A Combinatorial Distinction between Unit Circles and Straight Lines: How Many Coincidences can they Have?

György Elekes,* Miklós Simonovits† and Endre Szabó‡
Mathematical Institute of Eötvös University, Hungary
and Alfréd Rényi Mathematical Institute, Hungary
elekes@cs.elte.hu, miki@renyi.hu and endre@renyi.hu

June 26, 2009

Abstract

We give a very general sufficient condition for a one–parameter family of curves *not to have* n members with "too many" (i.e. a near–quadratic number of) triple points of intersections. As a special case, a combinatorial distinction between straight lines and unit circles will be shown. (Actually, this is more than just a simple application; originally this motivated our results.)

1. Introduction

The (very) general problem

Let Γ be a family of continuous curves in $\mathbb{R}^d$. We pick a set of n curves $\mathcal{G} = \{\gamma_1, \ldots, \gamma_n\} \subset \Gamma$ and a set of m points $\mathcal{P} = \{P_1, \ldots, P_m\} \in \mathbb{R}^d$ and define a graph on $\mathcal{G} \cup \mathcal{P}$ by connecting γ_i to P_j if γ_i passes through P_j. We shall call this (bipartite) graph the *incidence graph* of $\mathcal{G}$ and $\mathcal{P}$.

Certain properties of such graphs, especially the maximum possible number of edges as a function of n and m (i.e. bounds on the number of incidences) play central role in Computational Geometry as well as in Discrete or Combinatorial Geometry.

In this paper we study a "reverse" question:

> *if we know only the incidence graph (or some of its properties), can we infer something about the properties of the family* Γ*?*

Apart from trivial observations like "if two curves share two common points then Γ cannot be the family of straight lines", very little is known. (Actually, [5] contains a result that points to this direction, see Theorem 2 below.)

*Research partially supported by OTKA grants T 42750 and T 47056
†Research partially supported by OTKA grants T 69062, T 62321, T 48826
‡Research partially supported by OTKA grants K61116 and NK 72523

Many triple points

In terms of incidence graphs, a point P_j is a *triple point* if it is connected to at least three of the n curves in $\mathcal{G}$. Since three general curves do not pass through a common point, triple points can be considered as interesting coincidences.

Given a family Γ and a positive integer $n \in \mathbb{N}^+$, we select n curves $\gamma_1, \dots, \gamma_n \in \Gamma$ so that the number of triple points is maximized, and denote this maximum by $\mathcal{T}_\Gamma(n)$. More generally, for three (not necessarily distinct) families $\Gamma_1, \Gamma_2, \Gamma_3$, we select n curves from each Γ_i $(i = 1, 2, 3)$ and call a point P a triple point if, for $i = 1, 2, 3$, there exist distinct $\gamma_i \in \Gamma_i$ that pass through P. (Usual bipartite graphs cannot represent such structures; certain "four–partite" graphs can, but we do not need them.) We denote the maximum number of such triple points by $\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n)$, taken over all possible selections of the $n + n + n$ curves. We must emphasize that, even in this general case, we require that a triple point be the intersection of three *distinct* curves.

If any two curves intersect in at most B points (where B is a constant while n is large) then the maxima defined above really exist; in particular

$$\mathcal{T}_\Gamma(n) \le B\binom{n}{2} \qquad \text{and} \qquad \mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n) \le Bn^2,$$

since already the number of *pairwise* intersections in Γ (or between, say, Γ_1 and Γ_2) cannot exceed the claimed bound.

If no such B exists then no bound can be found for the $\mathcal{T}$ (e.g., if, for $i = 1 \dots 3$, Γ_i consists of the graphs of $y = i \cdot \sin x + t$, for $t \in \mathbb{R}$).[1] That is why, in what follows, we shall always assume the existence of such a B, i.e. that

$$\textit{no two curves intersect in more than } B \textit{ points.} \tag{1}$$

On the other hand, the number of "double" points can really attain this quadratic order of magnitude if the curves we select are in "sufficiently general position", e.g., if any two share a common point and these points are all distinct. This observation indicates that the "magic multiplicity" 3 is the smallest interesting value. In some cases even the number of triple points can be of order cn^2, e.g., for straight lines like those in Figure 4(c). However, as we shall see, in many cases the number of triple points is only $O(n^{2-\eta})$ for some constant $\eta \in (0, 1)$.

Problem 1. Characterize those families Γ, or triples of families $\Gamma_1, \Gamma_2, \Gamma_3$, for which $\mathcal{T}_\Gamma(n)$ or $\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n)$, respectively, attains a quadratic order of magnitude (i.e. at least cn^2, for a fixed $c > 0$ and infinitely many n).

If the function $\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n)$ for certain families $\Gamma_1, \Gamma_2, \Gamma_3$ attain a quadratic order of magnitude, a simple way to prove this is to exhibit n (or $n + n + n$) curves — for all $n \in \mathbb{N}$ — that have this many triple points.

The converse is harder: if a quadratic order of magnitude is impossible, how to demonstrate this? That is why our main result Theorem 14 concerns a sufficient condition for not having many triple points.

[1] It is perhaps unfortunate but we use the word "graph" in two completely different ways: until this point it was used to represent/emphasize the incidences of geometric curves. From now on graph theory is forgotten and the graph means the graph of a function.

The main result at a "philosophical" level

Roughly speaking, we show the following (all notions will be defined rigorously, including "envelopes").

> *using suitable (slightly different from usual) definitions of "parameterized families" and "envelopes", if one of three algebraically parameterized families has an envelope which is not an envelope for any of the other two families, then*
>
> $$\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n) = \mathcal{O}\big(n^{2-\eta}\big),$$
>
> *for a positive $\eta > 0$ that depends only on the degree of the families.*

Since we do not want to spoil the Introduction with a lot of technical details, we must, for the time being, postpone the exact formulation of our main result; see Theorem 14 for a precise statement.

On our results and proof methods

While writing this paper, the authors had to make two important decisions.

On the one hand, we had to choose between an analytic and an algebraic (or, rather, an algebro–geometric) approach. To make our results accessible for a wider audience, we chose the analytic point of view.

On the other hand, we decided to present a slightly restricted result (i.e. one with slightly more technical assumptions than necessary). This allows for a not-too-long proof but, at the same time, it is still sufficiently general for applications to other problems of Combinatorial Geometry. We are planning to publish another paper where we state our Main Theorem 14 in a more general form — with a more involved proof, of course.

Earlier results for straight lines

Studying the incidence structures of points and straight lines (more generally, of points and certain curves) has been one of the fundamental tasks of Combinatorial Geometry for a long time.

About 140 years ago Sylvester [12] posed his famous "Orchard Problem" which, in an equivalent (dual) form, asks for an arrangement of n straight lines in the Euclidean plane so that the number of triple points be maximized. Sylvester showed that if $\mathcal{L}$ denotes the family of all straight lines, then $\mathcal{T}_{\mathcal{L}}(n) = n^2/6 + \mathcal{O}(n)$ (cf. [7]).

The study of general "k–orchards" for $k \geq 4$ was initiated by Erdős. [2]

One of his conjectures resulted in a beautiful and widely applicable upper bound proven by Szemerédi and Trotter [14]. The most interesting special case of this bound asserts that

> *the number of incidences between n points and n straight lines in the Euclidean plane is at most $Cn^{4/3}$, for some absolute constant C.*

[2] The "k–orchard" problem asks: Given n points in the plane, how many straight lines can contain k points of them if no r of them are on a straight line ($r > k$). See [1], p315.

Since then, various proof techniques have been found, some of them even extending the Szemerédi–Trotter bound to "pseudo–lines" (i.e. curves with the property that any two intersect in at most one point) and "families with two degrees of freedom" (i.e. through any two given points there pass at most a bounded number of curves), see [11], [10], [13], and also the excellent monographs [8], [9].

Earlier results on unit circles

Another "orchard–like" problem was posed by Erdős in [6]: arrange n unit circles in the Euclidean plane so that the number of triple points be maximized. Denoting the family of all unit circles by $\mathcal{U}$, an upper bound of $\mathcal{T}_{\mathcal{U}}(n) \leq n(n-1)$ is obvious (since, as before, already the number of *pairwise* intersections obeys this bound). A lower bound of $\mathcal{T}_{\mathcal{U}}(n) \geq cn^{3/2}$ was proved in [2]. The gap between these two estimates is still wide open.

Also from another point of view, unit circles play a special role in Combinatorial Geometry. One of the most challenging unproved conjectures of Erdős concerns the maximum possible number of unit distances between n points in $\mathbb{R}^2$, and this can be bounded from above by half the number of incidences between the n points and n unit circles around them.

Since such circles obviously form a family with two degrees of freedom, they obey the aforementioned Szemerédi–Trotter bound — and it readily implies the best currently known upper bound on the number of unit distances [11].

The Szemerédi–Trotter bound is known to give the best order of magnitude for point-and-straight-line configurations, which is not the case for points and unit circles (let alone more general families with two degrees of freedom). Actually, it is widely believed that for unit circles and points much better upper bounds hold on the number of incidences. Thus, according to the famous Erdős conjecture on unit distances, n points and n unit circles cannot have more than $n^{1+\varepsilon}$ incidences, for any $\varepsilon > 0$ and $n > n_0(\varepsilon)$.

However, to the best of our knowledge, no such bound has been found so far, since all existing methods consider the set of unit circles just as a family with two degrees of freedom. That is why the known tools cannot distinguish them from straight lines — for which the bound cannot be improved.

As an application of our Main Theorem 14, we show a combinatorial distinction between families of straight lines and families of unit circles in Section 5.

An outline of what is coming

Assume we have an algebraically parameterized family $\Gamma = \{\gamma^{(t)} \ : \ t \in T\}$ of curves, i.e. there is a polynomial $p \in \mathbb{R}[x,y,t]$ or $p \in \mathbb{C}[x,y,t]$ such that $\gamma^{(t)} = \{(x,y) \ : \ p(x,y,t) = 0\}$, for all t in the parameter domain T. Here we do not care whether the points of the individual curves are parameterized somehow; rather, *curves* are assigned to each parameter $t \in T$.

If three such curves, say $\gamma^{(t_1)}$, $\gamma^{(t_2)}$, $\gamma^{(t_3)}$ pass through a common point (x,y), then three equations $p(x,y,t_i) = 0$ are satisfied. Eliminating x and y we get another polynomial equation

$$F(t_1,t_2,t_3) = 0. \tag{2}$$

It was shown in [5] that, if some n elements of Γ determine $> cn^2$ triple points, then the surface $S_F := \{F = 0\}$ must be very special: *there exist three independent univariate coordinate transforms on the three axes which, together, transform S_F into a plane* — unless S_F is a cylinder. The details are given in the forthcoming Surface Theorem 2.

Unfortunately, that theorem does not provide a "good characterization" in the sense that it only states the equivalence of *existence* assumptions. (A "really good" and efficient tool would be one that says: "structure A exists if and only if structure B does not"; this would allow for an easy proof of "A does not exist" by simply exhibiting a B.)

Fortunately, a good characterization was also found in [5]: if we express, say, parameter t_3 from equation (2) then the implicit function $t_3(t_1, t_2)$ must satisfy a partial differential equation of order three. Theoretically this allows for proving subquadratic upper bounds on $\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n)$ via elementary calculations, by showing that the differential equation is not satisfied.

In practice, however, even in simple, natural cases, these calculations may be impossible to carry out, even for powerful computers (see Section 5).

Our Main Theorem 14 becomes useful under such circumstances: it allows for similar bounds, based upon simple geometric considerations.

In Section 2 we present one of the most important tools for the proof of our Main Theorem: the Theorem 2, also called "Surface Theorem", proven in [5].

In order to prepare for the proof of our main result, we define partial envelopes and present some of their properties in Section 3. The main proof itself comes in Section 4.

In Sections 5–6 we state and prove our motivating Theorem 18: a combinatorial distinction between unit circles and straight lines.

Finally, we make some concluding remarks and formulate some conjectures.

2. Special surfaces

The first main ingredient of our proof is Theorem 2 below, proven in [5].

Assume we consider a plane $\alpha x+\beta y+\gamma z = \delta$, intersecting the cube $[0, n]^3$. If the coefficients $\alpha, \beta, \gamma, \delta$ are rationals with small numerators and denominators then this plane will contain $\sim n^2$ lattice points. If we apply independent univariate transformations in the three coordinates, x, y, z, then we can easily produce 2-dimensional surfaces — described by some equation $f(x) + g(y) + h(z) = \delta$ — containing a quadratic number of points from a product set $X \times Y \times Z$, where $|X| = |Y| = |Z| = n$. The main result of [5] asserts that if some appropriate algebraicity conditions hold, then (apart from being a cylinder) this is the only way for a surface $F(x, y, z) = 0$ to contain a near–quadratic number of points from such a product set $X \times Y \times Z$.

As usual, we call a (real or complex) function in one or two variable(s) *analytic* at a point if it can be expressed as a convergent power series in a neighborhood. Also, it is analytic on an open set if it is analytic at each point of the open set.

A *cylinder* over a curve $f(x, y) = 0$ is the surface

$$S := \{(x, y, z) \in \mathbb{C}^3 \ : \ f(x, y) = 0, \ z \in \mathbb{C}\}.$$

The definitions of cylinders over $g(x,z)=0$ or $h(y,z)=0$ are similar. It is worth noting that such cylinders always contain n^2 points of suitable $(\leq n)\times(\leq n)\times(\leq n)$ Cartesian products. To see this, just pick n arbitrary points on the curve $f(x,y)=0$ and n arbitrary values $z_1, z_2, \ldots, z_n \in \mathbb{C}$. Denote the x and y coordinates of the points by X and Y, respectively, and let $Z := \{z_1, z_2, \ldots, z_n\}$. Then $|X|$, $|Y| \leq |Z| = n$ and $X \times Y \times Z$ contains at least n^2 points of S.

Theorem 2 ("Surface Theorem", see [5], Theorem 3.). *For any positive integer d there exist positive constants $\eta = \eta(d) \in (0,1)$ and $n_0 = n_0(d)$ with the following property.*
If $V \subset \mathbb{C}^3$ is an algebraic surface (i.e. each component is two dimensional) of degree $\leq d$ then the following are equivalent:

(a) For at least one $n > n_0(d)$ there exist $X, Y, Z \subset \mathbb{C}$ such that $|X| = |Y| = |Z| = n$ and

$$|V \cap (X \times Y \times Z)| \geq n^{2-\eta};$$

(b) Let $\mathbb{D} \subset \mathbb{C}$ denote the open unit disc. Then either V contains a cylinder over a curve $F(x,y)=0$ or $F(x,z)=0$ or $F(y,z)=0$ or, otherwise, there are one-to-one analytic functions $g_1, g_2, g_3 : \mathbb{D} \to \mathbb{C}$ with analytic inverses such that V contains the $g_1 \times g_2 \times g_3$-image of a part of the plane $x+y+z=0$ near the origin:

$$V \supseteq \left\{ \Big(g_1(x), g_2(y), g_3(z)\Big) \in \mathbb{C}^3 \ : \ x, y, z \in \mathbb{D}, \ x+y+z=0 \right\}.$$

(c) For all positive integers n there exist $X, Y, Z \subset \mathbb{C}$ such that $|X| = |Y| = |Z| = n$ and $|V \cap (X \times Y \times Z)| \geq (n-2)^2/8$.

(d) Both (b) and (c) can be localized in the following sense. There is a finite subset $H \subset \mathbb{C}$ and an irreducible component $V_0 \subseteq V$ such that whenever $P \in V_0$ is a point whose coordinates are not in H and $U \subseteq \mathbb{C}^3$ is any neighborhood of P, then one may require that $\Big(g_1(0), g_2(0), g_3(0)\Big) = P$ in (b), and the Cartesian product $X \times Y \times Z$ in (c) lies entirely inside U. Furthermore, P has a neighborhood U' such that each irreducible component W of the analytic set $V_0 \cap U'$, with appropriate g_1, g_2 and g_3, can be written in the form

$$W = \left\{ \Big(g_1(x), g_2(y), g_3(z)\Big) \in \mathbb{C}^3 \ : \ x, y, z \in \mathbb{D}, \ x+y+z=0 \right\}.$$

If $V \subset \mathbb{R}^3$ then the equivalence of (a), (b), (c) and (d) still holds true with real *analytic functions g_1, g_2, g_3 defined on the interval $(-1,1)$.*

Remark 3. This version of (d) is in fact stronger than the original one in [5], but the proof given there applies without change to the stronger statement.

This result indicates a significant "jump": either V has the special form described in (b), in which case a quadratic order of magnitude is possible, by (b)$\Rightarrow$(c); or else we cannot even exceed $n^{2-\eta}$, by (a)$\Rightarrow$(b).

3. Implicitly vs explicitly parameterized families and their envelopes

Definition 4. Let G be an open domain in $\mathbb{R}^2$ or $\mathbb{C}^2$. A *curve* in the closure $\mathrm{cl}(G)$ is a level set of a continuous function $\mathrm{cl}(G) \to \mathbb{C}$ which is analytic inside G.

Remark 5. We note, that these kind of curves are not necessarily connected, and they may have isolated points. However, this will not cause any trouble.

We consider families Γ of curves in $\mathbb{R}^2$ or $\mathbb{C}^2$, parameterized by the elements of a "parameter space" $T \subset \mathbb{R}$ or $T \subset \mathbb{C}$, like

$$\Gamma = \{\gamma^{(t)} \ : \ t \in T\}. \tag{3}$$

The parameterization is an "implicit analytic parameterization" if there exists a trivariate function f, analytic on an open domain $G \subset \mathbb{R}^3$ or $G \subset \mathbb{C}^3$ and continuous on its closure $\mathrm{cl}(G)$, such that

$$\gamma^{(t)} = \{(u,v) \ : \ f(u,v,t) = 0\}, \quad \text{for all} \quad t \in T.$$

As opposed to implicit ones, we prefer explicit parameterizations.

Definition 6. Γ in (3) is "explicitly analytically parameterized" if there exists a *bivariate* function f, analytic on an open domain $G \subset \mathbb{R}^2$ or $G \subset \mathbb{C}^2$ and continuous on its closure $\mathrm{cl}(G)$, such that

$$\gamma^{(t)} = \{(u,v) \in \mathrm{cl}(G) \ : \ f(u,v) = t\} \quad \text{for all} \quad t \in T.$$

Remark 7. Curves of an implicitly analytically parameterized family can usually be cut into sub–arcs that can be parameterized explicitly — though we do not need this fact.

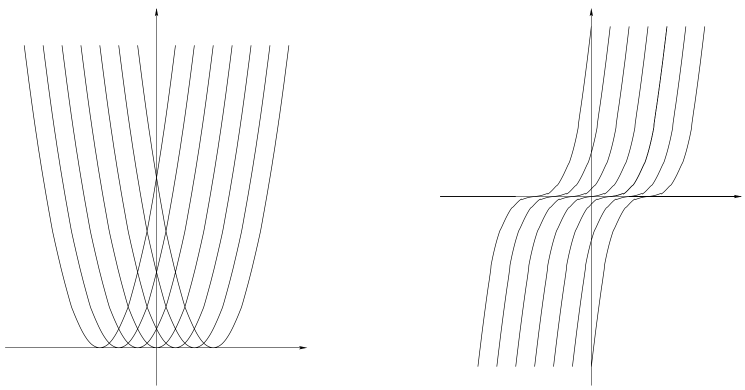

Figure 1: Implicitly analytically parameterized families: **(a)** $y - (x-t)^2 = 0$ and **(b)** $y - (x-t)^3 = 0$.

The parabolas in Figure 1(a) cannot be parameterized explicitly since more than one curve passes through any point above the x–axis. As for the cubics in Figure 1(b), $t = x - \sqrt[3]{y}$ is a *continuous* parameterization but it is not differentiable at any point of the x–axis (and so not analytic either). However, it is an explicit analytic parameterization for suitable closed sub–arcs, say those in Figure 2(b).

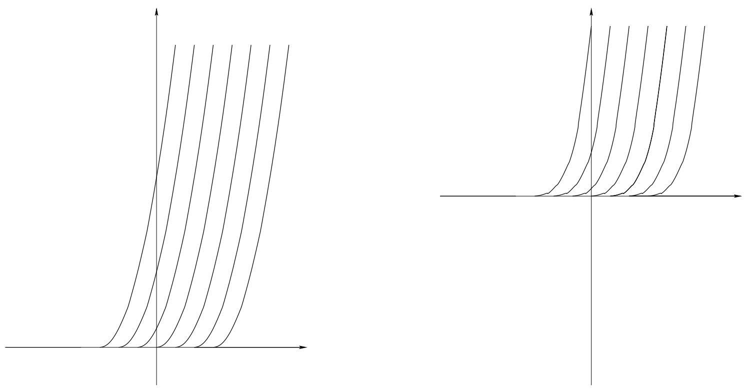

Figure 2: Explicitly analytically parameterized families:
(a) $t = x - \sqrt{y}$ and **(b)** $t = x - \sqrt[3]{y}$.

Envelopes of explicitly parameterized families

Usually in Differential Geometry an envelope of a family Γ of curves is a smooth curve that is tangent to each $\gamma \in \Gamma$. For *explicitly* parameterized families the situation is not that simple. E.g., in Figure 2(a)-(b), the x–axis is not a proper tangent line of the curves; rather, it only is a "half–tangent". Since this is typical in the case of sub–arcs of explicitly parameterized families, we shall use this general definition.

Definition 8. Let G be an open domain in the real or complex plane and let $\gamma \subset \mathrm{cl}(G)$ be a curve. A line L is the *half–tangent* of γ at a point P of the boundary $\mathrm{bd}(G)$ if $P \in \gamma \cap L$, P is not an isolated point of γ, and the following estimate holds:

$$\mathrm{dist}(Q, L) = o\left(\mathrm{dist}(Q, P)\right) \qquad \text{for} \qquad Q \in \gamma \ .$$

Definition 9. Two plane curves *touch* each other at a point P if there exists a straight line through P that is a tangent or half–tangent of both of the curves at P.

Definition 10. A smooth (open or closed) curve $\mathcal{E}$ is a *partial envelope* for an explicitly analytically parameterized family Γ, if

(i) $\mathcal{E}$ is the graph of an analytic real or complex function, say $y = h(x)$ or $x = h(y)$, defined on an open or closed interval or disk, respectively (i.e. $\mathcal{E} = \{(x, y) \ : \ y = h(x)\}$ or $\mathcal{E} = \{(x, y) \ : \ x = h(y)\}$);

(ii) no (non–empty open) sub–arc of $\mathcal{E}$ is contained in any $\gamma^{(t)} \in \Gamma$;

(iii) for each point $P \in \mathcal{E}$, there exists a t for which the curve $\gamma^{(t)} \in \Gamma$ touches $\mathcal{E}$ at P.

The adjective "partial" refers to the fact that we do not require that *each* $\gamma^{(t)} \in \Gamma$ touches $\mathcal{E}$.

Remark 11. (a) As we shall see in Lemma 13(ii), for explicitly analytically parameterized families, $\mathcal{E}$ must be a subset of $\mathrm{bd}(G)$. (Here $\mathcal{E} \subset \mathrm{cl}(G)$ is obvious since $\gamma^{(t)} \subset \mathrm{cl}(G)$ for all $\gamma^{(t)} \in \Gamma$.)

(b) Any non-trivial sub–arc of a partial envelope is a partial envelope;

(c) It is also worth noting that if a real $\mathcal{E}$ is a partial envelope for a family of analytically parameterized real curves then h can be extended to a complex analytic function whose graph defines a partial envelope for the family of the naturally extended, analytically parameterized complex curves.

The technical problems caused by *explicit* parameterization may be tedious but, in general, they are not too difficult to manage.

Example 12. The unit circles through a given point, say the origin, form a family of implicitly analytically parameterized curves. Indeed, if (t, u) is the center of such a circle, then we can eliminate, say, u from the equations

$$(x-t)^2 + (y-u)^2 = 1 = t^2 + u^2, \tag{4}$$

and get a polynomial equation

$$4(x^2+y^2)t^2 - 4x(x^2+y^2)t + (x^2+y^2)^2 - 4y^2 = 0.$$

Moreover, the circle $x^2 + y^2 = 4$ is obviously an envelope for them, in the usual Differential Geometric sense.
In order to get *explicitly* parameterized families, we express, say,

$$t = \frac{x}{2} \pm \frac{y}{2}\sqrt{\frac{4-x^2-y^2}{x^2+y^2}}. \tag{5}$$

(Equivalently, we could express u in a symmetric manner.) Since the right hand side of (5) has no limit at the origin, we exclude a neighborhood of it, of a small radius δ, and consider the open set given by $x^2 + y^2 < 4$, $x^2 + y^2 > \delta^2$, $y < \sqrt{1-(x-1)^2}$ and $x > \sqrt{1-(y+1)^2}$ as G (see the left hand side of Figure 5, where this domain is labelled as G_i^1, and the excluded neighbourhood is labelled as $B_\delta(a_i, b_i)$). Then the appropriate arcs of the unit circles are *explicitly* analytically parameterized on G by (5) with $+$ on the right hand side. We need four rotated copies of the domain G (labelled by $G_i^1, \ldots G_i^4$ on Figure 5) to cover all "right-banding" semi-circles, and we need four more mirrored and rotated copies (labelled by $G_i^5, \ldots G_i^8$ on Figure 5) to cover the "right-banding" semi-circles. Thus the whole family can be decomposed into eight explicitly parameterized (sub)families this way, four of them parameterized by t and four by u.
Moreover, each family has a quarter of the large circle as a partial envelope. (No portion of the small "inner circle" is an envelope since the unit circles do not touch it.)

A lemma on envelopes.

In the proof of the Main Theorem 14, the following statement will play an important role.

Lemma 13. *Let Γ be a family of curves, explicitly analytically parameterized by $f : cl(G) \to \mathbb{C}$ or $\to \mathbb{R}$, as in Definition 6, and let $\mathcal{E}$ be a partial envelope. Then the following hold.*

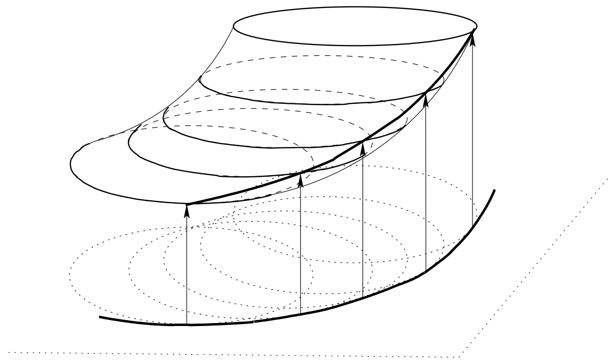

Figure 3: An envelope $\mathcal{E}$ (dashed) and its "lifting" by g on the cylinder over $\mathcal{E}$.

(i) There are no points of $\mathcal{E}$ to which f can be extended analytically;

(ii) Consequently, we have $\mathcal{E} \subset bd(G)$.

Proof. To prove (i), we assume that f can be extended analytically to an open set $\tilde{G}$ which contains G and intersects $\mathcal{E}$. This means, that there is an analytic function $\tilde{f} : \tilde{G} \to \mathbb{C}$ which agrees with f on G. We replace $\mathcal{E}$ with $\tilde{G} \cap \mathcal{E}$, so from now on $\tilde{f}$ is defined and analytic at each point of $\mathcal{E}$. Also, let us define the extended curves $\tilde{\gamma}^{(t)} = \{(u, v) \ : \ t = \tilde{f}(u, v)\}$ for all t.
The function $f(x, y)$, if restricted to $\mathcal{E}$, gives, by definition, the parameter t of the curve $\gamma^{(t)} \in \Gamma$ that touches $\mathcal{E}$ at (x, y). Also by definition, $\mathcal{E}$ is the graph of an analytic function, say $y = h(x)$, on an interval or disk I (the case of $x = h(y)$ is similar). We consider the composition

$$g(x) := f(x, h(x)) : I \to \mathbb{C}.$$

This g is clearly continuous on I; moreover, since we assumed that f can be extended analytically to every $(x, h(x)) \in \mathcal{E}$, it is also differentiable, as an univariate function, in the interior $\mathrm{int}(I)$, by the Chain Rule for the derivative of compositions of type $\mathbb{R} \to \mathbb{R}^2 \to \mathbb{R}$ or $\mathbb{C} \to \mathbb{C}^2 \to \mathbb{C}$.

Also, g cannot be a constant on $\mathcal{E}$ since $\mathcal{E}$ is not a subset of any $\gamma \in \Gamma$; thus there must exist a point $P_0(x_0, h(x_0)) \in \mathrm{int}(\mathcal{E})$ where $g'(x_0) \neq 0$. We are going to get the required contradiction by showing that the tangent plane of the graph of $\tilde{f}$ above P_0, i.e. at point $P_0^+ := \big(x_0, h(x_0), f(x_0, h(x_0))\big)$, is vertical — which is impossible.

To this end, we define two spatial curves on the graph of $\tilde{f}$ that pass through P_0^+ such that, at that point, the tangent lines of the two curves will both exist but will not coincide — hence they must span the tangent plane in question. Specifically, we consider the curves

$$\begin{aligned} &\{\big(x, h(x), g(x)\big) \ : \ x \in I\}; \qquad \text{and} \\ &\{\big(x, y, g(x_0)\big) \ : \ (x, y) \in \tilde{G}, \ \tilde{f}(x, y) = g(x_0)\}; \end{aligned}$$

the former one is the "lifting of $\mathcal{E}$ by function g" while the latter the lifting of the $\tilde{\gamma}^t$ that touches $\mathcal{E}$ at P_0 (i.e. it is $\tilde{\gamma}^{g(x_0)}$) to the fixed height $g(x_0)$. By assumption, there is a line L which is tangent to $\mathcal{E}$ and half–tangent to γ^t at

P_0, hence must be tangent to the extended curve $\tilde{\gamma}^t$ at P_0. Hence both lifted curves have, indeed, tangent lines at P_0^+; that of the latter curve is obviously horizontal while that of the former one is not, by $g'(x_0) \neq 0$. Since both lines project to L in the base plane, we conclude that the tangent plane at P_0^+ must be vertical — the required contradiction to the assumption that f can be extended analytically to $\tilde{G}$.

Now (ii) follows from (i) since it implies that $\mathcal{E}$ can contain no (interior) point of the open set G.

This completes the proof of Lemma 13. ∎

4. The Main Theorem

The following is our main result. Though it concerns families of analytically parameterized curves, we need the technical assumption that there is an *algebraic*, i.e. polynomial relation between the families (the reason being that the Surface Theorem 2 works only for this case).

Theorem 14 (Main Theorem). *Let $\Gamma_1, \Gamma_2, \Gamma_3$ be families explicitly parameterized by the functions f_1, f_2, f_3, analytic on open domains G_1, G_2, G_3 and continuous on $cl(G_1), cl(G_2), cl(G_3)$, respectively, and with the property that $\mathbf{G} = G_1 \cap G_2 \cap G_3$ is connected. Assume that any two curves intersect in at most B points, and the concurrency of three curves $\gamma^{(t_i)} \in \Gamma_i$ $(i = 1, 2, 3)$ is described by a polynomial relation in the sense that, denoting a triple point where they intersect by (u, v), the three parameters $t_i = f_i(u, v)$ satisfy a polynomial relation $F(t_1, t_2, t_3) = 0$, or, more explicitly*

$$F\big(f_1(u,v), f_2(u,v), f_3(u,v)\big) = 0 \tag{6}$$

identically on $cl(\mathbf{G})$, for a polynomial $F \in \mathbb{C}[t_1, t_2, t_3]$. Assume, moreover, that

(i) Γ_3 *has a partial envelope* $\mathcal{E}$*;*

(ii) $\mathcal{E} \subseteq G_1 \cap G_2$*;*

(iii) No f_i $(i = 1, 2, 3)$ is a constant on any non–empty open sub–arc of $\mathcal{E}$. (Intuitively: no non–empty open sub–arc of $\mathcal{E}$ is contained in any $\gamma \in \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$.)

Then

$$\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n) < B \cdot n^{2-\eta},$$

for a suitable $\eta = \eta(\deg(F))$ — provided that $n > n_0 = n_0(\deg(F))$.

Remark 15. The existence of an envelope $\mathcal{E}$ is sufficient but *not necessary* to make $\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n)$ subquadratic. Actually, if no such envelope exists, then anything can happen. To see this, consider the three families of concentric circles about three points $P_1, P_2, P_3 \in \mathbb{R}^2$, respectively. (Obviously, none of these families possesses an envelope.) On the one hand, the method shown in [3] gives that, if the P_i are collinear, then $\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n) \geq cn^2$. On the other hand, if they are non–collinear, then $\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n)$ is subquadratic (see [5], Theorem 33).

Remark 16. The applicability of Theorem 14 is limited to *one–parameter* families Γ_i (the reason, again, being that the Surface Theorem 2 works only for such families).

Remark 17. It is worth noting that requirement (iii) in Theorem 14 is not just a technical assumption. E.g., the $n+n$ straight lines and n parabolas

$$\begin{aligned}\Gamma_1 &:= \{y = t_1^2 \ : \ t_1 = 0, 1, \ldots, n-1\};\\ \Gamma_2 &:= \{x = t_2 \ : \ t_2 = 0, 1, \ldots, n-1\};\\ \Gamma_3 &:= \{y = (x-t_3)^2 \ : \ t_3 = 0, 1, \ldots, n-1\};\end{aligned}$$

have n^2 triple points — three curves of parameter t_1, t_2, t_3, respectively, pass through a common point if and only if $t_1 = |t_2 - t_3|$ — while the x–axis as $\mathcal{E}$ and the polynomial $F(t_1, t_2, t_3) := t_1^2 - (t_2 - t_3)^2$ satisfy all requirements but (iii).

Proof of the Main Theorem

(I) Without loss of generality we may assume that both the polynomial F and the surface $S_F = \{F = 0\}$ are irreducible. Indeed, the open domain $\mathbf{G}$ is connected, hence irreducible (as an analytic set). Therefore its image under the mapping

$$\mathbf{f} = f_1 \times f_2 \times f_3 \ : \ \mathrm{cl}(\mathbf{G}) \to S_F \subset \mathbb{R}^3,$$

defined by

$$(u, v) \mapsto \Big(f_1(u, v), f_2(u, v), f_3(u, v)\Big)$$

is, again, irreducible. Then $\mathbf{f}(\mathbf{G})$ must be contained in a single irreducible component of the surface S_F, and one can simply throw away all other components. Moreover, the analytic functions f_i are nonconstant, hence the polynomial F must depend on all three variables, and the surface S_F does not contain a cylinder over a curve (see Theorem 2(b)). Let $\eta = \eta(\deg(F)) \in (0, 1)$ and $n_0 = n_0(\deg(F))$ be the constants the existence of which is stated in Theorem 2. We want to show that $\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n) < B \cdot n^{2-\eta}$, for $n > n_0$.

(II) Assume for a contradiction that $n + n + n$ curves with parameter sets T_1, T_2, T_3, respectively, determine $\geq B \cdot n^{2-\eta}$ triple points for an $n > n_0$. Any three curves, say of parameter t_1, t_2, t_3, respectively, share at most B common points. Therefore, the surface S_F passes through $\geq n^{2-\eta}$ points of the $n \times n \times n$ Cartesian product $T_1 \times T_2 \times T_3$. In other words, $V = S_F$ and the T_i as X, Y, Z satisfy Theorem 2(a).

(III) Consequently, we can use Theorem 2(d) and localize Theorem 2(b). This gives us a finite subset $H \subset \mathbb{R}$ of "exceptional" or "forbidden" values, and after picking a point P and a surface W in (IV) below, we shall also obtain three analytic functions $g_1, g_2, g_3 : \mathbb{D} \to \mathbb{C}$. Without loss of generality, we may assume that the partial envelope $\mathcal{E}$ of Γ_3 whose existence we assumed in the Main Theorem 14, has the property that

$$\forall P \in \mathcal{E} \text{ and } i = 1, 2, 3, \quad g_i(P) \notin H. \tag{7}$$

Indeed, this only excludes finitely many points from any closed sub–arc of $\mathcal{E}$ — since the g_i are nowhere constant by assumption (iii) — thus, if necessary, $\mathcal{E}$ can be restricted to a suitable open sub–arc.

(IV) Now we pick an arbitrary point $Q \in \mathcal{E}$. Clearly, $\mathbf{f}(Q) \in S_F$, since $\mathcal{E} \subset \mathrm{cl}(\mathbf{G})$ by assumption (ii) and S_F is closed. Recall that $V_0 = V = S_F$ by the irreducibility assumption in (I), and $f_i(Q) \notin H$ for $i = 1, 2, 3$, by the assumption we made in (III), equation (7), so we can apply Theorem 2(d) and (b) to the point $P = \mathbf{f}(Q)$. Then we get a neighbourhood U' of $\mathbf{f}(Q)$, and the promised one-to-one analytic functions (with analytic inverses), $g_1, g_2, g_3 \; : \; (-1, 1) \to \mathbb{R}$ or $\mathbb{D} \to \mathbb{C}$ with the following property: The function $\mathbf{g} = g_1 \times g_2 \times g_3$ maps the origin $(0, 0, 0)$ to $\mathbf{f}(Q)$, and maps an open subset of the plane $x + y + z = 0$ onto the irreducible component of $W \subset S_F \cap U'$ containing $\mathbf{f}(\mathbf{G}) \cap U'$. This latter set is nonempty, since P lies inside U' and in the closure of $\mathbf{f}(\mathbf{G})$.

(V) Denote the inverses of the g_i by φ_1, φ_2, φ_3, respectively. Then the "coordinate-wise inverse" $\mathbf{g}^{-1} = \varphi_1 \times \varphi_2 \times \varphi_3$ maps W into the plane $x + y + z = 0$. In other words, for $(t_1, t_2, t_3) \in W$ we have

$$\varphi_1(t_1) + \varphi_2(t_2) + \varphi_3(t_3) = 0,$$

since the three quantities on the left hand side are coordinates of a point in the plane $x + y + z = 0$. But $\mathbf{f}(\mathbf{G}) \cap U' \subseteq W$, hence

$$\varphi_1\Big(f_1(u, v)\Big) + \varphi_2\Big(f_2(u, v)\Big) + \varphi_3\Big(f_3(u, v)\Big) = 0 \tag{8}$$

identically, in a neighborhood $\mathcal{U} \subset \mathrm{cl}(\mathbf{G})$ of Q. (This $\mathcal{U}$ is open inside $\mathrm{cl}(\mathbf{G})$ but not open in the plane, as Q is a boundary point.)

(VI) According to Lemma 13(i), f_3 cannot be extended analytically to any neighborhood of Q. On the other hand, re–writing (8) as

$$\varphi_3\big(f_3(x, y)\big) = -\varphi_1\big(f_1(x, y)\big) - \varphi_2\big(f_2(x, y)\big),$$

we get an explicit formula for f_3 in $\mathcal{U}$:

$$f_3(x, y) = g_3\Big(- \varphi_1\big(f_1(x, y)\big) - \varphi_2\big(f_2(x, y)\big)\Big) \; .$$

By assumption (ii) the right hand side is defined beyond Q, hence provides an analytic extension of f_3. This is the required contradiction. ∎

5. A Combinatorial Distinction between Unit Circles and Straight Lines

In this chapter we restrict our attention to the real plane $\mathbb{R}^2$. Recognizing unit circles (and, especially, distinguishing them from straight lines) does not seem to be difficult. E.g., anyone can tell that in Figure 4(a)–(b), there can only be found circles and no straight lines. Similarly, few people would doubt that there is no unit circle in Figure 4(c), just straight lines. However, one should be more careful. How do we know that the lines are *really* straight? Perhaps they may be (arcs of) unit circles, provided that our "unit" is very large — so huge that their tiny little arcs do not even seem to be "bent". This is the moment when the points of the 5×5 lattice become important:

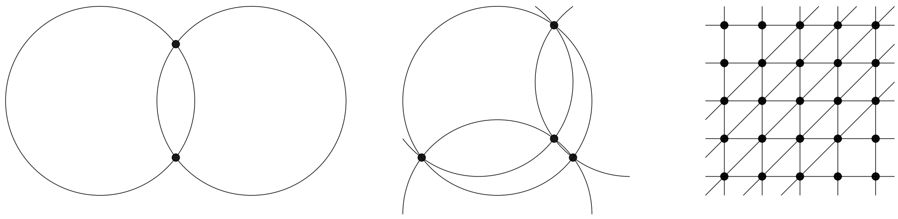

Figure 4: (a)–(b) unit circles; (c) straight lines?

> *is it possible that 25 points and 15 unit circles are incident upon each other just like in Figure 4(c)?*

Unfortunately, we do not know the answer to this simple question. However, we are going to show that, for any $n > n_0$, the n^2 points of an $n \times n$ lattice and $3n$ lines in a similar grid–like configuration (n horizontal, n vertical and n "diagonal" ones) can only have this prescribed incidence pattern if the lines are really straight and cannot if they are (arcs of) unit circles — and this holds even if we only require a near–quadratic number of incidences.

Theorem 18. *There exist an absolute constant $\eta \in (0,1)$ and a threshold n_0 with the following property.*
Let (a_1, b_1), (a_2, b_2), (a_3, b_3) be three distinct points in the Euclidean plane and Γ_1, Γ_2, Γ_3 be three families of unit circles, such that, for each $i \leq 3$, all circles of Γ_i pass through the common point (a_i, b_i). Then

$$\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n) \leq 2^{10} \cdot n^{2-\eta} + 3, \tag{9}$$

provided that $n > n_0$.

Remark 19. The conjecture that in this case $\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n) = o\big(n^2\big)$, originates from Székely (see [4], Conjecture 3.41).

Remark 20. For straight lines the situation is quite different from the one described in Theorem 18. A configuration like the one in Figure 4(c) gives $\approx 3n^2/4$ triple points — where the three points (a_1, b_1), (a_2, b_2), and (a_3, b_3) which are common to the corresponding families of curves, — can be considered as points on the line at infinity.
Similarly, if we allow *arbitrary* (i.e. not just unit) circles then they can produce any incidence pattern that straight lines can: just apply a suitable inversion to any configuration of points and straight lines. Even certain other conic sections have this property, e.g., shifted copies $y = x^2 + ax + b$ of the parabola $y = x^2$: just apply the diffeomorphism $(x, y) \mapsto (x, x^2 + y)$ to any configuration of points and straight lines.

6. Proof of Theorem 18

Assume we are given three families Γ_1, Γ_2, Γ_3 of unit circles and three distinct points (a_1, b_1), (a_2, b_2), $(a_3, b_3) \in \mathbb{R}^2$, with the property that all curves in Γ_i pass through (a_i, b_i), for $i \leq 3$.

1. During the proof we do not consider the three points (a_i, b_i) as triple points (though they might be). This will only add a "+3" at the end.
2. Pick a sufficiently small positive δ so that the δ–neighborhoods $B_\delta(a_i, b_i)$ do not contain any triple point.
3. We subdivide each Γ_i in the way described in Example 12, this subdivision is pictured in Figure 5. Thus we get three times eight subfamilies denoted by G_i^k with $i = 1, 2, 3$ and $k = 1, 2, \ldots 8$. This subdivision will effect the bound on $\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n)$ only by a factor of 8^3.

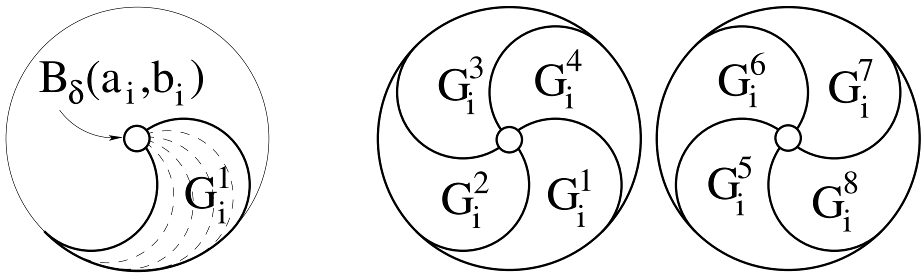

Figure 5: Subdivision of the family Γ_i into eight subfamilies. Left hand side: one of the subfamilies, G_i^1 (dotted arcs). Right hand side: all eight subfamilies.

4. Each such $\Gamma_i^{(k)}$ only covers G_i^k once. Thus, as in Example 12, the family can be explicitly analytically parameterized.
5. It is not difficult to find a trivariate polynomial equation $F(t_1, t_2, t_3) = 0$ that is satisfied by the parameters corresponding to any triple point. This is a rather straightforward calculation, our earlier manuscript had it. However, wishing to emphasize that we do not care for its actual form, we deleted it. (Actually, such polynomials can always be found in case of three *algebraically* parameterized families.)
6. Each $\Gamma_i^{(k)}$ ($i \le 3$, $k \le 8$) possesses an envelope (a quarter circle of radius 2) that is not an envelope for any of the $\Gamma_j^{(l)}$ for $j \neq i$. Thus each triple $\langle \Gamma_1^{(k)}, \Gamma_2^{(l)}, \Gamma_3^{(m)} \rangle$, for $k, l, m \le 8$, satisfies the assumptions of the Main Theorem 14. Thus they cannot have more than $2n^{2-\eta}$ triple points for $n > n_0$ — where $\eta = \eta(\deg(F))$ and $n_0 = n_0(\deg(F))$ are as in Theorem 14.
7. We conclude that, indeed, $\mathcal{T}_{\Gamma_1,\Gamma_2,\Gamma_3}(n) \le 2^{10} n^{2-\eta} + 3$. ∎

Concluding remarks

We have given a sufficient condition for three one–parameter families of curves (or for three copies of a single family) to have "few", more specifically at most $n^{2-\eta}$ triple intersections.

How far below quadratic should it be? Since we have no reasonable estimate for $\eta > 0$, nothing is known about the exact order of magnitude. It may well be that the number of triple points is at most $n^{1+\epsilon}$, for any $\epsilon > 0$. We do not even know any families that satisfy the assumptions of Theorem 14 and can produce a super–linear number of triple points, say $n \log n$.

Which more-than-one parameter families of curves can determine a quadratic number of triple points? Our methods do not work in this generality, since Theorem 2 only applies to 1–parameter families.

We cannot help mentioning a related, beautiful, unsolved problem of Erdős. Assume that, in the projective plane, n straight lines define at least cn^2 *quadruple* points, i.e. points where at least four lines meet. Is it true that, for sufficiently large $n > n_0(c)$, there must exist a point where at least *five* of them intersect?

References

[1] Peter Brass, William Moser, and János Pach, Research problems in discrete geometry. Springer, New York, 2005.

[2] György Elekes, n points in the plane can determine $n^{3/2}$ unit circles. *Combinatorica*, **4** (2-3):131, 1984.

[3] György Elekes, Circle grids and bipartite graphs of distances. *Combinatorica*, **15** 167–174, 1995.

[4] György Elekes, Sums versus products in number theory, algebra and Erdős geometry — a survey. *in: Paul Erdős and his Mathematics II Bolyai Math. Soc. Stud* **11**, pages 241–290, 2002.

[5] György Elekes and Endre Szabó, How to find groups? (And how to use them in Erdős geometry?), *(submitted to Combinatorica).*

[6] Paul Erdős, Some applications of graph theory and combinatorial methods to number theory and geometry, *Algebraic methods in Graph Theory, Coll. Math. Soc. J. Bolyai*, **25** 137–148, 1981.

[7] Zoltán Füredi and Ilona Palásti, Arrangements of lines with a large number of triangles. *Proc. AMS*, **92** 561–566, 1984.

[8] Jiři Matoušek, *Lectures on Discrete Geometry.* Springer–Verlag, Berlin, Heidelberg, New York, 2002.

[9] János Pach and Pankaj K. Agarwal, *Combinatorial Geometry.* J. Wiley and Sons, New York, 1995.

[10] János Pach and Micha Sharir, Repeated angles in the plane and related problems. *Journal of Combinatorial Theory, series A*, **59** 12–22, 1990.

[11] Joel Spencer, Endre Szemerédi, and W.T. Trotter Jr, Unit distances in the Euclidean plane. In *Graph theory and combinatorics (Cambridge, 1983)*, pages 293–303. Academic Press, London, 1984., 1983.

[12] J. J. Sylvester, Problem 2473. *Math. Questions from the Educational Times*, **8** 106–107, 1867.

[13] László A Székely, Crossing numbers and hard Erdős problems in discrete geometry. *Combinatorics, Probability and Computing*, **6**(3) 353–358, 1997.

[14] Endre Szemerédi and W. T. Trotter Jr. Extremal problems in Discrete Geometry. *Combinatorica*, **3** (3–4) 381–392, 1983.