) after Corollary 5.6).

Our bounds for spheres in general position improves the $O(m^{4/5}n^{4/5} + m^{1/2} + m)$ bound of Chung [10]. A lower bound of $\Omega(m^{4/3} \log \log m)$ can be obtained in the case $m = n$ from the currently best lower bound for the unit-distance problem in three dimensions discussed below. The most difficult step in the proof of the upper bound is the triangulation of an arrangement of spheres into a number of funnels which is only slightly supercubic in the number of spheres (see Section 6.3). This decomposition (Theorem 6.6) is of independent interest.

For the case of spheres where the points must be vertices of the arrangement our bound is the first known nontrivial upper bound. No matching lower bound is known if $m \geq cn^{5/4}$ for some constant c . To prove the upper bound for spheres and

Table 2.3. Summary of results on repeated distances.

	Bound	Section
Unit-distance in the plane	$O(m^{4/3})$	5.4
Unit-distance on a sphere	$\Theta(m^{4/3})$	5.4
Unit-distance in space	$O(m^{3/2}\beta(m))$	6.5
Bichromatic maximum distance	$\Theta(m)$	6.5
Bichromatic minimum distance	$O(m^{3/2}\beta(m))$	6.5

vertices we derive an extension of an extremum result for bipartite graphs with certain complete subgraphs prohibited (see Lemma 6.2); this result is of independent interest.

Distance Problems. Using the bounds for incidence problems summarized in Table 2.2 we derive new bounds for a variety of combinatorial distance problems in two and three dimensions. In each such problem we consider a set of m points and the multiset of $\binom{m}{2}$ distances, one for each point pair. For example, we ask how often a certain distance can be repeated and how many different distances the multiset must contain. We first list our results on repeated distances (see Table 2.3).

The upper bound on unit-distances in the plane is not new and dates back to Spencer *et al.* [55]—however, their constant of proportionality is much larger than the one that follows from our proof. Our upper bound proof for unit-distances in the plane carries over to points on a sphere in three dimensions (see remark (4) after Theorem 5.4) in which case a matching lower bound can be constructed from an arrangement of $m/2$ lines and $m/2$ points with $\Omega(m^{4/3})$ incidences (see [29]). It is interesting that for this lower bound it is necessary that the unit-distance is exactly one-fourth of the length of a great-circle. For other distances the currently best lower bound is $\Omega(m \log^* m)$ [29]. The upper bound for unit-distances in three dimensions improves the best previous bound which was $O(m^{8/5})$ [10]. We consider this the most important new result on counting distances derived in this paper. The currently best lower bound for this problem is $\Omega(m^{4/3} \log \log m)$ [25].

In the bichromatic case we have a total of m red and blue points in three-dimensional space and we consider only distances between points of different color. The linear upper bound on the number of times the maximum bichromatic distance can occur appears to be new although the proof is only a minor extension of the proofs of the same asymptotic bound for the number of diameters in the monochromatic case (see [33], [39], and [56]). For the minimum bichromatic distance our upper bound is new; however, no superlinear lower bound is known.

A problem related to repeated distances is that of the maximum number of furthest neighbor pairs in a set of m points in three dimensions. If no three points are collinear we show that $O(m^{3/2}\beta(m))$ is an upper bound which improves the $O(m^{8/5})$ bound of [10] (see Section 6.5). No superlinear lower bound is known.

Finally, we have some new results on different distances. Let $P = \{p_1, p_2, \dots, p_m\}$ be a set of points either in two or in three dimensions. For

Table 2.4. Summary of results on different distances.

	Bound	Section
$g(m)$ in the plane	$\Omega(m^{7/4})$	5.4
$g(m)$ in space (no collinearity)	$\Omega(m^{5/3}/\beta(m))$	6.5

$1 \leq i \leq m$ define g_i as the number of different distances from p_i . Furthermore, define $g(P) = \sum_{i=1}^m g_i$, and $g(m) = \min_{|P|=m} \{g(P)\}$. Our results are listed in Table 2.4.

The lower bound in two dimensions improves the $\Omega(m^{5/3})$ bound which follows from the bounds in [5] and [10] on the number of incidences between points and circles in the plane. In three dimensions our lower bound is established only for the case where no three points are collinear; it improves the $\Omega(m^{3/2})$ bound which follows from the incidence bound on points and spheres in [10].

3. Arrangements of Lines—An Example

In this section we show that the maximum number of edges, $K(m, n)$, bounding m cells in an arrangement of n lines is $O(m^{2/3}n^{2/3} + n)$. This improves the known upper bounds $O(m^2 + n)$ [6] and $O(mn^{1/2} + n)$ and $O(m^{1/2}n + m)$ [23]. The proof illustrates the use of the general proof components listed in the introduction. Indeed, this section serves as an introduction to our techniques before they are discussed in greater generality and different contexts in Sections 4–6. We also extend the proof to the related incidence problem and show that $I(m, n) = O(m^{2/3}n^{2/3} + m + n)$. Another purpose of this section is to demonstrate that our techniques yield a proof for this result which is simpler than that in [57]. In addition, we get the extra benefit that the constant of proportionality that we get is reasonably small whereas the constant in [57] is astronomically large.

Let $L = \{\ell_1, \ell_2, \dots, \ell_n\}$ be the given collection of lines, and assume that the m desired cells are designated by a collection of points $P = \{p_1, p_2, \dots, p_m\}$, where each p_i lies in a unique cell. The proof proceeds through the following six steps.

(1) *Establishing a Canham Threshold.* Canham's upper bound of $O(m^2 + n)$ on $K(m, n)$ can readily be transformed (by breaking the given cells into groups of $n^{1/2}$ cells each) to $K(m, n) = O(mn^{1/2} + n)$ which is the form that we use. An alternative proof of the latter bound can be derived from a well-known extremal theorem on graphs (see [45], [26], and [52]). We restate this theorem as Lemma 4.1 in a form that is convenient and allows us to derive immediately many similar bounds for curves other than lines and for spheres in three dimensions.

(2) *Sampling and Triangulating.* We choose a subset $R \subseteq L$ of size r , where r is an integer to be specified later. Let $\mathcal{A}(R)$ be the arrangement defined by R and assume that each cell is further decomposed into trapezoids by drawing a maximal vertical (relatively open) line segment through each vertex of $\mathcal{A}(R)$ so that the only intersection of the line segment with the lines in R is this vertex (see Fig. 3.1). This

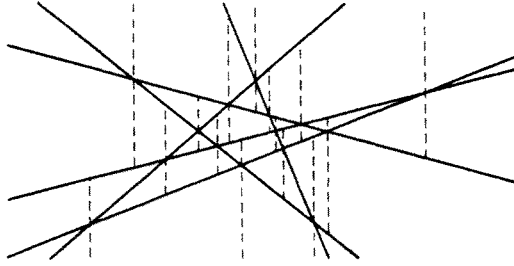


Fig. 3.1. Each cell is decomposed into trapezoids.

results in a collection of $k = O(r^2)$ (open) trapezoids,³ the aforementioned funnels. For each funnel or trapezoid Δ_i , $1 \leq i \leq k$, define $P_i = P \cap \Delta_i$ and let L_i be the subset of lines that have a nonempty intersection with Δ_i . Set $m_i = |P_i|$ and $n_i = |L_i|$ and note that $\sum_{i=1}^k m_i$ is equal to m , the number of points.⁴

(3) *Probabilistic Counting.* Using the method of probabilistic counting we can show that there exists a subset $R \subseteq L$ of size r so that $\sum_{i=1}^k m_i n_i^{1/2} = O(m(n/r)^{1/2})$. The idea of the proof of this claim is that if R is chosen at random, then the expected value of $\sum_{i=1}^k m_i n_i^{1/2}$ is $O(m(n/r)^{1/2})$. It follows that there exists a sample with this property (in fact, most samples have this property). To compute the expected value of the above random variable, we observe that the sum is the same as $\sum_{j=1}^m q_j^{1/2}$, where q_j is the number of lines intersecting the funnel that contains $p_j \in P$. Clearly, $q_j = n_i$ if $p_j \in \Delta_i$. Since the expectation is additive, even if events are dependent, we can concentrate on showing that the expected value of $q_j^{1/2}$, denoted by $E[q_j^{1/2}]$, is $O((n/r)^{1/2})$. $E[q_j^{1/2}]$ is bounded from above by $E[q_j]^{1/2}$ since $E[q_i] - E[q_j^{1/2}]^2$ is the variance of $q_j^{1/2}$ and therefore always nonnegative; so it suffices to show that $E[q_j] = O(n/r)$.

The crucial step in proving $E[q_j] = O(n/r)$ is the decomposition of the cells into regions of constant size, such as trapezoids which are determined by four or fewer lines (see Fig. 3.2 for the various ways how four or fewer lines can define a trapezoid). Such a trapezoid Δ has the property that it is in the triangulation of $\mathcal{A}(R)$ if and only if Δ is in the triangulation of $\mathcal{A}(R')$ for some $R' \subseteq R$, $|R'| \leq 4$, and Δ does not intersect any line in $R - R'$. Based on this observation the following intuitively plausible statement can be made concrete: the fact that Δ meets few lines of the random sample means that it probably meets few lines of L . The rest of the argument is technical and the details can be found in Section 5.2.

³ A tight upper bound on k is $3\binom{r}{2} + r + 1$ since every one of the at most $\binom{r}{2}$ vertices gives rise to two cell splits and the initial number of cells is at most $\binom{r}{2} + r + 1$.

⁴ Equality holds because we can assume that the points in P are in general position so that none of them lies on the boundary of a funnel.

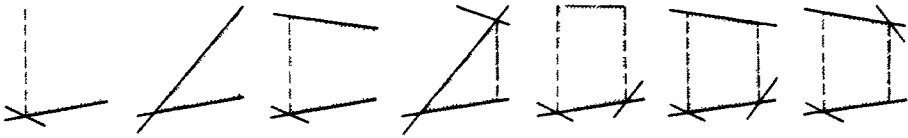


Fig. 3.2. Every trapezoid is defined by four or fewer lines.

(4) *Divide-and-Conquer.* Using the triangulation of the sample arrangement $\mathcal{A}(R)$ we solve a subproblem in each funnel. Specifically, within Δ_i we consider the n_i lines of L_i and count the edges of $\mathcal{A}(L_i)$ bounding the m_i cells marked by the points in P_i . The division of the original problem into $k = O(r^2)$ subproblems, however, is not as straightforward as might be hoped. Take, for example, a point $p \in P_i$ and let $c_i(p)$ be the cell in $\mathcal{A}(L_i)$ not necessarily restricted to Δ_i , that contains p . If $c_i(p)$ is contained in Δ_i , then $c_i(p) = c(p)$, the cell in the full arrangement $\mathcal{A}(L)$ that contains p . In this case the edges of $c(p)$ are all accounted for in the subproblem for Δ_i . But what if $c_i(p)$ intersects the boundary of Δ_i ? In this case we have $c_i(p) \cap \Delta_i = c(p) \cap \Delta_i$. The edges of $c(p)$ that are or contain edges of $c_i(p) \cap \Delta_i$ are accounted for, but it is possible that $c(p)$ also has other edges that lie fully outside Δ_i . For example, there can be a funnel $\Delta_j, j \neq i$, and a cell c_j in $\mathcal{A}(L_j)$ so that $c_j \cap \Delta_j = c(p) \cap \Delta_j$. The difficulty with this case comes from the fact that Δ_j does not contain p and c_j is therefore not marked as a cell whose edges must be counted. An important observation which facilitates a way to overcome this difficulty is that c_j intersects the boundary of Δ_j since $c(p) \subseteq c_j$ and $c(p)$ is not contained in Δ_j . Thus, c_j belongs to what we call the “inner zone” of Δ_j (more about this concept in the next step of the proof). To be on the safe side we count the edges (within Δ_j) of all inner zone cells for all $1 \leq j \leq k$.

(5) *Complexity of Zones.* The inner zone of Δ_j in $\mathcal{A}(L_j)$ is defined as the collection of $c \cap \Delta_j$ for all cells c in $\mathcal{A}(L_j)$ that intersect the boundary of Δ_j (see Fig. 3.3). The combinatorial complexity of the inner zone of Δ_j is defined as the total number of edges bounding its cells.⁵

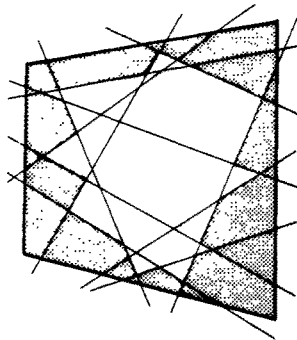


Fig. 3.3. The inner zone of a funnel.

⁵ Not that it matters much but an edge is counted twice if it bounds two cells of the inner zone.

The above discussion implies that the total number of edges bounding the m cells in $\mathcal{A}(L)$ is thus bounded from above by $\sum_{i=1}^k O(m_i n_i^{1/2} + n_i)$ (using the Canham threshold for the cells contained in their funnels) plus the combinatorial complexity of all inner zones, one per funnel. In the case of line arrangements, the complexity of a zone is linear in the number of lines (see [8], [21], or [15]).⁶ Therefore, the complexity of the inner zone of Δ_j is $O(n_j)$ which implies that all inner zones together have combined combinatorial complexity $\sum_{j=1}^k O(n_j)$.

This and the preceding sum of Canham bounds make it necessary to bound the sum of the n_j taken over all funnels Δ_j . Of course, $\sum_{j=1}^k n_j = \sum_{i=1}^n l_i$, where l_i is the number of funnels intersected by line ℓ_i .⁷ Since ℓ_i intersects the lines of R in at most r points it intersects at most $r + 1$ cells of $\mathcal{A}(R)$. By the zone result, invoked again, the total number of edges bounding those cells is $O(r)$. Since a cell with t edges is decomposed into at most t funnels, this implies that ℓ_i intersects $O(r)$ funnels of $\mathcal{A}(R)$. Therefore $\sum_{j=1}^k n_j = O(rn)$ which implies $\sum_{j=1}^k O(n_j) = O(rn + k) = O(rn)$.

(6) *Wrapping It Up.* By what we have said so far, assuming R is a sample that satisfies the property discussed in (3),

$$K(m, n) = \sum_{i=1}^k O(m_i n_i^{1/2} + n_i) = O\left(m \left(\frac{n}{r}\right)^{1/2} + rn\right).$$

We choose r so as to balance the two terms of the above bound, namely $r = \Theta(m^{2/3} n^{-1/3})$. Note that this is meaningful only if m is larger than $n^{1/2}$; fortunately for us, Canham's bound implies $K(m, n) = O(n)$ in the other case. Putting the two cases together we thus obtain the asserted bound

$$K(m, n) = O(m^{2/3} n^{2/3} + n).$$

Remark. The upper bound on $K(m, n)$ whose proof is sketched above is tight. The lower bound example is based on arranging the points in a square grid and choosing the lines close to highly populated (not necessarily horizontal or vertical) rows of points (see [15] for details).

Computing the Constant of the Upper Bound. The above proof is simple enough so that we can compute a reasonable constant of proportionality without adding too much complication to the proof. We do this by computing in turn the constants in the Canham threshold, the probabilistic counting result, and the complexity of zones. For this we use results that are derived later in this paper, specifically in Sections 4.1 and 5.2. The reader may choose to skip the rest of this section on first reading and return to it after these results are presented.

⁶ In these sources a *zone* is defined as the collection of cells in a line arrangement that intersect another line. It is plain that the maximum number of edges bounding the inner zone of a trapezoid Δ_j (not counting the edges that lie on sides of Δ_j) is at most four times as large as the maximum number of edges bounding a (standard) zone in $\mathcal{A}(L_j)$.

⁷ If $\ell_i \in R$, then $l_i = 0$ since we treat funnels as open sets.

Any two cells in a line arrangement can have at most four lines contributing edges to both cells. Using remark (1) after the bipartite graph lemma (Lemma 4.1) we can set $s = 2$ and $t = 5$ and thus get $2m_i n_i^{1/2} + n_i$ as an upper bound on the number of edges bounding m_i cells in an arrangement of n_i lines. When we take the sum of the first term over all funnels Δ_i we can use the probabilistic counting result (to be established in Section 5.2) which says that there exists a sample $R \subseteq L$ of size r such that $\sum_{\Delta_i} m_i n_i^{1/2} \leq 2m/n(r - 4)^{1/2}$. To get a good constant in the bound for the sum of the n_i is more difficult; it is treated implicitly when we get a bound on the number of edges contributed by the zones of the funnels.

First we need to observe that the inner zone of a half-plane (a degenerate trapezoid) in an arrangement of n_i lines has at most $4n_i - 2$ edges that do not also lie on the bounding line of the half-plane (see [15]). Here and later we count an edge twice (each side once) if the cells on both sides belong to the inner zone. Second, notice that a cell of $\mathcal{A}(L)$ that lies in more than one funnel must intersect at least one vertical side of each funnel it meets. This is because the top and bottom sides of a funnel are part of two original lines. Thus, we count at most $8n_i - 4$ inner zone edges for Δ_i . If such a counted edge meets a vertical side of Δ_i (and there are at most $4n_i - 4$ edges which do not⁸), then this edge is double counted because it reaches into the neighboring funnel where it is part of the inner zone again. This shows that the contribution of all inner zones is at most

$$\sum_{\Delta_i} (4n_i - 4) + \frac{1}{2} \sum_{\Delta_i} 4n_i = \sum_{\Delta_i} (6n_i - 4).$$

To get a bound on this sum we use $\sum_{\Delta_i} n_i = \sum_{i=1}^n l_i$, where l_i is the number of funnels intersected by line ℓ_i . We now show $l_i \leq 5r - 5$. Line ℓ_i intersects at most $r + 1$ cells of the sample arrangement, and by the zone theorem (see [15]) these cells have a total of at most $6r - 6$ vertices (where we count a vertex once for each such cell to which it belongs). The two endpoints of each edge of $\mathcal{A}(R)$ that intersects ℓ_i are counted at least twice which shows that there are at most $4r - 6$ distinct vertices. Each of these vertices is an endpoint of a vertical funnel side that potentially intersects ℓ_i . Adding to this the at most r intersections with lines in R gives at most $5r - 6$ intersections with sides of funnels and therefore $l_i \leq 5r - 5$.

We now put these partial results together and choose r to minimize the resulting constant. We have

$$\begin{aligned} K(m, n) &\leq \sum_{\Delta_i} (2m_i n_i^{1/2} + n_i + 6n_i - 4) \leq 4m \left(\frac{n}{r - 4} \right)^{1/2} + \sum_{\Delta_i} 7n_i \\ &\leq 4m \left(\frac{n}{r - 4} \right)^{1/2} + 35rn - 35n. \end{aligned}$$

⁸ We arrive at the $4n_i - 4$ as follows. First assume that all n_i lines intersect both vertical sides of the trapezoid. In this case, there are $2n_i$ edges meeting each side; thus at most $4n_i - 4$ edges remain. Without this assumption the number of edges that do not meet a vertical side can only be smaller since we could extend the trapezoid vertically until all lines that intersect it also meet its two vertical sides.

If we choose $r = cm^{2/3}n^{-1/3} + c'$, with $c = (\frac{2}{35})^{2/3}$ and $4 \leq c' < 5$ so that r is an integer, we get

$$K(m, n) \leq \frac{4}{c^{1/2}}m^{2/3}n^{2/3} + 35cm^{2/3}n^{2/3} + 35(c' - 1)n \leq 3\sqrt[3]{140}m^{2/3}n^{2/3} + 140n,$$

where $3\sqrt[3]{140}$, the constant factor of the leading term, is approximately 15.58. The above choice of r is appropriate as long as $m \geq \frac{35}{2}n^{1/2}$. In the other case the Canham threshold implies $K(m, n) \leq 36n$ which shows that the above bound is correct without any restriction on the number of cells.

Extension to Incidence Problem for Lines. It is rather straightforward to prove $2I(m, n) - 2m \leq \bar{K}(m, 2n)$, where $\bar{K}(m, 2n) = K(m, 2n)$ if $m \leq \binom{2n}{2} + 2n + 1$, the maximum number of cells in an arrangement of $2n$ lines, and $\bar{K}(m, 2n) = 4n^2$, otherwise. Start with a configuration that maximizes $I(m, n)$ and remove all points that do not lie on any vertex of the arrangement; this decreases the number of incidences by at most m . Now, replace each of the remaining $m' \leq m$ points by a sufficiently small disk, each with the same radius, and replace each line by two parallel lines that are tangent to all disks corresponding to points on the original line (see [23]). We thus arrive at an arrangement of $2n$ lines with m' cells bounded by a total of at least $2(I(m, n) - m)$ edges. This implies $I(m, n) \leq \frac{1}{2}K(m', 2n) + m$ and, using the above bound on $K(m, n)$, $I(m, n) \leq 3\sqrt[3]{70}m^{2/3}n^{2/3} + m + 140n$. We improve the constants of this bound below.

We follow the analysis given above for the many-faces problem but modify it in two significant ways. The first change addresses the fact that we can no longer assume that the lines are in general position. More specifically, we have to permit parallel and concurrent lines but we may still assume that no line is vertical and that no point is vertically aligned with another vertex of the arrangement, since we can rotate the configuration of lines and points without changing their incidences. The second assumption turns out to be very convenient since it implies that no point lies on a vertical side of a trapezoid defined by four or fewer lines. The second change has to do with the fact that in order to maximize the number of incidences all points lie on the lines rather than in the cells of the arrangement. Since we do not count edges of the cells we ignore the contribution of zones altogether, but now we have to take into account the number of incidences that happen on the lines of the sample arrangement. We address the first issue first and then derive the upper bound with the improved constants of proportionality.

A crucial assumption in the probabilistic counting argument is that a trapezoid is a funnel of a sample arrangement if and only if it is a trapezoid defined by a certain unique set of four or fewer lines of the sample and it does not intersect the other lines of the sample. This is because two such trapezoids are necessarily disjoint. Notice, however, that this disjointness property is lost if the lines are not in general position. To avoid the technical difficulties in dealing with lines in special

position we perturb the lines ever so slightly so that they are in general position. The trapezoid defined by four or fewer original lines is just the unperturbed version of the trapezoid of the corresponding perturbed lines, and we say that it does not “intersect” any sample lines if the corresponding perturbed trapezoid does not intersect any of the perturbed sample lines. Observe that the perturbation is only used to define and classify trapezoids—the arrangement of the problem is left unperturbed so no incidences are changed. Thus the claim about the existence of a sample of r lines with certain properties discussed in (3) above holds for degenerate cases as well.

We can now go ahead with the analysis. For the Canham threshold we get $I(m, n) \leq mn^{1/2} + n$ since no two points lie on two common lines; setting $s = t = 2$ and using remark (1) after the bipartite graph lemma (Lemma 4.1) implies the bound. Now take a “good” sample $R \subseteq L$ of size r , triangulate $\mathcal{A}(R)$, and consider the funnels Δ_i separately. This gives

$$I(m, n) \leq \sum_{\Delta_i} (m_i n_i^{1/2} + n_i) + m + rn,$$

where m_i is the number of points in the i th funnel and n_i is the number of lines that cut it. The term $m + rn$ on the right-hand side accounts for possible incidences involving points that lie on sample lines.⁹ Using the bound from the probabilistic counting argument and $\sum_{\Delta} n_i \leq (5r - 5)n$ as derived earlier we thus get

$$I(m, n) \leq 2m \left(\frac{n}{r-4} \right)^{1/2} + 6rn - 5n + m.$$

If we now choose $r = 6^{-2/3} m^{2/3} n^{-1/3} + c'$ (assuming $6n^{1/2} \leq m \leq \binom{n}{2}$), with $4 \leq c' < 5$ so that r is integer, we get

$$I(m, n) \leq 2\sqrt[3]{6m^{2/3}n^{2/3}} + \sqrt[3]{6m^{2/3}n^{2/3}} + (6c' - 5)n + m.$$

Thus,

$$I(m, n) \leq 3\sqrt[3]{6m^{2/3}n^{2/3}} + 25n + m.$$

The multiplicative constant of the leading term, $3\sqrt[3]{6}$, is approximately 5.45. When $m < 6n^{1/2}$ we get $I(m, n) \leq 7n$ from the Canham threshold, and we get $I(m, n) \leq \binom{n}{2} + m < 2m$ if m exceeds $\binom{n}{2}$ since at most $\binom{n}{2}$ points can be incident with more

⁹ Each line that does not belong to the sample intersects the r sample lines in at most r points and therefore contributes at most r incidences. Each sample line has at most r incidences with points that lie at vertices of the sample arrangement. Finally, each point that lies on a sample line but is not a vertex of the sample arrangement is incident to at most one sample line.

than one line each. Thus, if we replace the additive term m by $2m$ in the above upper bound for $I(m, n)$, then it holds without any restriction on the value of m relative to n . In summary we have shown

Theorem 3.1. *For arrangements of lines in the plane we have*

$$K(m, n) \leq 3\sqrt[3]{140m^{2/3}n^{2/3}} + 140n,$$

and

$$I(m, n) \leq 3\sqrt[3]{6m^{2/3}n^{2/3}} + 25n + 2m.$$

4. Canham-like Upper Bounds

Notice that the nature of our partitioning process (exemplified in Section 3) is such that each point is passed to exactly one subproblem, while a manifold may be distributed to several subproblems, according to the number of funnels it cuts. Thus we can expect that in a subproblem the ratio of manifolds to points increases, on the average. We use this effect to our advantage when we choose the size of the sample which determines the number of subproblems created. The reason is that when the number of points is small compared with the number of manifolds, we can use a different type of argument to bound the complexity of the cells containing the points.

These other arguments employ techniques from extremal graph theory on the bipartite graph defined by the incidences between the points or cells and the given manifolds. They usually yield bounds of the form $I(m, n)$ or $K(m, n) = O(mn^{1-1/k} + n)$ for some small integer k (Canham's original bound for lines can also be cast in this form). Thus if the number of points is really small, say $m = O(n^{1/k})$, compared with the number of manifolds, then these arguments show that $I(m, n)$ or $K(m, n) = O(n)$ and provide the basis for our reduction.

In this section we extend Canham's bound for lines to arrangements of more general curves such as circles and pseudocircles. These results will serve as bases which can be lifted to bounds for general values of m using the techniques exemplified in Section 3 and described fully in Section 5. Section 4.1 presents the extremal graph lemma which will be sufficient to derive the Canham thresholds for incidences (see Section 4.2). Additional tools from elementary geometry and topology are needed to generalize the bounds to edges bounding cells (see Sections 4.3-4.5).

4.1. A Lemma on Bipartite Graphs

Let $\mathcal{G} = (M \cup N, A)$ be a directed bipartite graph with sources in M and sinks in N , that is, $M \cap N = \emptyset$ and $(\mu, \nu) \in A$ only if $\mu \in M$ and $\nu \in N$. Let \mathcal{G} be *complete* if $A = M \times N$ in which case we write $\mathcal{G} = K_{m,n}$, $m = |M|$ and $n = |N|$. Another bipartite digraph $\mathcal{G}' = (M' \cup N', A')$ is a *subgraph* of \mathcal{G} if there are one-to-one

functions $f: M' \rightarrow M$ and $g: N' \rightarrow N$ so that $(f(\mu'), g(\nu')) \in A$ if $(\mu', \nu') \in A'$. Clearly, if \mathcal{G}' is a subgraph of \mathcal{G} , then $|M'| \leq |M|$ and $|N'| \leq |N|$. We also say that \mathcal{G} contains no $K_{s,t}$ if $K_{s,t}$ is not a subgraph of \mathcal{G} .

The following lemma establishes upper bounds on the size of A depending on the values of s and t such that \mathcal{G} contains no $K_{s,t}$. The lemma is a variant of various graph-theoretic extremal results obtained by Kővári *et al.* [45] and Erdős [26]. For completeness we include a proof of the lemma.

Lemma 4.1 (Bipartite Graph Lemma). *Let s and t be constants and let $\mathcal{G} = (M \cup N, A)$ be a bipartite digraph with $m = |M|$ sources and $n = |N|$ sinks. If \mathcal{G} contains no $K_{s,t}$, then $|A| = O(mn^{1-1/s} + n)$ and, symmetrically, $|A| = O(nm^{1-1/t} + m)$.*

Proof. We just prove the first upper bound—the second follows by symmetry. Let us define an s -regular hypergraph¹⁰ $\mathcal{H} = (M, H)$ as follows:

for every $v \in N$ adjacent to $\mu_1, \mu_2, \dots, \mu_s \in M$ we add $\{\mu_1, \mu_2, \dots, \mu_s\}$ to H .

With this definition the same s -tuple can be added to H two or more times, thus \mathcal{H} is really a multihypergraph. On the other hand, any s nodes in M are adjacent to at most $t - 1$ common nodes in N by assumption which limits the number of s -tuples in H to at most $(t - 1) \binom{m}{s}$.

Now let n_i be the number of nodes in M adjacent to $v_i \in N$ for $1 \leq i \leq n$. By definition of \mathcal{H} , v_i gives rise to $\binom{n_i}{s}$ s -tuples in H (which is 0 when $n_i < s$). Thus,

$$\sum_{i=1}^n \binom{n_i}{s} \leq (t - 1) \binom{m}{s}.$$

From this we get

$$\sum (n_i - (s - 1))^s \leq (t - 1)m^s,$$

where the summation extends over all i such that $n_i \geq s$. We can now use the Hölder inequality¹¹ to derive

$$\sum_{i=1}^n n_i \leq (s - 1)n + \sum_{n_i \geq s} (n_i - (s - 1)) \leq (s - 1)n + n^{1-1/s}(t - 1)^{1/s}m.$$

This proves the lemma since $|A| = \sum_{i=1}^n n_i$ and s and t are constants by assumption. □

¹⁰ A hypergraph is a pair (X, Y) , where X is a set of nodes and Y is a set of subsets of X . It is s -regular if every subset in Y contains exactly s nodes.

¹¹ In its full generality the Hölder inequality says that $(\sum_{i=1}^n a_i^p)^{1/p} (\sum_{i=1}^n b_i^q)^{1/q} \geq \sum_{i=1}^n a_i b_i$ assuming the a_i and b_i are nonnegative, $p > 1$, and $1/p + 1/q = 1$ [36]. In all applications we set $b_i = 1$ for all i .

Remarks. (1) The final inequality in the proof of the bipartite graph lemma holds independent of whether or not s and t are constants. We thus get the following slightly stronger upper bounds on the cardinality of A :

$$|A| \leq \min\{mn^{1-1/s}(t-1)^{1/s} + n(s-1), nm^{1-1/t}(s-1)^{1/t} + m(t-1)\}.$$

(2) For $s = t = 2$ it is known that the bound in the bipartite graph lemma is asymptotically tight—the best constants have been derived by Reiman [52]. This implies that the first bound in the lemma is asymptotically tight if $s = 2$ as long as t is a constant. No matching lower bounds are known for the case $s > 2$.

(3) In Section 6.2 we show a generalization of the bipartite graph lemma where \mathcal{G} is allowed to contain some copies of $K_{s,t}$ but not too many. This generalization will prove useful when we count incidences between points and spheres in three dimensions.

4.2. Counting Incidences

The bipartite graph lemma can be used to obtain upper bounds on the maximum number of incidences between m points and n curves of certain kinds. The curves that we consider are lines, pseudolines, unit-circles, (general) circles, and pseudocircles. We do not want to imply, however, that the methods of this paper are restricted to these curves only. The notions “pseudoline” and “pseudocircle” need some explanation. A *pseudoline* is a simple curve¹² unbounded at both ends, and we also require that it intersects any vertical line in exactly one point. A *pseudoline family* has the property that any two pseudolines in the set intersect in exactly one point where they cross. A *pseudocircle* is a simple closed (and thus bounded) curve,¹³ and we require that it intersects any vertical line in at most two points. A *pseudocircle family* is a set of such curves so that any two intersect in at most two points and if they intersect in two points then they cross there.

We do not claim that all of the following bounds are original—in fact some of them appear at various places in the literature. We include them because they are needed for our developments in Section 5.

Canham Threshold 4.2. *The maximum number of incidences between a set M of m points and a family N of n curves is*

- (i) $O(mn^{1/2} + n)$ and $O(nm^{1/2} + m)$ if N is a family of lines, pseudolines, or unit-circles, and
- (ii) $O(mn^{2/3} + n)$ and $O(nm^{1/2} + m)$ if N is a family of circles or pseudocircles.

¹² A *simple curve* is the image of a continuous one-to-one mapping f from an open interval to the plane. It is *unbounded at both ends* if $f(x)$ tends to infinity as x approaches either endpoint of the interval.

¹³ A *simple closed curve* is the image of a continuous one-to-one mapping from a circle to the plane.

Proof. We discuss lines and pseudolines, unit-circles, and circles and pseudo-circles in this sequence. In each case we consider the directed graph $\mathcal{G} = (M \cup N, A)$ with $(\mu, \nu) \in A$ if $\mu \in \nu$.

If N is a set of lines or pseudolines, then no two points lie on two or more common lines or pseudolines which implies that \mathcal{G} contains no $K_{2,2}$. The bounds follow from the bipartite graph lemma.

The case where N is a set of n unit-circles is slightly less trivial. Any two points lie on at most two common unit-circles which shows that \mathcal{G} contains no $K_{2,3}$. In addition, any two unit-circles intersect in at most two points which implies that \mathcal{G} contains no $K_{3,2}$ either. The bounds follow readily from the bipartite graph lemma.

Finally, when N is a family of circles or pseudocircles, then \mathcal{G} contains no $K_{3,2}$ because no three points can have two common (pseudo-) circles. \square

Remarks. (1) Using the bipartite graph lemma we can also show that the number of incidences is $O(nm^{1/2} + m)$ if N is a family of arbitrary curves, provided any two curves intersect in at most some constant number, $s - 1$, of points; so the corresponding \mathcal{G} contains no $K_{s,2}$.

(2) It is interesting to observe that the asymptotic bounds for lines and pseudolines also hold for unit-circles in spite of the fact that \mathcal{G} (defined for unit-circles) may contain many $K_{2,2}$'s. This shows that it is useful to formulate the bipartite graph lemma for forbidden $K_{s,t}$ in the directed sense (that is, without implicitly forbidding $K_{t,s}$ as well if $s \neq t$).

(3) The arguments used in the proof of Canham Threshold 4.2 are purely combinatorial—it is therefore not surprising that there is no difference between lines and pseudolines and between circles and pseudocircles.

(4) The only obstacle to proving $O(mn^{1/2} + n)$ for general circles is that two points can lie on an arbitrarily large number of common circles.¹⁴ Besides unit-circles there are a couple of other natural cases where such configurations are prohibited thereby leading to this improved bound. Those include sets of circles where the maximum number of collinear centers or the number of different radii is bounded from above by some constant. (See also remark (3) in Section 5.4.)

(5) Canham Threshold 4.2 implies that the number of incidences is $O(n)$ as long as m , the number of points, is $O(n^{1/2})$ in the case of lines, pseudolines, and unit-circles, and as long as m is $O(n^{1/3})$ in the case of (general) circles and pseudocircles.

4.3. Cells in (Pseudo-) Line Arrangements

An asymptotic version of Canham's bound on the number of edges bounding m cells in an arrangement of n lines can be obtained from the bipartite graph lemma using the fact that for any two cells there are at most four lines that contain edges of

¹⁴ In fact, a result of Erdos [24] on counting the number of different distances in a set of points shows that this improvement is impossible (see remark (2) in Section 5.4).

both cells. This is because any two disjoint convex sets have at most four common supporting lines, and cells in a line arrangement are certainly convex and disjoint. It is also true that any two cells in a pseudoline arrangement share at most four pseudolines supporting edges of both cells—but this is more difficult to show. We prove a bound on the number of edges bounding m cells in a pseudoline arrangement without showing that two cells share at most four pseudolines. The following straightforward topological lemma will be handy. We say that a region *touches* a curve if a nonzero length subset of the curve belongs to the boundary of the region. If the region is a cell in an arrangement including the curve, then this means that the curve contains an edge of the cell.

Lemma 4.3. *Let C be a simple curve (either closed or unbounded at both ends) that is partitioned into three (connected) pieces. There exist no two disjoint connected regions so that both lie on the same side of C and touch all three pieces of C .*

Proof. Assume that two such regions, R_1 and R_2 , exist. We can draw a third region, R_3 , on the other side of C so that R_3 touches all three pieces of C (see Fig. 4.1). But this implies a plane embedding of $K_{3,3}$ since the three pieces of C can be contracted to points and connected to a point each in R_1 , R_2 and R_3 without creating an intersection. \square

Lemma 4.3 is repeatedly used in arguments about (pseudo-) circles and also unit-circles in Sections 4.4 and 4.5. Again, we do not claim originality for the following result.

Canham Threshold 4.4. *The maximum number of edges bounding m cells in an arrangement of n lines or pseudolines is $O(mn^{1/2} + n)$.*

Proof. We only need to prove the statement for pseudolines. First, we replace each pseudoline, ℓ , by two pseudo-half-planes, the connected components of $E^2 - \ell$. Let N be the set of $2n$ pseudo-half-planes and consider $\mathcal{G} = (M \cup N, A)$, where M is the collection of m cells and $(\mu, \nu) \in A$ if cell μ is contained in pseudo-half-plane ν and the boundary of ν contains an edge of μ .

We show that \mathcal{G} contains no $K_{2,3}$. Assume \mathcal{G} does contain a $K_{2,3}$. Then there are three pseudo-half-planes h_1, h_2, h_3 and two cells in $h_1 \cap h_2 \cap h_3$ so that each

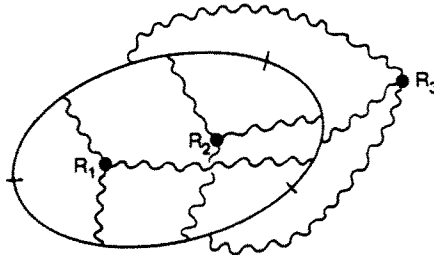


Fig. 4.1. Reduction to a planarity argument.

cell touches the boundary of each pseudo-half-plane. By definition of a family of pseudolines, $h_1 \cap h_2 \cap h_3$ is connected, its boundary is a simple curve, and the pseudoline ℓ_i bounding h_i intersects the simple curve in a connected piece for $i = 1, 2, 3$. We thus have a decomposition of this simple curve into three connected pieces so that both cells touch all pieces. But this contradicts Lemma 4.3. The assertion now follows because each $(\mu, \nu) \in A$ contributes one edge to the boundary of cell μ . \square

Remark. Notice that the graph \mathcal{G} defined in the above proof can contain subgraphs $K_{s,2}$ for s up to $n - 1$ (even in the case of lines). This keeps us from using the second bound in the bipartite graph lemma to prove $O(nm^{1/2} + m)$ as a bound on the number of edges bounding m cells in an arrangement of n lines or pseudolines. Nevertheless, this bound is correct and a proof can be found in [15].

4.4. Cells in (Pseudo-) Circle Arrangements

In this section we establish Canham thresholds for cells in arrangements of circles and pseudocircles. The bounds are the same in both cases so we can restrict ourselves to pseudocircles which are more general than circles. In the proof we replace each pseudocircle, c , by the two so-called *pseudodisks* that are the connected components of $E^2 - c$. First, we prove a technical lemma about three pseudodisks and three connected regions.

Lemma 4.5. *It is impossible to have three pseudodisks d_1, d_2, d_3 and three pairwise disjoint, connected regions contained in $d_1 \cap d_2 \cap d_3$ so that each region touches the boundary of each $d_i, i = 1, 2, 3$.*

Proof. We can restrict our attention to simple arrangements since a nonsimple arrangement can be made simple by adding vertices, edges, and cells, and thus making it easier to find three such regions. Consider all topologically different simple arrangements of three pseudocircles as shown in Fig. 4.2. We consider three cases depending on the shape of $D = d_1 \cap d_2 \cap d_3$.

Case 1. In this case D is connected and has three or fewer edges. (This includes cases where D is not simply connected.) Examples are all cells in Fig. 4.2(a)–(n) not explicitly mentioned in Cases 2 and 3. If D has less than three edges than there is nothing to prove. Otherwise, the existence of three disjoint regions in D that touch all three edges contradicts the nonplanarity of $K_{3,3}$.

Case 2. Here D is disconnected. Examples are the shaded parts in Fig. 4.2(l)–(n). In each such case, D consists of two simply connected components bounded by three edges each. By Lemma 4.3 each component can contain only one region touching all three edges which makes at most two regions altogether.

Case 3. In this case D is connected and bounded by four edges. There are 13 such D 's indicated by dashed boundary lines in Fig. 4.2. In each case, D is simply connected. Two opposite edges belong to a common pseudocircle (bounding

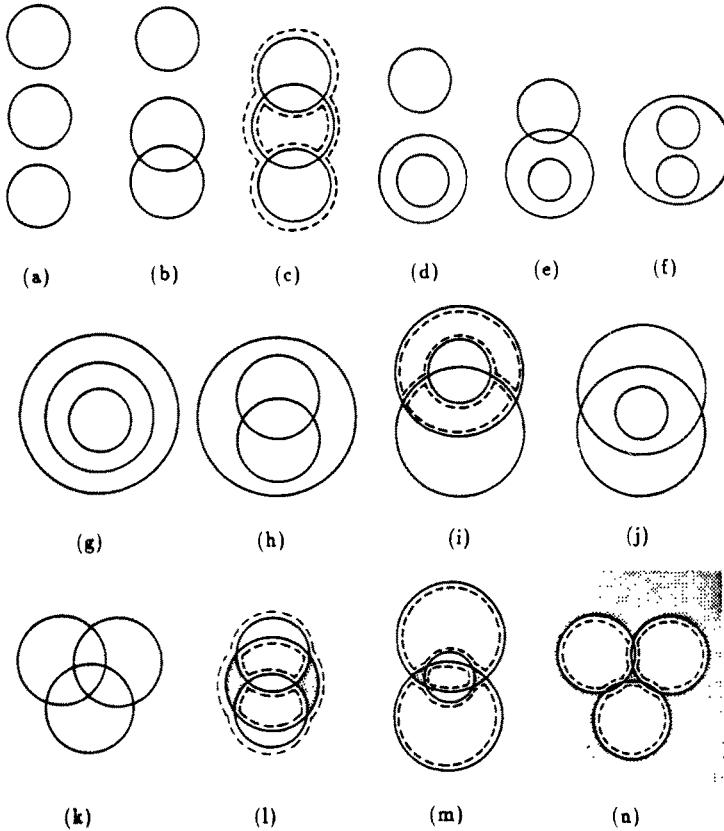


Fig. 4.2. The fourteen different simple arrangements of three pseudocircles. The dashed lines indicate regions with four edges, shading is used for disconnected regions.

a d_i) and can be connected by a simple curve that lies completely outside D . Thus, the existence of three regions in D each touching one of the two edges that belong to the same pseudocircle as well as both of the other two edges contradicts the nonplanarity of $K_{3,3}$.

The three cases exhaust all possibilities for $d_1 \cap d_2 \cap d_3$. □

Remarks. (1) Note that all fourteen different arrangements of three pseudocircles can be realized by three circles as shown in Fig. 4.2.

(2) A possible negative point of the above proof is that it relies on a (presumably) exhaustive enumeration of all simple arrangements of three pseudocircles. Alternatively, we could argue directly that D can only assume the three types of shapes considered in Cases 1-3.

We use Lemma 4.5 to prove the main result of this section.

Canham Threshold 4.6. *The maximum number of edges bounding m cells in an arrangement of n circles or pseudocircles is $O(mn^{2/3} + n)$.*

Proof. Besides Lemma 4.5 we need the following result proved in [41]:

the intersection of $k \geq 3$ pseudodisks is bounded by at most $6k - 12$ edges.

This implies that the number of edges bounding a cell in an arrangement is less than six times the number of (pseudo-) circles that contribute edges to the boundary of the cell.

Now let M be the set of m cells and let N be the set of $2n$ pseudodisks (two for each pseudocircle), and consider the graph $\mathcal{G} = (M \cup N, A)$ with $(\mu, \nu) \in A$ if cell μ is contained in pseudodisk ν and the boundary of ν contains an edge of μ . By Lemma 4.5 \mathcal{G} contains no $K_{3,3}$. The bound stated in the assertion now follows from the bipartite graph lemma (Lemma 4.1). \square

Remark. Since $O(nm^{1/2} + m)$ is an upper bound on the number of incidences between m points and n (pseudo-) circles, it might be conjectured that the same bound holds for the number of edges bounding m cells. This may be true but the currently best upper bound of this form is $O(\lambda_4(n)m^{1/2})$ which follows from the investigations in [18]. Here $\lambda_4(n)$ is superlinear only by a factor that depends on the inverse of Ackermann's function (see Section 5.5).

4.5. Cells Defined by Unit-Circles

For m cells in an arrangement of n unit-circles we can prove an upper bound on the total number of edges bounding the m cells that is asymptotically smaller than the bound derived for general circles (Canham Threshold 4.6). The special property of unit-circles that admits this improvement is that no unit-circle encloses an antipodal point pair¹⁵ of another unit-circle. This property implies that of the two arcs into which a unit-circle is cut by another unit-circle, only the shorter arc can be enclosed by the other unit-circle. The upper bound on the number of edges that we are going to establish is $O(mn^{1/2} + n)$ (Canham Threshold 4.7 below) which indicates that we use a forbidden $K_{2,t}$ argument, that is, no two cells touch t common unit-circles (we choose $t = 3$). This is, however, not true in general for any t as shown in Fig. 4.3. We get around this difficulty by considering a constant number of different cases.

Canham Threshold 4.7. *The maximum number of edges bounding m cells in an arrangement of n unit-circles in the plane is $O(mn^{1/2} + n)$.*

¹⁵ Two points of a circle are *antipodal* if they are collinear with the center of the circle.

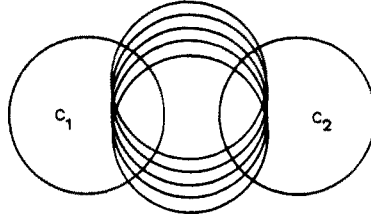


Fig. 4.3. Cells c_1 and c_2 touch five common unit-circles.

Proof. Let M denote the set of m cells, let N be the set of n unit-circles, and define $\mathcal{G} = (M \dot{\cup} N, A)$, with $(c, u) \in A$ if cell c touches unit-circle u . As mentioned above, we do a case-analysis which, in effect, considers \mathcal{G} as the union of a constant number of subgraphs. The cases that we distinguish are

- Case 1. Graph $\mathcal{G}_1 = (M \dot{\cup} N, A_1)$ where $(c, u) \in A_1$ if c touches u and is enclosed by u .
- Case 2. $\mathcal{G}_0 = (M \dot{\cup} N, A_0)$, where $(c, u) \in A_0$ if c touches u and is not enclosed by u . Note that $A = A_1 \cup A_0$. We distinguish two subcases.
 - Case 2.1. Cells that lie outside all unit-circles of N .
 - Case 2.2. Cells enclosed by unit-circles.

We next discuss these cases in sequence. Cases 1 and 2.1 are fairly straightforward and Case 2.2 is more complicated than one might hope. To simplify the discussion we define a *unit-disk* as the closed disk bounded by a unit-circle.

Case 1. Here we show that \mathcal{G}_1 contains no $K_{2,3}$, that is, it is not possible that two cells contained in three common unit-disks both touch all three bounding unit-circles. To see why this is true for unit-circles (it is not true for general circles as can be shown for the arrangement of Fig. 4.2(m)) observe that the boundary of the intersection of three unit-disks consists of at most three edges. For assume that $d_1 \cap d_2 \cap d_3$ (d_i bounded by the unit-circle u_i) has two edges contributed by u_2 . Then u_2 is contained in $d_1 \cap d_3$ which implies that one of the two unit-disks contains an antipodal point pair of u_2 , a contradiction. Thus, Lemma 4.3 applies and proves the rest.

Case 2.1. As proved in [41] the union of n pseudodisks is bounded by at most $6n - 12$ edges if $n \geq 3$. Unit-disks are just special pseudodisks and thus the bound applies here too. Since $6n - 12$ is certainly within $O(mn^{1/2} + n)$ we can ignore such cells from hereon altogether.

Case 2.2. Before we discuss this case in its generality we show that if two cells, c_1 and c_2 , can be enclosed by a common unit-circle, not necessarily in N , then they cannot both touch three common unit-circles from the outside. In contrast to Case 2.1, both cells are fairly small and therefore easier to control. To prove this we assume the contrary, that is, there are three unit-circles u_1, u_2, u_3 so that c_1 and c_2 touch and lie outside all three. Let u_0 be the unit-circle that encloses c_1 and c_2 , and let d_i denote the unit-disk bounded by u_i , $0 \leq i \leq 3$.

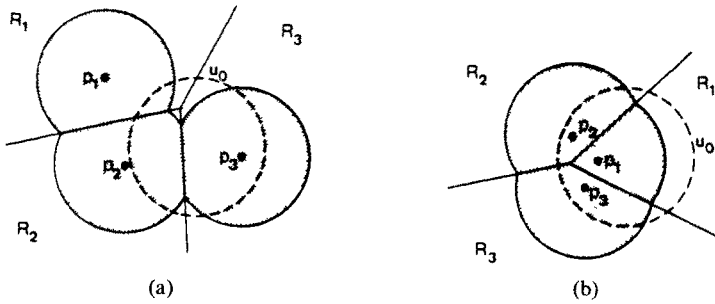


Fig. 4.4. The Voronoi diagram of the centers of u_1, u_2, u_3 . In (a) the union of the unit-disks has four edges, in (b) it has three.

Consider the Voronoi diagram of the centers p_1, p_2, p_3 of u_1, u_2, u_3 . It consists of regions R_1, R_2, R_3 so that $R_i = \{x \in E^2 \mid d(x, p_i) \leq d(x, p_j), j \neq i\}$ for $i = 1, 2, 3$ (see Fig. 4.4). Each R_i is the intersection of two closed half-planes and thus convex. The nice property of the Voronoi diagram is that $d_1 \cup d_2 \cup d_3 = (d_1 \cap R_1) \cup (d_2 \cap R_2) \cup (d_3 \cap R_3)$ which amounts to a decomposition of $d_1 \cup d_2 \cup d_3$. We distinguish two subcases.

Case 2.2.1. Here $u_0 \cap (d_i \cap R_i) \neq \emptyset$ for $i = 1, 2, 3$ (see Fig. 4.4(a)). We show that if c_1 and c_2 exist with the required properties, then we have a plane embedding of a $K_{3,3}$ and thus a contradiction. By assumption, c_1 and c_2 are contained in $d_0 - (d_1 \cup d_2 \cup d_3)$. Since both cells touch u_1, u_2, u_3 they also touch $d_0 \cap d_1 \cap R_1, d_0 \cap d_2 \cap R_2, d_0 \cap d_3 \cap R_3$ which are three convex regions with pairwise disjoint interiors enclosed by u_0 . Since each $d_0 \cap d_i \cap R_i, i = 1, 2, 3$, touches u_0 by assumption we can construct a region outside d_0 that touches each $d_0 \cap d_i \cap R_i$. We arrive at a contraction because c_1, c_2 , and the four regions give a plane embedding of $K_{3,3}$.

Case 2.2.2. Here $u_0 \cap (d_1 \cap R_1) = \emptyset$ (see Fig. 4.4(b)). We argue that $d_1 \cup d_2 \cup d_3$ is simply connected and bounded by three edges only; Lemma 4.3 will prove the rest. As mentioned before, the union of the three unit-disks is equal to $(d_1 \cap R_1) \cup (d_2 \cap R_2) \cup (d_3 \cap R_3)$. Note that d_1 contains the only vertex of the Voronoi diagram. If this were not the case then $u_1 \cap R_1$ would contain two antipodal points of u_1 , for example the intersections of u_1 with the line through p_1 and the vertex, which is impossible because $u_1 \cap R_1$ is enclosed by u_0 . The vertex of the Voronoi diagram is equally far from p_1, p_2 , and p_3 which implies that it lies in all three disks d_i . Consequently, $u_i \cap R_i$ is connected for $i = 1, 2, 3$, which implies that $d_1 \cup d_2 \cup d_3$ is bounded by only three edges and therefore also simply connected.

Now we are ready to discuss Case 2.2 in its full generality. Our goal is to show that $\mathcal{G}'_0 = (M' \cup N, A'_0)$ has not too many arcs, where M' is set of cells in M enclosed by unit-circles in N , and $A'_0 = A_0 \cap M' \times N$. We do this by expanding the graph and proving a bound for the expanded graph.

Define $\mathcal{G}_0 = (M'' \cup N'', A''_0)$ as follows. Draw vertical and horizontal lines through the integer grid points. These lines decompose each cell in M' into smaller cells, but because each cell in M' is enclosed by a unit-circle it is decomposed into at most nine new cells. The lines cut each unit-circle into at most eight disjoint arcs, and N'' is the set of all such arcs. Finally, $(c'', u'') \in A''_0$ if $c'' \in M''$, $u'' \in N''$, and c'' touches u'' and lies outside the unit-circle that contains u'' . By construction, we have $|M''| \leq 9|M'|$, $|N''| \leq 8|N|$, and $|A''_0| \geq |A'_0|$. If two cells are separated by one of the lines, then they have no common neighbor in \mathcal{G}_0 . Otherwise, they can have at most two common neighbors by the analysis in Cases 2.2.1 and 2.2.2. It follows that \mathcal{G}_0 has no $K_{2,3}$ and therefore $|A''_0| = O(mn^{1/2} + n)$.

In summary, we showed that \mathcal{G}_i contains no $K_{2,3}$ (Case 1) and that after removing all cells that lie outside all unit-circles, \mathcal{G}_0 can be expanded to less than nine times as many nodes so that the expanded graph contains no $K_{2,3}$. The cells outside all unit-circles are bounded by at most $6n - 12$ edges (Case 2.1). The removal of these cells from \mathcal{G}_0 therefore decreases the number of arcs by at most this number. Finally, we note that the number of edges of a cell in the arrangement defined by N has fewer than six times as many edges as there are unit-circles that contain those edges¹⁶ (again by a result of [41]). This together with the bipartite graph lemma (Lemma 4.1) implies the upper bound of the assertion. \square

5. General Upper Bounds

This section presents the details of our approach to proving general bounds for the incidences problem and the many-faces problem which improve the Canham-like bounds obtained in Section 4. The method is fairly modular and consists of three major steps—not counting Canham thresholds necessary to get the method off the ground—which are triangulating, probabilistic counting, and analyzing zones. At the heart of the method is a probabilistic counting result which is presented in appropriate generality in Section 5.2. It relies on a decomposition of the cells in an arrangement into regions of constant description. We call the decomposition a triangulation although the regions (funnels) are not necessarily triangles. Triangulations are the topic of Section 5.1. The analysis of zones in arrangements is discussed in Section 5.5. Putting all three components together, we prove upper bounds on the number of edges of many cells in Section 5.6. Upper bounds for incidences can be established without the use of zones. These are discussed in Section 5.3 immediately followed by applications to counting distances given in Section 5.4.

¹⁶ As a matter of fact, the blow up factor is at most two which can be shown using Davenport-Schinzel sequences of order 2 (see, e.g., [54]).

5.1. Triangulating Arrangements

The first tool that we need in order to obtain the funneling partition is the ability to "triangulate" the cells of an arrangement, that is, to break them up into finer pieces so that each piece is bounded by only a constant number of curves or surfaces. As we have already remarked, the reason for this refinement is to justify the probabilistic analysis that we use to prove that the funneling subdivision does a good job of distributing the manifolds among the funnels. The refinement is, in general, accomplished by using auxiliary manifolds from the same class as the manifolds forming the arrangement, or from a reasonably related class. In doing so we wish to keep the total number of funnels we obtain fairly close to the number of cells of the sample arrangement.

This section discusses only the two-dimensional case, and we refer to Section 6.3 for an extension of our triangulation technique to three dimensions. Triangulating an arrangement of lines is fairly straightforward and is discussed in sufficient detail in Section 3. We therefore confine this section to pseudolines, circles, and pseudocircles. The main part consists of the analysis of a particular triangulation of a circle arrangement. Thereafter we extend such triangulations to pseudolines and pseudocircles.

Arrangements of Circles. Let C be a set of n circles in the plane.¹⁷ Each cell of the arrangement, $\mathcal{A}(C)$, is decomposed by drawing vertical line segments through the left- and rightmost points of all circles. These line segments are extended as far as possible so that they intersect no other circles. In addition, we draw a maximal vertical (relatively open) line segment through every vertex of $\mathcal{A}(C)$ so that the only intersection with the circles is this vertex (see Fig. 5.1). A funnel (also called a *trapezoid*) is thus bounded by four edges: two vertical edges to the left and right and two circle edges at the top and the bottom. Up to three of these four edges can be degenerate, that is, empty or at infinity (see Fig. 5.1).

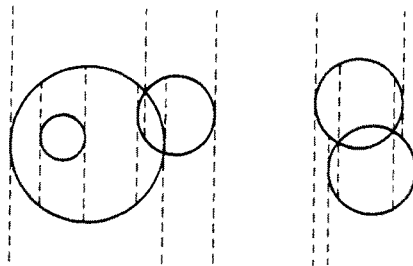


Fig. 5.1. Triangulating a circle arrangement using trapezoidal funnels.

¹⁷ We do not distinguish between the general case and the case of unit-circles.

In order to count the number of funnels generated, we perform a sweep-line argument where the line is vertical and sweeps from left to right. Initially (when the line is to the left of all circles), it intersects one funnel. Whenever the sweep-line passes the leftmost point of a circle it encounters three new funnels (assuming the circles are in general position), and when it passes a rightmost point it encounters one new funnel. When the line passes a vertex which is the intersection of two circles then it hits three new funnels. This gives an upper bound of

$$1 + 3n + n + 6\binom{n}{2} = 3n^2 + n + 1$$

for the number of funnels in the triangulation.¹⁸

Remarks. (1) The decomposition of the cells into “trapezoidal” funnels as described is not a cell complex since quite frequently we have a vertex of one funnel lie in the middle of an edge of an adjacent funnel. This has no consequences in our analysis.

(2) The line-sweep counting argument shows that the majority of the funnels are due to intersections of circles. In fact, if we define t as the number of intersecting pairs of circles in C , then the number of funnels is at most $6t + 4n + 1$.

Arrangements of Pseudolines and Pseudocircles. The decomposition of the cells into trapezoidal funnels as described for lines in Section 3 and for circles above can readily be extended to pseudolines as well as pseudocircles. This is in particular true because we define a pseudoline without x_1 -extreme points,¹⁹ and a pseudocircle with only two x_1 -extreme points.

Remarks. (1) For a pseudoline arrangement it is possible to avoid the additional assumption of no x_1 -extreme points using Levi’s lemma [46] which can be used to construct a different triangulation of the sample arrangement. We repeat the lemma in the language of this paper.

Let $\mathcal{A}(L)$ be a pseudoline arrangement and let p and q be any two points not both on the same pseudoline. Then there is a pseudoline, ℓ , passing through p and q so that $\mathcal{A}(L \cup \{\ell\})$ is still a valid pseudoline arrangement.

With this we can draw a piece of a pseudoline between any two nonadjacent vertices of a cell (of the partially triangulated sample arrangement) that has more

¹⁸ This analysis ignores cases such as vertices that lie on more than two circles and vertices that are, at the same time, left- and/or rightmost points of circles. Using a straightforward perturbation argument we can show, however, that the occurrence of such cases only decreases the number of funnels that are constructed.

¹⁹ A point p of a pseudoline or pseudocircle v is x_1 -extreme if all points on v in an ε -neighborhood of p have greater (or smaller) x_1 -coordinates than p .

than three edges. Every funnel of this triangulation is bounded by at most three edges and each edge intersects every pseudoline in at most one point where it crosses it. For reasons of uniformity we prefer to work with the triangulation using trapezoidal funnels, though.

(2) We pose the generalization of remark (1) to arrangements of pseudocircles as an open problem. Let C be a finite set of pseudocircles and let $\mathcal{A}(C)$ be the arrangement defined by C . Prove or disprove that for any two points p and q there is a pseudocircle c so that $p, q \in c$ and $\mathcal{A}(C \cup \{c\})$ is still a pseudocircle arrangement.

A crucial property of the triangulations described in this section is a certain independence property that allows us to decide whether a trapezoid is part of a triangulation without looking at its neighboring trapezoids. This property plays an important role in the probabilistic counting analysis of triangulations in the next section.

Lemma 5.1 (Independence Lemma). *Let R be a set of r curves (that is, lines, pseudolines, unit-circles, circles, or pseudocircles). A trapezoid Δ is in the triangulation of $\mathcal{A}(R)$ if and only if*

- (i) Δ is in the triangulation of $\mathcal{A}(R')$ for some $R' \subseteq R$, $|R'| \leq 4$, and
- (ii) Δ does not intersect any curve in R .

Proof. The “only if” part is straightforward since a trapezoid Δ in the triangulation of $\mathcal{A}(R)$ satisfies (ii) by definition. Furthermore, the top and bottom of Δ lie in two curves, by construction, and two additional curves suffice to determine the left and right sides of Δ . Thus, Δ is in the triangulation of these at most four curves. (There are less than four if a side is empty or at infinity; see also Fig. 3.2)

To see the “if” part notice that a trapezoid Δ that satisfies (i) and (ii) does not intersect any side used in the triangulation of $\mathcal{A}(R)$: no nonvertical side can intersect Δ because Δ is disjoint from all curves in R , and no vertical side can meet Δ because every vertex in $\mathcal{A}(R)$ is shielded from Δ by at least one of the curves in R' . Thus, Δ is contained in a trapezoid of the triangulation of $\mathcal{A}(R)$. But then this trapezoid must be equal to Δ since it intersects a line in R' , otherwise. \square

Remark. A trapezoid Δ satisfying (i) is said to be *defined* by R' . The smallest set R' that defines Δ is unique if we assume that the curves in R are in general position. This will be important later where we make crucial use of the fact that each trapezoid satisfying (i) and (ii) is defined by exactly one set R' .

5.2. Probabilistic Counting Results

The key intuition behind our partitioning scheme is that we are likely to do well if we partition based on sampling our own data. How well? For any given R of size r , on the average we expect roughly $O(n/r)$ manifolds to cut each funnel. Loosely

speaking, the reason is that in d dimensions a manifold will typically cut on the order of r^{d-1} funnels and there are on the order of r^d funnels altogether. The ε -net theory of Haussler and Welzl [38], as well as the probabilistic lemma of Clarkson [11] then guarantee that, with high probability over the random selection of R , every funnel in the funneling subdivision will behave like the “average” funnel, to within an $O(\log r)$ factor. Since we want a best-possible bound we need a tighter result here. We invoke another result also reported in more generality in Clarkson and Shor [12] that states that if we let the sample R vary, the expected number of manifolds cutting this funnel will be $O(n/r)$, with no logarithmic factor. We use this fact and the additivity of expectation to obtain our result. Recent results of Chazelle and Friedman [7] and Matoušek [48] show that for some types of input there exist triangulations for which every funnel is like the average to within a constant factor. This stronger bound is not necessary to obtain our combinatorial results.

The presentation of the probabilistic counting result in this section is general enough to cover the cases of lines, pseudolines, unit-circles, (general) circles, and pseudocircles. Even more cases (including spheres in three dimensions) are included in the more general but also more involved treatment in [12].

So let C be a set of n curves (of a type listed above) and let P be a set of m points. For a subset $R \subseteq C$ we triangulate the arrangement $\mathcal{A}(R)$ and denote the funnels by Δ_1 through Δ_k . Define $m_i = |P \cap \Delta_i|$ and let n_i be the number of curves that intersect Δ_i . For the time being we assume that the curves satisfy appropriate general position requirements and that the points avoid all curves as well as vertical funnel sides. With these definitions and assumptions we have the following preliminary result.

Lemma 5.2. *Let R be a set of r curves in the plane and let \mathcal{G}_R be the set of all trapezoids defined by four or fewer curves in R .*

- (i) *For each point p there are at most four trapezoids in \mathcal{G}_R that contain p and intersect exactly one curve each.*
- (ii) *The number of trapezoids in \mathcal{G}_R that intersect exactly one curve in R is $O(r^2)$.*

Proof. To show (i) notice that only one trapezoid in \mathcal{G}_R , we call it Δ_R , contains p and meets no curve in R . Similarly, define $\Delta_{R-\{c\}}$ for p and $R - \{c\}$, c an arbitrary curve in R . Let R' be the set of at most four curves that define Δ_R . A trapezoid Δ that satisfies the conditions in (i) has the property that it is $\Delta_{R-\{c\}}$ for exactly one curve $c \in R$. But $\Delta_{R-\{c\}} = \Delta_R$ unless $c \in R'$. The claim follows since there are at most four curves c in R' .

It is slightly more complicated to establish (ii). Take a curve $c \in R$ and consider the (unique) triangulation of $\mathcal{A}(R - \{c\})$. We count the number of trapezoids in this triangulation that meet c . This is equal to (or one more than) the number of points where c meets trapezoid sides. The number of such points that lie on nonvertical sides is $O(r)$ because c intersects the other curves in at most two points each by assumption. When we take the sum over all $c \in R$ we get $O(r^2)$ such intersections. Each vertical side of a trapezoid in the triangulation of $\mathcal{A}(R - \{c\})$

that intersects c overlaps with a vertical side in the triangulation of $\mathcal{A}(R)$ that has one endpoint on c at a point where c does not intersect any other curve. It is fairly difficult to bound the number of such vertical sides for an individual c , but it is easy to compute the sum over all $c \in R$. Namely, every vertex in $\mathcal{A}(R)$ and every point with vertical tangency generates two vertical sides each footing at a single curve c . By assumption, there are only $O(r^2)$ points generating vertical sides which implies that the total number of such intersections, summed over all $c \in R$, is $O(r^2)$. \square

Remark. With a little more care the above proof can be used to compute the constants hidden in the big-Oh notation. For lines and pseudolines we have $\binom{r}{2}$ vertices in $\mathcal{A}(R)$ and no points of vertical tangency. Thus there are at most $r^2 + 2\binom{r}{2} < 2r^2$ trapezoids in \mathcal{G}_R that meet exactly one line or pseudoline in R . For unit-circles, general circles, and pseudocircles we have $2\binom{r}{2}$ vertices in $\mathcal{A}(R)$ and $2r$ points of vertical tangency. It follows that the number of trapezoids in \mathcal{G}_R that intersect exactly one curve in R is at most $2r(r - 1) + 4\binom{r}{2} + 4r = 4r^2$.

We are now ready to present the probabilistic counting argument which leads to a result that says in a formal way that every funnel intersects $O(n/r)$ curves on the average.

Lemma 5.3 (Sampling Lemma). *Let C be a set of n curves in general position, triangulated into funnels Δ_1 through Δ_k , and let P be a set of m points avoiding the boundaries of the funnels. Define $m_i = |P \cap \Delta_i|$ and let n_i be the number of curves that intersect Δ_i . For every $0 \leq \alpha \leq 1$ and $4 < r < n$ there is a subset $R \subseteq C$ with $r = |R|$ so that*

- (i) $\sum_{i=1}^k m_i n_i^\alpha = O(m(n/r)^\alpha)$ and
- (ii) $\sum_{i=1}^k n_i = O(rn)$.

Proof. The proof consists of two steps each applying a probabilistic counting argument. The first step shows that the expectation of $\sum_{\Delta_i} m_i n_i^\alpha$, taken over all subsets R of size r , is bounded from above by $\delta_1 m(n/(r - 4))^\alpha$ for some constant δ_1 independent of α . It follows that there exists a subset R for which this inequality holds. The second step shows that the expected value of $\sum_{\Delta_i} n_i \leq \delta_2 (kn/r)$ for some constant δ_2 . Now, the subset that satisfies the first inequality may not satisfy the second and vice versa. However, by the nature of the argument we know that more than 50% of all subsets R of size r satisfy $\sum_{\Delta_i} m_i n_i^\alpha \leq 2\delta_1 m(n/(r - 4))^\alpha$, and, by the same token, more than 50% of the subsets R satisfy $\sum_{\Delta_i} n_i \leq 2\delta_2 (kn/r)$. It follows that there is at least one subset R that satisfies both inequalities.

Step 1. As indicated above, the idea of the proof is that if R is chosen at random, then the expected value of $\sum_{\Delta_i} m_i n_i^\alpha$ is $O(m(n/r)^\alpha)$. To compute the expected value of

the above random variable, we observe that the sum is the same as $\sum_{j=1}^m q_j^2$, where q_j is the number of curves intersecting the funnel that contains point $p_j \in P$. Clearly, $q_j = n_i$ if $p_j \in \Delta_i$. Since the expectation is additive we can concentrate on showing that the expected value of q_j^2 , denoted by $E[q_j^2]$, is $O((n/r)^2)$. By Jensen's inequality (see, for example, [30]), $E[q_j^2] \leq E[q_j]^2$, so it remains to show that $E[q_j] = O(n/r)$.

As before we write $C = \{c_1, c_2, \dots, c_n\}$ for the set of curves and we let R be a set of $r > 4$ sample curves, chosen randomly from C . Let also p be a fixed point in the plane, not on any of these curves and not vertically aligned with any vertex of the arrangement $\mathcal{A}(R)$. We use q for the random variable, depending on R , that counts the number of curves which intersect the funnel containing p .

As described in the previous section, we triangulate $\mathcal{A}(R)$, decomposing each cell into trapezoidal funnels. We are interested in the particular funnel Δ_R that contains p . As observed in Section 5.1 every funnel in the triangulation of $\mathcal{A}(R)$ is a funnel in the triangulation of some four or fewer curves of R . This observation suggests that we could find Δ_R in the following roundabout way: for all $R' \subseteq R$ with $|R'| \leq 4$, find the trapezoid of the triangulation of $\mathcal{A}(R')$ that contains p . The result is a set of $O(r^4)$ trapezoids, which we denote \mathcal{F}_R . (Observe that \mathcal{F}_R is a set which implies that it contains a trapezoid only once, even if it is defined by more than one set R' .) We know that $\Delta_R \in \mathcal{F}_R$, and by the independence lemma (Lemma 5.1) Δ_R is the unique trapezoid in \mathcal{F}_R that does not meet any curves of R .

To continue the proof define \mathcal{F}_C like \mathcal{F}_R so that it contains all trapezoids of the arrangement of some four or fewer curves of C that contain the point p . Furthermore, let $|\Delta|$ be the number of curves of C that meet trapezoid Δ , and for $\Delta \in \mathcal{F}_C$ let P_Δ be the probability that Δ is Δ_R , that is,

$$P_\Delta = \text{Prob}[\Delta = \Delta_R].$$

With these definitions $E[q] = \sum_{\Delta \in \mathcal{F}_C} |\Delta| P_\Delta$. To show that $E[q]$ is $O(n/r)$, we determine P_Δ . Since all subsets R of size r are equally likely, P_Δ is the number of subsets R such that $\Delta = \Delta_R$ divided by $\binom{n}{r}$, the number of subsets R of size r . We can thus compute P_Δ by determining how many subsets R have $\Delta = \Delta_R$. By the independence lemma, the following two conditions are necessary and sufficient for Δ to be Δ_R : Δ is in \mathcal{F}_R and none of the $|\Delta|$ curves meeting Δ are in R . Recall that Δ is in the triangulation of some four or fewer curves of C , and using the general position assumptions about C , there is a set $R_\Delta \subseteq C$ with $|R_\Delta| = r_\Delta \leq 4$ such that $\Delta \in \mathcal{F}_R$ if and only if $R_\Delta \subseteq R$. To put Δ in \mathcal{F}_R we must choose these r_Δ curves. To satisfy the second condition, the remaining $r - r_\Delta$ curves must be chosen from the $n - |\Delta| - r_\Delta$ curves that are not in R_Δ and do not meet Δ . Any such choice will do, so we have

$$P_\Delta = \binom{n - |\Delta| - r_\Delta}{r - r_\Delta} / \binom{n}{r}.$$

Now since

$$\binom{n - |\Delta| - r_\Delta}{r - r_\Delta} = \frac{n - |\Delta| - r + 1}{r - r_\Delta} \binom{n - |\Delta| - r_\Delta}{r - r_\Delta - 1} \leq \frac{n}{r - 4} \binom{n - |\Delta| - r_\Delta}{r - r_\Delta - 1},$$

we have

$$E[q] \leq \frac{n}{r - 4} \sum_{\Delta \in \mathcal{F}_C} |\Delta| \binom{n - |\Delta| - r_\Delta}{r - r_\Delta - 1} / \binom{n}{r}.$$

To complete the proof of the estimate for $E[q]$, it is enough to show that the sum in this expression is bounded from above by a constant. The key observation here is that the summand,

$$|\Delta| \binom{n - |\Delta| - r_\Delta}{r - r_\Delta - 1} / \binom{n}{r},$$

is the probability that Δ is a trapezoid in \mathcal{F}_R that meets exactly one curve of R . This observation follows from an argument much like that giving P_Δ : after choosing the r_Δ curves defining Δ , we have to choose exactly one curve from the set of $|\Delta|$ curves meeting Δ , and $r - r_\Delta - 1$ of the remaining ones. The sum is the expected number of trapezoids in \mathcal{F}_R that meet exactly one curve of R . By Lemma 5.2(i) the expectation is at most four.

This completes our proof that, for any point $p_j \in P$,

$$E[q_j^2] \leq E[q_j]^2 \leq \left(\frac{4n}{r - 4}\right)^2 \leq 4^2 \left(\frac{n}{r - 4}\right)^2.$$

Step 2. The main flow of the argument for part (ii) of the lemma is the same as in Step 1. We replace \mathcal{F}_R and \mathcal{F}_C by \mathcal{G}_R and \mathcal{G}_C , where \mathcal{G}_R is the set of all trapezoids defined by four or fewer curves in R , and \mathcal{G}_C is the same for C . Furthermore, we replace P_Δ by Q_Δ , where Q_Δ is the probability that $\Delta \in \mathcal{G}_C$ is a funnel in the triangulation of $\mathcal{A}(R)$. Finally, define $\chi_R = \sum |\Delta|$ where the sum is taken over all funnels Δ of the triangulation of $\mathcal{A}(R)$. With these definitions we have

$$E[\chi_R] = \sum_{\Delta \in \mathcal{G}_C} |\Delta| Q_\Delta.$$

We can compute Q_Δ by observing that Δ is a funnel in the triangulation of $\mathcal{A}(R)$ if and only if $\Delta \in \mathcal{G}_R$ and Δ does not meet any curve in R . So

$$Q_\Delta = \binom{n - |\Delta| - r_\Delta}{r - r_\Delta} / \binom{n}{r},$$

exactly the same P_Δ only that Q_Δ is defined for all trapezoids in \mathcal{G}_C . Thus we have

$$E[\chi_R] = \sum_{\Delta \in \mathcal{G}_C} |\Delta| Q_\Delta \leq \frac{n}{r-4} \sum_{\Delta \in \mathcal{G}_C} |\Delta| \binom{n-|\Delta|-r_\Delta}{r-r_\Delta-1} / \binom{n}{r}.$$

Here, the summand is the probability that Δ is a trapezoid in \mathcal{G}_R that meets exactly one curve in R . Thus, the sum on the right-hand side is the expected number of trapezoids in \mathcal{G}_R that intersect exactly one curve in R . By Lemma 5.2(ii) this number is $O(r^2)$ which implies part (ii) of the assertion. \square

Remark. The following stronger upper bounds on the sums in the sampling lemma are implied by the above proof and the remark after Lemma 5.2. The expression $\sum_{i=1}^k m_i n_i^\alpha \leq 4^\alpha m(n/(r-4))^\alpha$ holds for lines, pseudolines, unit-circles, general circles, and pseudocircles. The expression $\sum_{i=1}^k n_i \leq 2nr^2/(r-4)$ is true for lines and pseudolines, and $\sum_{i=1}^k n_i \leq 4nr^2/(r-4)$ holds for unit-circles, general circles, and pseudocircles.

5.3. Upper Bounds on Incidences in the Plane

Bounds on the maximum number of incidences between m points and n lines, pseudolines, unit-circles, general circles, and pseudocircles have been established in Section 4 (Canham Threshold 4.2). We now use the machinery provided in Sections 5.1 and 5.2 (as well as Canham Threshold 4.2 itself) to improve these bounds for a wide range of values of m depending on n . In each case we start with a sample of the curves, triangulate the sample arrangement, and finally do the necessary calculations to obtain the result. Beyond the fact that the argument is the same in each case, even the calculations are identical for lines, pseudolines, and unit-circles and for circles and pseudocircles since the Canham thresholds, the triangulation results, and the probabilistic counting bounds are asymptotically the same. In view of the fact that the argument and the calculation of the final result for lines has been given in Section 3, we omit the discussion of the first three cases altogether. The remainder of this section gives the complete derivation of the improved bounds for circles and pseudocircles and finally summarizes the results. We phrase the analysis for circles.

To get the analysis for circles started we randomly choose r of the n circles and triangulate the arrangement defined by the r circles. It consists of $k = O(r^2)$ funnels which we index from 1 through k . Let m_i be the number of points in the i th funnel and let n_i be the number of circles that intersect the i th funnel.²⁰ If $I(m, n)$ denotes the maximum number of incidences between m points and n circles, we have

$$I(m, n) \leq \sum_{i=1}^k I(m_i, n_i) + m + 2nr.$$

²⁰ Recall that a cell as well as a funnel is an open set. Since we can rotate a configuration of circles and points arbitrarily we may assume that no point lies on a vertical side of a funnel. Thus, if a point lies on the boundary of a funnel it must necessarily lie on at least one sample circle.

Here the sum covers the number of incidences within the funnels and the term $2nr + m$ bounds the number of incidences between points that lie on sample circles and all n given circles. Indeed, each circle c intersects the sample circles (other than c if it is itself in the sample) in at most $2r$ points, so it can contribute at most $2r$ incidences with points that lie on (other) sample circles. The only incidences that are not accounted for by this argument involve incidences between sample circles and points that are not vertices of the sample arrangement, but there are at most m such incidences total. The probabilistic counting argument in Section 5.2 implies that there is a sample so that the average n_i is $O(n/r)$ and $\sum_{i=1}^k m_i n_i^{2/3} = O(m(n/r)^{2/3})$. Using Canham Threshold 4.2(ii) we thus get

$$I(m, n) = \sum_{i=1}^k O(m_i n_i^{2/3} + n_i) + m + 2nr = O(mn^{2/3} r^{-2/3} + m + nr).$$

We now choose $r = \Theta(m^{3/5} n^{-1/5})$ which is feasible as long as $n^{1/3} \leq m \leq n^2$ and obtain

$$I(m, n) = O(m^{3/5} n^{4/5})$$

for this range of m . We have $I(m, n) = O(n)$ if $m < n^{1/3}$ and $I(m, n) = O(m)$ if $m > n^2$. A combination of these bounds gives the desired result. The above analysis applies almost verbatim to the case of lines, pseudolines, and unit-circles, the main difference being that we have to use Canham Threshold 4.2(i) in which case the choice of r is as in Section 3. We thus have

Theorem 5.4 (Planar Incidence Theorem). *The maximum number of incidences between a set M of m points and a set N of n curves is*

- (i) $O(m^{2/3} n^{2/3} + m + n)$ if N is a set of lines, pseudolines, or unit-circles, and
- (ii) $O(m^{3/5} n^{4/5} + m + n)$ if N is a set of n circles or pseudocircles.²¹

Remarks. (1) The bound for lines is not new and is originally due to Szemerédi and Trotter [57] who also give a number of applications to related geometric problems. This bound is tight.

(2) Also the upper bound for unit-circles is not new and can be found in [55]. In contrast to the case of lines, no matching lower bound is known and in fact it is conjectured that the bound in (i) is not tight (see [24]). It is worthwhile to remark that the same bound, $O(m^{2/3} n^{2/3} + m + n)$, also holds in the more general case when at most a constant number of circles intersect in two common points (see remark (4) after Canham Threshold 4.2).

²¹ It is implicit that we make the claim only for pseudolines that meet any vertical line in one point and for pseudocircles that meet any vertical line in at most two points. Since any pseudoline arrangement is combinatorially equivalent to one where all pseudolines satisfy this condition [32] this is no loss of generality in the case of pseudolines.

(3) The upper bound for circles is new and is an improvement of the $O(m^{3/4}n^{3/4} + m + n)$ bound observed in [5] and [10]. It can be used to improve known bounds for combinatorial distance problems in the plane (see Section 5.4). As for unit-circles, no matching lower bound is known and it seems unlikely that our bound is tight. The example of m points and n lines with $\Omega(m^{2/3}n^{2/3})$ incidences can be used to show the same lower bound for circles as follows. Choose an origin disjoint from all points and lines and perform an inversion with respect to this origin. This transform maps every point to a point and every line to a circle, and it preserves incidences. The weakness of this lower bound example is that every circle contains the origin which is an indication that not all degrees of freedom are properly used. Indeed, in the next section we will see a construction that gives a slightly higher lower bound for a small range of m .

(4) A stereographic projection of a sphere onto a plane maps a point to a point and a circle to a circle (or a line if the circle passes through the center of the projection). This fact can be used to extend the planar incidence theorem to the case where the points and circles lie on the surface of a sphere in three dimensions. Consequently, for m points and n circles on the sphere we have $O(m^{3/5}n^{4/5} + m + n)$ as an upper bound on the number of incidences in the general case and $O(m^{2/3}n^{2/3} + m + n)$ if no three circles intersect in two common points. The latter condition is satisfied, for example, if the circles are all congruent and are not great-circles of the sphere. If the n circles are great-circles, then the second bound still holds since two circles can intersect in only one point within any open hemisphere.

(5) The planar incidence theorem can be restated to bound the number of curves that contain at least some number of points and the number of points contained in at least some number of curves. Let k be larger than some constant that depends on the constants hidden in the big-Oh notation of the planar incidence theorem. Given a set of m points, the maximum number of lines (pseudolines, or unit-circles) containing at least k points each is $O(m^2/k^3 + m/k)$. Symmetrically, given n lines, pseudolines, or unit-circles, the maximum number of points each of which is incident to at least k of them is $O(n^2/k^3 + n/k)$. The corresponding bounds for circles and pseudocircles are $O(m^3/k^5 + m/k)$ and $O(n^2/k^{5/2} + n/k)$.

5.4. Applications to Distance Problems

As observed by Spencer *et al.* [55], the maximum number of unit-distance pairs²² in a set of m points in the plane is at most half the maximum number of incidences between m points and m unit-circles. To see this draw a unit-circle around each point and observe that $\{p, q\}$ is a unit-distance pair if and only if p lies on the circle around q and q lies on the circle around p . It follows that $O(m^{4/3})$ is an upper bound for the maximum number of unit-distance pairs. This result is not new but the proof in this paper is considerably simpler than the one in [55] and the

²² The pair $\{p, q\}$ is a *unit-distance pair* if $d(p, q) = 1$.

multiplicative constant implied by our proof is significantly smaller than theirs. Below we derive such an improved constant.

We start by considering the incidence problem for m points and n unit-circles in the plane. Since no two points are incident to three unit-circles (the corresponding bipartite digraph has no $K_{2,3}$) we get $I(m, n) \leq 2^{1/2}mn^{1/2} + n$ from remark (1) after the bipartite graph lemma (Lemma 4.1). From the remark after the sampling lemma (Lemma 5.3) we get $2^{3/2}m(n/(r-4))^{1/2}$ as an upper bound for the expected value of $\sum_{i=1}^k 2^{1/2}m_i n_i^{1/2}$ and $4nr^2/(r-4)$ as an upper bound for the expected value of $\sum_{i=1}^k n_i$. If we take both bounds times 2 we can be sure that there is a sample for which the (increased) bounds are true. So we get

$$I(m, n) \leq \sum_{i=1}^k (2^{1/2}m_i n_i^{1/2} + n_i) + 2nr + m \leq 2^{5/2}m \left(\frac{n}{r-4}\right)^{1/2} + \frac{8nr^2}{r-4} + 2nr + m.$$

Now choose $r = cm^{2/3}n^{-1/3} + c'$, with $c = 2^{1/3}/5^{2/3}$ and $4 \leq c' < 5$ so that r is an integer. Thus, we get

$$I(m, n) \leq 6\sqrt[3]{10m^{2/3}n^{2/3}} + 90n + m + 100\sqrt[3]{100n^{4/3}m^{-2/3}}.$$

If we set $m = n$ we get

$$I(m, m) \leq 6\sqrt[3]{10m^{4/3}} + 91m + 100\sqrt[3]{100m^{2/3}}.$$

Since the number of unit-distance pairs is at most half this quantity we have the following result.

Theorem 5.5. *The maximum number of unit-distance pairs in a set of m points in the plane is at most $3\sqrt[3]{10m^{4/3}} + O(m)$.*

The question of how often a given distance can occur in a set of points has also been asked for m points that lie on a sphere in three dimensions (see [29]). They use a geometric transform to show that the maximum number of unit-distance pairs in a set of m points on a sphere in three dimensions is $\Omega(m^{4/3})$. A matching upper bound follows from remark (4) after the planar incidence theorem (Theorem 5.4). The example in [29] consists of pairs at distance exactly one-fourth of a great-circle, and it seems likely that this is the only distance which can occur $\Omega(m^{4/3})$ times. The best lower bound for other distances is $\Omega(m \log^* m)$ [29].

Different Distances in the Plane. The remainder of this section discusses a problem about different distances. For a point set $P = \{p_1, p_2, \dots, p_m\}$ let g_i be the number of different distances from p_i , that is, $g_i = |\{d(p_i, p_j) \mid 1 \leq j \leq m, j \neq i\}|$. We define $g(P) = \sum_{i=1}^m g_i$ and

$$g(m) = \min\{g(P) \mid P \text{ set of } m \text{ points in the plane}\}.$$

Erdős [27] conjectures $g(m) = \Omega(m^2(\log m)^{-1/2})$ but proves only $\Omega(m^{3/2})$. The upper bound on the number of incidences between m points and n circles given in [5] and [10] implies $g(m) = \Omega(m^{5/3})$. We improve this bound using the planar incidence theorem. Around each p_i draw g_i circles that contain the other points.

This gives $2\binom{m}{2}$ incidences between m points and $g(P)$ circles. By the planar incidence theorem (ii) the number of incidences between that many points and circles is bounded by $O(m^{3/5}g(P)^{4/5})$ which implies $m^{3/5}g(P)^{4/5} = \Omega(m^2)$ for any set P of m points in the plane. Thus we have the following result.

Corollary 5.6. $g(m) = \Omega(m^{7/4})$.

As a consequence of Corollary 5.6, the average g_i and therefore also the maximum g_i is $\Omega(m^{3/4})$, always.

Remarks. (1) The above lower bound on $g(m)$ is new, but that the maximum g_i is $\Omega(m^{3/4})$ was known before (although only in unpublished form; see [28]). Recently, Chung, Szemerédi, and Trotter claimed to have improved the bound on the maximum g_i to $\Omega(m^{4/5})$.

(2) Let us indulge in an instructive exercise and assume that we can prove that the maximum number of incidences between m points and n circles is $O(m^{2/3}n^{2/3} + m + n)$ (which we cannot). This would imply that $m^{2/3}g(m)^{2/3} + m + g(m) = \Omega(m^2)$ and therefore $g(m) = \Omega(m^2)$. Consequently, the maximum g_i would be linear in m . This is indeed false as Erdős [24] proves that the number of different distances in the $\sqrt{m} \times \sqrt{m}$ grid is $O(m(\log m)^{-1/2})$. We conclude that the maximum number of incidences between m points and $m^2(\log m)^{-1/2}$ circles is $\Omega(m^2)$. In terms of $n = m^2(\log m)^{-1/2}$, the number of circles, the maximum number of incidences between $m = n^{1/2}(\log n)^{1/4}$ points and n circles is $\Omega(n(\log n)^{1/2})$. This improves the lower bound on the maximum number of incidences between m points and n circles mentioned in remark (3) after the planar incidence theorem for $n^{1/2}(\log n)^{1/4} \leq m \leq n^{1/2}(\log n)^{3/4}$. Of course, the lower bound on the number of incidences derived in this paragraph implies the same lower bound for the many-faces problem for circles.

(3) Remark (2) shows that $g(m) = \Theta(m^2)$ if $O(m^{2/3}n^{2/3} + m + n)$ is an upper bound on the number of incidences between the m points and the $g(P) = n$ circles whose centers are the m points. This bound holds if the maximum number of circles intersecting in two common points is at most some constant, which is true if the maximum number of collinear points in P is bounded by some constant. This result was originally obtained by Szemerédi in a direct and elegant manner (see [27]).

5.5. The Complexity of Zones

The *zone* of a manifold in an arrangement is the collection of cells of the arrangement crossed by the manifold. We need estimates for the number of faces

bounding the cells of a zone, the (*combinatorial*) *complexity* of a zone, in order to take care of the cells that spill outside of their subproblem funnel.

In Edelsbrunner *et al.* [18] upper bounds on the complexity of zones for rather general curves in the plane are derived. Their main theorem establishes upper bounds on the complexity of a zone in an arrangement of simple curves²³ that are only slightly superlinear in the number of curves. These bounds apply as long as there is some constant such that any two curves intersect in at most this constant number of points. We state their result after reviewing a related combinatorial concept.

The bounds on the complexity of zones are given in terms of maximum length functions of so-called Davenport–Schinzel sequences. These are sequences of symbols so that no two adjacent symbols are the same and there are no long alternations of any two symbols. More specifically, a sequence $a_1 a_2 \dots a_k$, with $a_i \in \{1, 2, \dots, n\}$ for $1 \leq i \leq k$, is a *Davenport–Schinzel sequence of order s* (for short an (n, s) -sequence) if $a_i \neq a_{i+1}$ for $1 \leq i \leq k - 1$ and there are no $s + 2$ indices $i_1 < i_2 < \dots < i_{s+2}$ so that $a_{i_1} = a_{i_3} = \dots \neq a_{i_2} = a_{i_4} = \dots$. Let $\lambda_s(n)$ be the length of the longest (n, s) -sequence. Then the following bounds on $\lambda_s(n)$ are known, where $\alpha(n)$ is the inverse of Ackermann’s function which is notorious for growing extremely slowly:

$$\lambda_1(n) = n \quad \text{and} \quad \lambda_2(n) = 2n - 1,$$

$$\lambda_3(n) = \Theta(n\alpha(n)) \quad (\text{see [37]}),$$

$$\lambda_4(n) = \Theta(n \cdot 2^{\alpha(n)}) \quad (\text{see [1]}),$$

$$\lambda_{2s}(n) = n \cdot 2^{\Theta(\alpha(n)^{s-1})} \quad (\text{see [1]}),$$

and

$$\lambda_{2s+1}(n) = n \cdot \alpha(n)^{O(\alpha(n)^{s-1})} \quad (\text{see [1]}).$$

Besides bounding the complexity of zones, bounds on the maximum length of Davenport–Schinzel sequences will also be used to analyze triangulations of sphere arrangements in space (see Section 6.3).

The theorem below summarizes the known results on the complexity of zones.

Theorem 5.7 (Zone Theorem). *Let $\Gamma \cup \{\gamma\}$ be a set of $n + 1$ simple curves in the plane with the property that any two intersect in at most s points where they must cross.*

- (i) *The complexity of the zone of γ in $\mathcal{A}(\Gamma)$ is $O(\lambda_{s+2}(n))$.*
- (ii) *If $\Gamma \cup \{\gamma\}$ is a set of lines or pseudolines, then the complexity of the zone of γ is $O(n)$.*
- (iii) *If Γ is a set of unit-circles and γ is a line or a circle, then the zone of γ has complexity $O(\lambda_3(n))$.*

²³ In this section we define a *simple curve* as the image of an interval or circle under a continuous one-to-one mapping to the Euclidean plane.

(Parts (i) and (iii) of the above theorem are taken from [18]. Part (ii) dates back to [8] and [21]; proofs can also be found in [15] and [18].)

Remark. Currently, there is only one nontrivial extension of the zone theorem known, namely for hyperplanes in $d \geq 3$ dimensions. The result is that the zone of a hyperplane in an arrangement of n other hyperplanes is bounded by at most $O(n^{d-1})$ faces of any dimension [21]. The lack of such results for other manifolds is the main reason for our lack of nontrivial bounds on the number of faces bounding many cells in three- and higher-dimensional arrangements. Bounds for planes and hyperplanes can be found in [19].

5.6. Upper Bounds on the Complexity of Many Cells

This section derives our final bounds on the maximum number of edges bounding m cells in an arrangement of n curves. As in earlier sections we consider lines, pseudolines, unit-circles, (general) circles, and pseudocircles.²⁴ In each case the proof follows a modular structure, and to go from one case to another only requires the adjustment of some parameters. As explained in Section 3, the major steps of the proof are

- (1) the establishment of a Canham threshold,
- (2) the triangulation of the arrangement defined by a sample of the curves,
- (3) the analysis of the triangulation by probabilistic counting,
- (4) the accounting done by divide-and-conquer,
- (5) the analysis of the combinatorial complexity of zones, and
- (6) the calculation of the final bound.

Steps (1)–(3) were discussed in appropriate detail in Sections 4, 5.1, and 5.2. To facilitate step (4) we mark each cell by a point somewhere within the cell. Using the appropriate Canham threshold we bound the number of edges bounding marked cells within the funnels; this covers all cells that do not cross the border between funnels. Now we add the complexity of the inner zones of the funnels to cover all edges that were not counted in step (4); these edges bound cells that reach into more than one funnel. Finally, we wrap it up by choosing the sample size to optimize the result.

The remainder of this section discusses all cases in turn and finally summarizes the results.

Lines and Pseudolines. Section 3 proved $K(m, n) = \Theta(m^{2/3}n^{2/3} + n)$ for m cells in an arrangement of n lines. Because of our assumption that a pseudoline does not have any x_1 -extreme points all steps are exactly the same as for lines. We thus obtain the same bound for pseudolines.

²⁴ Recall that we assume that a pseudoline intersects any vertical line in a point and that a pseudocircle intersects any vertical line in at most two points.

Unit-Circles. The major difference between the proofs for lines and for unit-circles is step (5), the complexity of zones. To estimate the contribution of the inner zones in the arrangement of sampled unit-circles we need bounds on the complexity of the zone of a unit-circle and the zone of a line in an arrangement of n unit-circles.²⁵ In both cases we talk about curves so that any two intersect in at most two points. By Theorem 5.7(iii), the number of edges bounding all zone cells is $O(\lambda_3(n))$. We can now do the necessary calculations to obtain a good bound on $K(m, n)$.

Let U be a set of n unit-circles and let $R \subseteq U$ be a sample of size r so that $\sum_{\Delta_i} m_i n_i^{1/2} = O(m(n/r^{1/2}))$ and $\sum_{\Delta_i} n_i = O(rn)$. Here, m_i is the number of points in funnel Δ_i and n_i is the number of unit circles intersecting Δ_i . Using part (iii) of the zone theorem we get

$$K(m, n) = \sum_{\Delta_i} O(m_i n_i^{1/2} + \lambda_3(n_i)) = O\left(m\left(\frac{n}{r}\right)^{1/2} + rn\alpha(n)\right).$$

We choose $r = \Theta(m^{2/3}n^{-1/3}\alpha(n)^{-2/3})$, which is meaningful as long as $m \geq n^{1/2}\alpha(n)$. With this choice of r we get

$$K(m, n) = O(m^{2/3}n^{2/3}\alpha(n)^{1/3}).$$

For the case $m \leq n^{1/2}\alpha(2)$ we still have Canham Threshold 4.7 which gives us the bound $O(mn^{1/2} + n) = O(m^{2/3}n^{2/3}\alpha(n)^{1/3} + n)$, thus covering all cases.

Circles and Pseudocircles. The arguments for pseudocircles and circles are identical which allows us to ignore pseudocircles altogether. In contrast to unit-circles we now have to deal with a weaker Canham threshold which bounds the number of edges of m cells only by $O(mn^{2/3} + n)$, see Section 4.4. All the other steps are the same as for unit-circles, except for the complexity of a zone which is now $O(\lambda_4(n))$ and that we need a sample R of size r so that $\sum_{\Delta_i} m_i n_i^{2/3} = O(m(n/r)^{2/3})$ and $\sum_{\Delta_i} n_i = O(rn)$. Such an R exists by the results in Section 5.2. We thus get

$$K(m, n) = \sum_{\Delta_i} O(m_i n_i^{2/3} + \lambda_4(n_i)) = O\left(m\left(\frac{n}{r}\right)^{2/3} + rn \frac{\lambda_4(n)}{n}\right).$$

We define $\beta_c(n) = \lambda_4(n)/n = \Theta(2^{\alpha(n)})$ and choose $r = \Theta(m^{3/5}n^{-1/5}\beta_c(n)^{-3/5})$, which is possible if $m \geq n^{1/3}\beta_c(n)$. This choice of r leads to

$$K(m, n) = O(m^{3/5}n^{4/5}\beta_c(n)^{2/5}).$$

For $m \leq n^{1/3}\beta_c(n)$ the Canham threshold for circles can be used to get the bound $O(mn^{2/3} + n) = O(m^{3/5}n^{4/5}\beta_c(n)^{2/5} + n)$ which thus covers all cases.

We finally summarize the results of this section.

²⁵ Since a cell can reach from one funnel to another only by intersecting the vertical side separating the two funnels it is actually sufficient to consider only zones of lines in unit-circle arrangements.

Theorem 5.8 (Planar Many Edges Theorem). *The maximum number of edges bounding a set of m cells in an arrangement $\mathcal{A}(N)$ of n curves is*

- (i) $\Theta(m^{2/3}n^{2/3} + n)$ if N is a set of lines or pseudolines,
- (ii) $O(m^{2/3}n^{2/3}\alpha(n)^{1/3} + n)$ if N is set of unit-circles, where $\alpha(n)$ is the extremely slowly growing inverse of Ackermann's function, and
- (iii) $O(m^{3/3}n^{4/5}\beta_c(n)^{2/5} + n)$ if N is a set of circles or pseudocircles, where $\beta_c = \Theta(2^{\alpha(n)})$.

Remarks. (1) For $n^{1/2} \leq m \leq n^{1/2}\alpha(n)$ the bound in (ii) can be improved to $O(mn^{1/2})$ using Canham Threshold 4.7. Similarly, for $n^{1/3} \leq m \leq n^{1/3}\beta_c(n)$ the bound in (iii) can be improved to $O(mn^{2/3})$ using Canham Threshold 4.6.

(2) The upper bound for unit-circles, $O(m^{2/3}n^{2/3}\alpha(n)^{1/3} + n)$, is almost tight which can be seen by extending the $\Omega(m^{2/3}n^{2/3} + n)$ lower bound for lines to unit-circles: take a line arrangement with m cells realizing the maximum and approximate each line by a large enough "unit-circle" so that no edges of the m cells are lost. It would be interesting to see whether or not the $\alpha(n)$ -factor is an artifact of the proof technique.

(3) No matching lower bound for the case of circles and pseudocircles is currently known. Except for a small range of values of m relative to n (see remark (2) after Corollary 5.6) the best known lower bound is the same as for unit-circles.

6. Spheres in Three Dimensions

This section demonstrates that the techniques laid out in Sections 4 and 5 can be used to derive bounds on the maximum number of incidences between m points and n spheres in three dimensions which improve upon previously known bounds. This problem is interesting because various combinatorial distance problems for points in three dimensions can be rephrased as point and sphere incidence problems. This relationship is discussed in Section 6.5.

Observe, however, that we cannot expect a nontrivial upper bound for the general problem since we can have all n spheres intersecting in a common circle and choose m points on this circle. This gives mn incidences which matches the trivial upper bound. We get interesting problems only after imposing some restrictions on the spheres or points. Section 6.1 imposes certain general position conditions on the spheres (and/or points). This line of investigation will lead to a new upper bound on the maximum number of unit-distances of a set of n points in three dimensions. In Section 6.2 we discuss the case where the points are required to be vertices of the arrangement defined by the spheres. For this case we need to prove an extension of the bipartite graph lemma (Lemma 4.1) which is of independent interest.

All steps of the general proof scheme developed in Section 5 go through without major difficulties except for the decomposition of sphere arrangements into cells of constant description, the funnels. This step is treated in Section 6.3. The results are summarized in Section 6.4 and applications to combinatorial distance problems in three dimensions are discussed in Section 6.5.

The reason for concentrating on spheres, rather than on planes, say, is that our techniques do not yield the best bounds for point/plane incidences and for the complexity of cells in arrangements of planes. These are best handled by dual techniques as demonstrated in the companion paper [19].

6.1. Canham-like Bounds Under General Position Assumptions

Assume that N , the set of n spheres, contains no three spheres that intersect in a common circle. Three points in three dimensions define a unique circle through them, unless they are collinear in which case no sphere can contain all three. This implies that any three points of the given set M belong to at most two common spheres. If we consider $\mathcal{G} = (M \cup N, A)$ with $(\mu, \nu) \in A$ if $\mu \in \nu$ this is equivalent to stating that \mathcal{G} contains no $K_{3,3}$. (Observe, however, that \mathcal{G} can contain $K_{2,t}$ for t up to n and $K_{s,2}$ for s up to m). Using the bipartite graph lemma we immediately get an upper bound on the number of incidences.

Canham Threshold 6.1. *The maximum number of incidences between m points and n spheres in three dimensions (assuming no three spheres intersect in a common circle) is $O(mn^{2/3} + n)$ and $O(nm^{2/3} + m)$.*

Remarks. (1) The first bound in Canham Threshold 6.1 implies that the number of incidences is $O(n)$ as long as $m \leq n^{1/3}$.

(2) Natural cases where the assumption in Canham Threshold 6.1 is satisfied include unit-spheres and spheres with no three collinear centers. If we relax the condition on the spheres and require that no t spheres intersect in a common circle, for some constant t , we still get the first upper bound of Canham Threshold 6.1.

(3) Suppose we require that no four points in M are cocircular and drop any conditions on N . In this case the graph \mathcal{G} that reflects the incidences contains no $K_{4,2}$ which implies that the maximum number of incidences is

$$O(mn^{3/4} + n) \quad \text{and} \quad O(nm^{1/2} + m).$$

Observe that the next natural but more restrictive condition on the points (no five points on a common sphere) trivializes the problem—in this case, each sphere can account for at most four incidences.

6.2. Vertices in Sphere Arrangements

In this section we assume that N is a set of n spheres in three dimensions whose arrangement has at least m vertices, and M is a subset of size m of the set of vertices. While the ramification for restricting the general position of spheres (as considered in Section 6.1) consists of applications to counting distances in three dimensions, the reason why choosing vertices only is interesting is that this problem comes

close to counting faces bounding m cells. No good bounds for this many-faces problem are known.

Before we state and prove an extension of the bipartite graph lemma let us understand why this lemma is not sufficient to yield any interesting results for the problem at hand. For one thing, an arbitrary number of spheres can intersect in a common circle, and up to $2n-4$ vertices can be cocircular. To see the lower bound take the circle of intersection of two spheres and choose the other $n-2$ spheres so that each cuts the circle in two unique points. The upper bound follows from the fact that a circle can intersect the n spheres in at most $2n$ points, but at least three spheres are needed to make each such point a vertex of the arrangement. It is thus easy to construct $\Omega(mn)$ incidences as long as $m \leq n$. This matches the trivial upper bound. We see that the problem is interesting only if m is superlinear in n . As far as the bipartite graph lemma is concerned, for every constant s (or t) we cannot find a t (or s) that is sublinear in n so that no s points are incident to t common spheres. So even the version (mentioned in remark (1) after its proof) that treats s and t as variables rather than constants does not generate any reasonable bounds.

Here is why there is still hope to prove some nontrivial bounds using a forbidden $K_{s,t}$ argument. Suppose a sphere v_i contains n_i (much larger than n) of the points in M . Such a sphere must exist unless the number of incidences is only $O(n^2)$. The number of cocircular 4-tuples among the n_i points is at most $O(n_i^3 n)$. The reason is that every three points define a unique circle on the sphere and the circle can contain at most $2n-4$ of the points altogether. This leaves a huge number of noncocircular 4-tuples. But each such 4-tuple belongs to at most one sphere and thus severely limits the number of possible incidences.

We next prove a generalization of the bipartite graph lemma which captures the idea indicated above in combinatorial terms. After that we return to the vertex and sphere incidence problem. We use “ s -tuple” as a synonym for “set of size s .”

Lemma 6.2 (Extended Bipartite Graph Lemma). *Let s and t be constants and let $\mathcal{G} = (M \cup N, A)$ be a bipartite digraph with $m = |M|$ sources and $n = |N|$ sinks. Let r be some number so that for any subset $\bar{M} \subseteq M$ with $|\bar{M}| \geq r$ there are at least $(|\bar{M}| - r)^s/s!$ s -tuples in \bar{M} with the property that all s nodes in every such s -tuple have at most $t - 1$ common adjacent nodes in N . Then $|A| = O(mn^{1-1/s} + rn)$.*

Proof. As in the proof of the bipartite graph lemma we consider an s -regular multihypergraph $\mathcal{H} = (M, H)$ —it is constructed as follows:

for every $v \in N$ adjacent to $\mu_1, \mu_2, \dots, \mu_s \in M$ we add $\{\mu_1, \mu_2, \dots, \mu_s\}$ to H —unless these s nodes in M are adjacent to more than $t - 1$ common nodes in N (in which case $\{\mu_1, \mu_2, \dots, \mu_s\}$ is not added to H).

Here H contains at most $t - 1$ copies of every possible s -tuple which implies

$$|H| \leq (t - 1) \binom{m}{s}.$$

Define n_i as the number of nodes in M adjacent to $v_i \in N$ for $1 \leq i \leq n$. By

assumption v_i gives rise to at least $(n_i - r)^s/s!$ s -tuples in H . (Of course, this statement is void if $n_i \leq r$.) This implies

$$\sum_{i=1}^n \frac{(n_i - r)^s}{s!} \leq (t - 1) \binom{m}{s} \leq \frac{t - 1}{s!} m^s,$$

where \sum^* indicates that we sum only over those indices i for which $n_i \geq r$. Using the Hölder inequality we thus get

$$\sum_{i=1}^n n_i \leq rn + \sum_{i=1}^n \frac{(n_i - r)^s}{s!} \leq rn + (t - 1)^{1/s} m^{1 - 1/s}$$

which implies the assertion since $|A| = \sum_{i=1}^n n_i$. □

We now apply the extended bipartite graph lemma to the point and sphere incidence problem where M is a subset of m vertices of the arrangement defined by the n spheres in N . Let us consider $\mathcal{G} = (M \dot{\cup} N, A)$ with the usual definition of A , and pick $s = 4$, $t = 2$, and $r = 2n - 4$. We argue that those parameters satisfy the assumptions of the lemma, provided $n \geq 4$. Take any subset \bar{M} of $\bar{m} > r$ points. Any three points in \bar{M} are cocircular with at most $2n - 7$ other points which implies that the number of cocircular 4-tuples in \bar{M} is at most $\binom{\bar{m}}{3}(2n - 7)/4$. Consequently, the number of noncocircular 4-tuples in \bar{M} is at least

$$\binom{\bar{m}}{4} - \binom{\bar{m}}{3} \frac{2n - 7}{4} = \frac{\bar{m}(\bar{m} - 1)(\bar{m} - 2)(\bar{m} - 2n + 4)}{4!} \geq \frac{(\bar{m} - r)^4}{4!}.$$

This proves the main result of this section.

Canham Threshold 6.3. *The maximum sum of degrees of m vertices²⁶ in an arrangement of n spheres in three dimensions is $O(mn^{3/4} + n^2)$.*

Remarks. (1) We noted before that the sum of degrees can be quadratic in n if m , the number of vertices, is linear in n . Canham Threshold 6.3 shows that the maximum sum does not increase beyond $O(n^2)$ as long as $m = O(n^{5/4})$.

(2) The extended bipartite graph lemma can be used to prove that the maximum sum of degrees of m vertices in an arrangement of n $(d - 1)$ -spheres in d dimensions is $O(mn^{1 - 1/(d+1)} + n^{d-1})$ using the same arguments as in three dimensions.

(3) Another application of the extended bipartite graph lemma is to bound the sum of degrees of m vertices in an arrangement of n hyperplanes in d dimensions. Using the same line of arguments as for spheres this gives $O(mn^{1 - 1/d} + n^{d-1})$ as an upper bound. Consequently, the maximum sum of vertex degrees is $\Theta(n^{d-1})$ for

²⁶ The *degree* of a vertex, in this case, is the number of spheres that contain the vertex.

$n^{d-2} \leq m \leq n^{d-2+1/d}$. This is, however, weaker than the bounds in [19] which show that the sum is $O(n^{d-1})$ for $m \leq n^{d-3/2}/\log n$.

6.3. Decomposition of a Sphere Arrangement

Let N be a set of n spheres s_i , $1 \leq i \leq n$, defining the arrangement $\mathcal{A}(N)$. The maximum number of vertices, edges, facets, and cells is cubic in the number of spheres.²⁷ Even if these numbers are maximized, the average number of vertices, edges, and facets per cell is constant. In contrast to the average case, the maximum number of vertices, edges, and facets bounding a single cell is quadratic in n . (To see the lower bound of this claim consider the unbounded cell of the arrangement defined by the following set of spheres. Half of the spheres are equally large and their centers lie on a short line segment; each pair intersects in a circle and all circles are contained in a torus that can be made arbitrarily narrow by decreasing the length of the line segment. The other half of the spheres are chosen pairwise disjoint so that each one intersects all circles inside the torus. The upper bound of this claim can be shown using a lifting transform that maps the spheres to hyperplanes in four dimensions; see [15].)

Our goal is to “triangulate” $\mathcal{A}(N)$, that is, decompose each cell of $\mathcal{A}(N)$ into subcells, the funnels, each with a constant size description. The main issues that arise are

- (i) how to define a funnel and a decomposition of $\mathcal{A}(N)$ into funnels, and
- (ii) how to bound the number of funnels that arise.

It goes without saying that our objective is to find a decomposition into as few funnels as possible. The type of funnel to be used will be chosen accordingly.

Figure 6.1 sketches the type of funnel we use. Its combinatorial structure is that of a cube. We refer to its front, back, top, bottom, left, and right facets in the sense suggested by Fig. 6.1. The front and back facets lie in two planes normal to the x_1 -axis. The top and bottom facets lie in two hemispheres²⁸ of the s_i . In Fig. 6.1 the top facet is part of a southern hemisphere and the bottom facet lies in a northern hemisphere, but all other combinations are allowed. The left and right facets are parts of two vertical elliptic cylinders.²⁹ The left (as well as the right) facet can be

²⁷ The exact bounds are $2\binom{n}{3}$ for the vertices, $6\binom{n}{3}$ for the edges, $6\binom{n}{3} + 2n$ for the 2-faces or *facets*, and $2\binom{n}{3} + 2n$ for the cells.

²⁸ A *hemisphere* is the part of a sphere that lies above (*northern hemisphere*) or below (*southern hemisphere*) the horizontal plane through the sphere's center. The circle that separates the northern from the southern hemisphere is the *equator* of the sphere.

²⁹ By this we mean a cylinder whose axis is parallel to the x_3 -axis and whose intersection with the x_1x_2 -plane is an ellipse. The tangent planes normal to the x_1 -axis meet the cylinder in two lines which partition it into its *left* and *right part*.

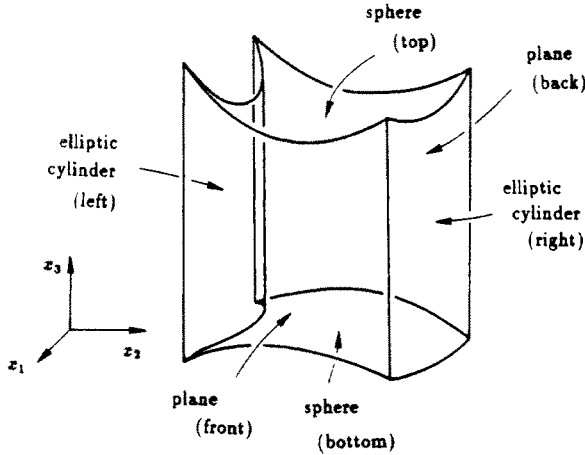


Fig. 6.1. A funnel is defined by two planes, two spheres, and two elliptic cylinders.

part of the left or the right part of such a cylinder; in Fig. 6.1 the left facet lies in the right part of a cylinder and so does the right facet.

Degenerate cases where some facets are empty or at infinity are allowed and, in fact, occur frequently in the decomposition of $\mathcal{A}(N)$ to be described. We now give a constructive definition of the decomposition of $\mathcal{A}(N)$. The construction takes four steps. We assume general position of the spheres when we describe the decomposition.

Step 1. From every point of each equator we draw a maximal vertical (relatively open) line segment that avoids all other spheres.³⁰ The union of all line segments drawn from a single equator forms a portion of a vertical circular cylinder.

After Step 1 every cell (of the refinement) has the property that it intersects every vertical line in a single (possibly empty) interval.

Step 2. Perform the same operation for all points of each circle $s_i \cap s_j, i \neq j$, that is, draw maximal vertical line segments that avoid all spheres except that they intersect s_i and s_j in a common point. (Figure 6.2(a) shows the line segments extended vertically upward from a circle; there is a sphere that blocks the line segments on the right-hand side). Since the vertical projection of a circle in three dimensions onto the $x_1 x_2$ -plane is an ellipse, the union of the line segments coming from a single circle forms a portion of a vertical elliptic cylinder.

The nice effect of Step 2 is that every cell (of the refinement) now has a unique top and a unique bottom facet, each lying in a hemisphere. However, the number of vertical walls (facets that are part of vertical circular or elliptic cylinders) of a single

³⁰ If there is no sphere above or below the point, then the line segment extends to infinity and is thus unbounded. If the point of the equator belongs to another sphere, then the line segment is empty.

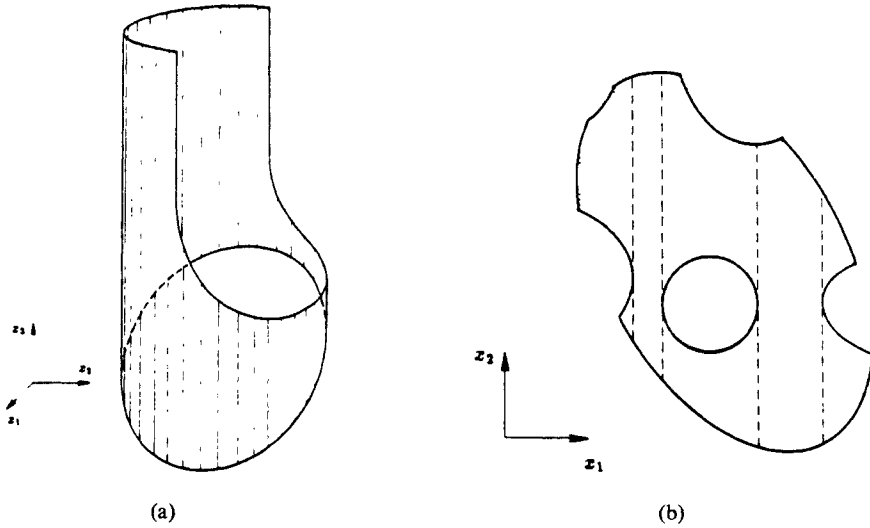


Fig. 6.2. Vertical walls are raised to decompose the cells into funnels as defined earlier.

cell can still be arbitrarily large. Since each cell intersects any vertical line in at most one interval, the vertical projection of a cell onto the x_1x_2 -plane is a connected (but not necessarily simply connected) region whose boundary edges are the vertical projections of the walls of the cell. Hence, the edges are portions of circles and ellipses (see Fig. 6.2(b)).

Step 3. Consider the vertical projection of a cell onto the x_1x_2 -plane. From the x_1 -extreme³¹ points of each edge we draw a maximal vertical line segment (parallel to the x_2 -axis) that lies inside the region (see the broken vertical line segments in Fig. 6.2(b)). The corresponding operation in three dimensions is to raise a wall vertically above the line segment inside the cell. We perform this operation for every cell created in Steps 1 and 2.

After Step 3, each cell (of the resulting refinement) has unique top, bottom, front, and back facets. Furthermore, it intersects any plane normal to the x_1 -axis in a single (possibly empty) simply connected region.

Step 4. Finally, from each vertex of the region (the vertical projection of a cell onto the x_1x_2 -plane) draw a maximal vertical line segment inside the region (see the dotted lines in Figs 6.2(b)). Again the corresponding operation in three dimensions is to raise a vertical wall. This is done for all cells.

After Step 4, each cell has unique top, bottom, front, back, left, and right facets; in fact, every cell is a funnel as defined earlier.

³¹ A point of a circle or ellipse is x_1 -extreme if the circle or ellipse lies on one side of the vertical line through the point. The x_1 -extreme points of an edge that is a subset of a circle or ellipse are the x_1 -extreme points of the circle or ellipse (if on the edge).

The remainder of this section bounds the maximum number of funnels created in Steps 1-4. We repeatedly use the fact that two (different) ellipses in the plane intersect in at most four points.

The main part of the analysis estimates the increase in the number of cells caused by Step 2. This is done by considering all vertical walls raised from a single circle of the form $s_i \cap s_j$. The analysis of Step 1 is a special case of that of Step 2 and can therefore be omitted. Step 3 adds at most four walls per circle: the projection of the circle onto the x_1x_2 -plane, which is an ellipse, has two x_1 -extreme points—from both we send a line segment upward and one downward (parallel to the x_2 -axis), and each of the four line segments in the x_1x_2 -plane is lifted to a wall in three dimensions. Since there are at most $\binom{n}{2} + n$ such circles, Step 3 increases the number of cells by at most $4\binom{n}{2} + 4n = O(n^2)$. Step 4 gives rise to two vertical walls per vertical edge created in Steps 1 and 2, so the number of such walls is proportional to the number of these edges.

In order to bound the increase in combinatorial complexity we bound the number of vertical edges created in Step 2. A vertical edge is created where two vertical walls intersect. Consider a circle $\gamma = s_1 \cap s_2$ and the vertical cylinder, C_γ , that is the union of all vertical lines intersecting γ . For $3 \leq i \leq n$, we define $\delta_i = C_\gamma \cap s_i$. Ignoring degenerate cases, δ_i is either a simple closed curve in C_γ (Fig. 6.3(a)), it consists of two simple closed curves separated by two vertical lines in C_γ (Fig. 6.3(b)), or it consists of two simple closed curves both meeting every vertical line in C_γ (Fig. 6.3(c)). This can be proved by considering the possible ways how the vertical projections of the cylinder and the sphere intersect. A vertical edge is created in C_γ where the intersection of two spheres meets C_γ (this is where two δ_i meet) or where one sphere meets γ (this is where one δ_i meets γ). We have the following properties.

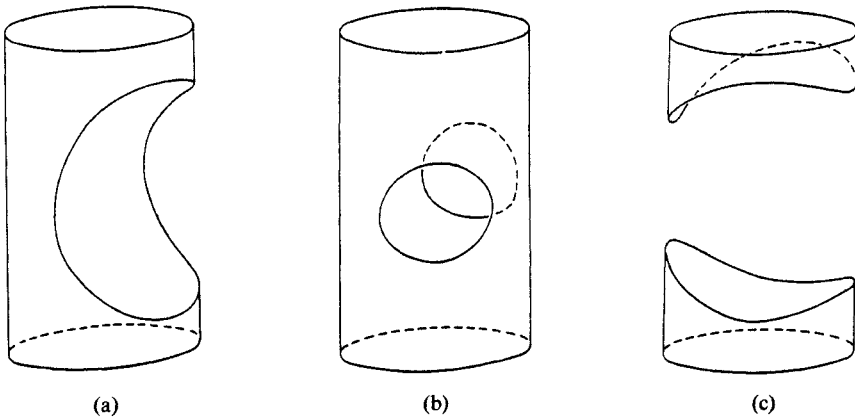


Fig. 6.3. A sphere can intersect an elliptic cylinder in three different ways.

Lemma 6.4. *Unless $\gamma \subseteq \delta_i$, $\gamma \cap \delta_i$ consists of at most two points for $3 \leq i \leq n$.*

Proof. Since $\gamma = s_1 \cap s_2$ and $\delta_i \subseteq s_i$ every point of $\gamma \cap \delta_i$ is also in $s_1 \cap s_2 \cap s_i$ (in fact, $\gamma \cap \delta_i = s_1 \cap s_2 \cap s_i$ but we do not need this to prove the lemma). Unless the three spheres intersect in a common circle (in which case $\gamma \subseteq \delta_i$), the intersection consists of at most two points. \square

Lemma 6.5. *Unless $s_i = s_j$ or $s_i \cap s_j \subseteq C_\gamma$, $\delta_i \cap \delta_j$ consists of at most four points for $3 \leq i, j \leq n$.*

Proof. We have $\delta_i \cap \delta_j \subseteq s_i \cap s_j$, which is a circle in three dimensions. The vertical projection of $s_i \cap s_j$ onto the x_1x_2 -plane is an ellipse and so is the intersection of C_γ with the x_1x_2 -plane. Unless the two ellipses are equal (in which case $s_i \cap s_j \subseteq C_\gamma$) they intersect in at most four points. Those points are the projections of the points of $\delta_i \cap \delta_j$. \square

A single δ_i has at most four points with vertical tangents in C_γ (the intersection points of C_γ with the equator of s_i). We remove from δ_i these at most four points as well as the at most two points of δ_i in γ . This leaves us with at most six connected curves for each δ_i . These curves have the following properties:

- (i) each curve intersects any vertical line in C_γ in at most one point,
- (ii) each curve lies either fully above or fully below γ , and
- (iii) any two curves intersect in at most four points.

Because of property (i) we can think of each curve as a partially defined continuous function from (a connected portion of) γ to the set of real numbers, where such a function, with image f , measures the difference in height of $\ell \cap f$ and $\ell \cap \gamma$, where ℓ is a vertical line in C_γ (if $\ell \cap f$ is below γ , then the difference is negative). Because of property (ii) each function f is either always positive or always negative. The *lower envelope* is the pointwise minimum of all positive functions; symmetrically, we define the *upper envelope* as the pointwise maximum of all negative functions. We only need to analyze the number of pieces³² of the lower envelope since the analysis of the upper envelope is completely symmetric. The number of pieces of the lower envelope is important since it is the same as the number of breakpoints between adjacent pieces. It is at the breakpoints where Step 2 creates vertical edges.

The number of curves, images of functions, above γ is at most $5n - 9$, five per each s_i plus one for the curve at infinity. We assign to each curve a unique integer between 1 and $5n - 9$, and write down a sequence by taking the pieces of the lower envelope and replacing each piece by the number of the curve containing it. Let (a_1, a_2, \dots, a_k) be the resulting sequence. We have $a_i \neq a_{i+1}$ since the pieces are maximal, and there is no alternation of length eight or more between two integers

³² Here we define a *piece* as a maximal connected subset of the image of the lower envelope that belongs to a single curve. In order to avoid intervals over which the lower envelope is not defined we add a function whose image is at infinity—it can also contain pieces of the envelope.

(see, e.g., [54]). It follows that we have a $(5n - 9, 6)$ -sequence and its length is at most

$$\lambda_6(5n - 9) = O(\lambda_6(n)) = n \cdot 2^{O(\alpha(n)^2)}$$

(see Section 5.5). Steps 1 and 2 operate on $O(n^2)$ circles which implies that the number of vertical edges and facets created is $O(n^2 \lambda_6(n))$. As mentioned earlier, the increase in the number of cells caused by Steps 3 and 4 is at most $O(n^2)$ plus a term that is proportional to the number of vertical edges created in Steps 1 and 2 which is $O(n^2 \lambda_6(n))$. We conclude with the main result of this section.

Theorem 6.6 (Sphere Triangulation Theorem). *The cells in an arrangement of n spheres in three dimensions can be decomposed into $O(n^2 \lambda_6(n))$ funnels.*

Remarks. (1) The decomposition of a sphere arrangement described in this section is not a cell complex since a facet of one funnel can contain an edge (or part thereof) of another funnel. This is, however, irrelevant for our use of such decompositions.

(2) We do not know whether or not $O(n^2 \lambda_6(n))$ is tight for the decomposition described. In fact, we do not even have an example where the number of funnels is more than cubic in n . This suggests two open problems. First, give the correct asymptotic order of the maximum number of funnels constructed by the above algorithm. Second, if that bound fails to be cubic, find a decomposition into $O(n^3)$ funnels, possibly with a different definition of “funnel” than used in this section.

(3) We believe that the above decomposition can be generalized to apply to any collection of n algebraic surfaces of fixed maximum degree in three dimensions, leading to $O(n^2 \lambda_s(n))$ appropriately defined funnels, where s depends on the type of surfaces and their maximum degree. This result is significant as we can compare it with the standard Collins’s cylindrical algebraic decomposition techniques [13], [53], [2]. Collins’s technique also produces a decomposition of an arrangement of n such algebraic surfaces into cells of constant description, but the number of such cells is enormous—doubly exponential in the number of dimensions. For example, in three dimensions the number of Collins cells that are produced in general is $O(n^7)$ (assuming constant algebraic degree of the surfaces).

(4) A challenging open problem is to generalize the sphere triangulation theorem to arrangements of algebraic surfaces in higher dimensions, or even just to arrangements of spheres in four or higher dimensions. An equally challenging problem is to obtain a decomposition of any collection C of cells in an arrangement of spheres or other surfaces in three dimensions, into funnels of constant description, whose total number is roughly proportional to the complexity of the cells in C plus, say a quadratic or slightly superquadratic overhead term. Some initial investigations of these problems in the polyhedral case are reported in [9], and [3].

(5) Section 6.4 uses the decomposition technique presented in this section also for degenerate arrangements of spheres. As explained in Section 3 for the case of lines, our policy is to perturb the spheres conceptually so that they form a nondegenerate arrangement and to triangulate this arrangement. The resulting

triangulation corresponds to a triangulation of the unperturbed arrangement which contains many zero-measure funnels.

(6) The triangulation of a sphere arrangement as defined in this section satisfies the independence lemma (Lemma 5.1) with the size of sets R' bounded from above by six. Based on this observation we can now extend the sampling lemma (Lemma 5.3) to the case of spheres. This can indeed be done without any difficulty, except for one step, namely Lemma 5.2(ii), which is necessary to bound the sum of the n_i . This step can be replaced by a tail estimation as described in [12] which then shows that for every $6 < r \leq n$ there exists a sample of size r so that the average number of spheres intersecting any one funnel is $O(n/r)$. An even stronger result, that the expected value of the square of the number of such spheres is $O((n/r)^2)$, can also be derived. We refer to [12] for details. The extension of the sampling lemma to spheres is used in the next section.

6.4. General Bounds on the Number of Incidences

Using the results of Sections 6.1–6.3 we can now derive improved upper bounds on the maximum number of incidences between m points and n spheres in three dimensions. We do this for the cases where no three spheres intersect in a common circle and where the points are vertices of the arrangement defined by the spheres. In both cases, the analysis starts with a random sample of r of the spheres and with a decomposition of the cells in the arrangement as described in Section 6.3. The sampling lemma in Section 5.2 implies that there exists such a sample with the desired properties detailed there.

Spheres in General Position. As mentioned before, we choose r of the n spheres and triangulate the arrangement defined by the r spheres. This is a decomposition of the cells into a total of $k = O(r^2 \lambda_6(r))$ funnels.³³ We number the funnels from 1 through k , let m_i denote the number of points contained in the i th funnel, and let n_i be the number of spheres intersection this funnel.³⁴ If $I(m, n)$ is the maximum number of incidences between m points and n spheres in general position, then

$$I(m, n) \leq \sum_{i=1}^k I(m_i, n_i) + J(m, n, r),$$

where $J(m, n, r)$ is the maximum number of incidences between m points and n spheres under the restriction that the m points lie on r spheres. The second term of the right-hand side bounds the number of incidences that occur on the sample spheres, the first term takes care of all other incidences.

³³ Recall that a funnel is an open set.

³⁴ Since we can rotate the configuration of spheres and points we can assume that a point lies on the boundary of a funnel only if it belongs to at least one sample sphere.

An upper bound on $\sum_{i=1}^k I(m_i, n_i)$ can be obtained using the first bound in Canham Threshold 6.1 and the sampling lemma for spheres (see remark (6) at the end of Section 6.3). In particular, this lemma implies that there is a sample of r spheres so that the average n_i is $O(n/r)$ and $\sum_{i=1}^k m_i n_i^{2/3} = O(m(n/r)^{2/3})$. We thus get

$$\sum_{i=1}^k I(m_i, n_i) = \sum_{i=1}^k O(m_i n_i^{2/3} + n_i) = O\left(m\left(\frac{n}{r}\right)^{2/3} + k\frac{n}{r}\right) = O(mn^{2/3}r^{-2/3} + r\lambda_6(r)n).$$

The two terms of the final bound balance at

$$r = \Theta\left(m^{3/8}n^{-1/8}\left(\frac{\lambda_6(m^3/n)}{m^3/n}\right)^{-3/8}\right)$$

(which can be shown using the extremely slow growth of $\lambda_6(t)/t$ in which case we have

$$\sum_{i=1}^k I(m_i, n_i) = O(m^{3/4}n^{3/4}\beta_s(m, n)^{1/4}),$$

where we define $\beta_s(m, n) = (\lambda_6(m^3/n))(m^3/n)$. (This choice of r makes sense as long as $m = \Omega(n^{1/3})$; see below for the other case.)

In order to bound $J(m, n, r)$ we index the r sample spheres from 1 through r and let \bar{m}_j denote the number of points on the j th sample sphere. The sum of the \bar{m}_j can exceed m since one point can lie on several sample spheres. Nevertheless we have

$$\sum_{j=1}^r \bar{m}_j \leq I(m, r) = O(rm^{2/3} + m)$$

using the second bound of Canham Threshold 6.1. Plugging in the value of r as chosen above gives

$$\sum_{j=1}^r \bar{m}_j = O(m^{25/24}n^{-1/8}\beta_s(m, n)^{-3/8} + m) = O(m)$$

if $m \leq n^3$, which we can assume since any point that is not a vertex of the arrangement of spheres is incident to at most two spheres by assumption. The number of incidences counted by $J(m, n, r)$ can now be bounded by considering the r sample spheres in turn. The j th sample sphere contains \bar{m}_j points and intersects the other spheres in less than n circles all of which are distinct by the general position assumption. Thus, the problem becomes an incidence counting problem for \bar{m}_j points and n circles on a sphere. For this we have an upper bound of the form

$$O(\bar{m}_j^{3/5}n^{4/5} + \bar{m}_j + n),$$

see remark (4) after the planar incidence theorem (Theorem 5.4). Therefore

$$J(m, n, r) = O\left(\sum_{j=1}^r (\bar{m}_j^{3/5} n^{4/5} + \bar{m}_j + n)\right) = O(m^{3/5} n^{4/5} r^{2/5} + m + rn).$$

After plugging in the above value of r we get

$$\begin{aligned} J(m, n, r) &= O(m^{3/4} n^{3/4} \beta_s(m, n)^{-3/20} + m + m^{3/8} n^{7/8} \beta_s(m, n)^{-3/8}) \\ &= O(m^{3/4} n^{3/4} + m) \end{aligned}$$

as long as $m = \Omega(n^{1/3})$. Finally, for $m = O(n^{1/3})$ Canham Threshold 6.1 implies that $I(m, n) = O(n)$. This completes the analysis of incidences for spheres with no three intersecting in a common circle. The result is summarized in the following theorem which improves the bounds given in Canham Threshold 6.1.

Theorem 6.7. *The maximum number of incidences between m points and n spheres in three dimensions (assuming that no three intersect in a common circle) is $O(m^{3/4} n^{3/4} \beta_s(m, n)^{1/4} + m + n)$, where*

$$\beta_s(m, n) = \frac{\lambda_6(m^3/n)}{m^3/n} = 2^{\Theta(\alpha(m^3/n^2))}.$$

Remarks. (1) There is a trivial upper bound of $O(n^3 + m)$ on the number of incidences between m points and n spheres no three of which intersect in a common circle. The bound given in Theorem 6.7 is slightly worse than this bound for $n^3 \beta_s(m, n)^{-1/3} < m < n^3 \beta_s(m, n)$. In view of the fact that $\beta_s(m, n)$ is small unless m is unreasonably large, the bound in Theorem 6.7 is not far off the insignificantly more refined bound

$$O(\min\{m^{3/4} n^{3/4} \beta_s(m, n)^{1/4}, n^3\} + m + n).$$

We view this as a strong indication that the $\beta_s(m, n)$ term is an artifact of our analysis. Indeed, it is inherited from the analysis of the triangulation method given in Section 6.3.

(2) It is not likely that the bound in Theorem 6.7 turns out to be tight, even if we ignore the $\beta_s(m, n)$ term. A lower bound construction which achieves $\Omega(n^{4/3} \log \log n)$ incidences for $m = n$ can be found in [25] (see also Section 6.5 below).

(3) The analysis used to derive Theorem 6.7 can also be used to improve upon the bound on the maximum number of incidences between m points and n spheres where we assume that no four of the points are cocircular (see remark (3) after Canham Threshold 6.1). The bounds that we obtain are $O(m^{8/11} n^{9/11} \beta_x(m, n)^{3/11} + m + n)$ and $O(nm^{1/2} + m)$, where $\beta_x(m, n) = \lambda_6(m^4/n)/(m^4/n)$.

(4) Let k be greater than some constant. Theorem 6.7 implies that, given a set of m points, the maximum number of spheres that contain at least k points each is $O((m^3/k^4)\beta(m) + m/k)$, where β is a generic and extremely slowly growing function. Symmetrically, given a set of n spheres, the maximum number of points that lie on at least k spheres each is $O((n^3/k^4)\beta(n) + n/k)$. Of course, here we talk only of spheres such that no three intersect in a common circle.

(5) The method in [10] can be used to extend the bound in Theorem 6.7 to four and higher dimensions (using our three-dimensional bound as the base case). The result is then that the maximum number of incidences between m points and n $(d - 1)$ -spheres in d dimensions is

$$O(m^{d/(d+1)}n^{d/(d+1)}\beta(m, n) + m + n)$$

if no d of our $(d - 1)$ -spheres intersect in a common 1-sphere (a circle). This improves the original bound in [10] which is $O(m^{(d+1)/(d+2)}n^{(d+1)/(d+2)} + mn^{(d-2)/(d-1)} + n)$. It is interesting to compare this with the “in-between” bound $O(m^{(2d-1)/(2d+1)}n^{(2d)/(2d+1)} + m + n)$ which we obtain if we base the approach in [10] on our bound for points and circles.

Vertices in Sphere Arrangements. The analysis of this case is similar to that of spheres in general position, the main difference being that we now use Canham Threshold 6.3 in place of Canham Threshold 6.1 to bound the number of incidences in each of the $k = O(r^2\lambda_6(r))$ funnels. Again we write $I(m, n)$ for the maximum number of incidences between m points and n spheres, where now the points are assumed to be vertices of the arrangement. Furthermore, we write m_i for the number of points in the i th funnel, n_i for the number of spheres intersecting the i th funnel, and $J(m, n, r)$ for the maximum number of incidences if the m points are restricted to lie on r spheres. Thus, we have

$$I(m, n) \leq \sum_{i=1}^k I(m_i, n_i) + J(m, n, r).$$

Using Canham Threshold 6.3 and the sampling lemma for spheres (see above) we bound the first term on the right-hand side as follows:

$$\begin{aligned} \sum_{i=1}^k I(m_i, n_i) &= \sum_{i=1}^k O(m_i n_i^{3/4} + n_i^2) = O\left(m\left(\frac{n}{r}\right)^{3/4} + k\left(\frac{n}{r}\right)^2\right) \\ &= O(mn^{3/4}r^{-3/4} + n^2\lambda_6(r)). \end{aligned}$$

The two terms in the last bound balance if $r = \Theta(m^{4/7}n^{-5/7}\beta_v(m, n)^{-4/7})$, where $\beta_v(m, n) = (\lambda_6(m^4/n^5))(m^4/n^5)$. In this case we have

$$\sum_{i=1}^k I(m_i, n_i) = O(m^{4/7}n^{9/7}\beta_v(m, n)^{3/7}).$$

Below we show $J(m, n, r) = O(rn^2)$ which, by the above choice of r , is asymptotically less than the bound on the sum of the $I(m_i, n_i)$ and therefore does not affect the result.

Instead of considering only m vertices let us take all vertices of the sphere arrangement and define $J(n, r)$ as the maximum number of incidences involving the vertices on some r of the n spheres. Clearly, $J(n, r) \leq rJ(n, 1)$. So consider a single sphere, s_0 , which meets the other spheres in circles. We count each circle once for each sphere that intersects s_0 in this circle. The vertices of the sphere arrangement that lie on s_0 are the vertices of the circle arrangement on s_0 . Since there are at most $n - 1$ circles, considering multiplicities, the number of incidences between vertices and circles is at most

$$4\binom{n-1}{2} \leq 2n^2.$$

Because each vertex is also incident to s_0 we get $J(n, 1) \leq 3n^2$ and therefore $J(n, r) \leq 3rn^2$. This completes the analysis for m vertices in an arrangement of n spheres.

Theorem 6.8. *The maximum sum of degrees of m vertices in an arrangement of n spheres in three dimensions is $O(m^{4/7}n^{9/7}\beta_v(m, n)^{3/7} + n^2)$, with*

$$\beta_v(m, n) = \frac{\lambda_6(m^4/n^5)}{m^4/n^5} = 2^{\Theta(\alpha(m^4/n^5))}.$$

Remarks. (1) Clearly the sum of degrees cannot exceed $O(mn)$ and $O(n^3)$; the first bound improves Theorem 6.8 if $m = O(n)$, the second bound does so if $m = \Omega(n^3\beta_v(m, n)^{-3/4})$. A refinement of the bound in Theorem 6.8 is therefore

$$O(\min\{mn, m^{4/7}n^{9/7}\beta_v(m, n)^{3/7}, n^3\} + \min\{mn, n^2\}).$$

While the improvement for small m is substantial but trivial, the improvement for large m is insignificant and due to an artifact of our analysis which gives rise to the nearly constant $\beta_v(m, n)$ -factor in the bound.

(2) It is likely that that bound in Theorem 6.8 is not tight, even if we ignore the $\beta_v(m, n)$ -factor. A further improvement hinges on an improvement of Canham Threshold 6.3.

(3) Let k be greater than some constant. By Theorem 6.8 the maximum number of vertices in an arrangement of n spheres that have degree k or higher is $O((n^3/k^{7/3})\beta(n) + n^2/k)$.

6.5. Applications to Distance Problems

Theorem 6.7 can be used to improve known upper bounds for various distance problems in three dimensions. We briefly discuss the history of two such problems,

state our improvements, and remark on additional distance problems. In each case, the reduction to Theorem 6.7 is done by drawing spheres of appropriate radii around the points so that the counted distances correspond to incidences.

Unit-Distance Pairs. This problem was originally posed by Erdős who reports initial results in [24] and [25]. The problem is to bound the maximum number of unit-distance pairs³⁵ in a set of m points in three dimensions. Erdős [25] proves that $\Omega(m^{4/3} \log \log m)$ is a lower bound and $O(m^{5/3})$ is an upper bound for this problem. This was subsequently improved by Beck [5] to $O(m^{13/8+\epsilon})$, $\epsilon > 0$, and by Chung [10] to $O(m^{8/5})$. All three upper bound proofs draw a unit-sphere around each point and base the argument on the fact that $\{p, q\}$ is a unit-distance pair if and only if p lies on the unit-sphere around q which is equivalent to q lying on the unit-sphere around p . The number of incidences for the m points and the m unit spheres is thus twice the number of unit-distance pairs. Theorem 6.7 applies since no three equally large spheres in three dimensions intersect in a common circle.

Corollary 6.9 (Unit-Distance Theorem). *The maximum number of unit-distance pairs in a set of m points in three dimensions is $O(m^{3/2}(\lambda_6(m)/m)^{1/4})$.*

Remark. (1) Without fear of redundancy we point out that the multiplicative factor, $(\lambda_6(m)/m)^{1/4}$, is small unless m is unreasonably large. It is probably an artifact of our proof technique, more particularly of the triangulation of a sphere arrangement described in Section 6.3.

(2) It is interesting to observe that the method of Chung [10] which improves Erdős' initial upper bound to $O(m^{8/5})$ can be further improved to $O(m^{11/7})$ just by using the planar incidence theorem for points and circles in the plane (Theorem 5.4(ii)). This bound is still inferior to the bound in the unit-distance theorem (see also remark (5) after Theorem 6.7).

Furthest Neighbor Pairs. The problem here is to bound the maximum number of furthest neighbor pairs³⁶ in a set of m points in three dimensions. It is easy to see that the maximum number of furthest neighbor pairs in a set of m points in three dimensions is $\Omega(m^2)$; the constant in front of the m^2 and lower-order terms can be found in [4]. As observed by Edelsbrunner and Skiena [22], the worst case cannot be realized if no three points are collinear; for this case they prove $O(m^{5/3})$ as an upper bound. This was improved to $O(m^{8/5})$ by Chung [10]. The relation to incidence counting becomes clear when we draw around each point a sphere that goes through all furthest neighbors of the point. Now, (p, q) is a furthest neighbor pair if and only if q lies on the sphere around p . No three of the m spheres intersect

³⁵ A unit-distance pair consist of two points that are one unit of distance apart. Since any point set can be increased or decreased by scaling, one unit is just any fixed positive distance. Thus, the problem is really to bound the maximum number of times the most popular distance can occur.

³⁶ A directed pair (p, q) is a furthest neighbor pair of a finite point set P if $d(p, q) = \max\{d(p, x) | x \in P\}$.

in a common circle since their centers are not collinear by assumption. Using Theorem 6.7 we thus get the following result.

Corollary 6.10 (Furthest Neighbor Theorem). *The maximum number of furthest neighbor pairs in a set of m points in three dimensions (no three of which are collinear) is $O(m^{3/2}(\lambda_6(m)/m)^{1/4})$.*

Remarks. (1) In contrast to the unit-distance problem, no superlinear lower bound on the maximum number of furthest neighbors pairs is known. A related result is that the maximum number of diameter pairs³⁷ in a set of m points in three dimensions is $2m - 2$ (see [39], [33], and [56]). This bound also holds for the maximum number of symmetric furthest neighbor pairs.³⁸ These results can be taken as indications that the maximum number of furthest neighbor points is also linear in the number of points but no proof is known.

(2) Of course, Theorem 6.7 can also be used to bound the maximum number of minimum distance pairs³⁹ in three dimensions, but a better (linear) upper bound can be proven. This proof (which uses a packing argument) cannot be extended to the case where each point is colored either red or blue and only distances between differently colored points are considered. Theorem 6.7 still applies and shows that $O(m^{3/2}(\lambda_6(m)/m)^{1/4})$ is also an upper bound for the bichromatic version of the minimum distance problem. It would be desirable to find a superlinear lower bound if there is one. (This problem was also proposed by János Pach at the Second Computational Geometry Day in New York, 1986.) In the two-dimensional case, the maximum number of minimum bichromatic distances is $O(m)$ because the graph defined by minimum distance pairs is planar (see [29]).

(3) For the bichromatic *maximum* distance problem the proof for the monochromatic case (see remark (1) above) can be extended to show a linear upper bound as follows. Around each of the m red point we draw a closed ball with a radius equal to the maximum bichromatic distance. Call this distance δ . For $m \geq 3$ the intersection of the m balls has the structure of a convex polytope with at most m (spherical) facets, at most $3m - 6$ edges, and at most $2m - 4$ vertices. The sum of all vertex degrees is therefore at most $6m - 12$. By construction, this polytope contains all blue points. If a red point does not generate a facet of the polytope, then the sphere that bounds the ball of the point can touch the polytope in at most one point. It follows that the red point can realize the maximum distance to at most one blue point. We count the other maximum distance pairs for the blue points. If a blue point lies on a facet it has one red point at distance δ , if it lies on an edge it has two red points at distance δ , and if it is a vertex of the polytope, then the number of such red points is equal to the degree of the vertex—in the latter two cases we ignored red points that do not generate facets. Blue points in the interior of the polytope have no red point at distance δ which implies the claimed linear bound.

³⁷ A pair $\{p, q\}$ is a *diameter pair* of a set P if $d(p, q) = \max\{d(x, y) \mid x, y \in P\}$.

³⁸ A pair $\{p, q\}$ is a *symmetric furthest neighbor pair* if (p, q) and (q, p) are furthest neighbor pairs.

³⁹ A pair $\{p, q\}$ is a *minimum distance pair* of a set P if $d(p, q) = \min\{d(x, y) \mid x, y \in P\}$.

(4) As in Section 5.4, we can use upper bounds on the number of incidences to derive lower bounds on the number of different distances defined by m points. Let p_1, p_2, \dots, p_m be the m points of set P and let g_i be the number of different distances from point p_i , that is, g_i is the cardinality of $\{d(p_i, p_j) \mid 1 \leq j \leq m, j \neq i\}$. Around each p_i we draw g_i spheres through the other points. Thus, we have m points, $n = \sum_{i=1}^m g_i$ spheres, and $2 \binom{m}{2}$ incidences. If we assume that no three points are collinear we can use Theorem 6.7 and get

$$2 \binom{m}{2} = O(m^{3/4} n^{3/4} \beta_s(m, n)^{1/4})$$

which implies that

$$n = \Omega\left(m^{5/3} \left(\frac{\lambda_6(m)}{m}\right)^{-1/3}\right).$$

In words, the average (and therefore also the maximum) g_i is $\Omega(m^{2/3} (\lambda_6(m)/m)^{-1/3})$. Strangely enough, the m points arranged in a cubic grid realize $O(m^{2/3})$ different distances each, but they violate the collinearity restriction under which the lower bound is obtained.

7. Discussion and Open Problems

The main contribution of this paper is a uniform and modular approach to a variety of combinatorial incidence and many-faces problems. Some of the bounds obtained with this approach are tight (for example, the bounds for the incidence and many-faces problems for lines and pseudolines), some are tight up to a $\beta(n)$ -factor (the bound for the many-faces problem for unit-circles), and some still leave large gaps to be closed. Bounds on incidence problems can be used to get bounds on combinatorial distance problems, such as the problem to count the maximum number of unit-distance pairs in a set of n points in three dimensions. For this and for other distance problems we obtain improved bounds, but except for the unit-distance problem for points on a sphere in three dimensions, no such bound seems to be tight.

Narrowing or closing the gaps that the approach of this paper leaves is a challenging open problem (see Section 2 for a summary of the results of this paper). More specialized open problems are mentioned throughout the paper and we hope that some will be the starting point for future developments in this area.

The reader of this paper might wonder why this paper makes no attempt to solve incidence and many-faces problems that seem closely related to problems successfully tackled in this paper. Examples are the incidence and many-faces problems for planes and hyperplanes. The generic answer is that some step of our modular proof technique breaks down or gives results that are too weak for any

new bounds. Below, we list more specific reasons why our method fails for a few interesting cases:

- (1) The Canham threshold obtained for m vertices in an arrangement of n (hyper-) planes (using the extended bipartite graph lemma (Lemma 6.2)) is weaker than the bounds derived in [19] using a different approach altogether.
- (2) The bipartite graph lemmas (Lemmas 4.1 and 6.2) seem to be of little use for establishing a Canham threshold for many faces in a line segment arrangement. Again, a quite different approach (see [20]) leads to bounds that are almost tight.
- (3) The bounds on the incidence problem for spheres in three dimensions fail to generalize to the corresponding many-faces problem because of the lack of a zone result for spheres.
- (4) The major obstacle in generalizing the incidence bound for three-dimensional spheres to four and higher dimensions is the lack of appropriate triangulation results.

Acknowledgments

The authors would like to thank the DEC Systems Research Laboratory at Palo Alto, where a good portion of the reported research was conducted, for the generous support. They also thank János Pach and Richard Pollack for useful discussions and moral encouragement. The second author wishes to thank Atul Jain for helpful suggestions on an earlier version of the paper. Last but not least the authors are indebted to two anonymous referees for their careful reading of the paper and for their constructive criticism.

References

- [1] P. Agarwal, M. Sharir, and P. Shor. Sharp upper and lower bounds on the length of general Davenport-Schinzel sequences. *J. Combin. Theory Ser. A*, to appear.
- [2] D. S. Arnon, G. E. Collins, and S. McCallum. Cylindrical algebraic decomposition: I. The basic algorithm. *SIAM J. Comput.* **13** (1984), 865–877.
- [3] B. Aronov and M. Sharir. Triangles in space or building (and analyzing) castles in the air. In *Proc 4th Ann. Sympos. Comput. Geom.*, 1988, pp. 381–391.
- [4] D. Avis, P. Erdős, and J. Pach. Repeated distances in space. *Graphs Combin.* **4** (1988), 207–217.
- [5] J. Beck. On the lattice property of the plane and some problems of Dirac, Motzkin and Erdős in combinatorial geometry. *Combinatorica* **3** (1983), 281–297.
- [6] R. J. Canham. A theorem on arrangements of lines in the plane. *Israel J. Math.* **7** (1969), 393–397.
- [7] B. Chazelle and J. Friedman. A deterministic view of random sampling and its use in geometry. In *Proc. 29th IEEE Sympos. Found Comput. Sci.*, 1988, pp. 539–549.
- [8] B. Chazelle, L. J. Guibas, and D. T. Lee. The power of geometric duality. *BIT* **25** (1985), 76–90.
- [9] B. Chazelle and L. Palios. Triangulating a non-convex polytope. In *Proc. 5th Ann. Sympos. Comput. Geom.*, 1989, pp. 393–400.
- [10] F. R. K. Chung. Sphere-and-point incidence relations in high dimensions with applications to unit distances and furthest-neighbor pairs. *Discrete Comput. Geom.* **4** (1988), 183–190.
- [11] K. L. Clarkson. New applications of random sampling in computational geometry. *Discrete Comput. Geom.* **2** (1987), 195–222.

- [12] K. L. Clarkson and P. W. Shor. Applications of random sampling in computational geometry, II. *Discrete Comput. Geom.* **4** (1989), 387–421.
- [13] G. Collins. Quantifier elimination for real closed fields by cylindrical algebraic decompositions. In *Proceedings of the 2nd GI Conference on Automata Theory and Formal Languages*. Lecture Notes in Computer Science, Vol. **35**. Springer-Verlag, Berlin, 1975, pp. 134–183.
- [14] D. P. Dobkin and M. J. Laszlo. Primitives for the manipulation of three-dimensional subdivisions. In *Proc. 3rd Ann. Sympos. Comput. Geom.*, 1987, pp. 86–99.
- [15] H. Edelsbrunner. *Algorithms in Combinatorial Geometry*. Springer-Verlag, Heidelberg, 1987.
- [16] H. Edelsbrunner and L. J. Guibas. Topologically sweeping an arrangement. *J. Comput. System Sci.* **38** (1989), 165–194.
- [17] H. Edelsbrunner, L. J. Guibas, J. Hershberger, R. Seidel, M. Sharir, J. Snoeyink, and E. Welzl. Implicitly representing arrangements of lines or segments. *Discrete Comput. Geom.* **4** (1989), 433–466.
- [18] H. Edelsbrunner, L. J. Guibas, J. Pach, R. Pollack, R. Seidel, and M. Sharir. Arrangements of curves in the plane—topology, combinatorics, and algorithms. *Theoret. Comput. Sci.*, to appear.
- [19] H. Edelsbrunner, L. J. Guibas, and M. Sharir. The complexity of many cells in arrangements of planes and related problems. *Discrete Comput. Geom.*, this issue, 197–216.
- [20] H. Edelsbrunner, L. J. Guibas, and M. Sharir. The complexity and construction of many faces in arrangements of lines and of segments. *Discrete Comput. Geom.*, this issue, 161–196.
- [21] H. Edelsbrunner, J. O’Rourke, and R. Seidel. Constructing arrangements of lines and hyperplanes with applications. *SIAM J. Comput.* **15** (1986), 341–363.
- [22] H. Edelsbrunner and S. S. Skiena. On the number of furthest neighbour pairs in a points set. *Amer. Math. Monthly*, **96** (1989), 614–618.
- [23] H. Edelsbrunner and E. Welzl. On the maximal number of edges of many faces in arrangements. *J. Combin. Theory Ser. A* **41** (1986), 159–166.
- [24] P. Erdős. On sets of distances of n points. *Amer. Math. Monthly* **53** (1946), 248–250.
- [25] P. Erdős. On sets of distances of n points in Euclidean space. *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **5** (1960), 165–169.
- [26] P. Erdős. On extremal problems of graphs and generalized graphs. *Israel J. Math.* **2** (1964), 183–190.
- [27] P. Erdős. On some problems of elementary and combinatorial geometry. *Ann. Mat. Pura Appl. (IV)* **103** (1975), 99–108.
- [28] P. Erdős. Extremal problems in number theory, combinatorics and geometry. In *Proc. ICM*, 1983.
- [29] P. Erdős, D. Hickerson, and J. Pach. A problem of Leo Moser about repeated distances on the sphere. *Amer. Math. Monthly*, to appear.
- [30] W. Feller. *An Introduction to Probability Theory and Its Applications*, Vol. II, Second edition. Wiley, New York, 1971.
- [31] S. J. Fortune. A sweepline algorithm for Voronoi diagrams. *Algorithmica* **2** (1987), 153–174.
- [32] J. E. Goodman and R. Pollack. Proof of Grünbaum’s conjecture of the stretchability of certain arrangements of pseudolines. *J. Combin. Theory Ser. A* **29** (1980), 385–390.
- [33] B. Grünbaum. A proof of Vázsonyi’s conjecture. *Bull. Res. Council Israel Sect. A* **6** (1956), 77–78.
- [34] B. Grünbaum. *Convex Polytopes*. Wiley, Chichester, 1967.
- [35] L. J. Guibas and J. Stolfi. Primitives for the manipulation of general subdivisions and the computation of Voronoi diagrams. *ACM Trans. Graphics* **4** (1985), 74–123.
- [36] G. H. Hardy, J. E. Littlewood and G. Pólya. *Inequalities*, Second edition. Cambridge, 1952.
- [37] S. Hart and M. Sharir. Nonlinearity of Davenport–Schinzel sequences and of generalized path compression schemes. *Combinatorica* **6** (1986), 151–177.
- [38] D. Haussler and E. Welzl. ϵ -nets and simplex range queries. *Discrete Comput. Geom.* **2** (1987), 127–151.
- [39] A. Heppes. Beweis einer Vermutung von A. Vázsonyi. *Acta Math. Acad. Sci. Hungar.* **7** (1956), 463–466.
- [40] S. Józsa and E. Szemerédi. The number of unit distances in the plane: Infinite and finite sets. *Colloq. Math. Soc. János Bolyai* **10** (1975), 939–950.
- [41] K. Kedem, R. Livne, J. Pach, and M. Sharir. On the union of Jordan regions and collision-free translational motion amidst polygonal obstacles. *Discrete Comput. Geom.* **1** (1986), 59–71.

- [42] D. E. Knuth. *Fundamental Algorithms: The Art of Computer Programming I*. Addison-Wesley, Reading, Mass., 1968.
- [43] D. E. Knuth. *Seminumerical Algorithms: The Art of Computer Programming II*. Addison-Wesley, Reading, Mass., 1969.
- [44] D. E. Knuth. *Sorting and Searching: The Art of Computer Programming III*. Addison-Wesley, Reading, Mass., 1973.
- [45] T. Kövári, V. T. Sós, and P. Turán. On a problem of K. Zarankiewicz. *Colloq. Math.* 3 (1954), 50–57.
- [46] F. Levi. Die Teilung der projektiven Ebene durch Gerade oder Pseudogerade. *Ber. Math.-Phys. Kl. Sächs. Akad. Wiss. Leipzig* 78 (1926), 256–267.
- [47] M. Mäntylä. *An Introduction to Solid Modeling*. Computer Science Press, Rockville, Md., 1988.
- [48] J. Matoušek. Construction of ε -nets. In *Proc. 5th Ann. Sympos. Comput. Geom.*, 1989, pp. 1–10.
- [49] W. O. J. Moser and J. Pach. *Research Problems in Discrete Geometry*. Academic Press, New York, to appear.
- [50] F. P. Preparata and M. I. Shamos. *Computational Geometry—an Introduction*. Springer-Verlag, New York, 1985.
- [51] G. Purdy and P. Erdős. Some extremal problems in combinatorial geometry. Manuscript, 1987.
- [52] I. Reiman. Über ein Problem von K. Zarankiewicz. *Acta Math. Hungar. Acad. Sci.* 9 (1958), 269–273.
- [53] J. T. Schwartz and M. Sharir. On the “piano movers” problem. II. General techniques for computing topological properties of real algebraic manifolds. *Adv. in Appl. Math.* 4 (1983), 298–351.
- [54] M. Sharir. Davenport–Schinzel sequences and their geometric applications. In *Theoretical Foundations of Computer Graphics and CAD*, R. A. Earnshaw (ed.), NATO ASI Series, Vol. F-40, Springer-Verlag, Berlin, 1988, pp. 253–278.
- [55] J. Spencer, E. Szemerédi, and W. T. Trotter, Jr. Unit distances in the Euclidean plane. In *Graph Theory and Combinatorics*, Academic Press, London, 1984, pp. 293–303.
- [56] S. Straszewicz. Sur un problème géométrique de P. Erdős. *Bull. Acad. Polon. Sci. Cl. III* 5 (1957), 39–40.
- [57] E. Szemerédi and W. T. Trotter, Jr. Extremal problems in discrete geometry. *Combinatorica* 3 (1983), 381–392.

Received November 21, 1988, and in revised form September 13, 1989.