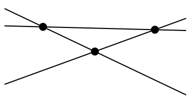
4

Incidence Problems

In this chapter we study a very natural problem of combinatorial geometry: the maximum possible number of incidences between m points and n lines in the plane. In addition to its mathematical appeal, this problem and its relatives are significant in the analysis of several basic geometric algorithms. In the proofs we encounter number-theoretic arguments, results about graph drawing, the probabilistic method, forbidden subgraphs, and line arrangements.

4.1 Formulation

Point–line incidences. Consider a set P of m points and a set L of n lines in the plane. What is the maximum possible number of their *incidences*, i.e., pairs (p, ℓ) such that $p \in P$, $\ell \in L$, and p lies on ℓ? We denote the number of incidences for specific P and L by $I(P, L)$, and we let $I(m, n)$ be the maximum of $I(P, L)$ over all choices of an m-element P and an n-element L. For example, the following picture illustrates that $I(3, 3) \geq 6$,



and it is not hard to see that actually $I(3, 3) = 6$.

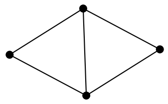
A trivial upper bound is $I(m, n) \leq mn$, but it it can never be attained unless $m = 1$ or $n = 1$. In fact, if m has a similar order of magnitude as n then $I(m, n)$ is asymptotically much smaller than mn. The order of magnitude is known exactly:

4.1.1 Theorem (Szemerédi–Trotter theorem). *For all $m, n \geq 1$, we have $I(m, n) = O(m^{2/3}n^{2/3} + m + n)$, and this bound is asymptotically tight.*

We give two proofs in the sequel, one simpler and one including techniques useful in more general situations. We will mostly consider only the most interesting case $m = n$. The general case needs no new ideas but only a little more complicated calculation.

Of course, the problem of point–line incidences can be generalized in many ways. We can consider incidences between points and hyperplanes in higher dimensions, or between points in the plane and some family of curves, and so on. A particularly interesting case is that of points and unit circles, which is closely related to counting unit distances.

Unit distances and distinct distances. Let $U(n)$ denote the maximum possible number of pairs of points with unit distance in an n-point set in the plane. For $n \leq 3$ we have $U(n) = \binom{n}{2}$ (all distances can be 1), but already for $n = 4$ at most 5 of the 6 distances can be 1; i.e., $U(4) = 5$:



We are interested in the asymptotic behavior of the function $U(n)$ for $n \to \infty$.

This can also be reformulated as an incidence problem. Namely, consider an n-point set P and draw a unit circle around each point of p, thereby obtaining a set C of n unit circles. Each pair of points at unit distance contributes two point–circle incidences, and hence $U(n) \leq \frac{1}{2} I_{1\mathrm{circ}}(n,n)$, where $I_{1\mathrm{circ}}(m,n)$ denotes the maximum possible number of incidences between m points and n unit circles.

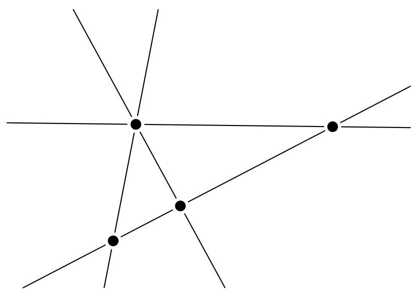
Unlike the case of point–line incidences, the correct order of magnitude of $U(n)$ is not known. An upper bound of $O(n^{4/3})$ can be obtained by modifying proofs of the Szemerédi–Trotter theorem. But the best known lower bound is $U(n) \geq n^{1+c_1/\log\log n}$, for some positive constant c_1; this is superlinear in n but grows more slowly than $n^{1+\varepsilon}$ for every fixed $\varepsilon > 0$.

A related quantity is the minimum possible number of distinct distances determined by n points in the plane; formally,

$$g(n) = \min_{P \subset \mathbf{R}^2 \colon |P| = n} |\{\mathrm{dist}(x,y) \colon x, y \in P\}|.$$

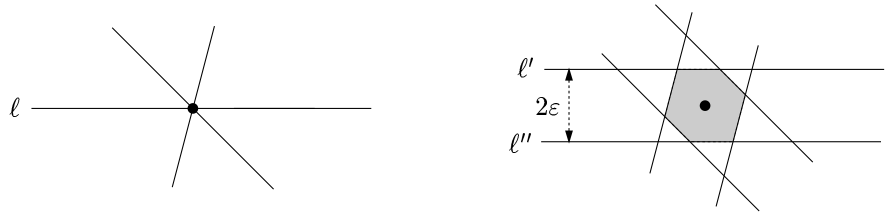
Clearly, $g(n) \geq \binom{n}{2} / U(n)$, and so the bound $U(n) = O(n^{4/3})$ mentioned above gives $g(n) = \Omega(n^{2/3})$. This has been improved several times, and the current best lower bound is approximately $\Omega(n^{0.863})$. The best known upper bound is $O(n/\sqrt{\log n})$.

Arrangements of lines. We need to introduce some terminology concerning line arrangements. Consider a finite set L of lines in the plane. They divide the plane into convex subsets of various dimensions, as is indicated in the following picture with 4 lines:



The intersections of the lines, indicated by black dots, are called the *vertices*. By removing all the vertices lying on a line $\ell \in L$, the line is split into two unbounded rays and several segments, and these parts are the *edges*. Finally, by deleting all the lines of L, the plane is divided into open convex polygons, called the *cells*. In Chapter 6 we will study arrangements of lines and hyperplanes further, but here we need only this basic terminology and (later) the simple fact that an arrangement of n lines in general position has $\binom{n}{2}$ vertices, n^2 edges, and $\binom{n}{2}+n+1$ cells. For the time being, the reader can regard this as an exercise, or wait until Chapter 6 for a proof.

Many cells in arrangements. What is the maximum total number of vertices of m distinct cells in an arrangement of n lines in the plane? Let us denote this number by $K(m,n)$. A simple construction shows that the maximum number of incidences $I(m,n)$ is asymptotically bounded from above by $K(m,n)$; more exactly, we have $I(m,n) \leq \frac{1}{2}K(m,2n)$. To see this, consider a set P of m points and a set L of n lines realizing $I(m,n)$, and replace each line $\ell \in L$ by a pair of lines ℓ', ℓ'' parallel to ℓ and lying at distance ε from ℓ:

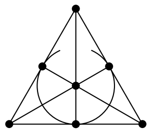


If $\varepsilon > 0$ is sufficiently small, then a point $p \in P$ incident to k lines in the original arrangement now lies in a tiny cell with $2k$ vertices in the modified arrangement.

It turns out that $K(m,n)$ has the same order of magnitude as $I(m,n)$, and the upper bound can be obtained by methods similar to those used for $I(m,n)$. In higher-dimensional problems, even determining the maximum possible complexity of a single cell can be quite challenging. For example, the maximum complexity of a single cell in an arrangement of n hyperplanes is described by the so-called upper bound theorem from the 1970s, which will be discussed in Chapter 5.

Bibliography and remarks. This chapter is partially based on a nice presentation of the discussed topics in the book by Pach and Agarwal [PA95], which we recommend as a source of additional information concerning history, bibliographic references, and various related problems. But we also include some newer results and techniques discovered since the publication of that book.

The following neat problem concerning point–line incidences was posed by Sylvester [Syl93] in 1893: Prove that it is impossible to arrange a finite number of points in the plane so that a line through every two of them passes through a third, unless they all lie on the same line. This problem remained unsolved until 1933, when it was asked again by Erdős and solved shortly afterward by Gallai. The solution shows, in particular, that it is impossible to embed the points of a finite projective plane $\mathcal{F}$ into $\mathbf{R}^2$ in such a way that points of each line of $\mathcal{F}$ lie on a straight line in $\mathbf{R}^2$. For example, the well-known drawing of the Fano plane of order 3 has to contain a curved line:



Recently Pinchasi [Pin02] proved the following conjecture of Bezdek, resembling Sylvester's problem: For every finite family of at least 5 unit circles in the plane, every two of them intersecting, there exists an intersection point common to exactly 2 of the circles.

The problems of estimating the maximum number of point–line incidences, the maximum number of unit distances, and the minimum number of distinct distances were raised by Erdős [Erd46]. For *point–line incidences*, he proved the lower bound $I(m,n) = \Omega(m^{2/3}n^{2/3} + m + n)$ (see Section 4.2) and conjectured it to be the right order of magnitude. This was first proved by Szemerédi and Trotter [ST83]. Simpler proofs were found later by Clarkson, Edelsbrunner, Guibas, Sharir, and Welzl [CEG$^+$90], by Székely [Szé97], and by Aronov and Sharir [AS01a]; they are quite different from one another, and we discuss them all in this chapter.

Tóth [Tót01a] proved the analogy of the Szemerédi–Trotter theorem for the complex plane; he used the original Szemerédi–Trotter technique, since none of the simpler proofs seems to work there.

A beautiful application of techniques of Clarkson et al. [CEG$^+$90] in geometric measure theory can be found in Wolff [Wol97]. This paper deals with a variation of the *Kakeya problem*: It shows that any Borel set in the plane containing a circle of every radius has Hausdorff dimension 2.

For *unit distances in the plane* Erdős [Erd46] established the lower bound $U(n) = \Omega(n^{1+c/\log\log n})$ (Section 4.2) and again conjectured it to be tight, but the best known upper bound remains $O(n^{4/3})$. This was first shown by Spencer, Szemerédi, and Trotter [SST84], and it can be re-proved by modifying each of the proofs mentioned above for point–line incidences. Further improvement of the upper bound probably needs different, more "algebraic," methods, which would use the "circularity" in a strong way, not just in the form of simple combinatorial axioms (such as that two points determine at most two unit circles).

For the analogous problem of *unit distances among n points in* $\mathbf{R}^3$, Erdős [Erd60] proved $\Omega(n^{4/3}\log\log n)$ from below and $O(n^{5/3})$ from above. The example for the lower bound is the grid $\{1,2,\ldots,\lfloor n^{1/3}\rfloor\}^3$ appropriately scaled; the bound $\Omega(n^{4/3})$ is entirely straightforward, and the extra $\log\log n$ factor needs further number-theoretic considerations. The upper bound follows by an argument with forbidden $K_{3,3}$; similar proofs are shown in Section 4.5. The current best bound is close to $O(n^{3/2})$; more precisely, it is $n^{3/2}2^{O(\alpha^2(n))}$ [CEG$^+$90]. Here the function $\alpha(n)$, to be defined in Section 7.2, grows extremely slowly, more slowly than $\log n$, $\log\log n$, $\log\log\log n$, etc. In *dimensions 4 and higher*, the number of unit distances can be $\Omega(n^2)$ (Exercise 2). Here even the constant at the leading term is known; see [PA95]. Among other results related to the unit-distance problems and considering point sets with various restrictions, we mention a neat construction of Erdős, Hickerson, and Pach [EHP89] showing that, for every $\alpha \in (0,2)$, there is an n-point set on the 2-dimensional unit sphere with the distance α occurring at least $\Omega(n\log^* n)$ times (the special distance $\sqrt{2}$ can even occur $\Omega(n^{4/3})$ times), and the annoying (and still unsolved) problem of Erdős and Moser, whether the number of unit distances in an n-point planar set in convex position is always bounded by $O(n)$ (see [PA95] for partial results and references).

For *distinct distances in the plane*, the best known upper bound, due to Erdős, is $O(n/\sqrt{\log n})$. This bound is attained for the $\sqrt{n}\times\sqrt{n}$ square grid. After a series of increases of the lower bound (Moser [Mos52], Chung [Chu84], Beck [Bec83], Clarkson et al. [CEG$^+$90], Chung, Szemerédi, and Trotter [CST92], Székely [Szé97], Solymosi and Tóth [ST01]) the current record is $\Omega(n^{4/(5-1/e)-\varepsilon})$ for every fixed $\varepsilon > 0$ (the exponent is approximately 0.863) by Tardos [Tar01], who improved a number-theoretic lemma in the Solymosi–Tóth proof. Aronov and Sharir [AS01b] obtained the lower bound of approximately $n^{0.526}$ for distinct distances in $\mathbf{R}^3$.

Another challenging quantity is the number $I_{\rm circ}(m,n)$ of *incidences of m points with n arbitrary circles* in the plane. The lower bound for point–line incidences can be converted to an example with m points, n circles, and $\Omega(m^{2/3}n^{2/3}+m+n)$ incidences, but in the case of $I_{\rm circ}(m,n)$, this lower bound is *not* the best possible for all m and n: Consider an example of an n-point set with $t = O(n/\sqrt{\log n}\,)$ distinct distances and draw the t circles with these distances as radii around each point; the resulting $tn = o(n^2)$ circles have $\Omega(n^2)$ incidences with the n points. The current record in the upper bound is due to Aronov and Sharir [AS01a], and for $m = n$ it yields $I_{\rm circ}(n,n) = O(n^{15/11+\varepsilon}) = O(n^{1.364})$. A little more about their approach is mentioned in the notes to Section 4.5, including an outline of a proof of a weaker bound $I_{\rm circ}(n,n) = O(n^{1.4})$. Two other methods for obtaining upper bounds are indicated in Exercises 4.4.2 and 4.6.4.

More generally, one can consider $I(P,\Gamma)$, the number of incidences between an m-point $P \subset \mathbf{R}^2$ and a family Γ of n planar curves. Pach and Sharir [PS98a] proved by Székely's method that if Γ is a *family of curves with k degrees of freedom and multiplicity type s*, meaning that for any k points there are at most s curves of Γ passing through all of them and no two curves intersect in more than k points, then $|I(P,\Gamma)| = O\left(m^{k/(2k-1)}n^{1-1/(2k-1)} + m + n\right)$, with the constant of proportionality depending on k and s. Earlier [PS92], they proved the same bound with some additional technical assumptions on the family Γ by the technique of Clarkson et al. [CEG$^+$90]. Most likely this bound is not tight for $k \geq 3$. Aronov and Sharir [AS01a] improved the bound slightly for Γ a family of graphs of univariate polynomials of degree at most k. The best known lower bound is mentioned in the notes to Section 4.2 below.

Point–plane incidences. Considering n points on a line in $\mathbf{R}^3$ and m planes containing that line, we see that the number of incidences can be mn without further assumptions on the position of the points and/or planes. Agarwal and Aronov [AA92] proved the upper bound $O(m^{3/5}n^{4/5} + m + n)$ for the number of incidences between m planes and n points in $\mathbf{R}^3$ if no 3 of the points are collinear, slightly improving on a result of Edelsbrunner, Guibas, and Sharir [EGS90]. In dimension d, the maximum number of incidences of n hyperplanes with m vertices of their arrangement is $O(m^{2/3}n^{d/3}+n^{d-1})$ [AA92], and this is tight for $m \geq n^{d-2}$ (for smaller m, the trivial $O(mn)$ bound is tight).

The complexity of *many cells in an arrangement of lines* was first studied by Canham [Can69], who proved $K(m,n) = O(m^2 + n)$, using the fact that two cells can have at most 4 lines incident to both of them (essentially a "forbidden $K_{2,5}$" argument; see Section 4.5). The tight bound $O(m^{2/3}n^{2/3}+m+n)$ was first achieved by Clarkson et al. [CEG$^+$90]. Among results for the complexity of m cells in other types of arrangements we mention the bound $O(m^{2/3}n^{2/3} + n\alpha(n) + n\log m)$ for segments by Aronov, Edelsbrunner, Guibas, and Sharir [AEGS92], $O(m^{2/3}n^{2/3}\alpha(n)^{1/3} + n)$ for unit circles [CEG$^+$90] (improved to $O(m^{2/3}n^{2/3}) + n)$ by Agarwal, Aronov, and Sharir [AAS01]), $O(m^{3/5}n^{4/5}2^{0.4\alpha(n)} + n)$ for arbitrary circles [CEG$^+$90] (also improved in [AAS01]; see the notes to Section 4.5), $O(m^{2/3}n + n^2)$ for planes in $\mathbf{R}^3$ by Agarwal and Aronov [AA92] (which is tight), and $O(m^{1/2}n^{d/2}(\log n)^{(\lfloor d/2\rfloor-1)/2})$ for hyperplanes in $\mathbf{R}^d$ by Aronov, Matoušek, and Sharir [AMS94]. If one counts only facets of m cells in an arrangement of n hyperplanes in $\mathbf{R}^d$, then the tight bound is $O(m^{2/3}n^{d/3} + n^{d-1})$ [AA92]. A few more references on this topic can be found in Agarwal and Sharir [AS00a].

The number of similar copies of a configuration. The problem of unit distances can be rephrased as follows. Let K denote a set consisting of two points in the plane with unit distance. What is the maximum number of congruent copies of K that can occur in an n-point set in the plane? This reformulation opens the way to various interesting generalizations, where one can vary K, or one can consider homothetic or similar copies of K, and so on. Elekes's survey [Ele01] nicely describes these problems, their relation to the incidence bounds, and other connections. Here we sketch some of the main developments.

Beautiful results were obtained by Laczkovich and Ruzsa [LR97], who investigated the maximum number of *similar* copies of a given finite configuration K that can be contained in an n-point set in the plane. Earlier, Elekes and Erdős [EE94] proved that this number is $\Omega(n^{2-(\log n)^{-c}})$ for all K, where $c > 0$ depends on K, and it is $\Omega(n^2)$ whenever all the coordinates of the points in K are algebraic numbers. Building on these results, Laczkovich and Ruzsa proved that the maximum number of similar copies of K is $\Omega(n^2)$ if and only if the *cross-ratio* of every 4 points of K is algebraic, where the cross-ratio of points $a, b, c, d \in \mathbf{R}^2$ equals $\frac{c-a}{c-b} \cdot \frac{d-b}{d-a}$, with a, b, c, d interpreted as complex numbers in this formula.

Their proof makes use of very nice results from the additive theory of numbers, most notably a theorem of Freiman [Fre73] (also see Ruzsa [Ruz94]): If A is a set of n integers such that $|A + A| \le cn$, where $A + A = \{a + b\colon a, b \in A\}$ and $c > 0$ is a constant, then A is contained in a d-dimensional generalized arithmetic progression of size at most Cn, with C and d depending on c only. Here a d-dimensional *generalized arithmetic progression* is a set of integers of the form $\{z_0 + i_1q_1 + i_2q_2 + \cdots + i_dq_d\colon i_1 = 0, 1, \ldots, n_1,\ i_2 = 0, 1, \ldots, n_2, \ldots, i_d = 0, 1, \ldots, n_d\}$ for some integers z_0 and $q_1, q_2, \ldots, q_d$. It is easy to see that $|A + A| \le C_d|A|$ for every d-dimensional generalized arithmetic progression, and Freiman's theorem is a sort of converse statement: If $|A + A| = O(|A|)$, then A is not too far from a generalized arithmetic progression. (Freiman's theorem has also been used for incidence-related problems by Erdős, Füredi, Pach, and Ruzsa [EFPR93], and Gowers's paper [Gow98] is an impressive application of results of this type in combinatorial number theory.)

Polynomials attaining $O(n)$ values on Cartesian products. Interesting results related to those of Freiman, as well as to incidence problems, were obtained in a series of papers by Elekes and his coworkers (they are described in the already mentioned survey [Ele01]). Perhaps even more significant than the particular results is the direction of research opened by them, combining algebraic and combinatorial tools. Let us begin with a conjecture of Purdy proved by Elekes and Rónyai [ER00] as a consequence of their theorems. Let P be a set of n distinct points lying on a line $u \subset \mathbf{R}^2$, let Q be a set of n distinct points lying on a line $v \subset \mathbf{R}^2$, and let $\mathrm{Dist}(P, Q) = \{\|p - q\|\colon p \in P, q \in Q\}$. If, for example, u and v are parallel and if both P and Q are placed with equal spacing along their lines, then $|\mathrm{Dist}(P, Q)| \le 2n$. Another such case is $P = \{(\sqrt{i}, 0)\colon i = 1, 2, \ldots, n\}$ and $Q = \{(0, \sqrt{j})\colon j = 1, 2, \ldots, n\}$: This time u and v are perpendicular, and again $|\mathrm{Dist}(P, Q)| \le 2n$. According to Purdy's conjecture, these are the only possible positions of u and v if the number of distances is linear: For every $C > 0$ there is an n_0 such that if $n \ge n_0$ and $|\mathrm{Dist}(P, Q)| \le Cn$, then u and v are parallel or perpendicular.

If we parameterize the line u by a real parameter x, and v by y, and denote the cosine of the angle of u and v by λ, then Purdy's conjecture can be reformulated in algebraic terms as follows: Whenever $X, Y \subset \mathbf{R}$ are n-point sets such that the polynomial $F(x, y) = x^2 + y^2 + 2\lambda xy$ attains at most Cn distinct values on $X \times Y$, i.e., $|\{F(x, y)\colon x \in X, y \in Y\}| \le Cn$, then necessarily $\lambda = 0$ or $\lambda = \pm 1$, provided that $n \geq n_0(C)$.

Elekes and Rónyai [ER00] characterized all bivariate polynomials $F(x, y)$ that attain only $O(n)$ values on Cartesian products $X \times Y$. For every C, d there exists an n_0 such that if $F(x, y)$ is a bivariate polynomial of degree at most d and $X, Y \subset \mathbf{R}$ are n-point sets, $n \geq n_0$, such that $F(x, y)$ attains at most Cn distinct values on $X \times Y$, then $F(x, y)$ has one of the two special forms $f(g(x) + h(y))$ or $f(g(x)h(y))$, where f, g, h are univariate polynomials. In fact, we need not consider the whole $X \times Y$; it suffices to assume that F attains at most Cn values on an arbitrary subset of δn^2 pairs from $X \times Y$ (with n_0 depending on δ, too). A similar result holds for a bivariate *rational function* $F(x, y)$, with one more special form to consider, namely $F(x, y) = f((g(x) + h(y))/(1 - g(x)h(y)))$.

We indicate a proof only for the special case of the polynomial $F(x, y) = x^2 + y^2 + 2\lambda xy$ from Purdy's conjecture (following Elekes [Ele99]); the basic idea of the general case is similar, but several more tools are needed, especially from elementary algebraic geometry. So let $Z = F(X, Y)$ be the set of values attained by F on $X \times Y$. For each $y_i \in Y$, put $f_i(x) = F(x, y_i)$, and define the family $\Gamma = \{\gamma_{ij}\colon i, j = 1, 2, \ldots, n,\ i \neq j\}$ of planar curves by $\gamma_{ij} = \{(f_i(t), f_j(t))\colon t \in \mathbf{R}\}$ (this is the key trick). Each γ_{ij} contains at least $\frac{n}{2}$ points of $Z \times Z$, since among the n points $(f_i(x_k), f_j(x_k))$, $x_k \in X$, no 3 can coincide, because the f_i are quadratic polynomials. Moreover, a straightforward (although lengthy) calculation using resultants verifies that for $\lambda \notin \{0, \pm 1\}$, at most 8 distinct curves γ_{ij} can pass through any two given distinct points $a, b \in \mathbf{R}^2$. Consequently, Γ contains at least $\frac{1}{9}n^2$ distinct curves. Using the bound of Pach and Sharir [PS92], [PS98a] on the number of incidences between points and algebraic curves mentioned above, with $Z \times Z$ as the points and the at least $\frac{1}{9}n^2$ distinct curves of Γ as the curves, we obtain that $|Z| = \Omega(n^{5/4})$. So there is even a significant gap: Either $\lambda \in \{0, \pm 1\}$, and then $F(X, Y)$ can have only $2n$ distinct elements for suitable X, Y, or $\lambda \notin \{0, \pm 1\}$ and then $|F(X, Y)| = \Omega(n^{5/4})$ for all X, Y.

Perhaps this latter bound can be improved to $\Omega(n^{2-\varepsilon})$ for every $\varepsilon > 0$ (so there would be an almost-dichotomy: either the number of values of F can be linear, or it has to be always near-quadratic). On the other hand, it is known that the polynomial $x^2 + y^2 + xy$ attains only $O(n^2/\sqrt{\log n})$ distinct values for x, y ranging over $\{1, 2, \ldots, n\}$, and so the bound need not always be linear or quadratic. It seems likely that in the general case of the Elekes–Rónyai theorem the number of values attained by F should be near-quadratic unless F is one of the special forms.

Further generalizations of the Elekes–Rónyai theorem were obtained by Elekes and Szabó; see [Ele01].

Exercises

1. Let $I_{1\text{circ}}(m, n)$ be the maximum number of incidences of m points with n unit circles and let $U(n)$ be the maximum number of unit distances for an n-point set.

   (a) Prove that $I_{1\text{circ}}(2n, 2n) = O(I_{1\text{circ}}(n, n))$. $\boxed{1}$

   (b) We have seen that $U(n) \leq \frac{1}{2} I_{1\text{circ}}(n, n)$. Prove that $I_{1\text{circ}}(n, n) = O(U(n))$. $\boxed{2}$

2. Show that an n-point set in $\mathbf{R}^4$ may determine $\Omega(n^2)$ unit distances. $\boxed{4}$

3. Prove that if $X \subset \mathbf{R}^d$ is a set where every two points have distance 1, then $|X| \leq d+1$. $\boxed{3}$

4. What can be said about the maximum possible number of incidences of n lines in $\mathbf{R}^3$ with m points? $\boxed{2}$

5. Use the Szemerédi–Trotter theorem to show that n points in the plane determine at most

   (a) $O(n^{7/3})$ triangles of unit area, $\boxed{3}$

   (b) $O(n^{7/3})$ triangles with a given fixed angle α. $\boxed{2}$

   The result in (a) was first proved by Erdős and Purdy [EP71]. As for (b), Pach and Sharir [PS92] proved the better bound $O(n^2 \log n)$; also see [PA95].

6. (a) Using the Szemerédi–Trotter theorem, show that the maximum possible number of distinct lines such that each of them contains at least k points of a given m-point set P in the plane is $O(m^2/k^3 + m/k)$. $\boxed{2}$

   (b) Prove that such lines have at most $O(m^2/k^2 + m)$ incidences with P. $\boxed{3}$

7. (Many points on a line or many lines)

   (a) Let P be an m-point set in the plane and let $k \leq \sqrt{m}$ be an integer parameter. Prove (using Exercise 6, say) that at most $O(m^2/k)$ pairs of points of P lie on lines containing at least k and at most $\sqrt{m}$ points of P. $\boxed{3}$

   (b) Similarly, for $K \geq \sqrt{m}$, the number of pairs lying on lines with at least $\sqrt{m}$ and at most K points is $O(Km)$. $\boxed{3}$

   (c) Prove the following theorem of Beck [Bec83]: There is a constant $c > 0$ such that for any n-point $P \subseteq \mathbf{R}^2$, at least cn^2 distinct lines are determined by P or there is a line containing at least cn points of P. $\boxed{2}$

   (d) Derive that there exists a constant $c > 0$ such that for every n-point set P in the plane that does not lie on a single line there exists a point $p \in P$ lying on at least cn distinct lines determined by points of P. $\boxed{1}$

   Part (d) is a weak form of the *Dirac–Motzkin conjecture*; the full conjecture, still unsolved, is the same assertion with $c = \frac{1}{2}$.

8. (Many distinct radii)

   (a) Assume that $I_{\text{circ}}(m,n) = O(m^\alpha n^\beta + m + n)$ for some constants $\alpha < 1$ and $\beta < 1$, where $I_{\text{circ}}(m,n)$ is the maximum number of incidences of m points with n circles in the plane. In analogy with to Exercise 7, derive that there is a constant $c > 0$ such that for any n-point set $P \subset \mathbf{R}^2$, there are at least cn^3 distinct circles containing at least 3 points of P each or there is a circle or line containing at least cn points of P. $\boxed{3}$

   (b) Using (a), prove the following result of Elekes (an answer to a question of Balog): For any n-point set $P \subset \mathbf{R}^2$ not lying on a common circle or line, the circles determined by P (i.e., those containing 3 or more points of P) have $\Omega(n)$ distinct radii. $\boxed{4}$

   (c) Find an example of an n-point set with only $O(n)$ distinct radii. $\boxed{3}$

9. (Sums and products cannot both be few) Let $A \subset \mathbf{R}$ be a set of n distinct real numbers and let $S = A + A = \{a + b\colon a, b \in A\}$ and $P = A \cdot A = \{ab\colon a, b \in A\}$.

   (a) Check that each of the n^2 lines $\{(x,y) \in \mathbf{R}^2\colon y = a(x-b)\}$, $a, b \in A$, contains at least n distinct points of $S \times P$. $\boxed{1}$

   (b) Conclude using Exercise 6 that $|S \times P| = \Omega(n^{5/2})$, and consequently, $\max(|S|, |P) = \Omega(n^{5/4})$; i.e., the set of sums and the set of products can never both have almost linear size. $\boxed{2}$ (This is a theorem of Elekes [Ele97] improving previous results on a problem raised by Erdős and Szemerédi.)

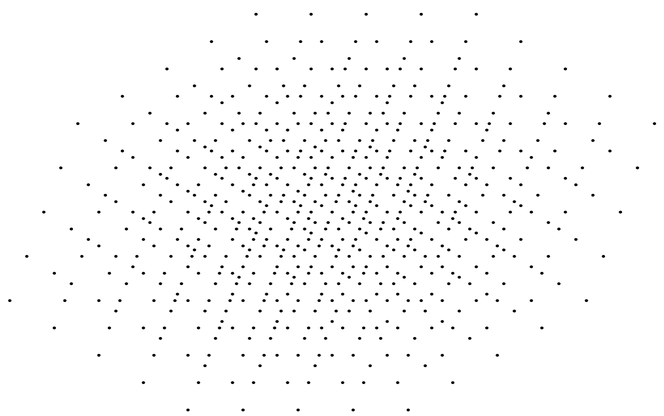
10. (a) Find n-point sets in the plane that contain $\Omega(n^2)$ similar copies of the vertex set of an equilateral triangle. $\boxed{1}$

(b) Verify that the following set P_m has $n = O(m^4)$ points and contains $\Omega(n^2)$ similar copies of the vertex set of a regular pentagon: Identify $\mathbf{R}^2$ with the complex plane $\mathbf{C}$, let $\omega = e^{2\pi \mathrm{i}/5}$ denote a primitive 5th root of unity, and put

$$P_m = \{i_0 + i_1\omega + i_2\omega^2 + i_3\omega^3 \colon i_0, i_1, i_2, i_3 \in \mathbf{Z},\ |i_j| \le m\}.$$

$\boxed{4}$

The example in (b) is from Elekes and Erdős [EE94], and the set P_∞ is called a *pentagonal pseudolattice*. The following picture shows P_2:



4.2 Lower Bounds: Incidences and Unit Distances

4.2.1 Proposition (Many point–line incidences). *We have $I(n,n) = \Omega(n^{4/3})$, and so the upper bound for the maximum number of incidences of n points and n lines in the plane in the Szemerédi–Trotter theorem is asymptotically optimal.*

It is not easy to come up with good constructions "by hand." Small cases do not seem to be helpful for discovering a general pattern. Surprisingly, an asymptotically optimal construction is quite simple. The appropriate lower bound for $I(m,n)$ with $n \neq m$ is obtained similarly (Exercise 1).

Proof. For simplicity, we suppose that $n = 4k^3$ for a natural number k. For the point set P, we choose the $k \times 4k^2$ grid; i.e., we set $P = \{(i,j)\colon i = 0,1,2,\ldots,k{-}1,\ j = 0,1,\ldots,4k^2{-}1\}$. The set L consists of all the lines with equations $y = ax + b$, where $a = 0,1,\ldots,2k{-}1$ and $b = 0,1,\ldots,2k^2{-}1$. These are n lines, as it should be. For $x \in [0,k)$, we have $ax+b < ak+b < 2k^2+2k^2 = 4k^2$. Therefore, for each $i = 0,1,\ldots,k{-}1$, each line of L contains a point of P with the x-coordinate equal to i, and so $I(P,L) \ge k \cdot |L| = \frac{1}{4}n^{4/3}$. $\square$

Next, we consider unit distances, where the construction is equally simple but the analysis uses considerable number-theoretic tools.

4.2.2 Theorem (Many unit distances). *For all $n \ge 2$, there exist configurations of n points in the plane determining at least $n^{1+c_1/\log\log n}$ unit distances, with a positive constant c_1.*

A configuration with the asymptotically largest known number of unit distances is a $\sqrt{n}\times\sqrt{n}$ regular grid with a suitably chosen step. Here unit distances are related to the number of possible representations of an integer as a sum of two squares. We begin with the following claim:

4.2.3 Lemma. *Let $p_1 < p_2 < \cdots < p_r$ be primes of the form $4k+1$, and put $M = p_1 p_2 \cdots p_r$. Then M can be expressed as a sum of two squares of integers in at least 2^r ways.*

Proof. As we know from Theorem 2.3.1, each p_j can be written as a sum of two squares: $p_j = a_j^2 + b_j^2$. In the sequel, we work with the ring $\mathbf{Z}[\mathrm{i}]$, the so-called Gaussian integers, consisting of all complex numbers $u + \mathrm{i}v$, where $u, v \in \mathbf{Z}$. We use the fact that each element of $\mathbf{Z}[\mathrm{i}]$ can be uniquely factored into primes. From algebra, we recall that a *prime* in the ring $\mathbf{Z}[\mathrm{i}]$ is an element $\gamma \in \mathbf{Z}[\mathrm{i}]$ such that whenever $\gamma = \gamma_1\gamma_2$ with $\gamma_1, \gamma_2 \in \mathbf{Z}[\mathrm{i}]$, then $|\gamma_1| = 1$ or $|\gamma_2| = 1$. Both existence and uniqueness of prime factorization follows from the fact that $\mathbf{Z}[\mathrm{i}]$ is a Euclidean ring (see an introductory course on algebra for an explanation of these notions).

Let us put $\alpha_j = a_j + \mathrm{i}\, b_j$, and let $\bar\alpha_j = a_j - \mathrm{i}\, b_j$ be the complex conjugate of α_j. We have $\alpha_j \bar\alpha_j = (a_j + \mathrm{i}\, b_j)(a_j - \mathrm{i}\, b_j) = a_j^2 + b_j^2 = p_j$. Let us choose an arbitrary subset $J \subseteq I = \{1, 2, \ldots, r\}$ and define $A_J + \mathrm{i}B_J = \left(\prod_{j\in J} \alpha_j\right)\left(\prod_{j\in I\setminus J} \bar\alpha_j\right)$. Then $A_J - \mathrm{i}B_J = \left(\prod_{j\in J} \bar\alpha_j\right)\left(\prod_{j\in I\setminus J} \alpha_j\right)$, and hence $M = (A_J + \mathrm{i}B_J)(A_J - \mathrm{i}B_J) = A_J^2 + B_J^2$. This gives one expression of the number M as a sum of two squares. It remains to prove that for two sets $J \neq J'$, $A_J + \mathrm{i}B_J \neq A_{J'} + \mathrm{i}B_{J'}$. To this end, it suffices to show that all the α_j and $\bar\alpha_j$ are primes in $\mathbf{Z}[\mathrm{i}]$. Then the numbers $A_J + \mathrm{i}B_J$ and $A_{J'} + \mathrm{i}B_{J'}$ are distinct since they have distinct prime factorizations. (No α_j or $\bar\alpha_j$ can be obtained from another one by multiplying it by a unit of the ring $\mathbf{Z}[\mathrm{i}]$: The units are only the elements $1, -1, \mathrm{i}$, and $-\mathrm{i}$.)

So suppose that $\alpha_j = \gamma_1\gamma_2$, $\gamma_1, \gamma_2 \in \mathbf{Z}[\mathrm{i}]$. We have $p_j = \alpha_j\bar\alpha_j = \gamma_1\gamma_2\bar\gamma_1\bar\gamma_2 = |\gamma_1|^2|\gamma_2|^2$. Now, $|\gamma_1|^2$ and $|\gamma_2|^2$ are both integers, and since p_j is a prime, we get that $|\gamma_1| = 1$ or $|\gamma_2| = 1$. □

Next, we need to know that the primes of the form $4k+1$ are sufficiently dense. First we recall the well-known *prime number theorem*: If $\pi(n)$ denotes the number of primes not exceeding n, then

$$\pi(n) = (1 + o(1))\frac{n}{\ln n} \quad \text{as } n \to \infty.$$

Proofs of this fact are quite complicated; on the other hand, it is not so hard to prove weaker bounds $cn/\log n < \pi(n) < Cn/\log n$ for suitable positive constants c, C.

We consider primes in the arithmetic progression $1, 5, 9, \ldots, 4k+1, \ldots$. A famous theorem of Dirichlet asserts that every arithmetic progression contains infinitely many primes unless this is impossible for a trivial reason, namely, unless all the terms have a nontrivial common divisor. The following theorem is still stronger:

4.2.4 Theorem. *Let d and a be relatively prime natural numbers, and let $\pi_{d,a}(n)$ be the number of primes of the form $a+kd$ ($k = 0, 1, 2, \dots$) not exceeding n. We have*

$$\pi_{d,a}(n) = (1 + o(1))\frac{1}{\varphi(d)} \cdot \frac{n}{\ln n},$$

where φ denotes the Euler function: $\varphi(d)$is the number of integers between 1 and d that are relatively prime to d.

For every $d \geq 2$, there are $\varphi(d)$ residue classes modulo d that can possibly contain primes. The theorem shows that the primes are quite uniformly distributed among these residue classes.

The proof of the theorem is not simple, and we omit it, but it is very nice and we can only recommend to the reader to look it up in a textbook on number theory.

Proof of the lower bound for unit distances (Theorem 4.2.2). Let us suppose that n is a square. For the set P we choose the points of the $\sqrt{n} \times \sqrt{n}$ grid with step $1/\sqrt{M}$, where M is the product of the first $r-1$ primes of the form $4k+1$, and r is chosen as the largest number such that $M \leq \frac{n}{4}$.

It is easy to see that each point of the grid participates in at least as many unit distances as there are representations of M as a sum of two squares of nonnegative integers. Since one representation by a sum of two squares of nonnegative integers corresponds to at most 4 representations by a sum of two squares of arbitrary integers (the signs can be chosen in 4 ways), we have at least $2^{r-1}/16$ unit distances by Lemma 4.2.3.

By the choice of r, we have $4p_1p_2\cdots p_{r-1} \leq n < 4p_1p_2\cdots p_r$, and hence $2^r \leq n$ and $p_r > (\frac{n}{4})^{1/r}$. Further, we obtain, by Theorem 4.2.4, $r = \pi_{4,1}(p_r) \geq (\frac{1}{2} - o(1))p_r/\log p_r \geq \sqrt{p_r} \geq n^{1/3r}$ for sufficiently large n, and thus $r^{3r} \geq n$. Taking logarithms, we have $3r\log r \geq \log n$, and hence $r \geq \log n/(3\log r) \geq \log n/(3\log\log n)$. The number of unit distances is at least $n\,2^{r-4} \geq n^{1+c_1/\log\log n}$, as Theorem 4.2.2 claims. Let us remark that for sufficiently large n the constant c_1 can be made as close to 1 as desired. □

Bibliography and remarks. Proposition 4.2.1 is due to Erdős [Erd46]. His example is outlined in Exercise 2 (also see [PA95]); the analysis requires a bit of number theory. The simpler example in the text is from Elekes [Ele01]. Its extension provides the best known lower bound for the number of incidences between m points and $n \geq m^{(k-1)/2}$ curves with k degrees of freedom: For a parameter $t \leq m^{1/k}$, let $P = \{(i,j)\colon 0 \leq i < t,\ 0 \leq j < \frac{n}{t}\}$, and let Γ consist of the graphs of the polynomials $\sum_{\ell=0}^{k-1} a_\ell x^\ell$ with $a_\ell = 0, 1, \dots, \lfloor \frac{n}{kt^{\ell+1}} \rfloor$, $\ell = 0, 1, \dots, k-1$.

Theorem 4.2.2 is due to Erdős [Erd46], and the proof uses ingredients well known in number theory. The prime number theorem (and also Theorem 4.2.4) was proved in 1896, by de la Valée Poussin and independently by Hadamard (see Narkiewicz [Nar00]).

Exercises

1. By extending the example in the text, prove that for all m, n with $n^2 \leq m$ and $m^2 \leq n$, we have $I(m,n) = \Omega(n^{2/3}m^{2/3})$. $\boxed{3}$

2. (Another example for incidences) Suppose that $n = 4t^6$ for an integer $t \geq 1$ and let $P = \{(i,j)\colon 0 \leq i,j < \sqrt{n}\}$. Let $S = \{(a,b), a,b = 1,2,\ldots,t, \gcd(a,b) = 1\}$, where $\gcd(a,b)$ denotes the greatest common divisor of a and b. For each point $p \in P$, consider the lines passing through p with slope a/b, for all pairs $(a,b) \in S$. Let L be the union of all the lines thus obtained for all points $p \in P$.

   (a) Check that $|L| \leq n$. $\boxed{2}$

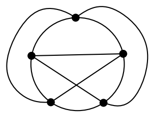
   (b) Prove that $|S| \geq ct^2$ for a suitable positive constant $c > 0$, and infer that $I(P,L) = \Omega(nt^2) = \Omega(n^{4/3})$. $\boxed{4}$

4.3 Point–Line Incidences via Crossing Numbers

Here we present a very simple proof of the Szemerédi–Trotter theorem based on a result concerning graph drawing. We need the notion of the crossing number of a graph G; this is the minimum possible number of edge crossings in a drawing of G. To make this rigorous, let us first recall a formal definition of a drawing.

An *arc* is the image of a continuous injective map $[0,1] \to \mathbf{R}^2$. A *drawing* of a graph G is a mapping that assigns to each vertex of G a point in the plane (distinct vertices being assigned distinct points) and to each edge of G an arc connecting the corresponding two (images of) vertices and not incident to any other vertex. We do not insist that the drawing be planar, so the arcs are allowed to cross. A *crossing* is a point common to at least two arcs but distinct from all vertices. In this section we will actually deal only with drawings where each edge is represented by a straight segment.

Let G be a graph (or multigraph). The *crossing number* of a drawing of G in the plane is the number of crossings in the considered drawing, where a crossing incident to $k \geq 2$ edges is counted $\binom{k}{2}$ times. So a drawing is planar if and only if its crossing number is 0. The *crossing number* of the graph G is the smallest possible crossing number of a drawing of G; we denote it by $\mathrm{cr}(G)$. For example, $\mathrm{cr}(K_5) = 1$:



As is well known, for $n > 2$, a planar graph with n vertices has at most $3n-6$ edges. This can be rephrased as follows: If the number of edges is at least $3n-5$ then $\mathrm{cr}(G) > 0$. The following theorem can be viewed as a generalization of this fact.

4.3.1 Theorem (Crossing number theorem). *Let $G = (V,E)$ be a simple graph (no multiple edges). Then*

$$\mathrm{cr}(G) \geq \frac{1}{64} \cdot \frac{|E|^3}{|V|^2} - |V|$$

(the constant $\frac{1}{64}$ can be improved by a more careful calculation).

The lower bound in this theorem is asymptotically tight; i.e., there exist graphs with n vertices, m edges, and crossing number $O(m^3/n^2)$; see Exercise 1. The assumption that the graph is simple cannot be omitted.

For a proof of this theorem, we need a simple lemma:

4.3.2 Lemma. *The crossing number of any simple graph $G = (V, E)$ is at least $|E| - 3|V|$.*

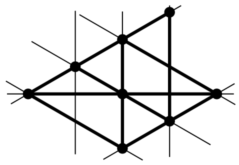
Proof. If $|E| > 3|V|$ and some drawing of the graph had fewer than $|E| - 3|V|$ crossings, then we could delete one edge from each crossing and obtain a planar graph with more than $3|V|$ edges. □

Proof of Theorem 4.3.1. Consider some drawing of a graph $G = (V, E)$ with n vertices, m edges, and crossing number x. We may assume $m \geq 4n$, for otherwise, the claimed bound is negative. Let $p \in (0, 1)$ be a parameter; later on we set it to a suitable value. We choose a random subset $V' \subseteq V$ by including each vertex $v \in V$ into V' independently with probability p. Let G' be the subgraph of G induced by the subset V'. Put $n' = |V'|$, $m' = |E(G')|$, and let x' be the crossing number of the graph G' in the drawing "inherited" from the considered drawing of G. The expectation of n' is $\mathbf{E}[n'] = np$. The probability that a given edge appears in $E(G')$ is p^2, and hence $\mathbf{E}[m'] = mp^2$, and similarly we get $\mathbf{E}[x'] = xp^4$. At the same time, by Lemma 4.3.2 we always have $x' \geq m' - 3n'$, and so this relation holds for the expectations as well: $\mathbf{E}[x'] \geq \mathbf{E}[m'] - 3\mathbf{E}[n']$. So we have $xp^4 \geq mp^2 - 3np$. Setting $p = \frac{4n}{m}$ (which is at most 1 since we assume $m \geq 4n$), we calculate that

$$x \geq \frac{1}{64} \frac{m^3}{n^2}.$$

The crossing number theorem is proved. □

Proof of the Szemerédi–Trotter theorem (Theorem 4.1.1). We consider a set P of m points and a set L of n lines in the plane realizing the maximum number of incidences $I(m, n)$. We define a certain topological graph $G = (V, E)$, that is, a graph together with its drawing in the plane. Each point $p \in P$ becomes a vertex of G, and two points $p, q \in P$ are connected by an edge if they lie on a common line $\ell \in L$ next to one another. So we have a drawing of G where the edges are straight segments. This is illustrated below, with G drawn thick:



If a line $\ell \in L$ contains $k \geq 1$ points of P, then it contributes $k-1$ edges to P, and hence $I(m, n) = |E| + n$. Since the edges are parts of the n lines, at most $\binom{n}{2}$ pairs may cross: $\mathrm{cr}(G) \leq \binom{n}{2}$. On the other hand, from the crossing number theorem 4.3.1 we get $\mathrm{cr}(G) \geq \frac{1}{64} \cdot |E|^3/m^2 - m$. So $\frac{1}{64} \cdot |E|^3/m^2 - m \leq \mathrm{cr}(G) \leq \binom{n}{2}$, and a calculation gives $|E| = O(n^{2/3}m^{2/3} + m)$. This proves the Szemerédi–Trotter theorem. $\square$

The best known upper bound on the number of unit distances, $U(n) = O(n^{4/3})$, can be proved along similar lines; see Exercise 2.

Bibliography and remarks. The presented proof of the Szemerédi–Trotter theorem is due to Székely [Szé97].

The crossing number theorem was proved by Ajtai, Chvátal, Newborn, and Szemerédi [ACNS82] and independently by Leighton [Lei84]. This result belongs to the theory of *geometric graphs*, which studies the properties of graphs drawn in the plane (most often with edges drawn as straight segments). A nice introduction to this area is given in Pach and Agarwal [PA95], and a newer survey is Pach [Pac99]. In the rest of this section we mention mainly some of the more recent results.

Pach and Tóth [PT97] improved the constant $\frac{1}{64}$ in Theorem 4.3.1 to approximately 0.0296, which is already within a factor of 2.01 of the best known upper bound (obtained by connecting all pairs of points of distance at most d in a regular $\sqrt{n} \times \sqrt{n}$ grid, for a suitable d). The improvement is achieved by establishing a better version of Lemma 4.3.2, namely, $\mathrm{cr}(G) \geq 5|E| - 25|V|$ for $|E| > 7|V| - 14$.

Pach, Spencer, and Tóth [PST00] proved that for graphs with certain forbidden subgraphs, the bound can be improved substantially: For example, if G has n vertices, m edges, and contains no cycle of length 4, then $\mathrm{cr}(G) = \Omega(m^4/n^3)$ for $m \geq 400n$, which is asymptotically tight. Generally, let $\mathcal{G}$ be a class of graphs that is monotone (closed under adding edges) and such that any n-vertex graph in $\mathcal{G}$ has at most $O(n^{1+\alpha})$ edges, for some $\alpha \in (0,1)$. Then $\mathrm{cr}(G) \geq cm^{2+1/\alpha}/n^{1+1/\alpha}$ for any $G \in \mathcal{G}$ with n vertices and $m \geq Cn\log^2 n$ edges, with suitable constants $C, c > 0$ depending on $\mathcal{G}$. The proof applies a generally useful lower bound on the crossing number, which we outline next. Let $\mathrm{bw}(G)$ denote the *bisection width* of G, i.e., the minimum number of edges connecting V_1 and V_2, over all partitions (V_1, V_2) of $V(G)$ with $|V_1|, |V_2| \geq \frac{1}{3}|V(G)|$. Leighton [Lei83] proved that $\mathrm{cr}(G) = \Omega(\mathrm{bw}(G)^2) - |V(G)|$ for any graph G of maximum degree bounded by a constant. Pach, Shahrokhi, and Szegedy [PSS96], and independently Sýkora and Vrťo [SV94], extended this to graphs with arbitrary degrees:

$$\mathrm{cr}(G) = \Omega\left(\mathrm{bw}(G)^2\right) - \frac{1}{16}\sum_{v \in V(G)} \deg_G(v)^2, \tag{4.1}$$

where $\deg_G(v)$ is the degree of v in G. The proof uses the following version, due to Gazit and Miller [GM90], of the well-known Lipton–Tarjan separator theorem for planar graphs: For any planar graph H and any nonnegative weight function $w\colon V(H) \to [0, \frac{2}{3}]$ with $\sum_{v \in V(H)} w(v) = 1$, one can delete at most $1.58\sqrt{\sum_{v \in V(H)} \deg_H(v)^2}$ edges in such a way that the total weight of vertices in each component of the resulting graph is at most $\frac{2}{3}$. To deduce (4.1), consider a drawing of G with the minimum number of crossings, replace each crossing by a vertex of degree 4, assign weight 0 to these vertices and weight $\frac{1}{|V(G)|}$ to the original vertices, and apply the separator theorem (see, e.g., [PA95] for a more detailed account). Djidjev and Vrťo [DV02] have recently strengthened (4.1), replacing $\mathrm{bw}(G)$ by the *cutwidth* of G. To define the cutwidth, we consider an injective mapping $f\colon V(G) \to \mathbf{R}$. Each edge corresponds to a closed interval, and we find the maximum number of these intervals with a common interior point. The cutwidth is the minimum of this quantity over all f.

To derive the result of Pach et al. [PST00] on the crossing number of graphs with forbidden subgraphs mentioned above from (4.1), we consider a graph $G \in \mathcal{G}$ with n vertices and m edges. If $\mathrm{cr}(G)$ is small, then the bisection width is small, so G can be cut into two parts of almost equal size by removing not too many edges. For each of these parts, we bisect again, and so on, until parts of some suitable size s (depending on n and m) are reached. By the assumption on $\mathcal{G}$, each of the resulting parts has $O(s^{1+\alpha})$ edges, and so there are $O(ns^{\alpha})$ edges within the parts. This number of edges plus the number of edges deleted in the bisections add up to m, and this provides an inequality relating $\mathrm{cr}(G)$ to n and m; see [PST00] for the calculations.

The notion of crossing number is a subtle one. Actually, one can give several natural definitions; a study of various notions and of their relations was made by Pach and Tóth [PT00]. Besides counting the crossings, as we did in the definition of $\mathrm{cr}(G)$, one can count the number of (unordered) pairs of edges that cross; the resulting notion is called the *pairwise crossing number* in [PT00] and we denote it by $\text{pair-cr}(G)$. We always have $\text{pair-cr}(G) \le \mathrm{cr}(G)$, but since two edges (arcs) are allowed to cross several times, it is not clear whether $\text{pair-cr}(G) = \mathrm{cr}(G)$ for all graphs G, and currently this seems to be a challenging open problem (see Exercise 4 for a typical false attempt at a proof). A simple argument shows that $\mathrm{cr}(G) \le 2\,\text{pair-cr}(G)^2$ (Exercise 4(c)). A stronger claim, proved in [PT00], is $\mathrm{cr}(G) \le 2\,\text{odd-cr}(G)^2$, where $\text{odd-cr}(G)$ is the *odd-crossing number* of G, counting the number of pairs of edges that cross an odd number of times. An inspiration for their proof is a theorem of Hanani and Tutte claiming that a graph G is planar if and only if $\text{odd-cr}(G) = 0$. In a drawing of G, call an edge e *even* if there is no edge crossed by e an odd number of times. Pach and Tóth show, by a somewhat complicated proof, that if we consider a drawing of G and let E_0 be the set of the even edges, then there is another drawing of G in which the edges of E_0 are involved in no crossings at all. The inequality $\mathrm{cr}(G) \le 2\,\text{odd-cr}(G)^2$ then follows by an argument similar to that in Exercise 4(c).

Finally, let us remark that if we consider rectilinear drawings (where each edge is drawn as a straight segment), then the resulting *rectilinear crossing number* can be much larger than any of the crossing numbers considered above: Graphs are known with $\mathrm{cr}(G) = 4$ and arbitrarily large rectilinear crossing numbers (Bienstock and Dean [BD93]).

Exercises

1. Show that for any n and m, $5n \le m \le \binom{n}{2}$, there exist graphs with n vertices, m edges, and crossing number $O(m^3/n^2)$. $\boxed{2}$
2. In a manner similar to the above proof for point–line incidences, prove the bound $I_{1\mathrm{circ}}(n,n) = O(n^{4/3})$, where $I_{1\mathrm{circ}}(m,n)$ denotes the maximum possible number of incidences between m points and n unit circles in the plane (be careful in handling possible multiple edges in the considered topological graph!). $\boxed{3}$
3. Let $K(n,m)$ denote the maximum total number of edges of m distinct cells in an arrangement of n lines in the plane. Prove $K(n,m) = O(n^{2/3}m^{2/3} + n + m)$ using the method of the present section (it may be convenient to classify edges into top and bottom ones and bound each type separately). $\boxed{3}$
4. (a) Prove that in a drawing of G with the smallest possible number of crossings, no two edges cross more than once. $\boxed{2}$

   (b) Explain why the result in (a) does not imply that $\text{pair-cr}(G) = \mathrm{cr}(G)$ (where $\text{pair-cr}(G)$ is the minimum number of pairs of crossing edges in a drawing of G). $\boxed{2}$

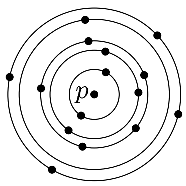
(c) Prove that if G is a graph with pair-cr$(G) = k$, then $\mathrm{cr}(G) \le \binom{2k}{2}$. $\boxed{4}$

4.4 Distinct Distances via Crossing Numbers

Here we use the methods from the preceding sections to establish a lower bound on the number of distinct distances determined by an n-point set in the plane. We do not go for the best known bound, whose proof is too complicated for our purposes, but in the notes below we indicate how the improvement is achieved.

4.4.1 Proposition (Distinct distances in $\mathbf{R}^2$). *The minimum number $g(n)$ of distinct distances determined by an n-point set in the plane satisfies $g(n) = \Omega(n^{4/5})$.*

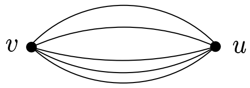
Proof. Fix an n-point set P, and let t be the number of distinct distances determined by P. This means that for each point $p \in P$, all the other points are contained in t circles centered at p (the radii correspond to the t distances appearing in P).



These tn circles obtained for all the n points of P have $n(n-1)$ incidences with the points of P. The first idea is to bound this number of incidences from above in terms of n and t, in a way similar to the proof of the Szemerédi–Trotter theorem in the preceding section, which yields a lower bound for t.

First we delete all circles with at most 2 points on them (the innermost circle and the second outermost circle in the above picture). We have destroyed at most $2nt$ incidences, and so still almost n^2 incidences remain (we may assume that t is much smaller than n, for otherwise, there is nothing to prove). Now we define a graph G: The vertices are the points of P and the edges are the arcs of the circles between the points. This graph has n vertices, almost n^2 edges, and there are at most t^2n^2 crossings because every two circles intersect in at most 2 points.

Now *if* we could apply the crossing number theorem to this graph, we would get that with n vertices and n^2 edges there must be at least $\Omega(n^6/n^2) = \Omega(n^4)$ crossings, and so $t = \Omega(n)$ would follow. This, of course, is too good to be true, and indeed we cannot use the crossing number theorem directly because our graph may have multiple edges: Two points can be connected by several arcs.



A multigraph can have arbitrarily many edges even if it is planar. But if we have a bound on the maximum edge multiplicity, we can still infer a lower bound on the crossing number:

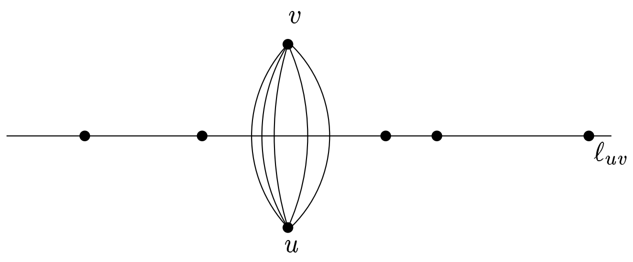
4.4.2 Lemma. *Let $G = (V, E)$ be a multigraph with maximum edge multiplicity k. Then*

$$\mathrm{cr}(G) = \Omega\left(\frac{|E|^3}{k|V|^2}\right) - O(k^2|V|).$$

We defer the proof to the end of this section.

In the graph G defined above, it appears that the maximum edge multiplicity can be as high as t. If we used Lemma 4.4.2 with $k = t$ in the manner indicated above, we would get only the estimate $t = \Omega(n^{2/3})$.

The next idea is to deal with the edges of very high multiplicity separately. Namely, we observe that if a pair $\{u, v\}$ of points is connected by k arcs, then the centers of these arcs lie on the symmetry axis ℓ_{uv} of the segment uv:



So the line ℓ_{uv} has at least k incidences with the points of P. But the Szemerédi–Trotter theorem tells us that there cannot be too many distinct lines, each incident to many points of P. Let us make this precise.

By a consequence of the Szemerédi–Trotter theorem stated in Exercise 4.1.6(b), lines containing at least k points of P each have altogether no more than $O(n^2/k^2 + n)$ incidences with P.

Let M be the set of pairs $\{u, v\}$ of vertices of G connected by at least k edges in G, and let E be the set of edges (arcs) connecting these pairs. Each edge in E connecting the pair $\{u, v\}$ contributes one incidence of the bisecting line ℓ_{uv} with a point $p \in P$. On the other hand, one incidence of such p with some ℓ_{uv} can correspond to at most $2t$ edges of E, because at most t circles are centered at p, and so ℓ_{uv} intersects at most $2t$ arcs with center p. So we have $|E| = O(tn^2/k^2 + tn)$.

Let us set k as large as possible but so that $|E| \le \frac{1}{2}n^2$, i.e., $k = C\sqrt{t}$ for a sufficiently large constant C. If we delete all edges of E, the remaining graph still has $\Omega(n^2)$ edges, but the maximum multiplicity is now below k. We can finally apply Lemma 4.4.2: With n vertices, $\Omega(n^2)$ edges, and edge multiplicity at most $k = O(\sqrt{t})$, we have at least $\Omega(n^4/\sqrt{t}\,)$ crossings. This number must be below t^2n^2, which yields $t = \Omega(n^{4/5})$ as claimed. □

Proof of Lemma 4.4.2. Consider a fixed drawing of G. We choose a subgraph G' of G by the following random experiment. In the first stage, we consider each edge of G independently, and we delete it with probability $1 - \frac{1}{k}$. In the second stage, we delete all the remaining multiple edges, and this gives G', which has n vertices, m' edges, and x' crossing pairs of edges. Consider the probability p_e that a fixed edge $e \in E$ remains in G'. Clearly $p_e \le \frac{1}{k}$. On the other hand, if e was one of $k' \le k$ edges connecting the same pair of vertices, then the probability that e survives the first stage while all the other edges connecting its two vertices are deleted is

$$\frac{1}{k}\left(1-\frac{1}{k}\right)^{k'-1} \geq \frac{1}{3k}$$

(since $(1-1/k)^{k-1} \geq \frac{1}{3}$). We get $\mathbf{E}[m'] \geq |E|/3k$ and $\mathbf{E}[x'] \leq x/k^2$. Applying the crossing number theorem for the graph G' and taking expectations, we have

$$\mathbf{E}[x'] \geq \frac{1}{64} \cdot \frac{\mathbf{E}[m'^3]}{n^2} - n.$$

By convexity (Jensen's inequality), we have $\mathbf{E}[m'^3] \geq (\mathbf{E}[m'])^3 = \Omega(|E|^3/k^3)$. Plugging this plus the bound $\mathbf{E}[x'] \leq x/k^2$ into the above formula, we get

$$\frac{x}{k^2} = \Omega\left(\frac{|E|^3}{k^3 n^2}\right) - O(n),$$

and the lemma follows. □

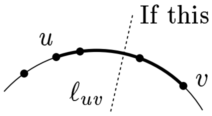
Bibliography and remarks. The proof presented above is, with minor modifications, that of Székely [Szé97]. The bound has subsequently been improved by Solymosi and Tóth [ST01] to $\Omega(n^{6/7})$ and then by Tardos [Tar01] to (approximately) $\Omega(n^{0.863})$.

The weakest point of the proof shown above seems to be the lower bound on the number of incidences between the points of P and the "rich" bisectors ℓ_{uv} ($\{u, v\}$ being the pairs connected by k or more edges). We counted as if each such incidence could be responsible for as many as t edges. While this does not look geometrically very plausible, it seems hard to exclude such a possibility directly. Instead, Solymosi and Tóth prove a better lower bound for the number of incidences of P with the rich bisectors differently; they show that if there are many edges with multiplicity at least k, then each of $\Omega(n)$ suitable points is incident to many (namely $\Omega(n/t^{3/2})$ in their proof) rich bisectors. We outline this argument.

We need to modify the definition of the graph G. The new definition uses an auxiliary parameter r (a constant, with $r = 3$ in the original Solymosi–Tóth proof). First, we note that by the theorem of Beck mentioned in Exercise 4.1.7, there is a subset $P' \subseteq P$ of $\Omega(n)$ points such that each $p \in P'$ sees the other points of P in $\Omega(n)$ distinct directions. For each $p \in P'$, we draw the t circles around p. If several points of P are visible from p in the same direction, we temporarily delete all of them but one. Then, on each circle, we group the remaining points into groups by r consecutive points, and on each circle we delete the at most $r-1$ leftover points fitting in no such group. This still leaves $\Omega(n)$ r-point groups on the circles centered at p.

Next, we consider one such r-point group and all the $\binom{r}{2}$ bisecting lines of its points. If at least one of these bisectors, call it ℓ_{uv}, contains fewer than k points of P (k being a suitable threshold), then we add the arc connecting u and v as an edge of G:

If this bisector has at most k points of P,



then the arc $\{u, v\}$ is added to G.

(This is not quite in agreement with our definition of a graph drawing, since the arc may pass though other vertices of G, but it is easy to check that if we permit arcs through vertices and modify the definition of the crossing number appropriately, Lemma 4.4.2 remains valid.) The groups where every bisector contains at least k points of P (call them *rich* groups) do not contribute any edges of G.

Setting $k = \alpha n^2/t^2$ for a small constant α, we argue by Lemma 4.4.2 that G has at most βn^2 edges for a small $\beta = \beta(\alpha) > 0$. It follows that most of the r-point groups must be rich, and so there is a subset $P'' \subseteq P'$ of $\Omega(n)$ points, each of them possessing $\Omega(n)$ rich groups on its circles. It remains to prove that each point $p \in P''$ is incident to many rich bisectors. We divide the plane around p into angular sectors such that each sector contains about $3rt$ points (of the $\Omega(n)$ points in the rich groups belonging to p). Each sector contains at least t complete rich groups (since there are t circles, and so the sector's boundaries cut through at most $2t$ groups), and we claim that it has to contain many rich bisectors. This leads to the following number-theoretic problem: we have tr distinct real numbers (corresponding to the angles of the points in the sector as seen from p), arranged into t groups by r numbers, and we form all the $\binom{r}{2}$ arithmetic averages of the pairs in each group (corresponding to the rich bisectors of the group). This yields $t\binom{r}{2}$ real numbers, and we want to know how many of them must be distinct.

It is not hard to see that for $r = 3$, there must be at least $\Omega(t^{1/3})$ distinct numbers, because the three averages $(a+b)/2$, $(a+c)/2$, and $(b+c)/2$ determine the numbers a, b, c uniquely. It follows, still for $r = 3$, that each of the $\Omega(\frac{n}{t})$ sectors has $\Omega(t^{1/3})$ distinct bisectors, and so each point in P'' has $\Omega(n/t^{2/3})$ incidences with the rich lines. Applying Szemerédi–Trotter now yields the Solymosi–Tóth bound of $t = \Omega(n^{6/7})$ distinct distances.

Tardos [Tar01] considered the number-theoretic problem above for larger r, and he proved, by a complicated argument, that for r large but fixed, the number of distinct pairwise averages is $\Omega(t^{1/e+\varepsilon})$, with $\varepsilon \to 0$ as $r \to \infty$. Plugging this into the proof leads to the current best bound mentioned above. An example by Ruzsa shows that the number of distinct pairwise averages can be $O(\sqrt{t}\,)$ for any fixed r, and it follows that the Solymosi–Tóth method as is cannot provide a bound better than $\Omega(n^{8/9})$. But surely one can look forward to the further continuation of the adventure of distinct distances.

Exercises

1. Let $I_{\rm circ}(m,n)$ be the maximum number of incidences between m points and n arbitrary circles in the plane. Fill in the details of the following approach to bounding $I_{\rm circ}(n,n)$. Let K be a set of n circles, C the set of their centers, and P a set of n points.

   (a) First, assume that the centers of the circles are mutually distinct, i.e., $|C| = |K|$. Proceed as in the proof in the text: Remove circles with at most 2 incidences, and let the others define a drawing of a multigraph G with vertex set P and arcs of the circles as edges. Handle the edges with multiplicity k or larger via Szemerédi–Trotter, using the incidences of the bisectors with the set C, and those with multiplicity $< k$ by Lemma 4.4.2. Balance k suitably. What bound is obtained for the total number of incidences? $\boxed{3}$

   (b) Extend the argument to handle concentric circles too. $\boxed{3}$

2. This exercise provides another bound for $I_{\rm circ}(n,n)$, the maximum possible number of incidences between n arbitrary circles and n points in the plane. Let K be the set of circles and P the set of points. Let P_i be the points with at least $d_i = 2^i$ and fewer than 2^{i+1} incidences; we will argue for each P_i separately.

Define the multigraph G on P_i as usual, with arcs of circles of K connecting neighboring points of P_i (the circles with at most 2 incidences with P_i are deleted). Let E be the set of edges of G. For a point $u \in P_i$, let $N(u)$ be the set of its neighboring points, and for a $v \in N(u)$, let $\mu(u,v)$ be the number of edges connecting u and v. For an edge e, define its *partner edge* as the edge following after e clockwise around its circle.

(a) Show that for each $u \in P_i$, $|\{v \in N(u)\colon \mu(u,v) \geq 4\sqrt{d_i}\}| < \sqrt{d_i}/2$. $\boxed{1}$

(b) Let $E_h \subseteq E$ be the edges of multiplicity at least $4\sqrt{d_i}$. Argue that for at least $\frac{1}{4}$ of the edges in E_h, their partner edges do not belong to E_h, and hence $|E \setminus E_h| = \Omega(|E|)$. $\boxed{3}$

(c) Delete the edges of E_h from the graph, and apply Lemma 4.4.2 to bound $|E \setminus E_h|$. What overall bound does all this give for $I_{\mathrm{circ}}(n,n)$? $\boxed{2}$

A similar proof appears in Pach and Sharir [PS98a] (for the more general case of curves mentioned in the notes to Section 4.1).

4.5 Point–Line Incidences via Cuttings

Here we explain another proof of the upper bound $I(n,n) = O(n^{4/3})$ for point–line incidences. The technique is quite different. It leads to an efficient algorithm and seems more generally applicable than the one with the crossing number theorem.

4.5.1 Lemma (A worse but useful bound).

$$I(m,n) = O(n\sqrt{m} + m), \tag{4.2}$$
$$I(m,n) = O(m\sqrt{n} + n). \tag{4.3}$$

Proof. There are at most $\binom{n}{2}$ crossing pairs of lines in total. On the other hand, a point $p_i \in P$ with d_i incidences "consumes" $\binom{d_i}{2}$ crossing pairs (their intersections all lie at p_i). Therefore, $\sum_{i=1}^{m} \binom{d_i}{2} \leq \binom{n}{2}$.

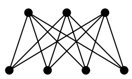
We want to bound $\sum_{i=1}^{m} d_i$ from above. Since points with no incidences can be deleted from P in advance, we may assume $d_i \geq 1$ for all i, and then we have $\binom{d_i}{2} \geq (d_i-1)^2/2$. By the Cauchy–Schwarz inequality,

$$\sum_{i=1}^{m}(d_i-1) \leq \sqrt{m\sum_{i=1}^{m}(d_i-1)^2} \leq \sqrt{2\binom{n}{2}m}\,,$$

and hence $\sum d_i = O(n\sqrt{m} + m)$.

The other inequality in the lemma can be proved similarly by looking at pairs of points on each line. Alternatively, the equality $I(n,m) = I(m,n)$ for all m,n follows using the geometric duality introduced in Section 5.1. $\square$

Forbidden subgraph arguments. For integers $r,s \geq 1$, let $K_{r,s}$ denote the complete bipartite graph on $r+s$ vertices; the picture shows $K_{3,4}$:



The above proof can be expressed using graphs with forbidden $K_{2,2}$ as a subgraph and thus put into the context of extremal graph theory.

A typical question in extremal graph theory is the maximum possible number of edges of a (simple) graph on n vertices that does not contain a given forbidden subgraph, such as $K_{2,2}$. Here the subgraph is understood in a non-induced sense: For example, the complete graph K_4 does contain $K_{2,2}$ as a subgraph. More generally, one can forbid all subgraphs from a finite or infinite family $\mathcal{F}$ of graphs, or consider "containment" relations other than being a subgraph, such as "being a minor."

If the forbidden subgraph H is not bipartite, then, for example, the complete bipartite graph $K_{n,n}$ has $2n$ vertices, n^2 edges, and no subgraph isomorphic to H. This shows that forbidding a nonbipartite H does not reduce the maximum number of edges too significantly, and the order of magnitude remains quadratic.

On the other hand, forbidding $K_{r,s}$ with some fixed r and s decreases the exponent of n, and forbidden bipartite subgraphs are the key to many estimates in incidence problems and elsewhere.

4.5.2 Theorem (Kővári–Sós–Turán theorem). *Let $r \leq s$ be fixed natural numbers. Then any graph on n vertices containing no $K_{r,s}$ as a subgraph has at most $O(n^{2-1/r})$ edges.*

If G is a bipartite graph with color classes of sizes m and n containing no subgraph $K_{r,s}$ with the r vertices in the class of size m and the s vertices in the class of size n, then

$$|E(G)| = O\left(\min(mn^{1-1/r} + n, m^{1-1/s}n + m)\right).$$

(In both parts, the constant of proportionality depends on r and s.)

Note that in the second part of the theorem, the situation is not symmetric: By forbidding the "reverse" placement of $K_{r,s}$, we get a different bound in general.

The upper bound in the theorem is suspected to be tight, but a matching lower bound is known only for some special values of r and s, in particular for $r \leq 3$ (and all $s \geq r$).

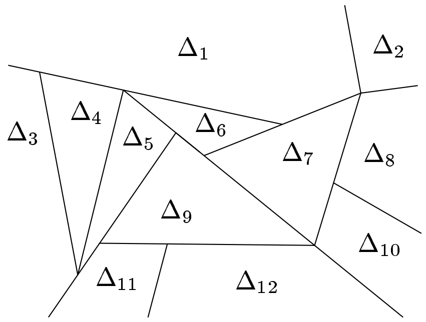
To see the relevance of forbidden $K_{2,2}$ to the point–line incidences, we consider a set P of points and a set L of lines and we define a bipartite graph with vertex set $P \cup L$ and with edges corresponding to incidences. An edge $\{p, \ell\}$ means that the point p lies on the line ℓ. So the number of incidences equals the number of edges. Since two points determine a line, this graph contains no $K_{2,2}$ as a subgraph: Its presence would mean that two distinct lines both contain the same two distinct points. The Kővári–Sós–Turán theorem thus immediately implies Lemma 4.5.1, and the above proof of this lemma is the usual proof of that theorem, for the special case $r = s = 2$, rephrased in terms of points and lines.

As was noted above, for arbitrary bipartite graphs with forbidden $K_{2,2}$, not necessarily being incidence graphs of points and lines in the plane, the bound in the Kővári–Sós–Turán theorem cannot be improved. So, in order to do better for point–line incidences, one has to use some more geometry than just the excluded $K_{2,2}$. In fact, this was one of the motivations of the problem of point–line incidences: In a finite projective plane of order q, we have $n = q^2+q+1$ points, n lines, and $(q+1)n \approx n^{3/2}$ incidences, and so the Szemerédi–Trotter theorem strongly distinguishes the Euclidean plane from finite projective planes in a combinatorial sense.

Proof of the Szemerédi–Trotter theorem (Theorem 4.1.1) for $m = n$. The bound from Lemma 4.5.1 is weaker than the tight Szemerédi–Trotter bound, but it is tight if $n^2 \le m$ or $m^2 \le n$. The idea of the present proof is to convert the "balanced" case (n points and n lines) into a collection of "unbalanced" subproblems, for which Lemma 4.5.1 is optimal. We apply the following important result:

4.5.3 Lemma (Cutting lemma). *Let L be a set of n lines in the plane, and let r be a parameter, $1 < r < n$. Then the plane can be subdivided into t* generalized triangles *(this means intersections of three half-planes) $\Delta_1, \Delta_2, \ldots, \Delta_t$ in such a way that the interior of each Δ_i is intersected by at most $\frac{n}{r}$ lines of L, and we have $t \le Cr^2$ for a certain constant C independent of n and r.*

Such a collection $\Delta_1, \ldots, \Delta_t$ may look like this, for example:



The lines of L are not shown.

In order to express ourselves more economically, we introduce the following terminology. A *cutting* is a subdivision of the plane into finitely many generalized triangles. (We sometimes omit the adjective "generalized" in the sequel.) A given cutting is a $\frac{1}{r}$*-cutting for* a set L of n lines if the interior of each triangle of the cutting is intersected by at most $\frac{n}{r}$ lines of L.

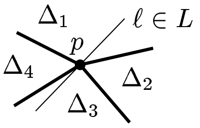
Proofs of the cutting lemma will be discussed later, and now we continue the proof of the Szemerédi–Trotter theorem.

Let P be the considered n-point set, L the set of n lines, and $I(P, L)$ the number of their incidences. We fix a "magic" value $r = n^{1/3}$, and we divide the plane into $t = O(r^2) = O(n^{2/3})$ generalized triangles $\Delta_1, \ldots, \Delta_t$ so that the interior of each Δ_i is intersected by at most $n/r = n^{2/3}$ lines of L, according to the cutting lemma .

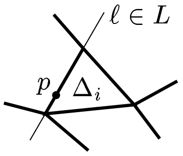
Let P_i denote the points of P lying inside Δ_i or on its boundary but not at the vertices of Δ_i, and let L_i be the set of lines of L intersecting the interior of Δ_i. The pairs (L_i, P_i) define the desired "unbalanced" subproblems. We have

$|L_i| \le n^{2/3}$, and while the size of the P_i may vary, the average $|P_i|$ is $\frac{n}{t} \approx n^{1/3}$, which is about the square root of the size of L_i.

We have to be a little careful, since not all incidences of L and P are necessarily included among the incidences of some L_i and P_i. One exceptional case is a point $p \in P$ not appearing in any of the P_i.



Such a point has to be the vertex of some Δ_i, and so there are no more than $3t$ such exceptional points. These points have at most $I(n, 3t)$ incidences with the lines of L. Another exceptional case is a line of L containing a side of Δ_i but not intersecting its interior and therefore not included in L_i, although it may be incident with some points on the boundary of Δ_i.



There are at most $3t$ such exceptional lines, and they have at most $I(3t, n)$ incidences with the points of P. So we have

$$I(L, P) \le I(n, 3t) + I(3t, n) + \sum_{i=1}^{t} I(L_i, P_i).$$

By Lemma 4.5.1, $I(n, 3t)$ and $I(3t, n)$ are both bounded by $O(t\sqrt{n} + n) = O(n^{7/6}) \ll n^{4/3}$, and it remains to estimate the main term. We have $|L_i| \le n^{2/3}$ and $\sum_{i=1}^{t} |P_i| \le 2n$, since each point of P goes into at most two P_i. Using the bound (4.2) for each $I(L_i, P_i)$ we obtain

$$\begin{aligned}\sum_{i=1}^{t} I(L_i, P_i) &\le \sum_{i=1}^{t} I(n^{2/3}, |P_i|) = \sum_{i=1}^{t} O(|P_i| n^{1/3} + n^{2/3}) \\ &= O(n^{1/3}) \Big(\sum_{i=1}^{t} |P_i| \Big) + O(tn^{2/3}) = O(n^{4/3}).\end{aligned}$$

This finally shows that $I(n, n) = O(n^{4/3})$. □

Bibliography and remarks. The bound in Lemma 4.5.1 using excluded $K_{2,2}$ is due to Erdős [Erd46].

Determining the maximum possible number of edges in a $K_{r,s}$-free bipartite graph with given sizes of the two color classes is known as the *Zarankiewicz problem.* The general upper bound given in the text was shown by Kővári, Sós, and Turán [KST54]. For a long time, matching lower bounds (constructions) were known only for $r \le 3$ and all $s \ge r$ (in these cases, even the constant in the leading term is known exactly; see Füredi [Für96] for some of these results and references). In particular, $K_{2,2}$-free graphs on n vertices with $\Omega(n^{3/2})$ edges are provided by incidence graphs of finite projective planes, and $K_{3,3}$-free graphs on n vertices with $\Omega(n^{5/3})$ edges were obtained by Brown [Bro66]. His construction is the "distance-k graph" in the 3-dimensional affine space over finite fields of order $q \equiv -1 \bmod 4$, for a suitable $k = k(q)$. Recently, Kollár, Rónyai, and Szabó [KRS96] constructed asymptotically optimal $K_{r,s}$-free graphs for s very large compared to r, namely $s \geq r!+1$, using results of algebraic geometry. This was slightly improved by Alon, Rónyai, and Szabó [ARS99] to $s \geq (r-1)!+1$. They also obtained an alternative to Brown's construction of $K_{3,3}$-free graphs with a better constant, and asymptotically tight lower bounds for some "asymmetric" instances of the Zarankiewicz problem, where one wants a $K_{r,s}$-free bipartite graph with color classes of sizes n and m (with the "orientation" of the forbidden $K_{r,s}$ fixed).

The approach to incidence problems using cuttings first appeared in a seminal paper of Clarkson, Edelsbrunner, Guibas, Sharir, and Welzl [CEG$^+$90], based on probabilistic methods developed in computational geometry ([Cla87], [HW87], and [CS89] are among the most influential papers in this development). Clarkson et al. did not use cuttings in our sense but certain "cuttings on the average." Namely, if n_i is the number of lines intersecting the interior of Δ_i, then their cuttings have $t = O(r^2)$ triangles and satisfy $\sum_{i=1}^{t} n_i^c \leq C(c) \cdot r^2 (\frac{n}{r})^c$, where $c \geq 1$ is an integer constant, which can be selected as needed for each particular application, and $C(c)$ is a constant depending on c. This means that the cth degree average of the n_i is, up to a constant, the same as if all the n_i were $O(\frac{n}{r})$. Technically, these "cuttings on the average" can replace the optimal $\frac{1}{r}$-cuttings in most applications. Clarkson et al. [CEG$^+$90] proved numerous results on various incidence problems and many-cells problems by this method; see the notes to Section 4.1.

The cutting lemma was first proved by Chazelle and Friedman [CF90] and, independently, by Matoušek [Mat90a]. The former proof yields an optimal cutting lemma in every fixed dimension and will be discussed in Section 6.5, while the latter proof applies only to planar cutting and is presented in Section 4.7. A third, substantially different, proof was discovered by Chazelle [Cha93a].

Yet another proof of the Szemerédi–Trotter theorem was recently found by Aronov and Sharir (it is a simplification of the techniques in [AS01a]). It is based on the case $d = 2$ of the following partition theorem of Matoušek [Mat92]: *For every n-point set $X \subset \mathbf{R}^d$, d fixed, and every r, $1 \leq r \leq n$, there exists a partition $X = X_1 \cup X_2 \cup \cdots \cup X_t$, $t = O(r)$, such that $\frac{n}{r} \leq |X_i| \leq \frac{2n}{r}$ for all i and no hyperplane h crosses more than $O(r^{1-1/d})$ of the sets X_i.* Here h crossing X_i means that X_i is not completely contained in one of the open half-spaces defined by h or in h itself.[1] This result is proved using the d-dimensional cutting lemma (see Section 4.6). The bound $O(r^{1-1/d})$ is asymptotically the best possible in general.

To use this result for bounding $I(L, P)$, where L is a set of n lines and P a set of n points in the plane, we let $X = \mathcal{D}_0(L)$ be the set of points dual to the lines of L (see Section 5.1). We apply the partition theorem to X with $r = n^{2/3}$ and dualize back, which yields a partition $L = L_1 \cup L_2 \cup \cdots \cup L_t$, $t = O(r)$, with $|L_i| \approx \frac{n}{r} = n^{1/3}$. The crossing condition implies that no point p is incident to lines from more than $O(\sqrt{r})$ of the L_i, not counting the pathological L_i where p is common to all the lines of L_i.

We consider the incidences of a point $p \in P$ with the lines of L_i. The i where p lies on at most one line of L_i contribute at most $O(\sqrt{r})$ incidences, which gives a total of $O(n\sqrt{r}) = O(n^{4/3})$ for all $p \in P$. On the other hand, if p lies on at least two lines of L_i then it is a vertex of the arrangement of L_i. As is easy to show, the number of incidences of k lines with the vertices of their arrangement is $O(k^2)$

[1] A slightly stronger result is proved in [Mat92]: For every X_i we can choose a relatively open simplex $\sigma_i \supseteq X_i$, and no h crosses more than $O(r^{1-1/d})$ of the σ_i.

(Exercise 6.1.6), and so the total contribution from these cases is $O(\sum |L_i|^2) = O(n^2/r) = O(n^{4/3})$. This proves the balanced case of Szemerédi–Trotter, and the unbalanced case works in the same way with an appropriate choice of r. Unlike the previous proofs, this one does not directly apply with pseudolines instead of lines.

Improved point–circle incidences. A similar method also proves that $I_{\rm circ}(n,n) = O(n^{1.4})$ (see Exercise 4.4.2 for another proof). Circles are dualized to points and points to surfaces of cones in $\mathbf{R}^3$, and the appropriate partition theorem holds as well, with no surface of a cone crossing more than $O(r^{2/3})$ of the subsets X_i.

Aronov and Sharir [AS01a] improved the bound to $I_{\rm circ}(m,n) = O(m^{2/3}n^{2/3} + m)$ for large m, namely $m \geq n^{(5-3\varepsilon)/(4-9\varepsilon)}$, and to $I_{\rm circ}(m,n) = O(m^{(6+3\varepsilon)/11}n^{(9-\varepsilon)/11} + n)$ for the smaller m (here, as usual, $\varepsilon > 0$ can be chosen arbitrarily small, influencing the constants of proportionality). Agarwal et al. [AAS01] obtained almost the same bounds for the maximum complexity of m cells in an arrangement of n circles.

A key ingredient in the Aronov–Sharir proof are results on the following question of independent interest. Given a family of n curves in the plane, into how many pieces ("pseudosegments") must we cut them, in the worst case, so that no two pieces intersect more than once? This problem, first studied by Tamaki and Tokuyama [TT98], will be briefly discussed in the notes to Section 11.1. For the curves being circles, Aronov and Sharir [AS01a] obtained the estimate $O(n^{3/2+\varepsilon})$, improving on several previous results.

To bound the number $I(P,C)$ of incidences of an m-point set P and a set C of n circles, we delete the circles containing at most 2 points, we cut the circles into $O(n^{3/2+\varepsilon})$ pieces as above, and we define a graph with vertex set P and with edges being the circle arcs that connect consecutive points along the pieces. The number of edges is at least $I(P,C) - O(n^{3/2+\varepsilon})$. The crossing number theorem applies (since the graph is simple) and yields $I(P,C) = O(m^{2/3}n^{2/3} + n^{3/2+\varepsilon})$, which is tight for m about $n^{5/4}$ and larger. For smaller m, Aronov and Sharir use the method with partition in the dual space outlined above to divide the original problem into smaller subproblems, and for these they use the bound just mentioned.

Exercises

1. Let $I_{\rm 1circ}(m,n)$ be the maximum number of incidences between m points and n unit circles in the plane. Prove that $I_{\rm 1circ}(m,n) = O(m\sqrt{n} + n)$ by the method of Lemma 4.5.1. $\boxed{2}$

2. Let $I_{\rm circ}(m,n)$ be the maximum possible number of incidences between m points and n arbitrary circles in the plane. Prove that $I_{\rm circ}(m,n) = O(n\sqrt{m} + n)$ and $I_{\rm circ}(m,n) = O(mn^{2/3} + n)$. $\boxed{2}$

4.6 A Weaker Cutting Lemma

Here we prove a version of the cutting lemma (Lemma 4.5.3) with a slightly worse bound on the number of the Δ_i. The proof uses the probabilistic method and the argument is very simple and general. We will improve on it later and obtain tight bounds in a more general setting in Section 6.5. In Section 4.7 below we give another, self-contained, elementary geometric proof of the planar cutting lemma .

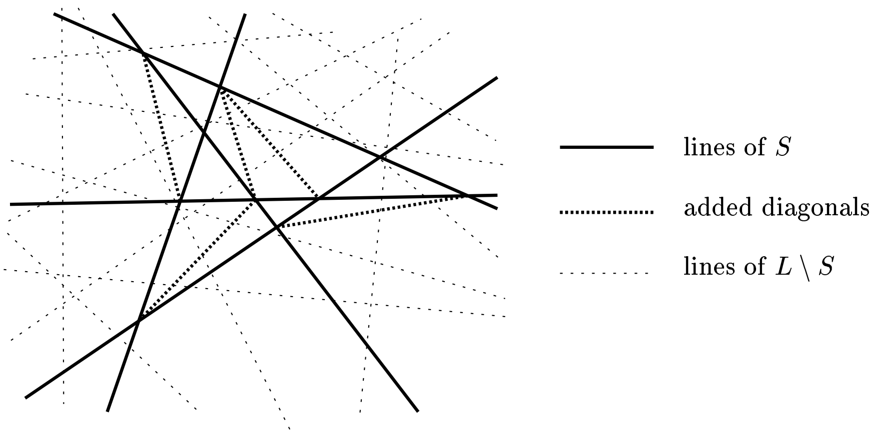
Here we are going to prove that every set of n lines admits a $\frac{1}{r}$-cutting consisting of $O(r^2 \log^2 n)$ triangles. But first let us see why at least $\Omega(r^2)$

triangles are necessary.

A lower bound. Consider n lines in general position. Their arrangement has, as we know, $\binom{n}{2}+n+1 \geq n^2/2$ cells. On the other hand, considering a triangle Δ_i whose interior is intersected by $k \leq \frac{n}{r}$ lines ($k \geq 1$), we see that Δ_i is divided into at most $\binom{k}{2}+k+1 \leq 2k^2$ cells. Since each cell of the arrangement has to show up in the interior of at least one triangle Δ_i, the number of triangles is at least $n^2/4k^2 = \Omega(r^2)$. Hence the cutting lemma is asymptotically optimal for $r \to \infty$.

Proof of a weaker version of the cutting lemma (Lemma 4.5.3). We select a random sample $S \subseteq L$ of the given lines. We make s independent random draws, drawing a random line from L each time. These are draws with replacement: One line can be selected several times, and so S may have fewer than s lines.

Consider the arrangement of S. Partition the cells that are not (generalized) triangles by adding some suitable diagonals, as illustrated below:



This creates (generalized) triangles $\Delta_1, \Delta_2, \ldots, \Delta_t$ with $t = O(s^2)$ (since we have a drawing of a planar graph with $\binom{s}{2}+1$ vertices; also see Exercise 2).

4.6.1 Lemma. *For $s = 6r \ln n$, the following holds with a positive probability: The Δ_i form a $\frac{1}{r}$-cutting for L; that is, the interior of no Δ_i is intersected by more than $\frac{n}{r}$ lines of L.*

This implies the promised weaker version of the cutting lemma: Since the probability of the sample S being good is positive, there exists at least one good S that yields the desired collection of triangles.

Proof of Lemma 4.6.1. Let us say that a triangle T is *dangerous* if its interior is intersected by at least $k = \frac{n}{r}$ lines of L. We fix some arbitrary dangerous triangle T. What is the probability that no line of the sample S intersects the interior of T? We select a random line s times. The probability that we never hit one of the k lines intersecting the interior of T is at most $(1 - k/n)^s$. Using the well-known inequality $1+x \leq e^x$, we can bound this probability by $e^{-ks/n} = e^{-6 \ln n} = n^{-6}$.

Call a triangle T *interesting* (for L) if it can appear in a triangulation for some sample $S \subseteq L$. Any interesting triangle has vertices at some three vertices of the arrangement of L, and hence there are fewer than n^6 interesting triangles.[2] Therefore, with a positive probability, a random sample S intersects the interiors of all the dangerous interesting triangles simultaneously. In particular, none of the triangles Δ_i appearing in the triangulation of such a sample S can be dangerous. This proves Lemma 4.6.1. $\square$

More sophisticated probabilistic reasoning shows that it is sufficient to choose $s = \text{const} \cdot r \log r$ in Lemma 4.6.1, instead of $\text{const} \cdot r \log n$, and still, with a positive probability no interesting dangerous triangle is missed by S (see Section 6.5 and also Exercise 10.3.4). This improvement is important for r small, say constant: it shows that the number of triangles in a $\frac{1}{r}$-cutting can be bounded independent of n.

To prove the asymptotically tight bound $O(r^2)$ by a random sampling argument seems considerably more complicated and we will discuss this in Section 6.5.

Bibliography and remarks. The ideas in the above proof of the weaker cutting lemma can be traced back at least to early papers of Clarkson (such as [Cla87]) on random sampling in computational geometry. The presented proof was constructed ex post facto for didactic purposes; the cutting lemma was first proved, as far as I know, in a stronger form (with $\log r$ instead of $\log n$).

Exercises

1. Calculate the exact expected size of S, a sample drawn from n elements by s independent random draws with replacements. $\boxed{3}$

2. Calculate the number of (generalized) triangles arising by triangulating an arrangement of n lines in the plane in general position. (First, specify how exactly the unbounded cells are triangulated.) $\boxed{2}$

3. (A cutting lemma for circles) Consider a set K of n circles in the plane. Select a sample $S \subseteq K$ by s independent random draws with replacement. Consider the arrangement of S, and construct its *vertical decomposition*; that is, from each vertex extend vertical segments upwards and downwards until they hit a circle of S (or all the way to infinity). Similarly extend vertical segments from the leftmost and rightmost points of each circle.

   (a) Show that this partitions the plane into $O(s^2)$ "circular trapezoids" (shapes bounded by at most two vertical segments and at most two circular arcs). $\boxed{2}$

   (b) Show that for $s = Cr \ln n$ with a sufficiently large constant C, there is a positive probability that the sample S intersects all the dangerous interesting circular trapezoids, where "dangerous" and "interesting" are defined analogously to the definitions in the proof of the weaker version of the cutting lemma . $\boxed{3}$

4. Using Exercises 3 and 4.5.1, show that the number of unit distances determined by n points in the plane is $O(n^{4/3} \log^{2/3} n)$. $\boxed{3}$

[2] The unbounded triangles have only 1 or 2 vertices, but they are completely determined by their two unbounded rays, and so their number is at most n^2.

5. Using Exercises 3 and 4.5.2, show that $I_{\mathrm{circ}}(n,n) = O(n^{1.4}\log^c n)$ (for some constant c), where $I_{\mathrm{circ}}(m,n)$ is the maximum possible number of incidences between m points and n arbitrary circles in the plane. $\boxed{3}$

4.7 The Cutting Lemma: a Tight Bound

Here we prove the cutting lemma in full strength. The proof is simple and elementary, but it does not seem to generalize to higher-dimensional situations.

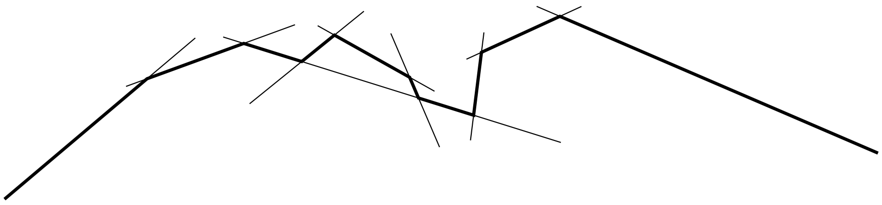
For simplicity, we suppose that the given set L of n lines is in general position. (If not, perturb it slightly to get general position, construct the $\frac{1}{r}$-cutting, and perturb back; this gives a $\frac{1}{r}$-cutting for the original collection of lines; we omit the details.) First we need some definitions and observations concerning levels.

Levels and their simplifications. Let L be a fixed finite set of lines in the plane; we assume that no line of L is vertical. The *level* of a point $x \in \mathbf{R}^2$ is defined as the number of lines of L lying strictly below x.

We note that the level of all points of an (open) cell of the arrangement of L is the same, and similarly for a (relatively open) edge. On the other hand, the level of an edge can differ from the levels of its endpoints, for example.

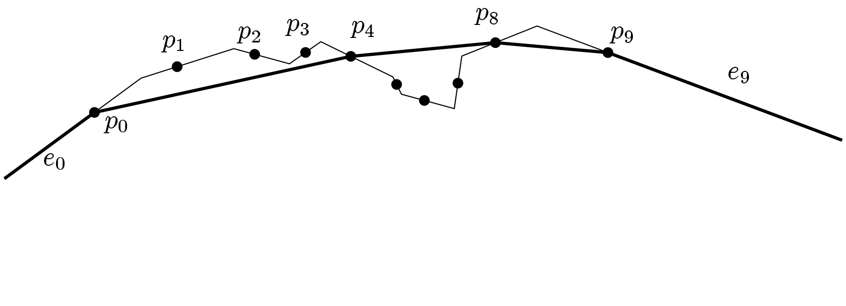
We define the *level* k of the arrangement of L, where $0 \le k < n$, as the set E_k of all edges of the arrangement of L having level exactly k. These edges plus their endpoints form an x-monotone polygonal line, where x-monotone means that each vertical line intersects it at exactly one point.

It is easy to see that the level k makes a turn at each endpoint of its edges; it looks somewhat like this:



The level k is drawn thick, while the thin segments are pieces of lines of L and they do not belong to the level k.

Let $e_0, e_1, \ldots, e_t$ be the edges of E_k numbered from left to right; e_0 and e_t are the unbounded rays. Let us fix a point p_i in the interior of e_i. For an integer parameter $q \ge 2$, we define the *q-simplification of the level k* as the monotone polygonal line containing the left part of e_0 up to the point p_0, the segments p_0p_q, $p_qp_{2q}, \ldots$, $p_{\lfloor (t-1)/q\rfloor q}p_t$, and the part of e_t to the right of p_t. Thus, the q-simplification has at most $\frac{t}{q}+2$ edges. Here is an illustration for $t = 9$, $q = 4$:



(We could have defined the q-simplification by connecting every qth vertex of the level, but the present way makes some future considerations a little simpler.)

4.7.1 Lemma.

(i) *The portion* Π *of the level* k *(considered as a polygonal line) between the points* p_j *and* p_{j+q} *is intersected by at most* $q+1$ *lines of* L.

(ii) *The segment* p_jp_{j+q} *is intersected by at most* $q+1$ *lines of* L.

(iii) *The* q*-simplification of the level* k *is contained in the strip between the levels* $k - \lceil q/2 \rceil$ *and* $k + \lceil q/2 \rceil$.

Proof. Part (i) is obvious: Each line of L intersecting Π contains one of the edges $e_j, e_{j+1}, \ldots, e_{j+q}$. As for (ii), Π is connected, and hence all lines intersecting its convex hull must intersect Π itself as well. The segment p_jp_{j+q} is contained in $\mathrm{conv}(\Pi)$.

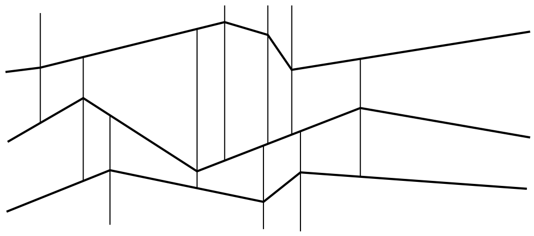
Concerning (iii), imagine walking along some segment p_jp_{j+q} of the q-simplification. We start at an endpoint, which has level k. Our current level may change only as we cross lines of L. Moreover, having traversed the whole segment we must be back to level k. Thus, to get from level k to $k+i$ and back to k we need to cross at least $2i$ lines on the way. From this and (ii), $2i \le q+1$, and hence $i \le \lfloor (q+1)/2 \rfloor = \lceil q/2 \rceil$. □

Proof of the cutting lemma for lines in general position. Let r be the given parameter. If $r = \Omega(n)$, then it suffices to produce a 0-cutting of size $O(n^2)$ by simply triangulating the arrangement of L. Hence we may assume that r is much smaller than n.

Set $q = \lceil n/10r \rceil$. Divide the levels $E_0, E_1, \ldots, E_{n-1}$ into q groups: The ith group contains all E_j with j congruent to i modulo q ($i = 0, 1, \ldots, q-1$). Since the total number of edges in the arrangement is n^2, there is an i such that the ith group contains at most n^2/q edges. We fix one such i; from now on, we consider only the levels $i, q+i, 2q+i, \ldots$, and we construct the desired $\frac{1}{r}$-cutting from them.

Let P_j be the q-simplification of the level $jq+i$. If E_{jq+i} has m_j edges, then P_j has at most $m_j/q + 3$ edges, and the total number of edges of the P_j, $j = 0, 1, \ldots, \lfloor (n-1)/q \rfloor$, can be estimated by $n^2/q^2 + 3(n/q+1) = O(n^2/q^2)$. We note that the polygonal chains P_j never intersect properly: If they did, a vertex of some P_j, which has level $qj+i$, would be above P_{j+1}, and this is ruled out by Lemma 4.7.1(iii).

We form the vertical decomposition for the P_j; that is, we extend vertical segments from each vertex of P_j upwards and downwards until they hit P_{j-1} and P_{j+1}:



This subdivides the plane into $O(n^2/q^2) = O(r^2)$ trapezoids.

We claim that each such trapezoid is intersected by at most $\frac{n}{r}$ lines of L. We look at a trapezoid in the strip between P_j and P_{j+1}. By Lemma 4.7.1(iii), it lies between the levels $qj+i-\lceil q/2\rceil$ and $q(j+1)+i+\lceil q/2\rceil$, and therefore, each of its vertical sides is intersected by no more than $3q$ lines. The bottom side is a part of an edge of P_j, and consequently, it is intersected by no more than $q+1$ lines; similarly for the top side. Hence the number of lines intersecting the considered trapezoid is certainly at most $10q \le \frac{n}{r}$. (A more careful analysis shows that one trapezoid is in fact intersected by at most $2q + O(1)$ lines; see Exercise 1.)

Finally, a $\frac{1}{r}$-cutting can be obtained by subdividing each trapezoid into two triangles by a diagonal. But let us remark that for applications of $\frac{1}{r}$-cuttings, trapezoids are usually as good as triangles. □

Bibliography and remarks. The basic ideas of the presented proof are from [Mat90a], and the presentation basically follows [Mat98]. The latter paper provides some estimates for the number of triangles or trapezoids in a $\frac{1}{r}$-cutting, as $r \to \infty$: For example, at least $2.54(1-o(1))r^2$ trapezoids are sometimes necessary, and $8(1+o(1))r^2$ trapezoids always suffice. The notion of levels and their simplifications, as well as Lemma 4.7.1, are due to Edelsbrunner and Welzl [EW86].

Exercises

1. (a) Verify that each trapezoid arising in the described construction is intersected by at most $2.5q + O(1)$ lines. Setting q appropriately, show that the plane can subdivided into $12.5r^2 + O(r)$ trapezoids, each intersected by at most $\frac{n}{r}$ lines, assuming $1 \ll r \ll n$. $\boxed{2}$

   (b) Improve the bounds from (a) to $2q + O(1)$ and $8r^2 + O(r)$, respectively. $\boxed{4}$