Distance Problems and Points on Curves

An Algebraic Approach

by

Ryan Schwartz

B.Sc. in Mathematics and Computer Science, The University of the Witwatersrand, 2005
B.Sc. (Honours) in Computer Science, The University of the Witwatersrand, 2005
B.Sc. (Honours) in Mathematics, The University of the Witwatersrand, 2006
M.Sc. in Mathematics, The University of British Columbia, 2009

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate and Postdoctoral Studies

(Mathematics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

April 2014

Abstract

In this thesis we give results on unit and rational distances, structure results for surfaces containing many points of a cartesian product and a survey of various, mainly combinatorial, applications of the Subspace Theorem. We take an algebraic approach to most of the problems and use techniques from commutative algebra, incidence geometry, number theory and graph theory.

In Chapter 2 we give extensions of a result of Elekes and Rónyai that says that if the graph of a real (or complex) polynomial contains many points of a cartesian product then the polynomial has a special additive or multiplicative form. We extend this to asymmetric cartesian products and to higher dimensions. Elekes and Rónyai's result was used to prove a conjecture of Purdy on the number of distinct distances between two collinear point sets in the plane. Our extensions give improved bounds for the conjecture.

In Chapter 3 we prove a special case of the Erdős unit distance problem. This problem asks for the maximum possible number of unit distances between n points in the plane in the form of an asymptotic upper bound. We provide an upper bound of $n^{1+7/\sqrt{\log n}}$ when we only consider unit distances corresponding to roots of unity and give a superlinear lower bound. We also consider related rational distance problems. We require an algebraic result of Mann on the number of solutions of linear equations of roots of unity.

In Chapter 4 we extend our result from the previous chapter to unit distances coming from a multiplicative subgroup of $\mathbb{C}^*$ of "low" rank. We use a corollary of the Subspace Theorem. In this case we get, for $\varepsilon > 0$, at most $cn^{1+\varepsilon}$ unit distances. We show that the well known lower bound construction of Erdős for the general unit distance problem consists of distances from such a subgroup and so our result applies to the best known maximal unit distance sets.

In Chapter 5 we give a survey of various applications of the Subspace Theorem including less well known combinatorial applications such as sum-product estimates and line configurations with few distinct intersections.

Preface

This thesis is based on four papers. The first three have been published and the last has been accepted for publication. The first three papers are given as published in this thesis with minor typographical and stylistic changes. The fourth paper is a survey paper and has been modified slightly in this thesis as parts of it appear in the other papers.

- Chapter 2 was published as:

  R. Schwartz, J. Solymosi and F. de Zeeuw, *Extensions of a result of Elekes and Rónyai,* Journal of Combinatorial Theory, Series A 120(7):1695-1713, 2013.

  The work was shared equally between the three authors.

- Chapter 3 was published as:

  R. Schwartz, J. Solymosi and F. de Zeeuw, *Rational distances with rational angles,* Mathematika 58(2):409-418, 2012.

  This paper also appeared as Chapter 4 of Frank de Zeeuw's PhD thesis. The work was shared equally between the three authors.

- Chapter 4 was published as:

  R. Schwartz, *Using the subspace theorem to bound unit distances,* Moscow Journal of Combinatorics and Number Theory 3(1):108-117, 2013.

  The idea of using the Subspace Theorem for this problem was given to me by József Solymosi.

- Chapter 5 is based on a survey talk given by József Solymosi at the workshop "Geometry, Structure and Randomness in Combinatorics"

in Pisa in September 2012. It is to appear in a special volume of papers from the workshop in the Edizioni Della Normale di Pisa (editors: Jirka Matoušek, Jarik Nešetřil and Marco Pellegrini).

The work of writing up the material was shared between the two authors.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank my research supervisor József Solymosi. I am indebted to him for his constant support, reassurance and wealth of ideas. I would also like to thank Frank de Zeeuw for many helpful discussions and distractions. I am grateful to my supervisory committee for their input, time and effort and to everyone in the UBC Mathematics Department for all their help over the years.

For the work in Chapter 2, the authors would like to thank the referees for their helpful comments and suggestions.

For the work in Chapter 3, the authors would like to thank Jirka Matoušek for making us aware of the lower bound for unit distances with no three points on a line and to the anonymous referee for helpful comments, suggestions and corrections.

For the work in Chapter 4, I am very grateful to József Solymosi for the idea of using the Subspace Theorem and other useful discussions. I would also like to thank Yann Bugeaud for making me aware of Amoroso and Viada's bound for the corollary of the Subspace Theorem giving the improved bound in Theorem 4.1.1. Lastly I would like to thank Christian Elsholtz for helpful comments and corrections in the last section and the referee for their helpful information.

For the work in Chapter 5, the authors are thankful to Jarik Nešetřil for the encouragement to write this survey. We are also thankful to the organizers of the workshop in Pisa, "Geometry, Structure and Randomness in Combinatorics", where the parts of this paper were presented.

Dedication

For my family who provide the foundation for my strength and the wings for my dreams.

Chapter 1

Introduction

Many problems in combinatorial geometry deal with points on curves or surfaces. Even grid points can be viewed as such arrangements by considering them as lying on collections of parallel lines. There are many useful results regarding the structure of such arrangements. Possibly the best known example, and a hallmark of the field, is the Szemerédi-Trotter Theorem which bounds the number of point-line incidences in the plane. A number of extensions of this result exist including analogues in higher dimensions and for curves instead of lines.

Distance problems are also common in combinatorial geometry. For example, it is natural to ask, given n points in the plane, how many times a certain distance appears or how many distinct distances appear. These problems are related to point-circle incidences. These two questions, first asked by Erdős, are called the unit distance problem and the distinct distances problem respectively and have been studied extensively with the first still open and the second being solved very recently. The first problem gets its name from the fact that by scaling we may assume that the distance in question is one.

These problems have a strong combinatorial flavour and can often be stated quite simply but for many of them elementary techniques fail. A range of results from various fields have been applied to such problems. These include analytic, probabilistic, topological, number-theoretic and algebraic methods. The geometric aspects of these problems make techniques from algebraic geometry useful and the combinatorial aspects make number-theoretic tools useful. Combining these a number of previously unreachable problems in the field have been solved.

In this thesis we give a number of results regarding points on curves and surfaces and distance problems using algebraic techniques. First we give extensions of various results of Elekes and Rónyai on the structure of hypersurfaces and collections of lines containing many points of a cartesian product. Next we give a number of special cases of the Erdős unit distance problem where we only consider distances from multiplicative groups with "low" rank. Finally we give an overview of the Subspace Theorem of Schmidt and

its extensions with particular emphasis on combinatorial applications. We will use tools from graph theory, incidence geometry, commutative algebra and number theory.

The remainder of the introduction will give an overview of the research area, describe the main results of the thesis and provide the necessary background to understand those results.

1.1 Some notation

Most of the results in this work are of an asymptotic nature so we will often use c, c', C, C', etc. to represent arbitrary constants. We may also use *big O notation* and *little-o notation*. Specifically,

- $f(x) = O(g(x))$ (or $f(x) \leq O(g(x))$) if there are positive constants M and x_0 such that $f(x) \leq Mg(x)$ for $x > x_0$.
- $f(x) = o(g(x))$ (or $f(x) \leq o(g(x))$) if $\lim_{x\to\infty} f(x)/g(x) = 0$.

We may use *grid* to mean cartesian product and *the plane* to mean $\mathbb{R}^2$ with the usual Euclidean distance.

1.2 Surfaces containing many points of a cartesian product

When dealing with points and curves it is natural to ask how they interact. For example, given a collection of curves, we may want to bound the number of intersections between pairs of curves. Bezout's Theorem, from algebraic geometry, bounds the number of intersections between two algebraic plane curves without a common factor by the product of their degrees. For details see [32]. Another example is, given a collection of points, to find the number of connecting lines between pairs of points. Beck's Theorem is a structure result regarding this problem. It says that, given a set of n points in the plane, either cn of the points lie on the same line or the points determine at least cn^2 distinct connecting lines [7]. A connecting line is a line between two points in the point set. Notice that in the first case of Beck's Theorem there are cn^2 pairs of points that give the same connecting line.

Another important interaction is a point-line incidence. Given a collection of points $\mathcal{P}$ and a collection of lines $\mathcal{L}$, a *point-line incidence* is a pair $(p, \ell) \in \mathcal{P} \times \mathcal{L}$ such that $p \in \ell$. See Figure 1.1 for an example. The more general point-curve incidences can be defined similarly. We will denote

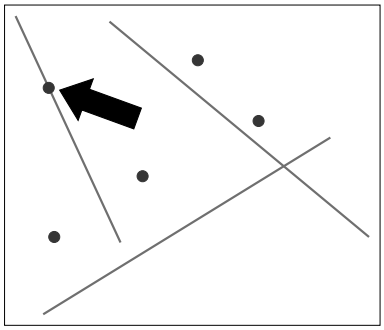


Figure 1.1: An example of a point-line incidence

the number of incidences between $\mathcal{P}$ and $\mathcal{L}$ (or a collection of curves $\mathcal{C}$) by $I(\mathcal{P}, \mathcal{L})$ (or $I(\mathcal{P}, \mathcal{C})$).

The Szemerédi-Trotter Theorem gives an upper bound on the number of point-line incidences in the plane. It was first proved by Szemerédi and Trotter in 1983 [64] using cell decomposition. A number of different proofs have been given with the simplest by Székely using the crossing inequality for graphs [62]. The Szemerédi-Trotter Theorem says that the number of point-line incidences between n points and m lines in the plane is at most $I(\mathcal{P}, \mathcal{L}) \leq C(m^{2/3}n^{2/3} + m + n)$.

Note that the $m^{2/3}n^{2/3}$ term above is dominant for $\sqrt{m} \leq n \leq m^2$ and most, if not all, applications of the Szemerédi-Trotter Theorem fall in this range. This theorem is tight up to the constant C with the lower bound coming from a suitably chosen grid. Such a grid was first given by Erdős [25] and a slightly simpler example was later given by Elekes [20].

An equivalent form of the Szemerédi-Trotter Theorem exists which deals with "rich" lines. Given a point set $\mathcal{P}$ and a positive integer k, a line is *k-rich* on $\mathcal{P}$ if the line contains at least k points from $\mathcal{P}$. The second formulation of the theorem states that, given a positive integer k and a point set $\mathcal{P}$ of size n, the number of k-rich lines on $\mathcal{P}$ is at most $C(n^2/k^3 + n/k)$.

The above two results hold for arbitrary point sets but more can be said about the structure of the lines when the points form a grid. If we have "many" lines and they all contain "many" points of the grid then the following result of Elekes gives that a positive fraction of the lines are parallel or have the same intersection point [18]. A collection of lines is said to be *concurrent* if they have the same intersection point.

Lemma 1.2.1. *Suppose $A, B \subset \mathbb{R}$ and $|A| = |B| = n$. For all $c_1, c_2 > 0$*

there exists $C > 0$, independent of n, such that if Cn lines in $\mathbb{R}^2$ are $c_1 n$-rich on $A \times B$, then at least $c_2 n$ of them must be parallel or concurrent.

Note that a line can contain at most n points of an $n \times n$ grid. So the above lemma deals with lines containing "many" grid points. Also, an $n \times n$ grid contains n^2 points. So by the second formulation of the Szemerédi-Trotter Theorem we cannot have more than $C'n$ $c_1 n$-rich lines. Thus the lemma deals with "many" rich lines. An extension of this result appears as Lemma 2.2.4 in Chapter 2. The extension deals with the case when we have fewer $c_1 n$-rich lines. Specifically, it says that given $2/3 \leq \alpha \leq 1$, if we have at least cn^α $c_1 n$-rich lines then at least $c' n^{3\alpha - 2}$ of the lines are parallel or concurrent. We retrieve Lemma 1.2.1 when $\alpha = 1$. At the other extreme, if $\alpha = 2/3$, then we only have a constant number of parallel or concurrent lines.

A collection of lines are in *general position* if no two are parallel and no three are concurrent. Our generalised line lemma only holds for $n^{2/3}$ or more lines but a conjecture of Solymosi says that the same should hold for a constant number or at most n^ε lines in general position. For more details see [20]. A recent result of Amirkhanyan et al. addresses this conjecture [2].

The theorems above give much information about points and lines but what if we have more complicated objects such as "well-behaved" curves? The following theorem of Pach and Sharir addresses this case [52, 53]. We define this so-called good behaviour as follows. A collection $\mathcal{C}$ of continuous real planar curves with no self-intersection has k *degrees of freedom* and *multiplicity-type* s if any pair of curves in Γ have at most s intersection points and for any k points there are at most s elements of Γ that contain all k points.

Theorem 1.2.2. *Let $\mathcal{P}$ be a set of n points and $\mathcal{C}$ be a set of m curves as defined above with k degrees of freedom and multiplicity type s. Then there is a constant $C > 0$ depending on k and s such that*

$$I(\mathcal{P}, \mathcal{C}) \leq C \left(m^{(2k-2)/(2k-1)} n^{k/(2k-1)} + m + n \right).$$

Note that this theorem holds for algebraic curves of bounded degree since these curves can be split up into a small number (depending on the degree) of continuous curves without self-intersection.

We already saw Elekes' result about rich lines. In 2000 Elekes and Rónyai used this result along with many other tools from algebra, graph theory and incidence geometry to prove a structure theorem about surfaces whose

graphs contain many points of a grid [21]. It was the inspiration for Chapter 2. The Elekes-Rónyai Theorem states that if the graph of a two variable polynomial $f \in \mathbb{R}[x, y]$ of bounded degree contains cn^2 points of an $n \times n \times n$ cartesian product then the polynomial has the form

$$f(x, y) = g(k(x) + l(y)), \quad \text{or} \quad f(x, y) = g(k(x) \cdot l(y)),$$

where $g, k, l \in \mathbb{R}[t]$.

So if the graph of a two variable real polynomial contains many points of a cartesian product then the polynomial should have a special *additive* or *multiplicative* form.

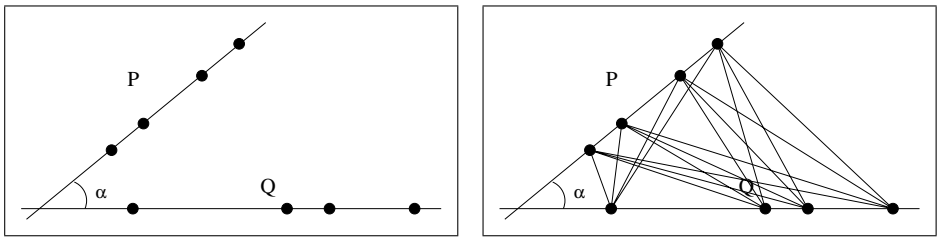
1.2.1 Results

In Chapter 2 we give a number of extensions of this result using our extension of Elekes' line lemma. Specifically, we get the result for a smaller assymetric cartesian product of size $n \times n^{5/6} \times n$ and two analogous results in one dimension higher for cartesian products of size $n \times n \times n \times n$ and $n \times n \times n^{5/6+\varepsilon} \times n$ respectively. Using the result of Amirkhanyan et al. in place of our line lemma even smaller cartesian products can be considered. We also provide a construction using translations of a parabola on a grid to show that one of our higher dimensional extensions is near optimal.

Our proofs of these results were based on the original proof of Elekes and Rónyai. The main difference was the generalised line lemma. The proofs proceed as follows:

1. break up the surface into fibres,
2. define curves from pairs of fibres and show that many of them are rich,
3. use the Pach-Sharir Theorem to show that many of the curves coincide,
4. show that many of the curves must share a common inner function,
5. reduce to the case of lines to get many lines containing many points of a cartesian product,
6. apply a line lemma to get many parallel or concurrent lines,
7. use some algebra to get the required form for the polynomial.

In this process the additive form comes from many parallel lines and the multiplicative form comes from many concurrent lines.



(a) Two collinear point sets. (b) All distances between the point sets.

Figure 1.2: Two collinear point sets P and Q. The distances between the two point sets are shown in the second figure. The angle between the lines containing the point sets is α.

For our extensions we require a number of algebraic results, the details of which can be found in Chapter 2. Most of these were given in Elekes and Rónyai's paper. For our higher dimensional extensions we require another algebraic result. Namely a special case of the Schwartz-Zippel Lemma which can be considered as a generalisation to higher dimensions of the fact that a degree d univariate polynomial can have at most d roots. Details and a proof of the lemma can be found in [36]. We also require a graph-theoretic result which states that if we colour the edges of a graph on n vertices with cn^2 edges such that a bounded number of colours meet at each vertex then the graph contains a monochromatic (one colour) subgraph containing $c'n^2$ edges. This result also appears in Elekes and Rónyai's work.

The Elekes-Ronyai Theorem and our extensions of it seem quite abstract but they can be applied to a number of combinatorial problems. They lend themselves especially well to distance problems. For instance, consider the following conjecture of Purdy. If the number of distinct distances between two collinear point sets in the plane of size n each is Cn then the lines containing the point sets must be parallel or orthogonal. Figure 1.2 contains such a configuration. This problem can be translated into a problem about the graph of a polynomial containing many points of a cartesian product. The Elekes-Rónyai Theorem then gives the stated conditions on the lines. Our extensions allow us to consider the same problem but with fewer points on one of the lines.

1.3 The Erdős unit distance problem

As mentioned above the unit distance problem asks for the maximum possible number $u(n)$ of times the distance one occurs between n points in the plane. This problem was first asked by Erdős in 1946 [25]. Erdős showed that $u(n) \geq n^{1+c/\log\log n}$ using a scaled portion of an integer lattice. The bound requires number-theoretic tools. Specifically, bounds on the number of representations of a number as a sum of two squares are used. More details of this can be found in Chapter 4. Erdős conjectured that the true magnitude of $u(n)$ is close to this lower bound. He was able to show that $u(n) \leq cn^{3/2}$ by using the fact that two unit circles intersect in at most two points. This bound was later improved to $cn^{4/3}$ by Spencer, Szemerédi and Trotter using a version of the Szemerédi-Trotter Theorem for unit circles [60]. A number of different proofs have been given, all providing different proofs of the Szemerédi-Trotter Theorem for unit circles, but the bound has not been improved.

In the same paper, Erdős also introduced the distinct distances problem which asks for the minimum possible number $g(n)$ of distinct distances between n points in the plane. He gave the bounds $cn^{1/2} \leq g(n) \leq cn/\sqrt{\log n}$. In this case the upper bound also comes from a suitably scaled portion of an integer lattice and the lower bound comes from a simple counting argument. Again, Erdős conjectured that the true magnitude is close to the upper bound. Unlike the unit distance problem the best bound for this problem has been improved many times over the years. The progression of these results is given in Table 1.1. For an in-depth account of this progression and a survey of the distinct distances and unit distance problems see [33] and [9].

We will mention the near-optimal solution of this problem due to Guth and Katz [34]. They attained the lower bound $cn/\log n$ which is very close to the upper bound $cn/\sqrt{\log n}$ given by Erdős. Before their proof the distinct distances problem was reduced to an incidence problem in higher dimensions by Elekes and Sharir using properties of rotations of the point set [22]. Guth and Katz were able to solve this problem using a number of powerful and ingenious techniques. Their use of the polynomial ham-sandwich theorem of Stone and Tukey [61] to provide a well-behaved partitioning of the space was perhaps the most remarkable. Their proof shows the power of using methods from various fields for problems in combinatorial geometry.

The unit distance and distinct distances problems are related in a very nice way. Specifically, a bound for the unit distance problem gives a bound for the distinct distances problem as follows: suppose we have a point set of

Lower bound	**Author [Publication]**	**Year**
$g(n) \geq cn^{1/2}$	P. Erdős [25]	1946
$g(n) \geq cn^{2/3}$	L. Moser [50]	1952
$g(n) \geq cn^{5/7}$	F.R.K. Chung [12]	1984
$g(n) \geq cn^{4/5}/\log n$	F.R.K. Chung, E. Szemerédi, W.T. Trotter [13]	1992
$g(n) \geq cn^{4/5}$	L. Székely [62]	1997
$g(n) \geq cn^{6/7}$	J. Solymosi, C.D. Toth [59]	2001
$g(n) \geq cn^{(4e/(5e-1))-\varepsilon}$	G. Tardos [66]	2003
$g(n) \geq cn^{((48-14e)/(55-16e))-\varepsilon}$	N.H. Katz, G. Tardos [38]	2004
$g(n) \geq cn/\log n$	L. Guth, N.H. Katz [34]	2010

Table 1.1: The progression of lower bounds for the Erdős distinct distances problem up to the near complete solution of Guth and Katz.

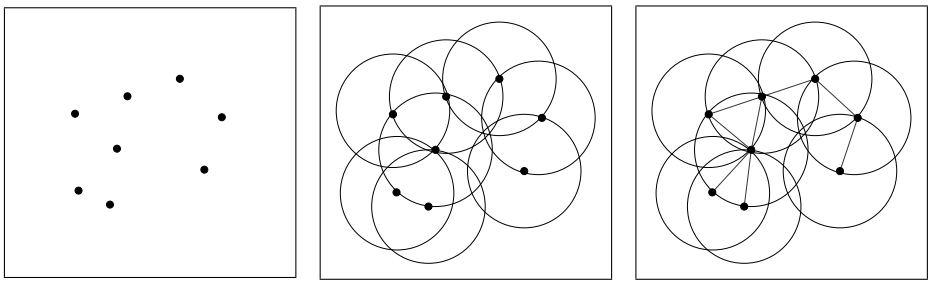
size n with the minimum number, $g(n)$, of distinct distances. There are $\binom{n}{2}$ pairs of points in total. So, by the pigeonhole principle some distance must appear at least $\binom{n}{2}/g(n)$ times. In particular this tells us that

$$u(n) \geq \frac{1}{g(n)}\binom{n}{2} \Rightarrow g(n) \geq \frac{1}{u(n)}\binom{n}{2}.$$

Thus if $u(n) \leq cn^{1+\varepsilon}$ then $g(n) \geq c'n^{1-\varepsilon}$.

The proof of the best upper bound for the unit distance problem uses the Szemerédi-Trotter Theorem for unit circles. Unit circles are similar to lines in a certain topological sense. Notice that two lines intersect in at most one point and two points uniquely define a line. Any collection of curves satisfying these properties are called *pseudolines*. For unit circles we have similar properties. In particular, any two unit circles intersect in at most two points and two points define at most two unit circles. By only considering the upper or lower halves of the circles the above properties of lines are satisfied and we get a collection of pseudolines. We choose the upper or lower halves according to which ones give us more incidences. The remarkable thing about the proof of the Szemerédi-Trotter Theorem is that it only uses the above properties of lines and so it also holds for pseudolines. So, for unit circles specifically, given a point set $\mathcal{P}$ of size n and a collection of unit circles $\mathcal{C}$ of size m in the plane the number of point-circle incidences is at most $I_U(\mathcal{P}, \mathcal{C}) \leq C(m^{2/3}n^{2/3} + m + n)$.

To prove the upper bound for the unit distance problem we consider n



(a) A collection of points in the plane.

(b) Unit circles centred at the points.

(c) Unit distances between the points.

Figure 1.3: A collection of points is given in (a). A unit circle is drawn around each point in (b). The correspondence between unit distances and point-circle incidences is shown in (c).

points in the plane and the n unit circles with the points as their centres. Note that two points have unit distance if each is on the unit circle centred at the other point. See Figure 1.3 for an example. So each unit distance corresponds to two point-circle incidences. Thus by the Szemerédi-Trotter Theorem for unit circles the number of unit distances is at most $Cn^{4/3}$.

Notice that given a pair of points in the plane we get a *distance* vector between them. This vector can be considered as a complex number and if the distance between the points is one then the complex number has modulus one.

Given a set of n points in the plane, the *unit distance graph* of the point set is the graph with vertices the n points and edges the unit distances. A path in this graph is call *irredundant* or *nondegenerate* if no subsum of the distance vectors equal zero. More details of this will be given in Chapters 3 and 4.

We will be looking at the unit distance problem when the complex numbers associated with the distance vectors come from a subgroup of $\mathbb{C}$ of low rank. A subgroup $\Gamma \subset \mathbb{C}^*$ has *rank* r (or just *finite rank*) if there exists a finitely generated subgroup $\Gamma_0 \subset \Gamma$ with r generators such that for every $x \in \Gamma$ there exists an integer $k \geq 0$ such that $x^k \in \Gamma_0$. Note that the group of roots of unity has rank 0 as every element of the group has finite order.

We will use a corollary of the Subspace Theorem. More details of the Subspace Theorem will be given in the next section but for now we just focus on the corollary and its use in the unit distance problem. The corollary bounds the number of certain solutions of linear equations coming from a

subgroup of a field with finite rank. The bound depends on the rank of the subgroup and the number of terms in the linear equation. This corollary was due originally to Evertse, Schlickewei and Schmidt [30]. We will use a recent variant of this result of Amoroso and Viada [3] that gives, to our knowledge, the best known bound.

Suppose Γ is a subgroup of $\mathbb{C}^*$ of finite rank r and $a_1, a_2, \ldots, a_k \in \mathbb{C}^*$. A solution of the equation

$$a_1 z_1 + a_2 z_2 + \cdots + a_k z_k = 1 \tag{1.1}$$

is called *nondegenerate* if no subsum of the left-hand side vanishes. That is $\sum_{j \in J} a_j z_j \neq 0$ for every nonempty $J \subset \{1, 2, \ldots, k\}$. The corollary bounds the number $A(k, r)$ of nondegenerate solutions of this equation with $z_i \in \Gamma$ a subgroup of $\mathbb{C}^*$ of rank r. Specifically, $A(k, r) \leq (8k)^{4k^4(k+kr+1)}$.

1.3.1 Results

We give bounds for two special cases of the unit distance problem and a number of similar results where we consider rational distances and not just unit distances. For our first result we only consider unit distances with angle to the x-axis a rational multiple of π. These distances correspond to roots of unity and so come from a multiplicative subgroup of $\mathbb{C}^*$ of rank 0. We call the maximum number of these unit distances $u_1(n)$. We give other results dealing with rational distances with angle to the x-axis a rational multiple of π. These results are similar to the first result but the bounds vary depending on how many points we allow on a line. Lastly we consider distances coming from a multiplicative subgroup of $\mathbb{C}^*$ of "low" rank. We require that the rank is at most $c \log n$ where n is the number of points. We call the maximum number of these unit distances $u_2(n)$.

In the first case we get the bounds

$$cn \log n \leq u_1(n) \leq n^{1+7/\sqrt{\log n}}.$$

The upper bound comes from an application of a result of Mann from 1965 on the number of solutions of linear equations of roots of unity [44]. From Mann's Theorem we get a corollary bounding the number of nontrivial solutions in roots of unity of a linear equation in k variables. The bound depends only on k. The lower bound construction of Erdős does not hold for this case as the unit distances in the scaled integer lattice are not generally roots of unity. The lower bound given comes from a modified version of a construction of Erdős and Purdy [27]. The construction is a projection of a hypercube.

For the problems with rational distances the upper bounds also come from Mann's Theorem and vary from $n^{1+7/\sqrt{\log n}}$ to cn^2 depending on how many points we allow on a line. Near optimal lower bounds are given using a similar construction to that of Erdős and Purdy above. We were able to get these results by showing that Mann's Theorem holds for sums of rational multiples of roots of unity.

For our last problem concerning distances from a group with low rank, given $\varepsilon > 0$, we get the bounds

$$n^{1+c/\log\log n} \leq u_2(n) \leq n^{1+\varepsilon}.$$

The upper bound comes from the corollary of the Subspace Theorem. Here the lower bound is the same as Erdős' construction as we show. This requires a careful analysis of the lower bound construction. The fact that our result works for the best known lower bound shows its usefulness and suggests that all maximal unit distance sets may have a special structure.

Mann's Theorem deals with linear equations of roots of unity and the corollary of the Subspace Theorem deals with linear equations of elements of a group of finite rank. So Mann's Theorem is a special case of the corollary and can be thought of as a starting point, at least from a combinatorial perspective, for the development of the Subspace Theorem.

The idea of all the proofs is quite similar. We will only highlight the proofs for unit distances but the proofs for rational distances follow the same lines. The proofs proceed as follows:

1. consider the unit distance graph where we only include edges that correspond to unit distances from the required group,

2. count irredundant/nondegenerate paths of length k where k is to be determined,

3. using simple degree arguments we get a lower bound for the number of such paths between a pair of vertices,

4. Mann's Theorem or the corollary of the Subspace Theorem give us an upper bound on the number of such paths between the same pair of vertices,

5. putting these together and optimizing for k we get the stated bounds.

Full details appear in Chapters 3 and 4.

1.4 Combinatorial applications of the Subspace Theorem

A version of the Subspace Theorem was first proved by Schmidt in 1972 [55]. We have already mentioned this theorem in the previous section with particular emphasis on its corollary and its subsequent use for the unit distance problem. It was a generalisation of the Thue-Siegel-Roth Theorem on the approximation of algebraic numbers to higher dimensions [54].

The Subspace Theorem says that the solutions of linear equations in a multiplicative subgroup of a field of finite rank come from a finite number of linear subspaces. Schmidt's original theorem did not give a quantitative bound on the number of linear subspaces but just guaranteed finiteness. A quantitative version, along with the corollary mentioned in the previous section, were subsequently given by Evertse, Schlickewei and Schmidt [30].

The Subspace Theorem is a powerful tool in number theory. It has appeared in various forms and been adapted and improved over time. Its more well known applications include diophantine approximation, results about integral points on algebraic curves and the construction of transcendental numbers. But its usefulness extends beyond the realms of number theory. Other applications of the Subspace Theorem include linear recurrence sequences and finite automata. The multitude of uses of the Subspace Theorem are elegantly highlighted in the surveys of Bilu [8], Evertse and Schlickewei [29] and Corvaja and Zannier [16]. These surveys give many proofs of results from number theory and algebraic geometry including those mentioned above.

The proof of the Subspace Theorem is very deep and beyond the scope of this work. We plan to focus on its use as a tool for combinatorial problems without delving into the background of the theorem very much. We will, however, mention its origins and present two of its more well known applications before proceeding to the combinatorial problems. The full details appear in Chapter 5.

A *linear recurrence sequence* is a sequence where the first few terms are given explicitly and the remaining terms are given via a linear recurrence relation. The Fibonacci numbers are probably the best known example of a linear recurrence sequence. The *zero set* of a linear recurrence sequence is the set of terms in the sequence that equal zero.

Given a set A of elements of a field we define the *sum set* and *product set* of A as

$$A + A = \{a + b : a, b \in A\}, \quad \text{and} \quad AA = \{ab : a, b \in A\}$$

respectively. These sets and their relationship are well studied in combinatorics and additive number theory. The idea behind the theory is that addition and multiplication are somehow incompatible so if one of these sets is small then the other should be large. Note that the size of the sum set and product set are between $|A|$ and $|A|^2$. It was conjecture by Erdős and Szemerédi that, given $\varepsilon > 0$, for any finite $A \subset \mathbb{Z}$, $|A + A| + |AA| \geq |A|^{2-\varepsilon}$. The best known bound for this problem is due to Solymosi with $|A + A| + |AA| \geq n^{4/3} - o(1)$ using an ingenious elementary method involving a cartesian product [58]. This bound holds for real numbers, not just integers, and a similar result has recently been given for complex numbers by Konyagin and Rudnev [40].

1.4.1 Results

First we give a few different formulations of the Subspace Theorem and state the corollary mentioned above. Then we give two well known (non-combinatorial) applications. The first is an example of proving that a number is transcendental. This involves considering the binary expansion of the number, assuming it is not transcendental and then using the Subspace Theorem to show that it must then be rational, a contradiction. The second example shows that the zero set of certain special linear recurrence sequences is finite. This involves considering the recurrence relation as a matrix and then using properties of the characteristic polynomial to apply the corollary of the Subspace Theorem.

For the combinatorial applications, first we present a result of Chang which states that if the product set is "very" small then the sum set is "very" large [10]. Specifically, if $|AA| \leq c|A|$ then $|A+A| \geq |A|^2/2+O(|A|)$. Lastly, we give a bound for complex line configurations with few distinct intersections due to Chang and Solymosi [11].

The main purpose of the chapter is to be an exposition of these results so that the methods may find more common use among discrete mathematicians.

1.5 Layout of this document

In Chapter 2 the extensions of the results of Elekes and Rónyai are given. The various extensions are stated in Section 2.1, the required background is given in Section 2.2, the proofs are given in Sections 2.3-2.4 and applications and lower bounds are given in Section 2.5.

In Chapters 3–4 the special cases of the Erdős unit distance problem are given. For the results using Mann's Theorem, the statements of our upper bounds are given in Section 3.1, Mann's result and its corollaries are given in Section 3.3, the proofs of the upper bounds are given in Section 3.4 and the lower bounds are given in Section 3.5. For the results using the Subspace Theorem, the statement of our upper bound is given in Section 4.1, the proof of the upper bound is given in Section 4.2 and the analysis of Erdős' lower bound is given in Section 4.3.

In Chapter 5 an exposition of the Subspace Theorem and its combinatorial applications is given. Background and various formulations and corollaries of the Subspace Theorem are given in Section 5.1, two well known applications of the Subspace Theorem are given in Section 5.2 and the applications of the Subspace Theorem to sum-product estimates and line configurations with few intersections are given in Section 5.3.

Chapter 6 highlights the conclusions of this thesis and possible future work. This includes the scope of the research and how it fits into the field and the limitations of the results and how they could be generalised or extended.

Chapter 2

Extensions of a result of Elekes and Rónyai

2.1 Introduction

2.1.1 Background

We are interested in polynomials on finite cartesian products, for instance of the form $f(x, y) \in \mathbb{R}[x, y]$ on $A \times B$, with $A, B \subset \mathbb{R}$ and $|A| = |B| = n$. We will focus on the question of how small the image $f(A, B)$ can be in terms of n.

For two basic examples, $x + y$ and xy, the image can be as small as cn, if A and B are chosen appropriately. For $f(x, y) = x + y$ one can take $A = B = [1, n]$ (or any other arithmetic progression of length n), so that $f(A, B) = A + B = [2, 2n]$; for $f(x, y) = xy$ one can take a geometric progression like $A = B = \{2^1, 2^2, \ldots, 2^n\}$, so that $f(A, B) = A \cdot B = \{2^2, 2^3, \ldots, 2^{2n}\}$. Similar small images can be obtained for polynomials of the form $f(x, y) = g(k(x) + l(y))$, for nonconstant polynomials g, k, l, by taking A so that $k(A) \subset [1, n]$, and B so that $l(B) \subset [1, n]$. A similar idea works for $f(x, y) = g(k(x) \cdot l(y))$.

For convenience, we will formulate the problem slightly differently: we consider the surface $z = f(x, y)$ in $\mathbb{R}^3$ and its intersection with a cartesian product $A \times B \times C$, with $|A| = |B| = |C| = n$. Then the image of f is 'small' if and only if that intersection is 'large'; for instance, $z = x + y$ has intersection with $[1, n]^3$ of size at least $\frac{1}{4}n^2$.

In 2000, Elekes and Rónyai [21] proved the following converse of the above observations.

Theorem 2.1.1 (Elekes-Rónyai Theorem)**.** *For every $c > 0$ and positive integer d there exists $n_0 = n_0(c, d)$ with the following property.*
Let $f(x, y)$ be a polynomial of degree d in $\mathbb{R}[x, y]$ such that for an $n > n_0$ the graph $z = f(x, y)$ contains cn^2 points of $A \times B \times C$, where $A, B, C \subset \mathbb{R}$

have size n. Then either

$$f(x,y) = g(k(x) + l(y)), \quad \text{or} \quad f(x,y) = g(k(x) \cdot l(y)),$$

where $g, k, l \in \mathbb{R}[t]$.

In fact, they proved that the same is true for rational functions, if one allows a third special form $f(x,y) = g((k(x) + l(y))/(1 - k(x)l(y)))$. Elekes and Szabó [23, 24] were able to extend this theorem to implicit surfaces $F(x,y,z) = 0$, and also showed that the surface need only contain $n^{2-\gamma}$ points of the cartesian product for the conclusion to hold, for some absolute 'gap' $\gamma > 0$.

Elekes and Rónyai used their result to prove a famous conjecture of Purdy. It says that given two lines in $\mathbb{R}^2$ and n points on each line, if the number of distinct distances between pairs of points, one on each line, is cn for some $c > 0$, then the lines are parallel or orthogonal. Elekes [19] also proved a 'gap version', only requiring the number of distances to be less than $cn^{5/4}$. For details and a proof of a variation of Purdy's conjecture, using our results below, see Section 2.5.2.

See [20, 46, 47] for more details and some related problems.

2.1.2 Results

In this paper we prove a number of extensions of Theorem 2.1.1. We extend the result to one dimension higher, to asymmetric cartesian products, and to both at the same time. The proofs are based on the proof of Theorem 2.1.1 by Elekes and Rónyai.

First we consider a less symmetric cartesian product.

Theorem 2.1.2. *For every $c > 0$ and positive integer d there exist $n_0 = n_0(c,d)$ and $\tilde{c} = \tilde{c}(c,d)$ with the following property.*
Let $f(x,y)$ be a polynomial of degree d in $\mathbb{R}[x,y]$ such that for an $n > n_0$ the graph $z = f(x,y)$ contains $cn^{11/6}$ points of $A \times B \times C$, where $A, B, C \subset \mathbb{R}$ and $|A| = n, |B| = \tilde{c}n^{5/6}$, and $|C| = n$. Then either

$$f(x,y) = g(k(x) + l(y)), \quad \text{or} \quad f(x,y) = g(k(x) \cdot l(y)),$$

where $g, k, l \in \mathbb{R}[t]$.

Using a recent result of Amirkhanyan, Bush, Croot and Pryby [2], regarding a conjecture of Solymosi about the number of lines in general position that can contain cn points of an $n \times n$ cartesian product (see Theorem

2.2.6), we get the following theorem. we note that only Theorems 2.1.3, 2.1.6, and 2.5.2 depend on it. Theorems 2.1.2 and 2.1.5 instead use our Lemma 2.2.4, a similar but weaker statement about lines on cartesian products, which we prove using the same technique as our other theorems.

Theorem 2.1.3. *For every $c > 0$ and positive integer d there exists $n_0 = n_0(c, d)$ with the following property.*
Let $f(x, y)$ be a polynomial of degree d in $\mathbb{R}[x, y]$ such that for an $n > n_0$ the graph $z = f(x, y)$ contains $cn^{3/2+\varepsilon}$ points of $A \times B \times C$, where $A, B, C \subset \mathbb{R}$ and $|A| = n, |B| = n^{1/2+\varepsilon}$ with $\varepsilon > 0$, and $|C| = n$. Then either

$$f(x, y) = g(k(x) + l(y)), \quad \text{or} \quad f(x, y) = g(k(x) \cdot l(y)),$$

where $g, k, l \in \mathbb{R}[t]$.

We also extend the Elekes-Rónyai Theorem to cartesian products of one dimension higher, i.e. to polynomials with one more variable.

Theorem 2.1.4. *For every $c > 0$ and positive integer d there exists $n_0 = n_0(c, d)$ with the following property.*
Let $f(x, y, z)$ be a polynomial of degree d in $\mathbb{R}[x, y, z]$ such that for an $n > n_0$ the graph $w = f(x, y, z)$ contains cn^3 points of $A \times B \times C \times D$, where $A, B, C, D \subset \mathbb{R}$ have size n each. Then either

$$f(x, y, z) = g(k(x) + l(y) + m(z)), \quad \text{or} \quad f(x, y, z) = g(k(x) \cdot l(y) \cdot m(z)),$$

where $g, k, l, m \in \mathbb{R}[t]$.

We can also prove a higher-dimensional version with a less symmetric cartesian product.

Theorem 2.1.5. *For every $c > 0$ and positive integer d there exists $n_0 = n_0(c, d)$ with the following property.*
Let $f(x, y, z)$ be a polynomial of degree d in $\mathbb{R}[x, y, z]$ such that for an $n > n_0$ the graph $w = f(x, y, z)$ contains $cn^{8/3+2\varepsilon}$ points of $A \times B \times C \times D$, where $A, B, C, D \subset \mathbb{R}$ and $|A| = n$, $|B| = |C| = n^{5/6+\varepsilon}$ with $\varepsilon > 0$, and $|D| = n$. Then either

$$f(x, y, z) = g(k(x) + l(y) + m(z)), \quad \text{or} \quad f(x, y, z) = g(k(x) \cdot l(y) \cdot m(z)),$$

where $g, k, l, m \in \mathbb{R}[t]$.

And using the previously mentioned result of Amirkhanyan et al. we get the following:

Theorem 2.1.6. *Given $c > 0$ and d a positive integer there exists $n_0 = n_0(c, d)$ with the following property.*
Let $f(x, y, z)$ be a polynomial of degree d in $\mathbb{R}[x, y, z]$ such that for an $n > n_0$ the graph $w = f(x, y, z)$ contains $cn^{2+2\varepsilon}$ points of $A \times B \times C \times D$, where $A, B, C, D \subset \mathbb{R}$ and $|A| = n$, $|B| = |C| = n^{1/2+\varepsilon}$ with $\varepsilon > 0$, and $|D| = n$. Then either

$$f(x, y, z) = g(k(x) + l(y) + m(z)), \quad \text{or} \quad f(x, y, z) = g(k(x) \cdot l(y) \cdot m(z)),$$

where $g, k, l, m \in \mathbb{R}[t]$.

In Section 2.5.3 we will give an example of a polynomial $f(x, y, z)$ whose graph contains cn^2 points of $A \times B \times C \times D$, where $|A| = |D| = n$ and $|B| = |C| = c'n^{1/2}$, but f does not have the required additive or multiplicative form of Theorem 2.1.6. This shows that Theorem 2.1.6 is near-optimal.

Note that as for the two-variable case, the converses of Theorems 2.1.4–2.1.6 all hold for some appropriately chosen cartesian products. Specifically, if $f(x, y, z) = g(k(x)+l(y)+m(z))$, one can choose A, B, and C so that $k(x)$, $l(y)$, and $m(z)$ have values in the same arithmetic progression. A similar construction works for the product case.

Theorems 2.1.1, 2.1.2, 2.1.4 and 2.1.5 would all hold if we consider functions over $\mathbb{C}$ instead of $\mathbb{R}$, but we will restrict ourselves to $\mathbb{R}$ here. The proofs for three-variable polynomials could be extended to $|B| \neq |C|$, at some cost to the exponents. It also seems possible to generalize our proofs to polynomials with even more variables.

In Section 2.1.3 we give a short outline of the proof of the Elekes-Rónyai Theorem, which provides a template for our subsequent proofs. Section 2.2 contains a number of concepts and results required throughout our proofs. In Section 2.3 we give the proofs of Theorems 2.1.2 and 2.1.3, while Section 2.4 contains the proofs of Theorems 2.1.4, 2.1.5, and 2.1.6. In Section 2.5 we describe a method of Elekes and Rónyai for checking whether a function has such an additive or multiplicative form. We also give an extension of the conjecture of Purdy, and an example showing the near-optimality of Theorem 2.1.6.

2.1.3 Outline of proofs

The following is an outline of the proof that Elekes and Rónyai gave in [21] of Theorem 2.1.1. Our theorems are obtained by adjusting this proof to three-variable f, and by using improved Line Lemmas (see Section 2.2.2) to get the asymmetric versions. All functions below are polynomials, and we

repeatedly recycle the positive constant c. We call a curve k-rich on a set if it contains at least k points from that set.

We split up the surface $z = f(x, y)$ into the n curves

$$z = f_i(x) = f(x, y_i)$$

with $y_i \in B$. We wish to decompose a cn-sized subset of the f_i as

$$f_i(x) = (p \circ \varphi_i \circ k)(x) = p(a_i k(x) + b_i),$$

where φ_i is linear and p and k are independent of i.

Then the cn lines $u = \varphi_i(t) = a_i t + b_i$ will be cn-rich on an $n \times n$ cartesian product. For such sets of lines we have various lemmas (2.2.3–2.2.7) that say that a cn-sized subset of them must be all parallel or all concurrent (have the same intersection point).

Given cn such decompositions with the lines φ_i all parallel, we can write $f(x, y_i) = p(ak(x) + b_i)$, and then conclude by an algebraic argument that there exists an $l(y)$ such that $f(x, y) = p(k(x) + l(y))$. If cn of the lines are concurrent, we can write $f(x, y_i) = p(a_i \cdot (k(x) + b))$, and then conclude that $f(x, y) = p(k(x) \cdot l(y))$.

To find the above decomposition of the f_i, we first remove their common inner functions (polynomials μ such that $f_i = \lambda_i \circ \mu$) up to linear equivalence. We can do this because the number of decompositions up to linear equivalence of a polynomial of degree d depends only on d (Lemma 2.2.10), so for large enough n there must be a cn-sized subset of the f_i that all have the same inner function of maximal degree. This maximal inner function will be the k above, and we remove it by writing $f_i = \widehat{f_i} \circ k$. Then we have a subset of $\widehat{f_i}$ with the property that if $\widehat{f_i} = \mu_i \circ \lambda$ and $\widehat{f_j} = \mu_j \circ \lambda$, then λ must be linear.

Now we combine pairs $\widehat{f_i}, \widehat{f_j}$ into new curves

$$\gamma_{ij}(t) = (\widehat{f_i}(t), \widehat{f_j}(t)).$$

We observe that these γ_{ij} are cn-rich on an $n \times n$ cartesian product, and that we have cn^2 of them. But by a theorem of Pach and Sharir (Lemma 2.2.2), such a set of rich curves can have size at most $c'n$.

This is not a contradiction: many of these γ_{ij} may coincide as sets in $\mathbb{R}^2$. But if for instance γ_{ij} and $\gamma_{i'j'}$ coincide, then by some algebra (Lemma 2.2.12) they must be reparametrizations of the same curve $(p(t), q(t))$, which means that we can write

$$\widehat{f_i} = p \circ \varphi, \qquad \widehat{f_j} = q \circ \varphi,$$
$$\widehat{f_{i'}} = p \circ \phi, \qquad \widehat{f_{j'}} = q \circ \phi.$$

We take one of these common decompositions, say with φ. Since we already removed all nonlinear common inner polynomials, φ must be linear. If we have enough such decompositions, we can ensure that they all have the form $f_i = p \circ \varphi_i$ for the same p. This gives us the desired decompositions

$$f_i = \widehat{f_i} \circ k = p \circ \varphi_i \circ k.$$

2.2 Preliminaries

2.2.1 Discrete geometry

We will make frequent use of the following well-known theorem, first proved in [64]. We say that a line (or any other curve) is *k-rich* on a point set $\mathcal{P}$ if it contains at least k points of $\mathcal{P}$.

Theorem 2.2.1 (Szemerédi-Trotter Theorem)**.** *There exists a positive constant $C_{ST} > 0$ such that given a set $\mathcal{P}$ of n points in $\mathbb{R}^2$, the number of lines k-rich on $\mathcal{P}$ is at most $C_{ST} \cdot (n^2/k^3 + n/k)$.*

This theorem was generalized by Pach and Sharir [52, 53] to continuous real planar curves without self-intersection. We will use the following corollary for algebraic curves, which follows quite easily since algebraic curves (of bounded degree) can be split up into a small number (depending on the degree) of curves without self-intersection. For details see Elekes and Rónyai [21].

Lemma 2.2.2 (Curve Lemma)**.** *Given $c > 0$ and a positive integer d, there exist $C_{CL} = C_{CL}(c, d)$ and $n_0 = n_0(c, d)$ such that the following holds.*
Given m distinct irreducible real algebraic curves of degree $\leq d$ that are cn-rich on A, where $A \subset \mathbb{R}^2$ and $|A| \leq n^2$, then for all $n > n_0$ we have

$$m < C_{CL} \cdot n.$$

2.2.2 Line lemmas

In the proof of Theorem 2.1.1 by Elekes and Rónyai, an important ingredient was the following result of Elekes [18] about lines containing many points from a cartesian product. We call a set of lines *concurrent* if they all have the same intersection point.

Lemma 2.2.3 (Line Lemma)**.** *Suppose $A, B \subset \mathbb{R}$ and $|A| = |B| = n$. For all $c_1, c_2 > 0$ there exists $C_{LL} > 0$, independent of n, such that if $C_{LL}n$ lines in $\mathbb{R}^2$ are $c_1 n$-rich on $A \times B$, then at least $c_2 n$ of them must be parallel or concurrent.*

We prove a generalization that will be crucial in Section 2.3. The proof is at the end of this section, and is modelled on that of Elekes.

Lemma 2.2.4 (Generalized Line Lemma)**.** *Suppose $A, B \subset \mathbb{R}$ and $|A| = |B| = n$. For all $c_1, c_2 > 0$, and $\beta \geq 0$ there exists $C_{GLL} > 0$, independent of n, such that if m lines in $\mathbb{R}^2$ are $c_1 n$-rich on $A \times B$, with no $c_2 n^\beta$ concurrent or parallel, then*

$$m < C_{GLL} \cdot n^{2/3+\beta/3}.$$

A collection of lines in $\mathbb{R}^2$ is said to be in *general position* if no two lines are parallel and no three lines are concurrent. The second author conjectured the following extension of the above result. For details see [20].

Conjecture 2.2.5. *Suppose $A, B \subset \mathbb{R}$ and $|A| = |B| = n$. For all $c > 0$ there exists $C_S > 0$ such that if m lines in general position are cn-rich on $A \times B$ then $m < C_S$.*

The following result of Amirkhanyan et al. is related to the above conjecture [2].

Theorem 2.2.6. *For every $\varepsilon > 0$ there exists $\delta > 0$ such that given n^ε lines in $\mathbb{R}^2$ in general position, they cannot all be $n^{1-\delta}$-rich on $A \times B$, where $|A| = |B| = n$.*

Thus if a collection of lines $\mathcal{L}$ in general position is cn-rich on $A \times B$ then $|\mathcal{L}| < n^\varepsilon$ for any $\varepsilon > 0$. We will use it in the form of the following corollary.

Corollary 2.2.7. *If m lines in $\mathbb{R}^2$ are cn-rich on $A \times B$, with $|A| = |B| = n$, such that no p are parallel and no q are concurrent, then*

$$m \leq (p+q)n^\varepsilon$$

for every $\varepsilon > 0$.

Proof. We show that the set of lines contains at least $k = \sqrt{m}/\sqrt{2(p+q)}$ lines in general position.

We pick any line, and then successively choose a new line that is not parallel to any of the previously chosen lines, and does not go through the intersection point of any pair of them. If we have chosen k such lines, then there are k slopes we may not choose, which excludes less than pk lines. And there are at most $\binom{k}{2}$ intersection points that we must avoid, so since there are less than q lines concurrent at a point, this excludes less than $q\binom{k}{2}$ lines. Hence we can continue in this way at least until $m \leq q\binom{k}{2} + pk + k$.

Then we also have $m \leq (p+q)k^2$, so we can get $k \geq \sqrt{m}/\sqrt{p+q}$ lines in general position. These lines are cn-rich on $A \times B$. Thus $k \leq n^{\varepsilon'}$ for every $\varepsilon' > 0$. This gives $m \leq (p+q)n^{\varepsilon}$ for every $\varepsilon > 0$. □

We give the proof of Lemma 2.2.4 below. We will use the dual of a theorem of Beck [7], which roughly states that given a collection of points, either 'many' of the points are on the same line, or pairs of the points determine 'many' distinct lines.

Theorem 2.2.8 (Dual of Beck's Theorem)**.** *There exists $C_{BT} > 0$ such that, given N lines in $\mathbb{R}^2$, either $C_{BT}N$ lines are concurrent or the lines determine $C_{BT}N^2$ distinct pairwise intersection points.*

Proof of Lemma 2.2.4. Let L be the set of lines and write $|L| = m = cn^{\alpha}$. For every pair $(\ell_i, \ell_j) \in L^2$ we define the linear functions $\gamma_{ij} = \ell_i \circ \ell_j^{-1}$ and $\Gamma_{ij} = \ell_j^{-1} \circ \ell_i$.

First we will prove that large subsets of the γ_{ij} and Γ_{ij} are also rich. Consider the tripartite graph H with vertex sets $A \cup L \cup B$. Given $a \in A$ and $\ell \in L$, (ℓ, a) is an edge in H if $\ell(a) \in B$. Similarly, given $\ell \in L$ and $b \in B$, (ℓ, b) is an edge if $\ell^{-1}(b) \in A$. Given $\ell \in L$, let $\deg_A(\ell)$ be the number of edges between ℓ and A and $\deg_B(\ell)$ the number of edges between l and B. Since the lines in L are $c_1 n$-rich on $A \times B$, we have $\deg_A(\ell) \geq c_1 n$ and $\deg_B(\ell) \geq c_1 n$ for each $\ell \in L$. Thus we have at least $cc_1 n^{1+\alpha}$ edges between A and L and at least $cc_1 n^{1+\alpha}$ edges between B and L.

We will count cycles of length four in H with one vertex in A and one vertex in B. Every such C_4 gives a point in $B \times B$ on $y = \gamma_{ij}(x)$ and a point in $A \times A$ on $y = \Gamma_{ij}(x)$ for some pair (i, j). The number of paths of length two with one endpoint in A and the other in B is at least

$$\#P_2 = \sum_{\ell \in L} \deg_A(\ell) \deg_B(\ell) \geq c_1^2 n^{2+\alpha}.$$

Let $p_{a,b}$ be the number of paths of length two between $a \in A$ and $b \in B$. Then

$$\#P_2 = \sum_{a \in A, b \in B} p_{a,b}.$$

Now, by Jensen's Inequality, the number of C_4's we are looking for is

$$\#C_4 = \sum_{a \in A, b \in B} \binom{p_{a,b}}{2} \geq |A \times B| \binom{\#P_2/|A \times B|}{2} \geq \frac{c_1^4 n^{2+2\alpha}}{4}.$$

Suppose there are fewer than $(c_1^4/8)n^{2\alpha}$ pairs $(\ell_i, \ell_j) \in L^2$ with at least $(c_1^4/8)n^2$ C_4's between them. Then H would have fewer than $(c_1^4/4)n^{2+2\alpha}$ C_4's, a contradiction. Thus, setting $c_3 = c_1^4/8$, we have at least $c_3 n^{2\alpha}$ pairs (i, j) for which γ_{ij} and Γ_{ij} are $c_3 n$-rich on $B \times B$ and $A \times A$ respectively.

Next we define a different graph G' and analyze it. The vertex sets of G' consist of those γ_{ij} that are $c_3 n$-rich on $B \times B$ and those $\Gamma_{i'j'}$ that are $c_3 n$-rich on $A \times A$. If γ_{ij} and $\gamma_{i'j'}$ coincide as point sets, we consider them as the same vertex. Similarly we identify any coinciding Γ_{ij} and $\Gamma_{i'j'}$, but we do not identify γ_{ij} and $\Gamma_{i'j'}$ should they coincide. We place an edge between γ_{ij} and Γ_{ij} for each pair (i, j), which means the graph is bipartite.

The graph may contain multiple edges, if we have $\ell_i, \ell_j, \ell_{i'}, \ell_{j'} \in L$ such that both $\gamma_{ij} = \gamma_{i'j'}$ and $\Gamma_{ij} = \Gamma_{i'j'}$. But this implies that the four lines are concurrent: If ℓ_i and ℓ_j intersect in (u, v), then $v = \ell_i \circ \ell_j^{-1}(v) = \ell_{i'} \circ \ell_{j'}^{-1}(v)$ and $u = \ell_j^{-1} \circ \ell_i(u) = \ell_{j'}^{-1} \circ \ell_{i'}(u)$, so (u, v) is also the intersection point of $\ell_{i'}$ and $\ell_{j'}$.

But with Beck's Theorem we can get a subgraph without multiple edges. We will assume that $\beta < \alpha$, and check it at the end of the proof. Then fewer than $c_2 n^\beta < C_{BT} c n^\alpha = C_{BT}|L|$ lines are concurrent, so by Theorem 2.2.8, the lines determine $C_{BT} n^{2\alpha}$ distinct intersection points. The corresponding lines span a subgraph G'' without multiple edges, and at least $C_{BT} n^{2\alpha}$ edges.

By the Szemerédi-Trotter Theorem, since all vertices are $c_3 n$-rich lines, the number of vertices is at most $\leq c_4 n$ for some constant $c_4 > 0$. The average degree in G'' is then $\geq (C_{BT}/c_4) n^{2\alpha-1}$. Thus G'' contains a connected component H containing at least $(C_{BT}/c_4) n^{2\alpha-1}$ vertices and at least $\frac{1}{2}(C_{BT}/c_4)^2 n^{4\alpha-2}$ edges.

Note that each γ_{ij} and Γ_{ij} have the same slope, so every vertex in H is a line with the same slope. If there are more than $c c_2 n^{\alpha+\beta}$ edges in H then we would have vertices $\gamma_{ij_1}, \gamma_{ij_2}, \ldots, \gamma_{ij_t}$ in this component with $t \geq c_2 n^\beta$. This implies that the lines $\ell_{j_1}, \ell_{j_2}, \ldots, \ell_{j_t}$ are all parallel, which is a contradiction. So we have a contradiction unless

$$\frac{C_{BT}^2}{2c_4^2} n^{4\alpha-2} < c c_2 n^{\alpha+\beta}.$$

From this we get that $\alpha \leq 2/3 + \beta/3$, if we choose $C_{GLL} = c > C_{BT}^2/(2c_2c_4^2)$. □

2.2.3 Algebra and graph theory

In the proofs of our higher-dimensional versions of the Elekes-Rónyai Theorem, we will need the following generalization of the fact that if a degree-d

polynomial of one variable has $d+1$ or more roots, then the polynomial is identically zero. It is a special case of the Schwartz-Zippel lemma (see for instance [36], Lemma 16.3).

Lemma 2.2.9 (Vanishing Lemma)**.** *Let K be a field, $F(y,z) \in K[y,z]$ with $\deg F = d$, and $B, C \subset K$ with $|B| = |C| = m$.*

If $F(y_i, z_j) = 0$ for $2dm$ of the pairs $(y_i, z_j) \in B \times C$, then $F(y,z) = 0$.

We will also need the following three algebraic lemmas, which appear with proofs in [21]. Let K be a field. We call two decompositions $f(x) = \varphi_1(\psi_1(x))$ and $f(x) = \varphi_2(\psi_2(x))$ of a polynomial $f \in K[x]$ into polynomials from $K[x]$ *equivalent* if $\psi_1(t) = a\psi_2(t) + b$ for some $a, b \in K$.

Lemma 2.2.10. *Let K be a field. Then no $f \in K[x]$ can have more than 2^d non-equivalent decompositions, where $d = \deg f$.*

Lemma 2.2.11.

1. *Let E be a field, $\varphi \in E[x]$ a polynomial of degree $d > 0$. Then every $F \in E[x]$ can be written in the form*

$$F = a_0 + a_1 x + \cdots + a_{d-1} x^{d-1},$$

   where $a_i \in E(\varphi)$, in a unique way.

2. *Suppose further that $E = L(y)$ is a rational function field over some field L, and $\varphi \in L[x]$. Let m be the degree of F in y. Then the degree of a_i in y is at most $m(d+1)$. (Here a_i is viewed as a polynomial of φ and y over L.)*

Lemma 2.2.12 (Reparametrization Lemma)**.** *Suppose that two parametric curves $(f_1(t), g_1(t))$ and $(f_2(t), g_2(t))$ coincide as sets, with $f_i, g_i \in K[t]$ for a field K. Then there are $p, q, \varphi_1, \varphi_2 \in K[t]$ such that*

$$f_1 = p \circ \varphi_1, \qquad g_1 = q \circ \varphi_1,$$
$$f_2 = p \circ \varphi_2, \qquad g_2 = q \circ \varphi_2.$$

Finally, we need the following graph-theoretic lemma, also proved in [21].

Lemma 2.2.13 (Graph Lemma)**.** *For every c and k there is a $C_{GL} = C_{GL}(c,k)$ with the following property.*
If a graph has N vertices and cN^2 edges, and the edges are colored so that at most k colors meet at each vertex, then it has a monochromatic subgraph with $C_{GL}N^2$ edges.

2.3 Proof of Theorems 2.1.2 and 2.1.3

Suppose $z = f(x, y)$ contains $cn^{\alpha+1}$ points of $A \times B \times C$, where $|A| = |C| = n$ and $|B| = \tilde{c}n^{\alpha}$; $\tilde{c}$ and α will be determined later. Throughout we will use $d = \deg f$. All functions will be polynomials.

2.3.1 Constructing $\widehat{f_i}$

For each of the $\tilde{c}n^{\alpha}$ $y_i \in B$ define

$$f_i(x) = f(x, y_i).$$

Then each f_i is a polynomial in $\mathbb{R}[x]$ of degree at most d.

Lemma 2.3.1. *If at least $d+1$ of the f_i are identical, then $f(x, y) = q(x)$. In particular, the conclusion of Theorems 2.1.2 and 2.1.3 holds.*

Proof. Suppose that $f_i(x) = q(x)$ at least $d+1$ times. Then considering $F(y) = f(x, y) - q(x)$ as a polynomial in y over the field $\mathbb{R}(x)$, we have $F(y)$ vanishing $d+1$ times, so $F(y) = 0$ identically. □

Assumption: Throughout the rest of this section we will assume that at most d of the f_i are identical.

Let $c_1 = \min(c/2, \tilde{c}/2)$. Then at least $c_1 n^{\alpha}$ of the f_i are $c_1 n$-rich on $A \times C$. Otherwise $z = f(x, y)$ would contain fewer than $cn^{\alpha+1}$ points of $A \times B \times C$.

We construct a graph G with the $c_1 n^{\alpha}$ $c_1 n$-rich f_i as vertices and edge set E consisting of all pairs (f_i, f_j).

Lemma 2.3.2. *There is a subgraph of G with edge set $\widehat{E} \subset E$ of size $|\widehat{E}| \geq c_2 n^{2\alpha}$, such that the following holds. There is a polynomial $k(x)$ such that for all $(f_i, f_j) \in \widehat{E}$ we can write*

$$f_i = \widehat{f_i} \circ k,$$

$$f_j = \widehat{f_j} \circ k,$$

and $\widehat{f_i}$ and $\widehat{f_j}$ share no non-linear common inner function.
The $\widehat{f_i}$ are also $c_1 n$-rich on $k(A) \times C$.

Proof. Color each edge (f_i, f_j) of G with the equivalence class of a common inner function φ of maximum degree, i.e. $f_i(x) = g(\varphi(x))$ and $f_j(x) = h(\varphi(x))$, and no such φ of higher degree exists; two such inner functions φ, ϕ are equivalent if $\phi(x) = a\varphi(x) + b$.

By Lemma 2.2.10, at every vertex there are at most 2^d colors, so by the Graph Lemma 2.2.13, with $N = c_1 n^\alpha$, there is a monochromatic subgraph with $C_{GL}N^2 = C_{GL}c_1^2 n^{2\alpha}$ edges. We take $\widehat{E}$ to be the edge set of this subgraph. This means that all the f_i involved in this subgraph have a common inner function $k(x)$ (actually up to equivalence, but by modifying the $\widehat{f_i}$ that is easily overcome), and no pair corresponding to an edge of $\widehat{E}$ has a common inner function of higher degree. That allows us to define the $\widehat{f_i}$ as in the theorem; they must be rich since otherwise the f_i could not be rich. □

2.3.2 Constructing γ_{ij}

For the $c_2 n^{2\alpha}$ pairs $\widehat{f_i}, \widehat{f_j}$ for which $(f_i, f_j) \in \widehat{E}$, we construct the curves

$$\gamma_{ij}(t) = \left(\widehat{f_i}(t), \widehat{f_j}(t)\right).$$

Lemma 2.3.3.

1. *At least $c_3 n^{2\alpha}$ γ_{ij} are $c_3 n$-rich on $C \times C$.*
2. *Each γ_{ij} is an irreducible algebraic curve of degree at most d^2.*

Proof. 1. We define a bipartite graph with vertex set $k(A) \cup \{\widehat{f_i}\}$, and we connect $t \in k(A)$ with $\widehat{f_i}$ if $\widehat{f_i}(t) \in C$. Since $|\widehat{E}| \geq c_2 n^{2\alpha}$, the number of $\widehat{f_i}$ is at least $\sqrt{c_2} n^\alpha$, each of them $c_2 n$-rich, so the bipartite graph has $m \geq c_2^{3/2} n^{\alpha+1}$ edges. We count the paths of length two between different $\widehat{f_i}$'s, using the fact that $k(A) \leq n$:

$$\#P_2 = \sum_{t \in k(A)} \binom{\deg(t)}{2} \geq |k(A)| \binom{m/|k(A)|}{2} \geq c' n^{2\alpha+1}.$$

Hence at least $c'' n^{2\alpha}$ pairs $(\widehat{f_i}, \widehat{f_j})$ share $c'' n$ common neighbors t in this graph. In other words, $c'' n^{2\alpha}$ of the γ_{ij} have a point in $C \times C$ for $c'' n$ different t.

It is possible that different t give the same point $\gamma_{ij}(t)$, so these γ_{ij} could have fewer than $c'' n$ points in $C \times C$. However, because $\deg \widehat{f_i} \leq d$, this can happen for at most d different t at a time, so each γ_{ij} will certainly be $(c''/d)n$-rich. Setting $c_3 = c''/d$ we are done.

2. We require the notion of the resultant of two polynomials; for proofs of the following facts see [48]. Let $R(x, y)$ be the resultant with respect to t (so considering x, y as coefficients) of the two polynomials $x - \widehat{f}_i(t)$ and $y - \widehat{f}_j(t)$. This is a polynomial of degree $\leq d^2$ with the property that $R(x, y) = 0$ if and only if there is a t such that $x = \widehat{f}_i(t)$ and $y = \widehat{f}_j(t)$; in other words, γ_{ij} is the algebraic curve $R(x, y) = 0$. By Theorem 1 in [48], $R(x, y)$ is a power of an irreducible polynomial, which means that the curve γ_{ij} defined by it is irreducible.

□

2.3.3 Decomposing $\widehat{f}_i$

Lemma 2.3.4. *There is a subset S of $c'n^{2\alpha-1}$ of the γ_{ij} that all coincide as point sets, and such that the set T of $\widehat{f}_i$ occurring in the first coordinate of a $\gamma_{ij} \in S$ has size $c_4 n^{2\alpha-1}$.*

Proof. Since the γ_{ij} are irreducible and have degree $\leq d^2$, we can apply the Curve Lemma 2.2.2. Thus there exists n_0 such that for $n > n_0$, there can be at most $C_{CL}n$ distinct $c_3 n$-rich curves on $C \times C$, so $c_3 n^{2\alpha}/C_{CL}n = c'n^{2\alpha-1}$ of them must coincide.

Set $c_4 = c'/d^2$. If fewer than $c_4 n^{2\alpha-1}$ of the $\widehat{f}_i$ occurred among these coinciding γ_{ij}, then some $\widehat{f}_i$ would have to occur at least $d+1$ times, say in the first coordinate. But if $(\widehat{f}_i, \widehat{f}_j)$ and $(\widehat{f}_i, \widehat{f}_{j'})$ coincide, then we must have $\widehat{f}_j = \widehat{f}_{j'}$. So we would have $d+1$ of the $\widehat{f}_i$ coinciding, hence also $d+1$ of the f_i, contradicting our Assumption after Lemma 2.3.1. □

Lemma 2.3.5. *There are $c_4 n^{2\alpha-1}$ f_i with*

$$f_i(x) = p(a_i k(x) + b_i)$$

where $a_i, b_i \in \mathbb{R}$ and $p \in \mathbb{R}[x]$.

Proof. By the Reparametrization Lemma 2.2.12, for each coinciding pair of curves γ_{ij} and $\gamma_{i'j'}$ from S, we can find p, φ_i, and $\varphi_{i'}$ such that

$$\widehat{f}_i = p \circ \varphi_i \quad \text{and} \quad \widehat{f}_{i'} = p \circ \varphi_{i'}.$$

Hence we have such decompositions for each pair of the $\widehat{f}_i \in T$.

The $\widehat{f}_i$ were constructed so that any pair corresponding to an edge of $\widehat{E}$ has no nonlinear common inner function. That implies that the φ_i are linear, hence invertible, which allows us to assume that all $\widehat{f}_i \in T$ can be decomposed using the same p. Indeed, if $\widehat{f}_i = p \circ \varphi_i = q \circ \phi_i$ and $\widehat{f}_j = q \circ \phi_j$,

then $q = p \circ (\varphi_i \circ \phi_i^{-1})$, so we can write $\widehat{f_j} = p \circ (\varphi_i \circ \phi_i^{-1} \circ \phi_j)$; by repeatedly modifying the φ_i this way we can reach all $\widehat{f_k} \in T$.

Write $\varphi_i(t) = a_i t + b_i$; then for the $c_4 n^{2\alpha-1}$ $f_i = \widehat{f_i} \circ k$ with $\widehat{f_i} \in T$ we have $f_i = p \circ \varphi_i \circ k$. □

2.3.4 Proof of Theorem 2.1.2

At this point we will apply the Generalized Line Lemma 2.2.4 with $\beta = 0$ to obtain Theorem 2.1.2. Then we need $2\alpha - 1 = 2/3$, so we set $\alpha = 5/6$.

Note that the $c_4 n^{2/3}$ lines $u = \varphi_i(t) = a_i t + b_i$ live on $k(A) \times p^{-1}(C)$, which is essentially an $n \times n$ cartesian product (both sets might be smaller than n, but we can just add arbitrary points to fill them out). They are $c_1 n$-rich there, since otherwise the f_i couldn't be $c_1 n$-rich.

We conclude that either $d+1$ of the lines $u = \varphi_i(t)$ are parallel, or $d+1$ are concurrent. Otherwise, by Lemma 2.2.4 with $\beta = 0$ there would be fewer than $C_{GLL} n^{2/3}$ lines. But we can take $\tilde{c}$ in Theorem 2.1.2 to be large enough so that $c_4 > C_{GLL}$. Indeed, one can easily check that each c_i was an increasing unbounded function of c_{i-1}.

By Lemma 2.3.6 below, if $d+1$ of the lines are parallel, then f has the additive form $f(x, y) = p(k(x) + l(y))$. By Lemma 2.3.7 below, if $d+1$ of the lines are concurrent, then f has the multiplicative form $f(x, y) = p(k(x) \cdot l(y))$. That finishes the proof of Theorem 2.1.2.

2.3.5 Proof of Theorem 2.1.3

We will now use Corollary 2.2.7, instead of Lemma 2.2.4 as above, which will result in Theorem 2.1.3.

We start with $\alpha = 1/2 + \varepsilon$. Then we end up with $c_4 n^{2\alpha-1} = c_4 n^{2\varepsilon}$ lines $u = a_i t + b_i$ which are $c_1 n$-rich on an $n \times n$ cartesian product. Certainly $c_4 n^{2\varepsilon} > 2(d+1) n^{\varepsilon'}$ for some $\varepsilon' > 0$, so by Corollary 2.2.7 with $p = q = d+1$ either $d+1$ of the lines are parallel or $d+1$ are concurrent.

By Lemma 2.3.6 below, the parallel case would give the additive form for f, and Lemma 2.3.7 below, the concurrent case would give the multiplicative form for f. That finishes the proof of Theorem 2.1.3.

2.3.6 The parallel case

Lemma 2.3.6. *If $d+1$ of the lines φ_i are parallel, then there is a polynomial $l(y)$ such that $f(x, y) = p(k(x) + l(y))$.*

Proof. The lines can be written as $\varphi_i(t) = at + b_i$, so (replacing k by k/a) we can write $f_i(x) = p(k(x)+b_i)$, for $d+1$ different f_i. We use the following two polynomial expansions of $f_i(x) = f(x, y_i) = p(k(x) + b_i)$:

$$\sum_{l=0}^{N} v_l \cdot (k(x) + b_i)^l = \sum_{m=0}^{N} w_m(y_i) \cdot k(x)^m.$$

The first is immediate from $p(k(x) + b_i)$; the second requires a little more thought.

By Lemma 2.2.11, there is a unique expansion of the polynomial f of the form $f(x, y) = \sum_{l=0}^{D-1} c_l(k(x), y)x^l$, where $D = \deg k$. By the same lemma, we have a unique expansion $f_i(x) = \sum_{l=0}^{D-1} d_l(k(x))x^l$, so that we have

$$\sum_{l=0}^{D-1} c_l(k(x), y_i)x^l = \sum_{l=0}^{D-1} d_l(k(x))x^l \quad \Rightarrow \quad c_l(k(x), y_i) = d_l(k(x)).$$

But since $f_i(x) = p(k(x) + b_i)$, uniqueness implies that $d_l = 0$ for $l > 0$, hence $c_l(k(x), y_i) = 0$ for $l > 0$. Since we have this for $d + 1$ different i, it follows that $c_l(k(x), y) = 0$ for $l > 0$, so $f(x, y) = c_0(k(x), y)$, which means there is an expansion $f(x, y) = \sum w_m(y)k(x)^m$. Now plugging in $y = y_i$ gives the required expansion.

Comparing the coefficients of $k(x)^{N-1}$ in the two expansions above, we get

$$v_{N-1} + (N - 1)v_N b_i = w_{N-1}(y_i),$$

which implies that $b_i = \frac{1}{(N-1)v_N}(w_{N-1}(y_i) - v_{N-1})$. If we now define the polynomial

$$l(y) = \frac{1}{(N - 1)v_N}(w_{N-1}(y) - v_{N-1}),$$

we have that for $d + 1$ of the y_i (note that v_l and w_m do not depend on the choice of y_i)

$$f(x, y_i) = p(k(x) + l(y_i)).$$

Since the degree of f is d, this implies that $f(x, y) = p(k(x) + l(y))$. □

2.3.7 The concurrent case

Lemma 2.3.7. *If $d + 1$ of the lines φ_i are concurrent, then there are polynomials $P(t)$, $K(x)$ and $L(y)$ such that*

$$f(x, y) = P(K(x) \cdot L(y)).$$

Proof. The lines can be written as $\varphi_i(t) = a_i t + b$, so (replacing $p(x)$ by $p(x-b)$) we can write $f_i(x) = p(a_i \cdot k(x))$, for $d+1$ different f_i. We again use two polynomial expansions of $f_i(x) = f(x, y_i) = p(a_i \cdot k(x))$:

$$\sum_{l=0}^{N} v_l \cdot (a_i \cdot k(x))^l = \sum_{m=0}^{N} w_m(y_i) \cdot k(x)^m.$$

Both are obtained in the same way as in the proof of Lemma 2.3.6.

We cannot proceed exactly as before, since a_i might occur here only with exponents, and we cannot take a root of a polynomial. But we can work around that as follows. Define M to be the greatest common divisor of all exponents m for which $w_m \neq 0$ in the second expansion; then we can write M as an integer linear combination of these m, say $M = \sum \mu_m m$. Comparing the coefficients of any $k(x)^m$ with $w_m \neq 0$ in the two expansions above, we get

$$a_i^m = \frac{1}{v_m} w_m(y_i),$$

which tells us that

$$a_i^M = \prod_{m=0}^{N} (a_i^m)^{\mu_m} = L(y_i),$$

where $L(y)$ is a rational function.

If we define $P(s) = p(s^{1/M})$, or equivalently $P(t^M) = p(t)$, then the definition of M gives that $P(s)$ is a polynomial. We also define $K(x) = k^M(x)$. Then

$$P(K(x) \cdot L(y_i)) = P(k^M(x) \cdot a_i^M) = p(k(x) \cdot a_i) = f(x, y_i).$$

Since we have this for $d+1$ of the y_i, we get that $f(x, y) = P(K(x)L(y))$. This also tells us that $L(y)$ is in fact a polynomial, since otherwise $f(x, y)$ could not be one. □

2.4 Proof of Theorems 2.1.4, 2.1.5, and 2.1.6

Suppose $w = f(x, y, z)$ contains $cn^{1+2\alpha}$ points of $A \times B \times C \times D$ and $|B| = |C| = n^\alpha$. For Theorem 2.1.4 we have $\alpha = 1$; for the other two theorems we will determine the right choice of α later. Throughout we will use $d = \deg f$. All functions will be polynomials. We will shorten or omit several of the proofs, because they are very similar to those in Section 2.3.

2.4.1 Constructing $\widehat{f}_{ij}$

For each of the $n^{2\alpha}$ points $(y_i, z_j) \in B \times C$, we cut a fibre out of the solid:

$$w = f_{ij}(x) = f(x, y_i, z_j).$$

Lemma 2.4.1. *If at least $2dn^{\alpha}$ of the f_{ij} are identical, then $f(x, y, z) = q(x)$.*
In particular, the conclusion of Theorems 2.1.4, 2.1.5, and 2.1.6 holds.

Proof. Suppose that $f_{ij}(x) = q(x)$ at least $2dn^{\alpha}$ times. Then for $F(y, z) = f(x, y, z) - q(x)$ and $K = \mathbb{R}(x)$, the Vanishing Lemma 2.2.9 with $b = c = n^{\alpha}$ gives $F(y, z) = 0$. □

Assumption: Throughout the rest of this proof we will assume that fewer than $2dn^{\alpha}$ of the f_{ij} are identical.

Let $c_1 = c/2$. Then at least $c_1 n^{2\alpha}$ of the f_{ij} are $c_1 n$-rich on $A \times D$. Otherwise $w = f(x, y, z)$ would contain fewer than $cn^{1+2\alpha}$ points of $A \times B \times C \times D$.

We construct a graph G with the $c_1 n^{2\alpha}$ f_{ij} as vertices and edge set E consisting of the pairs $(f_{ij}, f_{i'j'})$.

Lemma 2.4.2. *There is a subgraph of G with edge set $\widehat{E} \subset E$ of size $|\widehat{E}| \geq c_2 n^{4\alpha}$, such that the following holds. There is a polynomial $k(x)$ such that for all $(f_{ij}, f_{i'j'}) \in \widehat{E}$ we can write*

$$f_{ij} = \widehat{f}_{ij} \circ k,$$

$$f_{i'j'} = \widehat{f}_{i'j'} \circ k,$$

and $\widehat{f}_{ij}$ and $\widehat{f}_{i'j'}$ share no non-linear inner function.
The $\widehat{f}_{ij}$ are also $c_2 n$-rich on $k(A) \times D$.

2.4.2 Constructing $\gamma_{iji'j'}$

For the $c_2 n^{4\alpha}$ pairs $\widehat{f}_{ij}, \widehat{f}_{i'j'}$ for which $(f_{ij}, f_{i'j'}) \in \widehat{E}$ we construct the curves

$$\widehat{\gamma}_{iji'j'}(t) = \left(\widehat{f}_{ij}(t), \widehat{f}_{i'j'}(t)\right),$$

Lemma 2.4.3.

1. *At least $c_3 n^{4\alpha}$ of the $\gamma_{iji'j'}$ are $c_3 n$-rich on $D \times D$.*

2. Each $\gamma_{iji'j'}$ is an irreducible algebraic curve of degree at most d^2.

Proof. We omit the proof of the second part of the lemma as it is proved in the same way as the second part of Lemma 2.3.3. The proof of the first part follows.

We define a bipartite graph with vertex set $E = k(A) \cup \{\widehat{f_{ij}}\}$, and we connect $t \in k(A)$ with f_{ij} if $f_{ij}(t) \in D$. Then this graph has $m = c_2^{3/2} n^{1+2\alpha}$ edges. We count the 2-paths:

$$\#P_2 = \sum_{x \in k(A)} \binom{d(x)}{2} \geq |k(A)| \binom{m/|k(A)|}{2} \geq c' n^{1+4\alpha}.$$

Hence at least $c'' n^{4\alpha}$ pairs $\widehat{f_{ij}}, \widehat{f_{i'j'}}$ share $c''n$ common neighbors t in this graph. This implies that if $c_3 = c''/d$ then $c_3 n^{4\alpha}$ of the $\gamma_{iji'j'}$ have at least $c_3 n$ points in $D \times D$. □

2.4.3 Decomposing $\widehat{f_{ij}}$

Lemma 2.4.4. *There is a subset of $c_4 n^{4\alpha-1}$ of the $\gamma_{iji'j'}$ that all coincide, and such that $c_4 n^{3\alpha-1}$ of the $\widehat{f_{ij}}$ occur in these $\gamma_{iji'j'}$.*

Proof. By the Curve Lemma, for $n > n_0$, there can be at most $C_{CL} n$ distinct $c_3 n$-rich curves on $D \times D$, so $c' n^{4\alpha-1}$ must coincide. Setting $c_4 = c'/d^2$ gives that at least $c_4 n^{3\alpha-1}$ of the $\widehat{f_{ij}}$ occur.

□

Lemma 2.4.5. *There are $c_4 n^{3\alpha-1}$ pairs (i, j) for which*

$$f_{ij}(x) = p(a_{ij} k(x) + b_{ij})$$

where $a_{ij}, b_{ij} \in \mathbb{R}$ and $p \in \mathbb{R}[x]$.

Proof. For each coinciding pair of curves $\gamma_{iji'j'}$ and $\gamma_{aba'b'}$, we can write

$$\widehat{f_{ij}} = p \circ \varphi_{ij} \quad \text{and} \quad \widehat{f_{ab}} = p \circ \varphi_{ab}$$

by the Reparametrization Lemma. By construction of the $\widehat{f_{ij}}$, the φ_{ij} must be linear, which allows us to assume that all pairs use the same p. Write $\varphi_{ij}(t) = a_{ij} t + b_{ij}$; then for the $c_4 n^{3\alpha-1}$ corresponding f_{ij} we have $f_{ij} = p \circ \varphi_{ij} \circ k$. □

2.4.4 Proof of Theorem 2.1.4

Here we set $\alpha = 1$, so we have $c_4 n^2$ rich lines $u = \varphi_{ij}(t) = a_{ij}t + b_{ij}$ that are rich on the (essentially) $n \times n$ cartesian product $k(A) \times p^{-1}(D)$.

We claim that either $c_5 n^2$ of the lines $u = \varphi_{ij}(t)$ are parallel, or $c_5 n^2$ are concurrent, counting multiplicities. By the Szemerédi-Trotter Theorem (2.2.1), at most $C_{ST} n$ of the lines are distinct. By our Assumption after Lemma 2.4.1, fewer than $2dn$ are identical. This implies that for some c' we can split the lines into $c'n$ classes of size at least $c'n$, such that within each class the lines are identical, and between the classes the lines are distinct.

We take a representative of each class and apply the Line Lemma 2.2.3 to these $c'n$ representatives, telling us that $c''n$ are parallel or $c''n$ are concurrent. Taking all of the corresponding classes together gives $(c'' \cdot c')n^2$ lines that are all parallel or all concurrent.

By Lemma 2.4.6 below, we only need $2dn$ lines parallel, to show that f has the additive form $f(x, y, z) = p(k(x) + l(y) + m(z))$, so $c''c'n^2$ will certainly suffice. Similarly, by Lemma 2.4.6, if $c''c'n^2$ of the lines are concurrent, then f has the multiplicative form $f(x, y, z) = p(k(x) \cdot l(y) \cdot m(z))$. That finishes the proof of Theorem 2.1.4.

2.4.5 Proof of Theorem 2.1.5

We have $c_5 n^{3\alpha-1}$ $c_5 n$-rich lines, for an α to be determined below. Many of these lines may coincide, so let n^β be the number of classes of coinciding lines. The average size of a class is then $c_5 n^{3\alpha-1-\beta}$, so for some $c' > 0$ and $\varepsilon > 0$ we can find a subset of $c'n^\beta$ classes that all have size at least $c'n^{3\alpha-1-\beta-\varepsilon}$.

To apply Lemmas 2.4.6 and 2.4.8 and finish the proof, we will need $2dn^\alpha$ lines that are all parallel or concurrent. To obtain these we need $\frac{2d}{c'} n^{\alpha-(3\alpha-1-\beta-\varepsilon)} = \frac{2d}{c'} n^{1+\beta+\varepsilon-2\alpha}$ representatives of the coinciding classes that are all parallel or all concurrent, since each class has size at least $c'n^{3\alpha-1-\beta-\varepsilon}$. To get these representatives using Lemma 2.2.4, we need

$$c'n^\beta \geq \frac{2d}{c'} n^{2/3+(1+\beta+\varepsilon-2\alpha)/3},$$

for which it will suffice to have $3\beta - \varepsilon \geq 2 + 1 + \beta + \varepsilon - 2\alpha$, so we will need to choose α so that $\beta \geq 3/2 + \varepsilon - \alpha$.

On the other hand, if any of the n^β classes contains at least $2dn^\alpha$ lines, then also $2dn^\alpha$ of the f_{ij} would be identical, contradicting our assumption after Lemma 2.4.1. Hence all classes are smaller than $2dn^\alpha$, which implies

that

$$n^\beta \geq \frac{c_5}{2d} n^{2\alpha-1},$$

hence we have $\beta \geq 2\alpha - 1 - \varepsilon$.

The second inequality for β will imply the first if we choose α so that

$$2\alpha - 1 - \varepsilon \geq 3/2 + \varepsilon - \alpha,$$

hence $\alpha = 5/6 + \varepsilon$ will do.

2.4.6 Proof of Theorem 2.1.6

For Theorem 2.1.6, we do the same as for Theorem 2.1.5, except that instead of Lemma 2.2.4 we apply Corollary 2.2.7. To get the right number of parallel or concurrent lines, we set $p = q = 2dn^{1+\beta+\varepsilon-2\alpha}$ in the Corollary, so we require

$$c'n^\beta > (p+q)n^{\varepsilon'} = 4dn^{1+\beta+\varepsilon-2\alpha+\varepsilon'}$$

for some ε'. That will hold if $\beta > 1+\beta+\varepsilon-2\alpha+\varepsilon'$, or $\alpha \geq 1/2+\varepsilon/2+\varepsilon'/2$, which is satisfied for $\varepsilon' = \varepsilon$ and $\alpha = 1/2 + \varepsilon$ as in Theorem 2.1.6.

2.4.7 The parallel case

Lemma 2.4.6. *If $2dn^\alpha$ of the lines φ_{ij} are parallel, then there is a polynomial $r(y,z)$ such that $f(x,y,z) = p(k(x) + r(y,z))$.*

Proof. We can write $f_{ij}(x) = p(k(x) + b_{ij})$. We use the following two polynomial expansions of $f_{ij}(x) = f(x, y_i, z_j) = p(k(x) + b_{ij})$:

$$\sum_{l=0}^{N} v_l \cdot (k(x) + b_{ij})^l = \sum_{m=0}^{N} w_m(y_i, z_j) \cdot k(x)^m.$$

The first is immediate from $p(k(x) + b_{ij})$; the second requires a little more explanation.

By Lemma 2.2.11, there is a unique expansion of the polynomial f of the form $f(x,y,z) = \sum_{l=0}^{D-1} c_l(k(x), y, z)x^l$, where $D = \deg k$. By the same lemma, we have a unique expansion $f_{ij}(x) = \sum_{l=0}^{D-1} d_l(k(x))x^l$, so that we have

$$\sum_{l=0}^{D-1} c_l(k(x), y_i, z_j)x^l = \sum_{l=0}^{D-1} d_l(k(x))x^l \quad \Rightarrow \quad c_l(k(x), y_i, z_j) = d_l(k(x)).$$

But since $f_{ij}(x) = p(k(x) + b_{ij})$, uniqueness implies that $d_l = 0$ for $l > 0$, hence $c_l(k(x), y_i, z_j) = 0$ for $l > 0$. We have this for every y_i, z_j such that φ_{ij} is one of the parallel lines.

Then we have $2dn^\alpha$ zeroes of $c_l(k(x), y, z)$, so applying the Vanishing Lemma with $|B| = |C| = n^\alpha$ gives $c_l(k(x), y, z) = 0$ for $l > 0$. Thus $f(x, y, z) = c_0(k(x), y, z)$, which means there is an expansion $f(x, y, z) = \sum w_m(y, z)k(x)^m$. Now plugging in $y = y_i, z = z_j$ gives the expansion required above.

Comparing the coefficients of $k(x)^{N-1}$ in the two expansions above, we get

$$v_{N-1} + (N-1)v_N b_{ij} = w_{N-1}(y_i, z_j),$$

which implies that $b_{ij} = \frac{1}{(N-1)v_N}(w_{N-1}(y_i, z_j) - v_{N-1})$. If we now define the polynomial

$$r(y, z) = \frac{1}{(N-1)v_N}(w_{N-1}(y, z) - v_{N-1}),$$

we have that for our $2dn^\alpha$ pairs (y_i, z_j) (note that v_l and w_m do not depend on the choice of pair)

$$f(x, y_i, z_j) = p(k(x) + r(y_i, z_j)).$$

By the Vanishing Lemma with $|B| = |C| = n^\alpha$, applied to $F(y, z) = f(x, y, z) - p(k(x) + r(y, z))$ over $K = \mathbb{R}(x)$, we get the desired equality $f(x, y, z) = p(k(x) + r(y, z))$. □

Lemma 2.4.7. *There are polynomials l and m such that*

$$f(x, y, z) = p(k(x) + l(y) + m(z)).$$

Proof. By applying the above with the roles of x and y swapped, we can also write $f(x, y, z) = P(K(y) + R(x, z))$. Then we calculate the quotient f_x/f_y (using the notation $f_x = \partial f/\partial x$) for both forms,

$$\frac{f_x}{f_y} = \frac{k'(x)}{r_y(y, z)} = \frac{R_x(x, z)}{K'(y)},$$

which tells us that $r_y(y, z)$ (and $R_x(x, z)$) is independent of z. Integrating with respect to y then gives that $r(y, z) = l(y) + m(z)$, which proves our claim. □

2.4.8 The concurrent case

Lemma 2.4.8. *If $2dn^\alpha$ of the lines φ_{ij} are concurrent, There are polynomials $P(t)$, $K(x)$ and $R(y,z)$ such that*

$$f(x,y,z) = P(K(x) \cdot R(y,z)).$$

Proof. We can write $f_{ij}(x) = p(a_{ij} \cdot k(x))$. We again use two polynomial expansions of $f_{ij}(x) = f(x, y_i, z_j) = p(a_{ij} \cdot k(x))$:

$$\sum_{l=0}^{N} v_l \cdot (a_{ij} \cdot k(x))^l = \sum_{m=0}^{N} w_m(y_i, z_j) \cdot k(x)^m.$$

Both are obtained in the same way as in the proof of Lemma 2.4.6.

We cannot proceed exactly as before, since a_{ij} might only occur here with exponents, and we cannot take a root of a polynomial. But we can work around that as follows. Define M to be the greatest common divisor of all exponents m for which $w_m \neq 0$ in the second expansion; then we can write M as an integer linear combination of these m, say $M = \sum \mu_m m$. Comparing the coefficients of any $k(x)^m$ with $w_m \neq 0$ in the two expansions above, we get

$$a_{ij}^m = \frac{1}{v_m} w_m(y_i, z_j),$$

which tells us that

$$a_{ij}^M = \prod (a_{ij}^m)^{\mu_m} = R(y_i, z_j),$$

where $R(y,z)$ is a rational function.

If we define $P(s) = p(s^{1/M})$, or equivalently $P(t^M) = p(t)$, then the definition of M gives that $P(s)$ is a polynomial. We also define $K(x) = k^M(x)$. Then for each of the $2dn^\alpha$ pairs y_i, z_j we have

$$P(K(x) \cdot R(y_i, z_j)) = P(k^M(x) \cdot a_{ij}^M) = p(k(x) \cdot a_{ij}) = f(x, y_i, z_j).$$

Applying the Vanishing Lemma with $|B| = |C| = n^\alpha$ over $\mathbb{R}(x)$ to the numerator of $f(x,y,z) - P(K(x)R(y,z))$, we get that $f(x,y,z) = P(K(x)R(y,z))$. This also tells us that R is in fact a polynomial, since otherwise f could not be one. □

Lemma 2.4.9. *There are polynomials L and M such that $f(x,y,z) = P(K(x) \cdot L(y) \cdot M(z))$.*

Proof. By applying the above with the roles of x and y swapped, we can also write $f(x,y,z) = P^*(K^*(y) \cdot R^*(x,z))$. Then we calculate the quotient f_x/f_z for both forms,

$$\frac{f_x}{f_z} = \frac{K'(x)R(y,z)}{K(x)R_z(y,z)} = \frac{R^*_x(x,z)}{R^*_z(x,z)},$$

which tells us that

$$\frac{R_z(y,z)}{R(y,z)} = \frac{\partial}{\partial z}\log(R(y,z))$$

is independent of y. Integrating we get that $\log(R(y,z)) = \lambda(y) + \mu(z)$, hence

$$R(y,z) = e^{\lambda(y)} \cdot e^{\mu(z)} = L(y)M(z),$$

which also implies that $L(y)$ and $M(z)$ are polynomials, as desired. □

This finishes the proof.

2.5 Applications and limitations

In this section we give some applications and limitations of the main results. We start by giving a simple condition to check whether a function has the required additive or multiplicative form required in the main results. Then we give a proof of our variant of Purdy's conjecture. Finally we give a construction using parabolas that shows that the exponents in Theorem 2.1.6 cannot be improved significantly.

2.5.1 How to check if a function is additive or multiplicative

Here we extend a technique from [21] for checking whether a given function has the additive or multiplicative form in the conclusions of our theorems. Given a differentiable function $f(x,y) : \mathbb{R}^2 \to \mathbb{R}$, we define

$$q_f(x,y) = \frac{\partial^2}{\partial x \partial y}\log\left[\frac{\partial f/\partial x}{\partial f/\partial y}\right].$$

Suppose f is of the form $f(x,y) = p(k(x)+l(y))$ or $f(x,y) = p(k(x)l(y))$, where p, k and l are nonconstant. Then one can check that

$$q_f(x,y) = 0$$

identically.

So, if we have a differentiable function $f : \mathbb{R}^2 \to \mathbb{R}$, and q_f is not identically zero, then we know that the function does not have the additive or multiplicative form. The converse of this result also holds, although we do not need that fact here.

A similar condition holds for functions f of the form $f(x, y, z) = p(k(x)+l(y)+m(z))$ or $f(x, y, z) = p(k(x)l(y)m(z))$. If we define

$$q_f(x, y, z) = \frac{\partial^2}{\partial x \partial y} \log\left[\frac{\partial f/\partial x}{\partial f/\partial y}\right].$$

Then $q_f(x, y, z) = 0$.

Notice that with f in the form above,

$$\frac{\partial f/\partial x}{\partial f/\partial y} = \frac{k'(x)}{l'(y)} \quad \text{or} \quad \frac{\partial f/\partial x}{\partial f/\partial y} = \frac{k'(x)l(y)m(z)}{k(x)l'(y)m(z)} = \frac{k'(x)l(y)}{k(x)l'(y)}$$

is independent of z. This provides another way of checking whether a function does not have the additive or multiplicative form.

Similar conditions could be checked using partial derivatives with respect to z. If $f(x, y, z) = p(k(x) + l(y) + m(z))$ or $f(x, y, z) = p(k(x)l(y)m(z))$ we get

$$r_f(x, z) = \frac{\partial^2}{\partial x \partial z} \log\left[\frac{\partial f/\partial x}{\partial f/\partial z}\right] = 0$$

and

$$s_f(y, z) = \frac{\partial^2}{\partial y \partial z} \log\left[\frac{\partial f/\partial y}{\partial f/\partial z}\right] = 0.$$

Note that in this case the converse does not hold. In the example in Section 2.5.3 below $q_f = 0$, $r_f = 0$ and $s_f = 0$, but f does not have the required decomposition.

2.5.2 On a conjecture of Purdy

The following theorem was conjectured by Purdy (for example, see [9]) and proved by Elekes and Rónyai in [21]. We will use the notation $D(P, Q) = \{d(p, q) : p \in P, q \in Q\}$ for the set of distances between two point sets.

Theorem 2.5.1. *For all c there is an n_0 such that for $n > n_0$ the following holds for any two lines ℓ_1 and ℓ_2 in $\mathbb{R}^2$ and sets P_i of n points on ℓ_i.*

If $|D(P_1, P_2)| < cn$ then the two lines are parallel or orthogonal.

Using Theorem 2.1.3 we can extend it to the asymmetric case when we have fewer points on one of the lines. The proof is similar to that in [21].

Theorem 2.5.2. *For every $c > 0$ and $\varepsilon > 0$ there is an n_0 such that for $n > n_0$ the following holds for any two lines ℓ_1 and ℓ_2 in $\mathbb{R}^2$, P_1 a set of n points on ℓ_1, and P_2 a set of $n^{1/2+\varepsilon}$ points on ℓ_2.*

If $|D(P_1, P_2)| < cn$ then the two lines are parallel or orthogonal.

Proof. Assume the lines are not parallel. Parametrize the lines from their intersection point, l_1 by x_1 and l_2 by x_2, and let X_1 and X_2 represent P_1 and P_2 in this parametrization. Then the condition on the distances means by the Law of Cosines that the polynomial $f(x_1, x_2) = x_1^2 + 2\lambda x_1 x_2 + x_2^2$ assumes $< cn$ values on $X_1 \times X_2$, where $-\lambda$ is the cosine of the angle between the two lines.

Then $z = f(x_1, x_2)$ contains $> c' n^{3/2+\varepsilon}$ points of the cartesian product $X_1 \times X_2 \times E$ where $E = \{a^2 : a \in D(P_1, P_2)\}$. By Theorem 2.1.3, this implies that f has the additive or multiplicative form. Thus q_f, as defined in Section 2.5.1, should be identically zero. A quick calculation shows that this is only possible if $\lambda = -1, 0$, or 1, which means that the angle between the lines is 0, which is not possible, or $\pi/2$, which means the lines are orthogonal. □

2.5.3 Limits on the asymmetry of the cartesian product

In this section we show that Theorem 2.1.6 is near-optimal. We will use the notation $[a, b] = \{a, a+1, \ldots, b-1, b\}$.

Consider

$$f(x, y, z) = x + (y - z)^2,$$

and let $A = D = [1, k^2]$ and $B = C = [1, k]$ for an even integer k. If we set $n = k^2$, then $|A| = |D| = n$ and $|B| = |C| = n^{1/2}$. We can think of the solid $w = f(x, y, z)$ as consisting of translates of the parabola $w = y^2$ from the wy-plane.

We have $x + (y - z)^2 \in D$ when (for instance)

$$x \in [1, k^2/2], \quad y \in [1, k/2] \quad \text{and} \quad z \in [1, k/2].$$

Then the solid $w = f(x, y, z)$ contains at least $\frac{1}{8}k^4 = \frac{1}{8}n^2$ points of $A \times B \times C \times D$.

But the function $f(x, y, z) = x + (y - z)^2$ does not have one of the forms $p(k(x) + l(y) + m(z))$ or $P(K(x)L(y)M(z))$. Note that $q_f = 0$, $r_f = 0$ and $s_f = 0$, so we cannot use the method above to show that f does not have the additive or multiplicative form. Instead we consider a degree argument.

Suppose $f(x, y, z) = P(K(x)L(y)M(z))$. Since each of P, K, L and M must have degree at least one, we would have $\deg f \geq 3$, a contradiction. So f does not have the multiplicative form.

Now suppose that $f(x, y, z) = p(k(x) + l(y) + m(z))$. Then p, k, l and m have degree at least one and at most two. If $\deg p = 2$ then $\deg k = 1$, implying f has a term of the form cx^2, which it doesn't. If $\deg p = 1$, then f couldn't contain the term $-2yz$. So f does not have the additive form either.

Therefore the graph of $w = f(x, y, z) = x+(y-z)^2$ contains many points of $A \times B \times C \times D$, but f cannot be written in the additive or multiplicative form. Hence any extension of Theorem 2.1.4 with $|B| = |C|$ would have to have $|B| = |C| \geq cn^{1/2}$ for some constant $c > 0$. Theorem 2.1.6 supposes $|B| = |C| = n^{1/2+\varepsilon}$ for some $\varepsilon > 0$, so that condition cannot be improved significantly.

Chapter 3

Rational distances with rational angles

3.1 Introduction

In this chapter we prove our first special case of the unit distance problem along with similar results for rational distances.

We will show that the upper bound $n^{1+7/\sqrt{\log n}}$ holds if we only consider unit distances that have *rational angle*, by which we mean that the line through the pair of points makes a rational angle in degrees with the x-axis (or equivalently, its angle in radians, divided by π, is rational). Under this restriction, we can use an algebraic theorem of Mann [44] to get a uniform bound on the number of paths between two fixed vertices in the unit distance graph, which will lead to a contradiction if there are too many unit distances with rational angle between the points.

In fact, our proof also shows that the bound $n^{1+7/\sqrt{\log n}}$ holds for the number of rational distances with rational angles, if we have no three points on a line. The lower bound, $n^{1+c/\log\log n}$, of Erdős does not apply in this case as we are restricted to rational angles. But a construction of Erdős and Purdy gives a superlinear lower bound for unit (and hence rational) distances with rational angles.

If instead we allow up to n^α points on a line where $3/5 \leq \alpha \leq 1$, the number of rational distances with rational angles is bounded by $3n^{1+\alpha}$. This bound is tight up to a constant factor with the lower bound now coming from an $n^{1-\alpha} \times n^\alpha$ square grid. If we allow up to n^α points on a line where $0 < \alpha < 3/5$, the number of rational distances with rational angles is bounded above by $n^{1+\alpha+7/\sqrt{\log n}}$. We get a lower bound of $cn^{1+\alpha}$ from $n^{1-\alpha}$ horizontal lines each containing n^α rational points so that no three points on different lines are collinear.

In Section 3.2 we will state our main results and give an outline of the proof. Section 3.3 contains the algebraic tools that we will use, including, for completeness, a proof of Mann's Theorem. In Section 3.4 we use the

bounds obtained from Mann's theorem and some graph theory to prove our main results. In Section 3.5 we give lower bounds for the main results.

3.2 Main results and proof sketch

We will say that a pair of points in $\mathbb{R}^2$ has *rational angle* if the line segment between them, viewed as a complex number $z = re^{\pi i\gamma}$, has $\gamma \in \mathbb{Q}$. Our first result is the following.

Theorem 3.2.1. *Given n points in $\mathbb{R}^2$, the number of pairs of points with unit distance and rational angle is at most $n^{1+7/\sqrt{\log n}}$.*

Roughly speaking, our proof goes as follows. Given n points in the plane, we construct a graph with the points as vertices, and as edges the unit line segments that have rational angle. We can represent these unit line segments as complex numbers, which must be roots of unity because of the rational angle condition. Then if this graph has many edges, it should have many cycles of a given length k, and each such cycle would give a solution to the equation

$$\sum_{i=1}^{k} \zeta_i = 0,$$

with ζ_i a root of unity. Using an algebraic theorem of Mann from 1965 [44], we could give a uniform bound on the number of such solutions, depending only on k (under the non-degeneracy condition that no subsum vanishes). If the number of non-degenerate cycles goes to infinity with n, this would give a contradiction.

However, dealing with cycles of arbitrary length is not so easy, so instead in our proof we count non-degenerate paths (which we will call *irredundant* paths) of length k between two fixed vertices, which correspond to solutions of the equation

$$\sum_{i=1}^{k} \zeta_i = a,$$

where $a \in \mathbb{C}, a \neq 0$ corresponds to the line segment between the two points. We have extended Mann's theorem to this type of equation, giving a similar upper bound and proving our result.

In fact, in our proof it turns out that it is not necessary for the lengths to be 1, but that they only need to be rational. This is because our extension

of Mann's theorem also works for equations of the type

$$\sum_{i=1}^{k} a_i \zeta_i = a,$$

where $a_i \in \mathbb{Q}$ and $a \in \mathbb{C}, a \neq 0$. This leads to the following results.

Theorem 3.2.2. *Suppose we have n points in $\mathbb{R}^2$, no three of which are on a line. Then the number of pairs of points with rational distance and rational angle is at most $n^{1+7/\sqrt{\log n}}$.*

Theorem 3.2.3. *Suppose we have n points in $\mathbb{R}^2$, with no more than n^α on a line, where $0 < \alpha < 3/5$. Then the number of pairs of points with rational distance and rational angle is at most $n^{1+\alpha+7/\sqrt{\log n}}$.*

Theorem 3.2.4. *Suppose we have n points in $\mathbb{R}^2$, with no more than n^α on a line, where $3/5 \leq \alpha \leq 1$. Then the number of pairs of points with rational distance and rational angle is at most $3n^{1+\alpha}$.*

Remark: It has been brought to the attention of the authors that a result of Conway and Jones gives an improved bound in Theorem 3.3.1 from the following section [15]. Using this improved bound we can replace the $7/\sqrt{\log n}$ in the exponents in Theorems 3.2.1–3.2.3 with $c(\log\log n)^{1/3}/(\log n)^{2/3}$.

3.3 Mann's Theorem

For completeness we provide a proof of Mann's Theorem. We then prove the extension that we will need to prove the main result in the next section.

Theorem 3.3.1 (Mann)**.** *Suppose we have*

$$\sum_{i=1}^{k} a_i \zeta_i = 0,$$

with $a_i \in \mathbb{Q}$, the ζ_i roots of unity, and no subrelations $\sum_{i \in I} a_i \zeta_i = 0$ where $\emptyset \neq I \subsetneq [k]$. Then

$$(\zeta_i/\zeta_j)^m = 1$$

for all i, j, with $m = \prod_{\substack{p \leq k \\ p \text{ prime}}} p$.

The bound of Conway and Jones mentioned at the end of the last section gives m satisfying the inequality $m \leq e^{c\sqrt{k \log k}}$.

Proof. We can assume that $\zeta_1 = 1$ and $a_1 = 1$, so that we have $1 + \sum_{i=2}^{k} a_i \zeta_i = 0$. We take a minimal m such that $\zeta_i^m = 1$ for each i.
We will show that m must be squarefree, and that a prime p that divides m must satisfy $p \leq k$. Together these prove the theorem.
Let p be a prime dividing m. Write $m = p^j \cdot m^*$ with $(p, m^*) = 1$, and use that to factor each ζ_i as follows:

$$\zeta_i = \rho^{\sigma_i} \cdot \zeta_i^*,$$

with ρ a primitive p^jth root of unity so

$$\rho^{p^j} = 1, \quad (\zeta_i^*)^{p^{j-1} m^*} = 1, \quad 0 \leq \sigma_i \leq p - 1.$$

Now reorganize the equation as follows:

$$0 = 1 + \sum_{i=2}^{k} a_i \zeta_i = 1 + \sum_{\ell=0}^{p-1} \alpha_\ell \rho^\ell = f(\rho),$$

where the coefficients are of the form

$$\alpha_\ell = \sum_{i \in I_\ell} a_i \zeta_i^* \in \mathbb{Q}(\zeta_2^*, \ldots, \zeta_k^*) = K,$$

with $I_\ell = \{i \in [k] : \sigma_i = \ell\}$. So f is a polynomial over the field K of degree $\leq p - 1$ and $f(\rho) = 0$. The polynomial f is not identically zero, since that would give a subrelation containing strictly fewer than k terms. To see this, observe that we must have $\sigma_i \geq 1$ for at least one i, otherwise $\zeta_i^{m/p} = 1$ for each i, contradicting the minimality of m.
But we can compute the degree of ρ over K to be

$$\deg_K(\rho) = \frac{\phi(m)}{\phi(p^{j-1} m^*)} = \frac{\phi(p^j)}{\phi(p^{j-1})} = \begin{cases} p - 1 & \text{if } j = 1 \\ p & \text{if } j > 1. \end{cases}$$

This is a contradiction unless $j = 1$, which proves that m is squarefree.
Knowing that m is squarefree, we have $m = p \cdot m^*$ with $(p, m^*) = 1$, and

$$\zeta_i = \rho^{\sigma_i} \cdot \zeta_i^*, \quad \rho^p = 1, \quad (\zeta_i^*)^{m^*} = 1, \quad 0 \leq \sigma_i \leq p - 1.$$

Still $f(\rho) = 0$ for $f(x)$ a polynomial over K, not identically zero. But we know ([42], Ch. VI.3) that the minimal irreducible polynomial of ρ over K is $F(x) = x^{p-1} + x^{p-2} + \cdots + x + 1$, hence we must have $f(x) = cF(x)$ for some $c \in K$. In particular, f has p terms, which implies that our original relation had at least p terms, so $k \geq p$. □

Theorem 3.3.2. *Suppose we have*

$$\sum_{i=1}^{k} a_i\zeta_i = a, \quad \sum_{j=1}^{k} a_j^*\zeta_j^* = a,$$

with $a \in \mathbb{C}, a \neq 0$, $a_i \in \mathbb{Q}$, *roots of unity* ζ_i, *and no subrelations* $\sum_{i\in I} a_i\zeta_i = 0$ *or* $\sum_{j\in J} a_j^*\zeta_j^* = 0$ *where* $\emptyset \neq I \subsetneq [k]$ *and* $\emptyset \neq J \subsetneq [k]$. *Then for any* ζ_j^* *there is a* ζ_i *such that*

$$\left(\zeta_j^*/\zeta_i\right)^m = 1$$

with $m = \prod_{\substack{p\leq 2k \\ p \text{ prime}}} p$.

Proof. We have $\sum a_i\zeta_i = a = \sum a_j^*\zeta_j^*$, which gives the single equation

$$\sum_{i=1}^{k} a_i\zeta_i - \sum_{j=1}^{k} a_j^*\zeta_j^* = 0. \tag{3.1}$$

Mann's Theorem does not apply immediately, because there might be subrelations. But we can break the equation up into minimal subrelations

$$\sum_{i\in I_\ell} a_i\zeta_i - \sum_{j\in I_\ell^*} a_j^*\zeta_j^* = 0, \tag{3.2}$$

where each $I_\ell \neq \emptyset$, $I_\ell^* \neq \emptyset$, and there are no further subrelations.
Given ζ_j^*, there is such a minimal subrelation of length $\leq 2k$ in which it occurs, and which must also contain some ζ_i. Applying Mann's Theorem to this equation gives $\left(\zeta_j^*/\zeta_i\right)^m = 1$ with $m = \prod_{\substack{p\leq 2k \\ p \text{ prime}}} p$. □

Note that in the above proof we require $a \neq 0$. If $a = 0$ and there is no proper subrelation as in (3.2) then (3.1) still has the subrelations

$$\sum_{i=1}^{k} a_i\zeta_i = 0, \quad \sum_{j=1}^{k} a_j^*\zeta_j^* = 0,$$

so we cannot use Mann's Theorem to get a relation between a ζ_i and ζ_j^*.

For $a \in \mathbb{C}, a \neq 0, k \in \mathbb{Z}, k > 0$ we define Z_a^k to be the set of k-tuples of roots of unity $(\zeta_1, \ldots, \zeta_k)$ for which there are $a_i \in \mathbb{Q}$ such that $\sum_{i=1}^k a_i \zeta_i = a$ with no subrelations, i.e.:

$$Z_a^k = \{(\zeta_1, \ldots, \zeta_k) \mid \exists a_i \in \mathbb{Q} : \sum_{i=1}^{k} a_i \zeta_i = a, \sum_{i \in I} a_i \zeta_i \neq 0 \text{ for } \emptyset \neq I \subset [k]\}.$$

Corollary 3.3.3. *Let* $C(k) = \prod_{\substack{p \leq 2k \\ p \text{ prime}}} p$. *Given* $a \in \mathbb{C}, a \neq 0$, $|Z_a^k| \leq (k \cdot C(k))^k$.

Proof. Fix an element $(\zeta_1, \ldots, \zeta_k) \in Z_a^k$ and let $m = C(k)$ and $M_i = \zeta_i^{-m}$ for $1 \leq i \leq k$. Then for ζ_j^* in any element of Z_a^k, we have an i such that $M_i \left(\zeta_j^*\right)^m = 1$. In other words ζ_j^* is a solution of $M_i x^m = 1$. Each of these k equations has $m = C(k)$ solutions, hence there are at most $k \cdot m = k \cdot C(k)$ choices for each ζ_j^*. □

3.4 Rational distances and Mann's Theorem

We are now in a position to prove the main results. Suppose we have a graph $G = G(V, E)$ on $v(G) = n$ vertices and $e(G) = cn^{1+\alpha}$ edges. We will denote the minimum degree in G by $\delta(G)$. The following lemma assures us that we can remove low-degree vertices from our graph without greatly affecting the number of edges.

Lemma 3.4.1. *Let* G *be as above. Then* G *contains a subgraph* H *with* $e(H) = (c/2)n^{1+\alpha}$ *edges such that* $\delta(H) \geq (c/2)n^{\alpha}$.

Proof. We iteratively remove vertices from G of degree less than $(c/2)n^{\alpha}$. Then, the resulting subgraph H has $\delta(H) \geq (c/2)n^{\alpha}$ and we removed fewer than $(c/2)n^{1+\alpha}$ edges so H contains more than $(c/2)n^{1+\alpha}$ edges. □

Note that the subgraph H constructed above contains at least $v(H) = (c/2)^{1/2} n^{1/2+\alpha/2}$ vertices.

Suppose we are given a path on k edges $P_k = p_0 p_1 \ldots p_k$. We call this path *irredundant* if

$$\sum_{i \in I} \overrightarrow{p_i p_{i+1}} \neq 0$$

for any $\emptyset \neq I \subset \{0, 1, \ldots, k-1\}$.

Proof of Theorem 3.2.2. Let G be the graph with the n points in the plane as vertices and the rational distances with rational angles between pairs of points as edges. Suppose there are $n^{1+f(n)}$ such distances for some positive function f. Then $e(G) = n^{1+f(n)}$. We will count the number of irredundant paths P_k in G, for a fixed k that we will choose later. By Lemma 3.4.1 we can assume that $e(G) \geq (1/2)n^{1+f(n)}, v(G) \geq (1/2)n^{1/2+f(n)/2}$ and $\delta(G) \geq (1/2)n^{f(n)}$.

The number of irredundant paths P_k starting at any vertex v is at least

$$N = \prod_{\ell=0}^{k-1}(\delta(G) - 2^\ell + 1),$$

since, if we have constructed a subpath P_ℓ of P_k, then at most $2^\ell - 1$ of the at least $\delta(G)$ continuations are forbidden. Thus the total number of irredundant paths P_k is at least

$$\frac{n^{1/2+f(n)/2}N}{2} \geq \left(n^{1/2+f(n)/2}/2\right)\prod_{\ell=0}^{k-1}((1/2)n^{f(n)} - 2^\ell + 1) \geq \frac{n^{(k+1/2)f(n)+1/2}}{2^{2k+1}}$$

if $2^k \leq (1/2)n^{f(n)}$, which is true as long as $k < f(n)\log n/\log 2$. It follows that there are two vertices v and w with at least

$$\frac{n^{1/2+f(n)/2}N}{n^2} \geq \left(n^{f(n)/2-3/2}\right)\prod_{\ell=0}^{k-1}((1/2)n^{f(n)} - 2^\ell + 1) \geq \frac{n^{(k+1/2)f(n)-3/2}}{4^k}$$

irredundant paths P_k between them. We will call the set of these paths $\mathcal{P}_{vw}$, so that we have $|\mathcal{P}_{vw}| \geq \left(n^{(k+1/2)f(n)-3/2}\right)/4^k$.

Given $P_k \in \mathcal{P}_{vw}$, $P_k = p_0p_1 \dots p_k$, consider the k-tuple $(\zeta_1, \dots, \zeta_k)$ where ζ_i is the root of unity in the direction from p_{i-1} to p_i, i.e. $\zeta_i = \overrightarrow{p_{i-1}p_i}/|\overrightarrow{p_{i-1}p_i}|$. Note that $(\zeta_1, \dots, \zeta_k) \in Z_a^k$, because P_k is irredundant. Since there are no three points on a line, this process gives an injective map from $\mathcal{P}_{vw}$ to Z_a^k so $|\mathcal{P}_{vw}| \leq (k \cdot C(k))^k$ by Corollary 3.3.3. Thus

$$\frac{n^{(k+1/2)f(n)-3/2}}{4^k} \leq (k \cdot C(k))^k \Longrightarrow n^{(k+1/2)f(n)-3/2} \leq (4k \cdot C(k))^k.$$

But this gives

$$e^{((k+1/2)f(n)-3/2)\log n} \leq e^{k\log(4k\cdot C(k))} \Longrightarrow f(n) \leq \frac{\log(4k) + \log(C(k))}{\log n} + \frac{3}{2k}.$$

The term $\log(C(k))$ is the log of the product of the primes less than or equal to $2k$. This is a well known number-theoretic function called the Chebyshev function and denoted by ϑ, specifically $\vartheta(2k) = \log(C(k))$. We use the following bound on ϑ (for a proof see [5]):

$$\vartheta(x) < 4x \log 2 < 3x, \quad \text{for } x \geq 2.$$

This gives

$$f(n) < \frac{\log(4k) + 6k}{\log n} + \frac{1}{k} < \frac{7}{\log n}k + \frac{3}{2k}.$$

Let k be an integer such that $f(n)\log n/16 < k < f(n)\log n/14$, (possible since otherwise $f(n) = O(1/\log n)$ giving $n^{f(n)} = O(1)$). Then the condition that $k < f(n)\log n/\log 2$ is clearly satisfied, and we get

$$f(n) < \frac{7}{\log n} \cdot \frac{f(n)\log n}{14} + \frac{24}{f(n)\log n} \Longrightarrow f(n) < \frac{7}{\sqrt{\log n}}.$$

This completes the proof. □

Proof of Theorem 3.2.1. In the statement of Theorem 3.2.1, the requirement that there are no three points on a line is unnecessary. This is because, from any point, there is only one unit distance in any direction. Thus we can apply the same proof as in Theorem 3.2.2 to Theorem 3.2.1 without having to worry about multiple points on a line. Thus we also have a proof of Theorem 3.2.1. □

Consider a path $P_k = p_0p_1 \dots p_k$. If the distance from p_{i-1} to p_i is less than the distance from p_{i-1} to any vertex on the line connecting p_{i-1} and p_i and not in $P_{i-1} = p_0p_1 \dots p_{i-1}$ then P_k is called a *shortest path*.

Proof of Theorem 3.2.3. This proof is almost the same as the proof of Theorem 3.2.2 except that instead of considering all irredundant paths P_k, we only consider shortest irredundant paths. Suppose there are $n^{1+\alpha+f(n)}$ edges in the rational distance graph. Since there are at most n^α points on a line, we get that from any vertex v there are at least

$$N = \prod_{\ell=0}^{k-1}\left(\frac{\delta(G)}{n^\alpha} - 2^\ell + 1\right) \geq \frac{n^{kf(n)}}{4^k}$$

shortest irredundant paths P_k, if $k < f(n)\log n/\log 2$. For any two vertices v, w let $\mathcal{P}_{v,w}$ be the set of shortest irredundant paths P_k between v

and w. Then there are two vertices v, w such that the number of shortest irredundant paths between v and w is at least

$$|\mathcal{P}_{v,w}| \geq \frac{n^{(k+1/2)f(n)+\alpha/2-3/2}}{4^k}.$$

By Mann's Theorem, since we are looking at shortest irredundant paths, $|\mathcal{P}_{v,w}| \leq (k \cdot C(k))^k$. Let k be an integer such that $f(n)\log n/16 < k < f(n)\log n/14$. Then

$$\frac{n^{(k+1/2)f(n)+\alpha/2-3/2}}{4^k} \leq (k \cdot C(k))^k \Longrightarrow f(n) < \frac{7}{\sqrt{\log n}}.$$

□

Proof of Theorem 3.2.4. Assume we have a configuration of n points with at most n^α on a line, $3/5 \leq \alpha \leq 1$, and $n^{1+\alpha+f(n)}$ rational distances with rational angles, for some positive function $f(n)$.

The graph G on these points has $e(G) = n^{1+\alpha+f(n)}$. By Lemma 3.4.1 we can assume that $e(G) \geq n^{1+\alpha+f(n)}/2$, $v(G) \geq n^{1/2+\alpha/2+f(n)/2}/2$ and $\delta(G) \geq n^{\alpha+f(n)}/2$. We now count irredundant paths P_2 of length 2. Note that an irredundant path on two edges is just a noncollinear path.

For any vertex v, since we have at most n^α points on a line, v is the midpoint of at least

$$N = \delta(G)(\delta(G) - n^\alpha) \geq \frac{n^{2(\alpha+f(n))}}{8}$$

paths P_2 if $f(n) \geq \log 4/\log n$ (if $f(n) < \log 4/\log n$ then $n^{f(n)} < 4$, completing the proof.) Thus there are two vertices v and w with at least $(1/8)n^{(5/2)(\alpha+f(n))-3/2}$ noncollinear paths P_2 between them.

But by Corollary 3.3.3 there is a constant number of directions from each of v and w. Since we are looking at noncollinear paths P_2, the direction from v and the direction from w uniquely determine the midpoint for a path P_2. Thus there are at most $(k \cdot C(k))^k = 144$ noncollinear paths P_2 between v and w, since $k = 2$.

Putting these upper bounds and lower bounds together we get that $n^{(5/2)(\alpha+f(n))-3/2} \leq 2^7 3^2$. This gives

$$f(n) \leq \frac{14\log 2 + 4\log 3}{5\log n} + \frac{3}{5} - \alpha \leq \frac{14\log 2 + 4\log 3}{5\log n} < \frac{3}{\log n},$$

since $\alpha \geq 3/5$. But this gives $n^{f(n)} < 3$, completing the proof. □

3.5 Lower bounds

In this section we give lower bounds for the theorems given in Section 3.2.

The bounds in Theorems 3.2.1 and 3.2.2 are not far from optimal as the following construction of Erdős and Purdy [27] shows.

Suppose we have n points, no three on a line, with the maximum possible number of unit distances with rational angles; we call this number $f(n)$. Consider these points as the set $\{z_1, \ldots, z_n\}$ of complex numbers. For any $a \in \mathbb{C}$ with $|a| = 1$, $a \neq z_i - z_j$ for any $i \neq j$, the set $\{z_1, \ldots, z_n, z_1 + a, \ldots, z_n + a\}$ contains at least $2f(n) + n$ unit distances since there are $f(n)$ amongst each of the sets $\{z_1, \ldots, z_n\}$ and $\{z_1 + a, \ldots, z_n + a\}$ and $|z_i - (z_i + a)| = 1$ for each i. This new set may have three points on a line, but we show that we can choose a appropriately so this is not the case.

Consider a pair of points z_i and z_j. For each z_k, the set of points $\{z_k + a : |a| = 1\}$ intersects the line through z_i and z_j in at most two points. So there are at most two values of a that will give three points on a line. There are $\binom{n}{2}$ pairs of points and n choices for z_k so there are at most $2n\binom{n}{2} = n^2(n-1)$ values of a that make a point $z_k + a$ collinear with two points z_i and z_j. Similarly we have $n^2(n-1)$ values of a that make a point z_k collinear with two points $z_i + a$ and $z_j + a$. Thus there are only finitely many values of a that give three points on a line. There are infinitely many choices for a so we are done.

This shows that $f(2n) \geq 2f(n) + n$ for $n > 2$ and clearly $f(2) = 1$. From this we get that $f(2^k) \geq 2^{k-1}(k-1) = 2^{k-1}\log_2(2^{k-1})$. Taking $2^k \leq n < 2^{k+1}$ we get that $f(n) \geq cn \log n$ for all n. This construction gives a lower bound for Theorems 3.2.1 and 3.2.2.

The bound in Theorem 3.2.3 is not far from optimal. In fact we can get a lower bound of $cn^{1+\alpha}$. Consider $n^{1-\alpha}$ lines parallel to the x-axis, and choose n^α rational points on each line such that no three points on different lines are collinear (this can always be done since there are infinitely many rational points to choose from). There are $cn^{2\alpha}$ rational distances on each horizontal line and $n^{1-\alpha}$ such lines giving at least $cn^{1+\alpha}$ rational distances with rational angles (all the angles are zero).

The bound in Theorem 3.2.4 is tight up to a constant factor as can be seen by considering an $n^{1-\alpha} \times n^\alpha$ square grid. Then there are at least $cn^{2\alpha}$ rational distances on each of the $n^{1-\alpha}$ horizontal lines in the grid containing n^α points. This gives at least $cn^{1+\alpha}$ rational distances with rational angles (the angles are all zero).

Chapter 4

Using the Subspace Theorem to bound unit distances

4.1 Introduction

In this chapter we give our second special case of the Erdős unit distance problem and show that the unit distances in the lower bound construction of Erdős satisfy the conditions of our main result. In the previous chapter we considered unit distances that correspond to roots of unity. In this chapter we consider unit distances coming from a group with "low" rank—roots of unity come from a group with rank 0. Our main tool is a corollary of the Subspace Theorem which we state below. More details of the Subspace Theorem can be found in the next chapter.

Consider two points $p = (p_1, p_2), q = (q_1, q_2) \in \mathbb{R}^2$ with unit distance. Considering the vector $\overrightarrow{pq}$ between these two points we get the complex number

$$z = z(p, q) = (q_1 - p_1) + i(q_2 - p_2), \quad \text{with } |z| = 1.$$

We will restrict our attention to unit distances with z coming from a multiplicative subgroup of $\mathbb{C}^*$ (the multiplicative group of nonzero complex numbers) of finite rank. Recall that a subgroup $\Gamma \subset \mathbb{C}^*$ has rank r if there exists a finitely generated subgroup $\Gamma_0 \subset \Gamma$ with r generators such that for every $x \in \Gamma$ there exists an integer $k \geq 0$ such that $x^k \in \Gamma_0$.

Suppose Γ is a subgroup of $\mathbb{C}^*$ of finite rank r and $a_1, a_2, \ldots, a_k \in \mathbb{C}^*$. Recall that a solution of the equation $a_1z_1+a_2z_2+\cdots+a_kz_k = 1$ is called *nondegenerate* if no subsum of the left-hand side vanishes. That is $\sum_{j\in J} a_j z_j \neq 0$ for every nonempty $J \subset \{1, 2, \ldots, k\}$. We will consider the number $A(k, r)$ of nondegenerate solutions of this equation with $z_i \in \Gamma$. We now give the corollary of the Subspace Theorem with the best known bound due to Amoroso and Viada [3].

Theorem 4.1.1. *Suppose $a_1, a_2, \ldots, a_k \in \mathbb{C}^*$ and Γ has finite rank r. Then the number of nondegenerate solutions of the equation*

$$a_1z_1 + a_2z_2 + \cdots + a_kz_k = 1 \tag{4.1}$$

with $z_i \in \Gamma$ is at most

$$A(k,r) \leq (8k)^{4k^4(k+kr+1)}.$$

Amoroso and Viada proved Theorem 4.1.1 over an arbitrary algebraically closed field K of characteristic 0 but we only require it over $\mathbb{C}$.

We will use this to prove the following result.

Theorem 4.1.2. *Let $\varepsilon > 0$. Then there exist $n_0 = n_0(\varepsilon)$, a positive integer, and $c = c(\varepsilon) > 0$ such that given $n > n_0$ points in the plane, the number of unit distances with z coming from a subgroup $\Gamma \subset \mathbb{C}^*$ with rank $r < c \log n$ is at most $n^{1+\varepsilon}$.*

We will prove this theorem in the next section and in Section 4.3 we will show that the unit distances from the lower bound construction of Erdős satisfy the hypotheses of this theorem.

4.2 Proof of the main result

The proof of Theorem 4.1.2 is quite similar to the proof of Theorem 3.2.1 in Section 3.4. The main difference is that we use the Subspace Theorem instead of Mann's result to get an upper bound for paths in the unit distance graph.

Suppose $G = G(V, E)$ is a graph on $v(G) = n$ vertices and $e(G) = cn^{1+\alpha}$ edges. We denote the minimum degree in G by $\delta(G)$. The following lemma shows that we can remove low degree vertices from our graph without greatly affecting the number of edges.

Lemma 4.2.1. *Let G be as above. Then G contains a subgraph H with $e(H) \geq (c/2)n^{1+\alpha}$ edges such that $\delta(H) \geq (c/2)n^{\alpha}$.*

Proof. We remove vertices from G of degree less than $(c/2)n^{\alpha}$ one by one. After removing a vertex some of its neighbours may have their degree reduced to below $(c/2)n^{\alpha}$. If so then we add such vertices to the list of vertices to be removed. In total we remove fewer than $(c/2)n^{1+\alpha}$ edges so the resulting subgraph H has at least $(c/2)n^{1+\alpha}$ edges. Thus the above process must stop and the resulting graph has $\delta(H) \geq (c/2)n^{\alpha}$. □

Note that the subgraph H constructed above contains at least $v(H) = \sqrt{c}n^{1/2+\alpha/2}$ vertices.

Suppose we are given a path on k edges $P_k = p_0p_1 \dots p_k$. We call this path *irredundant* if

$$\sum_{i \in I} \overrightarrow{p_i p_{i+1}} \neq 0$$

for every $\emptyset \neq I \subset \{0, 1, \dots, k-1\}$.

Proof of Theorem 4.1.2. Let G be the graph with the n points in the plane as vertices and the unit distances with z coming from Γ as edges. Suppose there are $n^{1+\varepsilon}$ such distances. Then $e(G) = n^{1+\varepsilon}$. We will show that we can take ε as small as we like. We will count the number of irredundant paths P_k in G, for a fixed k that we will choose later. By Lemma 4.2.1 we can assume that $e(G) \geq (1/2)n^{1+\varepsilon}, v(G) \geq n^{1/2+\varepsilon/2}$ and $\delta(G) \geq (1/2)n^{\varepsilon}$.

The number of irredundant paths P_k starting at any vertex v is at least

$$N \geq \prod_{\ell=0}^{k-1} (\delta(G) - 2^{\ell} + 1) \geq \frac{n^{k\varepsilon}}{2^{2k}}.$$

The first inequality is true since if we have constructed a subpath P_ℓ of P_k, then at most $2^\ell - 1$ of the at least $\delta(G)$ possible continuations are forbidden. In the second inequality we have assumed that $2^k \leq (1/2)n^{\varepsilon}$, which is true as long as $k < \varepsilon \log n / \log 2 - 1$ (we will show that this holds at the end of the proof). Thus the total number of irredundant paths P_k is at least $Nn^{1/2+\varepsilon/2}/2 \geq n^{1/2+(k+1/2)\varepsilon}/2^{2k+1}$. It follows that there are two vertices v and w with at least $Nn^{1/2+\varepsilon/2}/n^2 \geq n^{(k+1/2)\varepsilon-3/2}/4^k$ irredundant paths P_k between them. We will call the set of these paths $\mathcal{P}_{vw}$, so that we have

$$|\mathcal{P}_{vw}| \geq \frac{n^{(k+1/2)\varepsilon-3/2}}{4^k}.$$

Given $P_k \in \mathcal{P}_{vw}$, $P_k = p_0p_1 \dots p_k$, consider the k-tuple $(z_1, \dots, z_k)$ where z_i is the complex number in the direction from p_{i-1} to p_i, i.e. $z_i = z(p_{i-1}, p_i) = \overrightarrow{p_{i-1}p_i}$. Let $a = z(v, w)$. Then $z_1 + z_2 + \dots + z_k = a$. Since the path is irredundant no subsum on the left vanishes. So P_k corresponds to a nondegenerate solution of Equation (4.1) with $a_i = 1/a$ for $i = 1, 2, \dots, k$. Thus, by Theorem 4.1.1,

$$|\mathcal{P}_{vw}| \leq (8k)^{4k^4(k+kr+1)}.$$

Putting these inequalities together and taking logarithms we get

$$\begin{aligned} ((k + 1/2)\varepsilon - 3/2) \log n &\leq k \log 4 + 4k^4(k + kr + 1) \log(8k) \\ &\leq 5rk^5 \log k, \end{aligned}$$

where the last inequality holds for large k. From this we get

$$\varepsilon \leq \frac{5rk^5 \log k}{(k+1/2)\log n} + \frac{3}{2(k+1/2)} \leq \frac{5rk^4 \log k}{\log n} + \frac{3}{2k}. \tag{4.2}$$

We consider the expression on the right hand side as a function of k. Optimizing this function we get

$$k \geq \exp\Big((1/5)W(5c_2 \log n/r)\Big)$$

for some constant $c_2 > 0$ where W is the positive real-valued function satisfying $x = W(x)e^{W(x)}$. This function is called the Lambert W function and was first studied by J.H. Lambert in 1758 [41]. The following asymptotic expression is due to N.G. de Bruijn [17]:

$$W(x) = \log(x) - \log\log(x) + O\big(\frac{\log\log\log x}{\log\log x}\big).$$

We don't require this much accuracy. One can easily check, and we will just use the fact, that $(1/2)\log x \leq W(x) \leq \log x$ for $x \geq e$.

Then we can take

$$c'\left(\frac{\log n}{r}\right)^{1/5} \leq k \leq c''\left(\frac{\log n}{r}\right)^{1/5}$$

for some constants $c', c'' > 0$.

For any $\varepsilon > 0$ there is a constant $c > 0$ such that if $r + 1 \leq c \log n$ then the inequality in (4.2) holds for large n. When counting P_k's we made the assumption that $k \leq \varepsilon \log n/\log 2 - 1$. Checking the above values of k and f we see that this holds for large n.

This completes the proof. □

4.3 Analysis of Erdős' lower bound

It would be interesting to analyze the possible group structure of point sets with the maximum (or near maximum) number of unit distances. We will now show that the lower bound configuration for the unit distance problem given by Erdős satisfies the hypotheses of Theorem 4.1.2. Matoušek has given a very in-depth account of Erdős' lower bound and we will follow that here [46].

We require the following number-theoretic functions:

$$\pi_{d,a}(x) = \sum_{\substack{p \le x \\ p \equiv a(d)}} 1, \quad \vartheta_{d,a}(x) = \sum_{\substack{p \le x \\ p \equiv a(d)}} \log p, \quad \psi_{d,a}(x) = \sum_{\substack{p^\ell \le x \\ p^\ell \equiv a(d)}} \log p,$$

where the first two sums are over primes less than x of the form $p = a + kd$ and the last sum is over primes p and positive integers ℓ such that $p^\ell = a + kd$ and $p^\ell \le x$. These are analogues of the prime counting function and Chebyshev functions for arithmetic progressions.

We will use the following results regarding these functions all of which are well known in number theory. For details see [35] and [49].

Theorem 4.3.1 (The Prime Number Theorem for Arithmetic Progressions)**.** *Suppose a and d are positive integers such that $(a, d) = 1$. Then*

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

A simple consequence of this result is that if a and d are positive integers such that $a < d$ and $(a, d) = 1$ then the kth prime of the form $p_i = a + k_i d$ satisfies $p_k = (1 + o(1))k \log k/\varphi(d)$.

Theorem 4.3.2. *Suppose a and d are positive integers such that $(a, d) = 1$. Then*

$$\psi_{d,a}(n) = (1 + o(1))\frac{n}{\varphi(d)}.$$

Theorem 4.3.2 can be deduced from Theorem 4.3.1 by partial summation.

Theorem 4.3.3. *Suppose a and d are positive integers such that $(a, d) = 1$. Then*

$$\vartheta_{d,a}(n) = (1 + o(1))\psi_{d,a}(n).$$

The above two theorems give $\vartheta_{d,a}(n) = (1 + o(1))n/\varphi(d)$.

We will also use the following fact. For details see [51].

Theorem 4.3.4. *The number of integer solutions, $R(m)$, of $x^2 + y^2 = m$ where $m = p_1 p_2 \dots p_r$ and the p_i are distinct primes of the form $p_i = 4k_i + 1$ is*

$$R(m) = 2^{r+2}.$$

The lower bound configuration consists of n points in a $\sqrt{n} \times \sqrt{n}$ grid. The step in the grid is chosen to be $1/\sqrt{m}$ where m is the product of the first $r-1$ primes of the form $4k+1$ and r is the largest number with $m \leq n/4$. We will in fact consider a $\sqrt{n} \times \sqrt{n}$ grid with step 1 and then count the distances of length $\sqrt{m}$. This gives a lower bound to the unit distance problem by scaling the point set by $1/\sqrt{m}$.

We have $4p_1p_2 \dots p_{r-1} \leq n < 4p_1p_2 \dots p_r$. From this the bound $r \geq \log n/(3 \log \log n)$ is found using the prime number theorem for arithmetic progressions. Distances equal to $\sqrt{m}$ in this configuration correspond to integer solutions of $x^2 + y^2 = m$. In the lower bound, the fact that there are at least $2^{r-1}/16$ such distances is used. But an upper bound on the number of such distances can also be found. By Theorem 4.3.4 there are at most $4 \cdot 2^{(r-1)+2} = 2^{r+3}$ such distances from any point so we have at most $2^{r+3}n$ such distances in total. Erdős' construction gives a lower bound for r. If we can find an upper bound for r then we are done as will be described below.

We will briefly explain the reason that m is defined as above as this highlights the generators to choose for a multiplicative subgroup of $\mathbb{C}^*$ containing the unit distances of the configuration. A prime p has a unique expression, up to the order of the terms, of the form $x^2 + y^2 = p$ with x and y positive integers if and only if $p = 2$ or $p = 4k + 1$ for some integer k. The Brahmagupta-Fibonacci identity says that the product of two numbers, each expressible as the sum of two squares, is itself expressible as the sum of two squares. Specifically

$$\begin{aligned}(a^2 + b^2)(c^2 + d^2) &= (ac - bd)^2 + (ad + bc)^2 \\ &= (ac + bd)^2 + (ad - bc)^2.\end{aligned}$$

So as we multiply more primes of the form $4k + 1$ together we get more expressions of the resulting number as a sum of two squares. So all solutions of $x^2 + y^2 = m$ can be described in terms of the solutions of $x_j^2 + y_j^2 = p_j$.

More formally, we consider the ring R of points in $\mathbb{Z}^2$ with addition defined coordinate-wise, so $(a, b) + (c, d) = (a + c, b + d)$, and multiplication defined as follows: $(a, b) \cdot (c, d) = (ac - bd, ad + bc)$. One can check that R is actually a ring and is in fact isomorphic to the Gaussian integers $\mathbb{Z}[i]$ since the operations correspond to complex addition and complex multiplication. But $\mathbb{Z}[i]$ is a unique factorization domain so R is also a unique factorization domain. We will consider the elements of R as distance vectors.

In our grid we are looking for the distance $m = p_1 \dots p_{r-1}$. By Theorem 4.3.4, the number of pairs $(x, y) \in \mathbb{Z}^2$ with $x^2 + y^2 = m$ is $R(m) = 2^{r+1}$. Suppose $x_j^2 + y_j^2 = p_j$. We consider the point $(x_j, y_j) \in R$. The product of

these $r-1$ points $(x,y) = (x_1,y_1)(x_2,y_2)\dots(x_{r-1},y_{r-1})$ has magnitude

$$|(x,y)| = |(x_1,y_1)|\dots|(x_{r-1},y_{r-1})| = \sqrt{p_1\dots p_{r-1}} = \sqrt{m}.$$

So this product gives a point with length $\sqrt{m}$.

Now, R is a unique factorization domain. That means that the point $(x,y) = (x_1,y_1)\dots(x_{r-1},y_{r-1})$ has unique factorization. Specifically, in any other factorization of $(x,y) = (x'_1,y'_1)\dots(x'_{r-1},y'_{r-1})$ there is a bijection ϕ of the factors such that $(x_j,y_j) = u_j(x'_{\phi(j)},y'_{\phi(j)})$ where u_j is a unit. The units in R correspond to the units in $\mathbb{Z}[i]$. In the latter these are $1,-1,i,-i$ so in the former they are $(1,0),(-1,0),(0,1)$ and $(0,-1)$. Two elements $(a,b),(c,d) \in R$ are called *associates* if $(a,b) = u(c,d)$ for some unit u. So in a unique factorization domain the factorization of an element is unique up to ordering and associates. Now, since the p_j's are odd primes we cannot have $x_j = \pm y_j$ for $1 \le j \le r-1$. One can check that (x_j,y_j) and $(x_j,-y_j)$ are not associates and $(x_j,y_j),(x_k,y_k)$ are not associates for $j \ne k$. So we have two points to choose from for each p_j, namely (x_j,y_j) and $(x_j,-y_j)$, giving 2^{r-1} choices for (x,y). None of the factors are associates so these choices for (x,y) are all distinct. If we multiply a given (x,y) by a unit then we get four different values. So we get $4.2^{r-1} = 2^{r+1}$ distinct points (x,y) each with length $\sqrt{m}$. So these give all possible required distances by Theorem 4.3.4. The units are torsion points of R (they have finite multiplicative order in R) so they don't affect the rank.

Going back to unit distances, if we take the complex numbers

$$z_j = m^{-1/(2r-2)}(x_j + iy_j), \qquad w_j = m^{-1/(2r-2)}(x_j - iy_j)$$

for $1 \le j \le r-1$ then these generate the multiplicative group of unit distances in the configuration. Thus the unit distances come from a multiplicative subgroup of $\mathbb{C}^*$ of rank at most $r-1$. So we just need to bound r from above.

We do this by looking at the inequality $p_1\dots p_{r-1} \le n/4$. Taking logarithms we get $\vartheta_{4,1}(p_{r-1}) \le \log(n/4)$. By Theorems 4.3.2 and 4.3.3 we get

$$\frac{p_{r-1}}{2\sqrt{2}} \le \log(n/4).$$

By the remark after Theorem 4.3.1 we get

$$\frac{(r-1)\log(r-1)}{2\sqrt{2}} \le 2\sqrt{2}\log(n/4).$$

Solving for r we get

$$r \leq \frac{16 \log n}{\log \log n}.$$

Thus the unit distances come from a multiplicative subgroup of rank at most $r - 1 \leq 16 \log n / \log \log n - 1$. Since

$$\frac{16 \log n}{\log \log n} - 1 \leq c \log n$$

for large n this configuration is covered by Theorem 4.1.2.

Chapter 5

Combinatorial applications of the Subspace Theorem

5.1 Introduction

The Subspace Theorem has appeared in various forms and been adapted and improved over time. Its applications include diophantine approximation, results about integral points on algebraic curves and the construction of transcendental numbers. But its usefulness extends beyond the realms of number theory. Other applications of the Subspace Theorem include linear recurrence sequences and finite automata. In fact, these structures are closely related to each other and the construction of transcendental numbers.

The Subspace Theorem also has a number of remarkable combinatorial applications. We have already seen its use for the unit distance problem. The purpose of this chapter is to give a survey of some other applications including sum-product estimates and bounds on line configurations with few distinct intersection points. The presentation will be from the point of view of a discrete mathematician. We will state a number of variants of the Subspace Theorem below but we will not prove any of them as the proofs are beyond the scope of this work. We have already seen a proof of Mann's Theorem and its corollary in Section 3.3. The corollary of Mann's Theorem also has applications to combinatorics and can be considered as a special case of the Subspace Theorem.

Wolfgang M. Schmidt was the first to state and prove a variant of the Subspace Theorem in 1972 [55]. His theorem has been extended multiple times and has played a very important role in modern number theory. Before we state the Subspace Theorem we need some definitions. A *linear form* is an expression of the form $L(x) = a_1x_1 + a_2x_2 + \cdots + a_nx_n$ where $a_1, \ldots, a_n$ are constants and $x = (x_1, \ldots, x_n)$. A collection of linear forms is *linearly independent* if none of them can be expressed as a linear combination of the others. A complex number is *algebraic* if it is a root of a univariate polynomial with rational coefficients. Given $x = (x_1, \ldots, x_n)$ we define the

maximum norm of x:

$$\|x\| = \max(|x_1|, \ldots, |x_n|).$$

Theorem 5.1.1 (Subspace Theorem I)**.** *Suppose we have n linearly independent linear forms $L_1, L_2, \ldots, L_n$ in n variables with algebraic coefficients. Given $\varepsilon > 0$, the non-zero integer points $x = (x_1, x_2, \ldots, x_n)$ satisfying*

$$|L_1(x)L_2(x)\ldots L_n(x)| < \|x\|^{-\varepsilon}$$

lie in finitely many proper linear subspaces of $\mathbb{Q}^n$.

This generalises the Thue-Siegel-Roth Theorem on the approximation of algebraic numbers [54] to higher dimensions.

Theorem 5.1.1 has been extended in various directions by many authors including Schmidt himself, Schlickewei, Evertse, Amoroso and Viada. Analogues have been proved using p-adic norms and over arbitrary number fields and bounds on the number of subspaces required have been found. These bounds depend on the degree of the number field and the dimension. For some of these results and more information see [29], [30] and [3].

Now we give a p-adic version of the Subspace Theorem that we will use in the next section. Given a prime p, the *p-adic* absolute value is denoted $|x|_p$ and satisfies $|p|_p = 1/p$. $|x|_\infty$ denotes the usual absolute value so $|x|_\infty = |x|$. We may refer to ∞ as the *infinite prime*.

Theorem 5.1.2 (Subspace Theorem II)**.** *Suppose $S = \{\infty, p_1, \ldots, p_t\}$ is a finite set of primes, including the infinite prime. For every $p \in S$ let $L_{1,p}, \ldots, L_{n,p}$ be linearly independent linear forms in n variables with algebraic coefficients. Then for any $\varepsilon > 0$ the solutions $x \in \mathbb{Z}^n$ of*

$$\prod_{p \in S} \prod_{i=1}^{n} |L_{i,p}(x)|_p \leq \|x\|^{-\varepsilon}$$

lie in finitely many proper linear subspaces of $\mathbb{Q}^n$.

The power and utility of the Subspace Theorem is already evident in the above forms but there is a corollary, often itself called the Subspace Theorem, which makes even more applications possible. We have already seen this version in Chapter 4 but we state it here for completeness. The bound given is due to Amoroso and Viada [3].

Theorem 5.1.3 (Subspace Theorem III)**.** *Given an algebraically closed field K and a multiplicative subgroup Γ of K^* of finite rank r, suppose $a_1, a_2, \ldots, a_n \in K^*$. Then the number of solutions of the equation*

$$a_1 z_1 + a_2 z_2 + \cdots + a_n z_n = 1 \tag{5.1}$$

with $z_i \in \Gamma$ and no subsum on the left-hand side vanishing is at most

$$A(n, r) \leq (8n)^{4n^4(n+nr+1)}.$$

The structure of this paper will be as follows. In the next section we give two well known applications of the Subspace Theorem. Namely, we give an example of showing that a number is transcendental and a result showing that certain linear recurrence sequences can only have finitely many zero terms. In Section 5.3 we give combinatorial applications. These include a sum-product estimate of Chang for sets with small product set and a bound on line configurations with few distinct intersections of Chang and Solymosi.

5.2 Number-theoretic applications

5.2.1 Transcendental numbers

Adamczewski and Bugeaud showed that all irrational automatic numbers are transcendental using the Subspace Theorem. An *automatic number* is a number for which there exists a positive integer b such that when the number is written in b-ary form it is the output of a finite automaton with input the nonnegative integers written from right to left. For more details see [1] or [8].

Here we will use a method similar to the proof of Theorem 3.3 in [8] to show:

Theorem 5.2.1. *The number α given by the infinite sum*

$$\alpha = \sum_{n \geq 1} \frac{1}{2^{2^n}}$$

is transcendental.

Kempner showed in the early twentieth century that a large class of numbers defined similarly to α are transcendental [39]. The Subspace Theorem provides a tidy proof of this fact.

Proof of Theorem 5.2.1. Consider the binary expansion:

$$\alpha = \frac{1}{4} + \frac{1}{16} + \frac{1}{256} + \frac{1}{65536} + \cdots = 0.0101000100000001\ldots_2\,.$$

So the binary expansion of α consists of sections of zeros of increasing length separating solitary ones. Thus the expansion is not periodic and hence α is not rational. We let b_n be the string given by the first n digits of this expansion. One can check that each b_n has two disjoint substrings of zeros of length $n/8$.

Assume α is not transcendental. Then it is algebraic. Now each b_n starts with a string $AOBO$, where O is a string of zeroes, the length of O is at least $n/8$ and A and B might have length zero. We will use the rational number represented in base 2 by $0.AOBOBO\ldots$ to approximate α. Call this number π. Then

$$\pi = \frac{M}{2^a(2^b-1)}$$

where $M \in \mathbb{Z}$ and a and b are the lengths of the strings A and OB respectively. Clearly $b \geq n/8$ and $a+b \leq n$ since AOB is a substring of b_n. Since α starts with b_n we have

$$|\alpha - \pi| \leq \frac{1}{2^{a+b+n/8}} \implies |2^{a+b}\alpha - 2^a\alpha - M| \leq \frac{1}{2^{n/8}}.$$

Now we apply the Subspace Theorem in the form given in Theorem 5.1.2. We let $S = \{2, \infty\}$ and

$$\begin{array}{lll} L_{1,\infty}(x) = x_1, & L_{2,\infty}(x) = x_2, & L_{3,\infty}(x) = \alpha x_1 - \alpha x_2 - x_3, \\ L_{1,2}(x) = x_1, & L_{2,2}(x) = x_2, & L_{3,2}(x) = x_3. \end{array}$$

Note that by our assumption that α is not transcendental the linear form $L_{3,\infty}$ has algebraic coefficients. Let $x = (2^{a+b}, 2^a, M)$. Now $|M| \leq 2^{a+b}$ since $0 < \pi < 1$. So $\|x\| \leq 2^{a+b} \leq 2^n$. Multiplying the absolute values of the linear forms together we get

$$\begin{aligned} \prod_{p\in S}\prod_{i=1}^{3} |L_{i,p}(x)| &= |2^a|_2|2^a|_\infty|2^{a+b}|_2|2^{a+b}|_\infty|M|_2|2^{a+b}\alpha - 2^a\alpha - M|_\infty \\ &\leq \frac{1}{2^{n/8}} \\ &\leq \frac{1}{\|x\|^{1/8}}. \end{aligned}$$

The first inequality holds because $|\alpha - \pi| \leq 2^{-a-b-n/8}$ and $|2|_2|2|_\infty = 1$.

We can do this for each n and $b = b(n)$ increases as n increases since $b \geq n/8$. Thus infinitely many of the vectors $x = x(n)$ are distinct. By Theorem 5.1.2 these vectors are contained in finitely many subspaces of $\mathbb{Q}^3$. Thus one of these subspaces contains infinitely many of them. That is, there exist $c, d, e \in \mathbb{Q}$ such that

$$c2^{a(n)} + d2^{a(n)+b(n)} + eM(n) = 0$$

for infinitely many n. The coefficient e cannot be zero since $b(n) \to \infty$ as $n \to \infty$. Dividing by $2^{a(n)}(2^{b(n)} - 1)$ and taking limits we get $\alpha = -d/e$ so α is rational. This is a contradiction. Thus α must be transcendental. □

5.2.2 Linear recurrence sequences

A linear recurrence sequence is a sequence of numbers where the first few terms are given and the higher order terms can be found by a recurrence relation. A famous example is the Fibonacci sequence $\{F_n\}$ where $F_0 = 0, F_1 = 1$ and $F_n = F_{n-1} + F_{n-2}$ for $n \geq 2$. More formally, a *linear recurrence sequence* consists of constants $a_1, \ldots, a_k$ in a field K for some $k > 0$ along with a sequence $\{R_m\}_{m=0}^{\infty}$ with $R_i \in K$ for $0 \leq i \leq k-1$ and

$$R_n = a_1 R_{n-1} + a_2 R_{n-2} + \cdots + a_k R_{n-k}, \quad \text{for } n \geq k.$$

If $\{R_m\}$ is not expressible by any shorter recurrence relation then it is said to have order k. In this case $a_k \neq 0$.

We are interested in the structure of the zero set of a linear recurrence sequence. This is the set

$$S(\{R_m\}) = \{i \in \mathbb{N} : R_i = 0\}.$$

The Skolem-Mahler-Lech Theorem states that this set consists of the union of finitely many points and arithmetic progressions [43]. Schmidt has given a quantitative bound for this theorem using various tools including the Subspace Theorem [56].

We will show a special case of this theorem using Theorem 5.1.3. We will restrict our attention to simple nondegenerate linear recurrence sequences. To define such sequences we need to define the companion polynomial of the recurrence sequence. If $\{R_m\}$ is given as above then the *companion polynomial* of $\{R_m\}$ is $C(x) = x^k - a_1 x^{k-1} - \cdots - a_{k-1}x - a_k$. Suppose the roots of this polynomial are $\alpha_1, \ldots, \alpha_\ell$ with multiplicity $b_1, \ldots, b_\ell$ respectively. Each α_i is nonzero since $a_k \neq 0$. If the companion polynomial has k distinct roots

it is called *simple*. If α_i/α_j is not a root of unity for any $i \neq j$ then the sequence is called *nondegenerate*. A version of this theorem was given in [29]. The improved bound given here is due to Amoroso and Viada [4].

Theorem 5.2.2. *Suppose $\{R_m\}$ is a simple nondegenerate linear recurrence sequence of order k with complex coefficients. Then*

$$|S(\{R_m\})| \leq (8k)^{8k^6}.$$

Proof. We can express the recurrence relation using a matrix equation:

$$\begin{bmatrix} a_1 & a_2 & \dots & a_{k-1} & a_k \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}^n \begin{bmatrix} R_{k-1} \\ R_{k-2} \\ \vdots \\ R_0 \end{bmatrix} = \begin{bmatrix} R_{k-1+n} \\ R_{k-2+n} \\ \vdots \\ R_n \end{bmatrix}. \tag{5.2}$$

We call the $k \times k$ matrix above A. The characteristic polynomial of A is given by

$$\chi(\lambda) = (-1)^k(\lambda^k - a_1\lambda^{k-1} - \cdots - a_k).$$

This is the same, up to sign, as the companion polynomial of $\{R_m\}$. Thus A has distinct nonzero eigenvalues and so can be diagonalized. So $A = PDP^{-1}$ for some invertible $k \times k$ matrix P and

$$D = \begin{bmatrix} \alpha_1 & 0 & \dots & 0 \\ 0 & \alpha_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & \alpha_k \end{bmatrix}.$$

So D^n is a diagonal matrix with α_i^n in the i-th row and column. Thus, multiplying out PD^nP^{-1} we see that every element of A^n is a linear combination of the n-th powers of the α_i's. Now, by the matrix equation in (5.2) we see that R_n is given by the k-th row of A^n multiplied by $[R_{k-1}, R_{k-2}, \dots, R_0]^T$ and so

$$R_n = c_1\alpha_1^n + c_2\alpha_2^n + \cdots + c_k\alpha_k^n \quad \text{for every } n \geq k.$$

Then applying Theorem 5.1.3 to the equation $c_1x_1 + c_2x_2 + \cdots + c_kx_k = 0$ with solutions from the group of rank at most k generated by $\{\alpha_1, \dots, \alpha_k\}$ we get that the number of solutions is at most

$$A(k,k) \leq (8k)^{4k^4(k+k^2+1)} \leq (8k)^{8k^6}.$$

Since the sequence is nondegenerate we cannot have two values n, n' giving the same value for α_i^n and $\alpha_i^{n'}$ for each i, hence each solution corresponds to a unique value from $S(\{R_m\})$. □

5.3 Combinatorial applications

The corollary of the Subspace Theorem given in Theorem 5.1.3 gives a bound on the number of nondegenerate solutions of a linear equation from a multiplicative group with rank not too large. What happens if the group in question has rank zero? This corresponds to solutions that are roots of unity. The corollary can then be seen as a generalisation of Corollary 3.3.3, which is the corollary of Mann's Theorem that we used to bound rational distances in Chapter 3.

We mentioned before that the Subspace Theorem can be considered as a generalisation of the Thue-Siegel-Roth Theorem but Mann's result provides another possible starting point for the development of the Subspace Theorem. The roots of unity give a relatively simple example of an infinite multiplicative group.

Although these results were developed in different fields and for very different problems they seem to provide just what is required for many combinatorial problems. We have seen one such application already and will provide a few more below.

5.3.1 Sum-product estimates

The theory of sum sets and product sets plays an important part in combinatorics and additive number theory. The goal of the field is to show that for any finite subset of a field either the sum set or the product set is large. We will focus on the complex numbers.

Formally, given a set $A \subset \mathbb{C}$, the sum set, denoted by $A+A$, and product set, denoted by AA, are

$$A + A := \{a + b : a, b \in A\}, \qquad AA := \{ab : a, b \in A\}.$$

Note that

$$|A| \leq |A + A|, |AA| \leq \binom{|A| + 1}{2} = \frac{|A|^2}{2} + \frac{|A|}{2}.$$

The following long standing conjecture of Erdős and Szemerédi [28] has led to much work in the field.

Conjecture 5.3.1. *Let $\varepsilon > 0$ and $A \subset \mathbb{Z}$ with $|A| = n$. Then*

$$|A + A| + |AA| \geq Cn^{2-\varepsilon}.$$

This conjecture is still out of reach. The best known bound, which holds for real numbers and not just integers, is $Cn^{4/3-o(1)}$ due to Solymosi [58]. A similar bound was proved recently by Konyagin and Rudnev in [40].

Chang showed that when the product set is small the Subspace Theorem can be used to show that the sum set is large [10]. The following reformulation of Chang's observation is due to Andrew Granville.

Theorem 5.3.2. *Let $A \subset \mathbb{C}$ with $|A| = n$. Suppose $|AA| \leq Cn$. Then there is a constant C' depending only on C such that*

$$|A + A| \geq \frac{n^2}{2} + C'n.$$

We will present a proof of Theorem 5.3.2 below. To use the Subspace Theorem we need a multiplicative subgroup with finite rank to work with. The following lemma of Freiman, which appears as Lemma 1.14 in [31], provides this.

Lemma 5.3.3 (Freiman)**.** *Let $A \subset \mathbb{C}$. If $|AA| \leq C|A|$ then A is a subset of a multiplicative subgroup of $\mathbb{C}^*$ of rank at most $r(C)$.*

Proof of Theorem 5.3.2. We consider solutions of $x_1 + x_2 = x_3 + x_4$ with $x_i \in A$. A solution of this equation corresponds to two pairs of elements from A that give the same element in $A + A$. Let us suppose that $x_1 + x_2 \neq 0$ (there are at most $|A| = n$ solutions of the equation $x_1 + x_2 = 0$ with $x_1, x_2 \in A$.)

First we consider the solutions with $x_4 = 0$. Then by rearranging we get

$$\frac{x_1}{x_3} + \frac{x_2}{x_3} = 1. \tag{5.3}$$

By Lemma 5.3.3 and Theorem 5.1.3 there are at most $s_1(C)$ solutions of $y_1+y_2 = 1$ with no subsum vanishing. Each of these gives at most n solutions of (5.3) since there are n choices for x_3. There are only two solutions of $y_1 + y_2 = 1$ with a vanishing subsum, namely $y_1 = 0$ or $y_2 = 0$, and each of these gives n solutions of (5.3). So we have a total of $(s_1(C) + 2)n$ solutions of (5.3).

For $x_4 \neq 0$ we get

$$\frac{x_1}{x_4} + \frac{x_2}{x_4} - \frac{x_3}{x_4} = 1. \tag{5.4}$$

Again by Freiman's Lemma and the Subspace Theorem, the number of solutions of this with no vanishing subsum is at most $s_2(C)n$. If we have a vanishing subsum then $x_1 = -x_2$ which is a case we excluded earlier or

$x_1 = x_3$ and then $x_2 = x_4$, or $x_2 = x_3$ and then $x_1 = x_4$. So we get at most $2n^2$ solutions of (5.4) with a vanishing subsum (these are the $x_1+x_2 = x_2+x_1$ identities.)

So, in total, we have at most $2n^2 + s(C)n$ solutions of $x_1 + x_2 = x_3 + x_4$ with $x_i \in A$. Suppose $|A + A| = k$ and $A + A = \{\alpha_1, \ldots, \alpha_k\}$. We may assume that $\alpha_1 = 0$. Recall that we ignore sums $a_i + a_j = 0$. Let

$$P_i = \{(a, b) \in A \times A : a + b = \alpha_i\}, \quad 2 \le i \le k.$$

Then

$$\sum_{i=2}^{k} |P_i| \ge n^2 - n = n(n-1).$$

Also, a solution of $x_1 + x_2 = x_3 + x_4$ corresponds to picking two values from P_i where $x_1 + x_2 = \alpha_i$. Thus

$$2n^2 + s(C)n \ge \sum_{i=2}^{k} |P_i|^2 \ge \frac{1}{k-1} \left(\sum_{i=2}^{k} |P_i| \right)^2 \ge \frac{n^2(n-1)^2}{k-1}$$

by the Cauchy-Schwarz inequality. The bound for k follows. □

5.3.2 Line configurations with few intersections

A number of other combinatorial results follow from the Subspace Theorem. We give one more of these, from combinatorial geometry. This is similar to a result due to Chang and Solymosi [11]. A *complex line* is a line in the complex plane. Specifically, a complex line is given by an equation $ax + by = c$ where $a, b, c \in \mathbb{C}$ and x and y are the (complex) coordinates in $\mathbb{C}^2$. Given two lines L and M we denote their intersection point by $L \cap M$.

Theorem 5.3.4. *Let $C > 0$. Then there exists $c > 0$ such that for any $n + 3$ lines $L_1, L_2, L_3, M_1, \ldots, M_n$ in $\mathbb{C}^2$, with the L_i not all parallel and $L_1 \cap L_2, L_1 \cap L_3$ and $L_2 \cap L_3$ distinct the following holds. If the number of distinct intersection points $L_i \cap M_j, 1 \le i \le 3, 1 \le j \le n$, is at most $C\sqrt{n}$ then any line $L \notin \{L_1, L_2, L_3\}$ has at least cn distinct intersection points $L \cap M_j, 1 \le j \le n$.*

This is a structure result. We have already mentioned other structure results in the form of Beck's Theorem and the results of Elekes and Rónyai and their extensions from Chapter 2.

We do not prove Theorem 5.3.4 completely but only give a sketch of how it follows from the Subspace Theorem. We don't try to find an efficient

quantitative version here and we don't explain the methods used in detail. The techniques applied are standard methods in additive combinatorics. All the details can be found in [65]. The proof requires the Szemerédi Regularity Lemma [63] and the Balog-Szemerédi Theorem [6].

Apply a projective transformation which moves L_1 to the x-axis, L_2 to the y-axis, and L_3 to the horizontal line $y = 1$. The three lines have distinct intersection points thus such a transformation exists. Let us denote the x-coordinates of $L_1 \cap M_i$ and $L_3 \cap M_j$ by x_i and y_j respectively. The two sets of x-coordinates are denoted by X and Y. Define a bipartite graph with vertices given by the x-coordinates of the intersection points of the lines M_i with L_1 and L_3 (with vertex sets X and Y without multiplicity.) Two points are connected by an edge in the graph if they are connected by a line M_j. This is a bipartite graph on at most $C\sqrt{n}$ vertices with n edges. Using Szemerédi's Regularity Lemma one can find a regular (random-like) bipartite graph, G, with at least $c'n$ edges and vertex sets $V_1 \subset X$ and $V_2 \subset Y$. If $M_i \cap L_2$ is the point $(0, \alpha)$ then $x_i/y_i = \alpha/(\alpha - 1)$, or equivalently $x_i = \alpha y_i/(\alpha - 1)$. The Balog-Szemerédi Theorem and Freiman's Lemma imply that there are large subsets $X' \subset V_1$ and $Y' \subset V_2$ so that X' and Y' are subsets of a multiplicative subgroup of $\mathbb{C}^*$ of rank at most $r(C)$. As G is regular, the subgraph spanned by X', Y' still has at least some $c''n$ edges. We show that the lines represented by these $c''n$ edges cannot have high multiplicity intersections outside of L_1, L_2, L_3. If (a, b) is a point of M_i connecting two points of X' and Y' then $(a - x_i)/(a - y_i) = b/(b - 1)$, which gives the solution (x_i, y_i) to the equation $cx + dy = 1$ if $a \neq 0, b \neq 0, 1$. Here c, d depend on a and b only. As x_i and y_i are from a multiplicative group of bounded rank, we have a uniform bound, B, on the number of lines between X' and Y' which are incident to (a, b). There are $c''n$ lines connecting at most $C\sqrt{n}$ points. No more than $C\sqrt{n}/2$ of them might be parallel to any given line. Any line intersects at least $c''n - C\sqrt{n}$ of them. Any intersection point outside of the lines L_1, L_2, and L_3 is incident to at most B lines, so there are at least cn distinct intersection points $L \cap M_j, 1 \leq j \leq n$, with any other line.

We are unaware of any proof of this fact without the Subspace Theorem.

Chapter 6

Conclusion

We have given a number of results in this thesis using a wide range of different techniques from various fields. All of the problems covered have a rich history and background and have been studied by many researchers. Strong progress has been made towards these problems but many questions still remain and a number of the results have the potential to be altered or improved.

6.1 Surfaces containing many points of a cartesian product

Our extensions of the Elekes-Rónyai Theorem give results on asymmetric cartesian products and in one dimension higher. For the first case one could ask if even more asymmetric cartesian products can be considered. We know that Theorem 2.1.6 cannot be improved by much because of the construction given in Section 2.5.3 but we could not find similar constructions for our other theorems and it is believed that these results are not tight. We know this is true for certain polynomials as can be seen in the best bounds for Purdy's Conjecture.

For our higher dimensional cases we used similar techniques. But we required an additional tool, namely the special case of the Schwartz-Zippel Lemma. We believe that a similar method should work for even higher dimensional versions but we did not perform the analysis.

Our main tool in improving Elekes and Rónyai's result was Lemma 2.2.4, our generalised line lemma. The improvement in this lemma gave the improved bounds. The rest of the proof remained the same. We could not find any other places in the proof to improve the bounds. Perhaps a different approach to the problem could give better bounds. Our proof of the generalised line lemma went along the same lines as that of Elekes. We just provided a more careful analysis.

The line lemma requires that all the lines are cn-rich on an $n \times n$ cartesian product. If we cannot guarantee this richness then none of the methods

work.

As mentioned in Chapter 2, our generalised line lemma and all the other elements of our proofs extend to curves over $\mathbb{C}$.

All our results hold for graphs of polynomials. But what if we instead consider implicit surfaces? Elekes and Szabó have given a version of the Elekes-Rónyai Theorem in this case when the surface contains $cn^{2-\gamma}$ points of an $n \times n \times n$ cartesian product for some small absolute constant γ [23, 24]. The proof requires tools from algebraic geometry. Using our generalised line lemma it should be possible to extend this result as well.

Regarding applications of the Elekes-Rónyai Theorem, much work has been done on Purdy's Conjecture. Around the same time as his work with Rónyai, Elekes showed using a similar method, but tailored for the specific polynomial from the problem, that Purdy's Conjecture holds if there are at most $cn^{5/4}$ distinct distances [19]. He conjectured that the same should hold for $cn^{2-\delta}$ distinct distances. Very recently Sharir, Sheffer and Solymosi have shown that the number of distinct distances can be raised to $cn^{4/3}$ [57]. Their method is very different. They show that the collinear point sets give rise to a collection of hyperbolas with two degrees of freedom and multiplicity-type 2 and then they use the Pach-Sharir Theorem.

6.2 The Erdős unit distance problem

We have dealt with two special cases of the unit distance problem, the second generalising the first. For our stronger version using the corollary of the Subspace Theorem we have shown that the lower bound construction is satisfied by our result. This is promising in that it suggests that perhaps the unit distances in any maximal configuration should have a special group structure. Possible future work could be to try and understand the structure of maximal unit distance sets and to show that the unit distances come from a group with "low" rank.

During the 27th European Workshop on Computational Geometry, 2011, Jiří Matoušek highlighted a number of results he believed to be important to the field of Discrete Geometry [47]. Our special case of the unit distance problem using Mann's Theorem was mentioned showing the relevance of this work, and its subsequent generalisation, to the field.

Using Mann's Theorem we were also able to get bounds when considering rational distances. We could not get similar results using the Subspace Theorem as it depends on the coefficients of the linear equation.

Instead of considering the Euclidean plane we could ask for the number

of unit distances in other spaces. Matoušek has given many results on the unit distance problem in different metric spaces [45]. In particular he has shown that there are at most $cn \log n \log \log n$ unit distances for "most" norms on the plane. So the Euclidean norm is special in a way.

Returning to the Euclidean distance but moving to higher dimensions the problem changes significantly. For $\mathbb{R}^4$ and higher the number of unit distances increases dramatically and the problem is much simpler. The following construction shows this and was originally communicated to Erdős by Lenz. Consider two orthogonal circles in $\mathbb{R}^4$ of radius $1/\sqrt{2}$ and with the same centre. For example

$$C_1 = \{(x_1, x_2, 0, 0) : x_1^2 + x_2^2 = 1/2\}, \quad C_2 = \{(0, 0, x_3, x_4) : x_3^2 + x_4^2 = 1/2\}.$$

The distance between $(x_1, x_2, 0, 0)$ and $(0, 0, x_3, x_4)$ is $\sqrt{1/2 + 1/2} = 1$. Thus every point on the first circle has unit distance to every point on the second circle. So taking $n/2$ points on each circle we see that we have at least $n^2/4$ unit distances. A similar argument can be used in higher dimensions.

In $\mathbb{R}^3$ the maximum number of unit distances is at least $cn^{4/3} \log \log n$. This was again shown by Erdős using a scaled portion of a three dimensional integer lattice and using number-theoretic techniques [26]. The best known upper bound is $cn^{3/2}$ due independently to Kaplan et al. [37] and Zahl [69] using the polynomial partitioning method of Guth and Katz. These slightly improved the previous bound due to Clarkson et al. [14]. This might be an interesting direction to consider.

The unit distance problem is also interesting in $\mathbb{C}^2$. The same upper bound is possible since the Szemerédi-Trotter Theorem holds over $\mathbb{C}^2$ [67], [68]. Theorem 4.1.2, our main result of Chapter 4, holds in the complex plane since all elements of the proof including the Subspace Theorem extend to the complex case.

6.3 Combinatorial applications of the Subspace Theorem

A number of applications of the Subspace Theorem have been given and it is believed that many more should be possible. It is an extremely powerful theorem that deserves to be recognised in yet another field. It has been used to give relatively simple proofs of previously known results in combinatorics and it has provided proofs to previously open problems. The generality of the Subspace Theorem means that it often gives results for complex numbers where many other methods in combinatorics are restricted to the real case.

It seems likely that the Subspace Theorem can shed more light on sum-product problems. Chang's result on the size of the sum set and product set given in Section 5.3.1 relies on Freiman's Lemma saying that if the product set is small then the elements of A should come from a group of bounded rank. If we have that the elements of A come from a group of rank at most $c \log |A|$ then $|A + A|$ is at least $|A|^{2-\varepsilon}$. So one could try and show that if the product set has size at most $|A|^{2-\delta}$ then A comes from a group with rank at most $c \log |A|$. This would prove Conjecture 5.3.1.

Bibliography

[1] B. Adamczewski and Y. Bugeaud. On the complexity of algebraic numbers I: Expansions in integer bases. *Annals of Mathematics*, 165:547–565, 2007.

[2] G. Amirkhanyan, A. Bush, E. Croot, and C. Pryby. On rich lines in general position in grids. Presented in "Additive Combinatorics in Paris 2012", Institut Henri Poincaré, July 2012. http://caparis2012en.files.wordpress.com/2011/05/ernie_croot.pdf.

[3] F. Amoroso and E. Viada. Small points on subvarieties of a torus. *Duke Mathematical Journal*, 150(3):407–442, 2009.

[4] F. Amoroso and E. Viada. On the zeros of linear recurrence sequences. *Acta Arithmetica*, 147:387–396, 2011.

[5] T.M. Apostol. *Introduction to Analytic Number Theory*, chapter 4.5: Inequalities for $\pi(n)$ and p_n, pages 82–85. Springer, 1976.

[6] A. Balog and E. Szemerédi. A statistical theorem of set addition. *Combinatorica*, 14:263–268, 1994.

[7] J. Beck. On the lattice property of the plane and some problems of Dirac, Motzkin, and Erdős in combinatorial geometry. *Combinatorica*, 3(3):281–297, 1983.

[8] Y. Bilu. The Many Faces of the Subspace Theorem (after Adamczewski, Bugeaud, Corvaja, Zannier...). Séminaire Bourbaki, Exposé 967, 59ème année (2006-2007); Astérisque 317 (2008), 1-38., May 2007.

[9] W. Brass, W. Moser, and J. Pach. *Research Problems in Discrete Geometry*. Springer, 2006.

[10] M.-C. Chang. Sum and product of different sets. *Contributions to Discrete Mathematics*, 1(1), 2006.

[11] M.-C. Chang and J. Solymosi. Sum-product theorems and incidence geometry. *Journal of the European Mathematical Society*, 9(3):545–560, 2007.

[12] F.K.R. Chung. The number of different distances determined by n points in the plane. *Journal of Combinatorial Theory, Series A*, 36:342–354, 1984.

[13] F.K.R. Chung, E. Szemerédi, and W.T. Trotter. The number of different distances determined by a set of points in the Euclidean plane. *Discrete and Computational Geometry*, 7:1–11, 1992.

[14] K. Clarkson, H. Edelsbrunner, M. Sharir, and E. Welzl. Combinatorial complexity bounds for arrangements of curves and surfaces. *Discrete and Computational Geometry*, 5:99–160, 1990.

[15] J.H. Conway and A.J. Jones. Trigonometric diophantine equations (on vanishing sums of roots of unity). *Acta Arithmetica*, 30:229–240, 1976.

[16] P. Corvaja and U. Zannier. Applications of the subspace theorem to certain diophantine problems. In H.P. Schlickewei, Schmidt K., and R.F. Tichy, editors, *Diophantine Approximation*, volume 16 of *Developments in Mathematics*, pages 161–174. Springer Vienna, 2008.

[17] N.G. de Bruijn. *Asymptotic Methods in Analysis*. North-Holland, 1961.

[18] Gy. Elekes. On linear combinatorics, I. *Combinatorica*, 17(4):447–458, 1997.

[19] Gy. Elekes. A note on the number of distinct distances. *Periodica Mathematica Hungarica*, 38:173–177, 1999.

[20] Gy. Elekes. Sums versus products in number theory, algebra and Erdős geometry. In *Paul Erdős and his Mathematics II*, volume 11 of *Bolyai Society Mathematical Studies*, pages 241–290. 2002.

[21] Gy. Elekes and L. Rónyai. A combinatorial problem on polynomials and rational functions. *Journal of Combinatorial Theory, Series A*, 89:1–20, 2000.

[22] Gy. Elekes and M. Sharir. Incidences in three dimensions and distinct distances in the plane. *Combinatorics, Probability and Computing*, 20(4):571–608, 2011.

[23] Gy. Elekes, M. Simonovits, and E. Szabó. A combinatorial distinction between unit circles and straight lines. *Combinatorics, Probability and Computing*, 18:691–705, 2009.

[24] Gy. Elekes and E. Szabó. How to find groups? (and how to use them in Erdős geometry?). *Combinatorica*, 32:537–571, 2012.

[25] P. Erdős. On sets of distances of n points. *American Mathematical Monthly*, 53(5):248–250, May 1946.

[26] P. Erdős. On sets of distances of n points in euclidean space. *Magy. Tud. Akad. Mat. Kut. Int. Közl.*, 5:165–169, 1960.

[27] P. Erdős and G.B. Purdy. Some extremal problems in geometry, IV. In *Proceedings of The Seventh Southeastern Conference on Combinatorics, Graph Theory, and Computing*, pages 307–322, 1976.

[28] P. Erdős and E. Szemerédi. On sums and products of integers. In *Studies in Pure Mathematics*, pages 213–218. Birkhäuser, 1983.

[29] J.-H. Evertse and H.P. Schlickewei. The absolute subspace theorem and linear equations with unknowns from a multiplicative group. In K. Győry, H. Iwaniec, and J. Urbanowicz, editors, *Number theory in progress: Proceedings of the international conference of number theory in honour of the 60th birthday of Andrzej Schinzel*, 1999.

[30] J.-H. Evertse, H.P. Schlickewei, and W.M. Schmidt. Linear equations in variables which lie in a multiplicative group. *Annals of Mathematics*, 155(3):807–836, 2002.

[31] G.A. Freiman. *Foundations of a Structural Theory of Set Addition*. Translations of Mathematical Monographs. American Mathematical Society, 1973.

[32] William Fulton. *Algebraic Curves: An Introduction to Algebraic Geometry*. 1969.

[33] J. Garibaldi, A. Iosevich, and S. Senger. *The Erdős Distance Problem*, volume 56 of *Student Mathematical Library*. American Mathematical Society Press, 2011.

[34] L. Guth and N.H. Katz. On the Erdős distinct distances problem in the plane. *Preprint*, arXiv:1011.4105, 2010.

[35] H. Iwaniec and E. Kowalski. *Analytic Number Theory*. Volume 53 of AMS Colloquium Publications, 2004.

[36] S. Jukna. *Extremal Combinatorics*. Springer, 2011.

[37] A. Kaplan, J. Matoušek, Z. Safernová, and M. Sharir. Unit distances in three dimensions. *Combinatorics, Probability and Computing*, 21:597–610, 2012.

[38] N.H. Katz and G. Tardos. A new entropy inequality for the Erdős distance problem. 342:119–126, 2004.

[39] A.J. Kempner. On transcendental numbers. *Transactions of the American Mathematical Society*, 17(4):476–482, 1916.

[40] S.V. Konyagin and M. Rudnev. On new sum-product type estimates. *Preprint*, arXiv:1207.6785, 2013.

[41] J.H. Lambert. Observationes variae in mathesin puram. *Acta Helvetica, physico-mathematico-anatomico-botanico-medica*, 3:128–168, 1758.

[42] S. Lang. *Algebra*. Addison-Wesley, 3rd edition, 1994.

[43] C. Lech. A note on recurring series. *Arkiv der Mathematik*, 2:417–421, 1953.

[44] H.B. Mann. On linear relations between roots of unity. *Mathematika*, 12:107–117, 1965.

[45] J. Matoušek. The number of unit distances is almost linear for most norms. *Advances in Mathematics*, 226:2618–2628, 2011.

[46] J. Matoušek. *Lectures on Discrete Geometry*. Springer, 2002.

[47] J. Matoušek. The dawn of an algebraic era in discrete geometry? *EuroCG 2011*, extended abstract, 2011.

[48] J. McKay and S. Wang. An inversion formula for two polynomials in two variables. *Journal of Pure and Applied Algebra*, 40:245–257, 1986.

[49] H.L. Montgomery and R.C. Vaughan. *Multiplicative number theory. I. Classical theory*. Cambridge Studies in Advanced Mathematics, 97. Cambridge University Press, 2007.

[50] L. Moser. On different distances determined by n points. *American Mathematical Monthly*, 59:85–91, 1952.

[51] I. Niven, H.S. Zuckerman, and H.L. Montgomery. *An introduction to the theory of numbers*. John Wiley & Sons, Inc., fifth edition, 1991.

[52] J. Pach and M. Sharir. Repeated angles in the plane and related problems. *Journal of Combinatorial Theory, Series A*, 59(1):12–22, 1992.

[53] J. Pach and M. Sharir. On the number of incidences between points and curves. *Combinatorics, Probability and Computing*, 7:121–127, 1998.

[54] K.F. Roth. Rational approximations to algebraic numbers. *Mathematika*, 2(1):1–20, 1955.

[55] W.M. Schmidt. Norm form equations. *The Annals of Mathematics, Second Series*, 96(3):526–551, 1972.

[56] W.M. Schmidt. The zero multiplicity of linear recurrence sequences. *Acta Mathematica*, 182(2):243–282, 1999.

[57] M. Sharir, A. Sheffer, and J. Solymosi. Distinct distances on two lines. *Journal of Combinatorial Theory, Series A*, 120(7):1732–1736, 2013.

[58] J. Solymosi. Bounding multiplicative energy by the sumset. *Advances in Mathematics*, 222(2):402–408, 2009.

[59] J. Solymosi and C.D. Tóth. Distinct distances in the plane. *Discrete and Computational Geometry*, 25:629–634, 2001.

[60] J. Spencer, E. Szemerédi, and W. Trotter. Unit distances in the Euclidean plane. In B. Bollobas, editor, *Graph Theory and Combinatorics: Proceedings of the Cambridge Combinatorial Conference, in Honour of Paul Erdős*, pages 293–303. Academic Press, 1984.

[61] A.H. Stone and J.W. Tukey. Generalized sandwich theorems. *Duke Mathematical Journal*, 9:356–359, 1942.

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

[63] E. Szemerédi. Regular partitions of graphs. In *Problèmes Combinatoires et Théorie des Graphes*, pages 399–401, Orsay, 1978. Colloques Internationaux C.N.R.S. (1976).

[64] E. Szemerédi and W.T. Trotter. Extremal problems in discrete geometry. *Combinatorica*, 3(3-4):381–392, 1983.

[65] T. Tao and V.H. Vu. *Additive Combinatorics*. Cambridge Studies in Advanced Mathematics 105. Cambridge University Press, 2006.

[66] G. Tardos. On distinct sums and distinct distances. *Advances in Mathematics*, 180:275–289, 2003.

[67] C.D. Toth. The Szemerédi-Trotter theorem in the complex plane. *Preprint*, arXiv:0305283, 2003.

[68] J. Zahl. A Szemeredi-Trotter type theorem in $\mathbb{R}^4$. *Preprint*, arXiv:1203.4600, 2012.

[69] J. Zahl. An improved bound on the number of point-surface incidences in three dimensions. *Contributions to Discrete Mathematics*, 8(1):100–121, 2013.