A survey of Elekes-Rónyai-type problems

Frank de Zeeuw

Abstract

We give an overview of recent progress around a problem introduced by Elekes and Rónyai. The prototype problem is to show that a polynomial \( f \in \mathbb{R}[x, y] \) has a large image on a Cartesian product \( A \times B \subset \mathbb{R}^2 \), unless \( f \) has a group-related special form. We discuss this problem and a number of variants and generalizations. This includes the Elekes-Szabó problem, which generalizes the Elekes-Rónyai problem to a question about an upper bound on the intersection of an algebraic surface with a Cartesian product, and curve variants, where we ask the same questions for Cartesian products of finite subsets of algebraic curves.

These problems lie at the crossroads of combinatorics, algebra, and geometry: They ask combinatorial questions about algebraic objects, whose answers turn out to have applications to geometric questions involving basic objects like distances, lines, and circles, as well as to sum-product-type questions from additive combinatorics. As part of a recent surge of algebraic techniques in combinatorial geometry, a number of quantitative and qualitative steps have been made within this framework. Nevertheless, many tantalizing open questions remain.

Contents

1 The Elekes-Rónyai problem
  1.1 Sums, products, and expanding polynomials .......................... 2
  1.2 Extensions ...................................................... 3
  1.3 Applications .................................................... 4
  1.4 About the proof of Theorem 1.1 .................................. 6

2 The Elekes-Szabó problem
  2.1 Intersecting varieties with Cartesian products .................... 8
  2.2 A derivative test for special \( F \) .................................. 10
  2.3 Applications ...................................................... 11
  2.4 About the proof of Theorem 2.1 .................................. 13

3 Elekes-Rónyai problems on curves
  3.1 Distances on curves ................................................ 14
  3.2 Other polynomials on curves ...................................... 16
  3.3 Elekes-Szabó problems on curves .................................. 17
  3.4 Collinear triples on curves ...................................... 19

4 Other variants
  4.1 Longer one-dimensional products ................................. 20
  4.2 Two-dimensional products ....................................... 21
  4.3 Expanding polynomials over other fields ........................ 23
1 The Elekes-Rónyai problem

1.1 Sums, products, and expanding polynomials

Erdős and Szemerédi [23] introduced the following problem in 1983. Given a finite set A
in some ring, is it true that the sumset A + A or the productset A · A must be large? The
rationale is that for an arithmetic progression the sumset is small, but the productset is
large, while for a geometric progression, the reverse is true. Erdős and Szemerédi proved
for A ⊂ Z that
\[ \max\{|A + A|, |A · A|\} = \Omega(|A|^{1+c}) \]  
(1)
for a very small c > 0. This statement was later generalized to \( \mathbb{R} \), and the constant has
over the years been improved to \( 4/3 + c' \) for a small c' > 0 [52, 28].

The intuition behind this statement is that a set cannot have many “coincidences”
for both addition and multiplication. A statement like (1) is not the only way to cap-
ture this intuition. Elekes [11] suggested that, since polynomials combine addition and
multiplication, for most polynomials \( f \in \mathbb{R}[x, y] \) we should have
\[ |f(A \times A)| = \Omega(|A|^{1+c}) \]  
(2)
for any finite A ⊂ \( \mathbb{R} \), with a constant c > 0 that may depend on f. More generally, we
should have
\[ |f(A \times B)| = \Omega(n^{1+c}) \]  
(3)
for A, B ⊂ \( \mathbb{R} \) with |A| = |B| = n, and a similar bound when A and B have different sizes.

Of course, (3) cannot hold for all polynomials, as it fails when \( f(x, y) = x + y \) and
A and B are arithmetic progressions with the same difference, or when \( f(x, y) = xy \) and
A and B are geometric progressions with the same ratio. More generally, if \( f \) has the
additive form
\[ f(x, y) = g(h(x) + k(y)) \]  
(4)
with univariate polynomials g, h, k, then one has |f(A × B)| = O(n) if one chooses A and
B in such a way that h(A) and k(B) are arithmetic progressions. Similarly, if \( f \) has the
multiplicative form
\[ f(x, y) = g(h(x) \cdot k(y)) \]  
(5)
then |f(A × B)| = O(n) if h(A) and k(B) are geometric progressions. We will call a
polynomial additive if it has the form in (4), and multiplicative if it has the form in (5).

Elekes [11] conjectured that the additive form (4) and the multiplicative form (5)
are the only exceptions to the bound (3). Elekes proved a weaker form of this statement in
[11], and he collaborated with Rónyai [16] to prove this conjecture in full. We state an
improved version of the result due to Raz, Sharir, and Solymosi [39].

**Theorem 1.1** (Elekes-Rónyai, Raz-Sharir-Solymosi). Let \( f \in \mathbb{R}[x, y] \) be a polynomial of
degree \( d \) that is not additive or multiplicative. Then for all A, B ⊂ \( \mathbb{R} \) with |A| = |B| = n
we have
\[ |f(A \times B)| = \Omega_d(n^{4/3}) \].
To be precise, the bound stated in 15 was \(|f(A \times B)| = \omega(n)| \), but inspection of the proof leads to a bound of the form \(|f(A \times B)| = \Omega(n^{1+c_d}) |\) with a constant \(c_d > 0 |\) depending on the degree of \(f |\). Raz, Sharir, and Solymosi 39 showed that this dependence on the degree of \(f |\) is not necessary, and moreover improved the constant significantly. Their proof used a setup inspired by Sharir, Sheffer, and Solymosi 41, which gave a similar improvement on a special case (see Theorem 1.3). The same bound \(|f(A \times B)| = \Omega(n^{4/3}) |\) was obtained by Hegyvári and Hennecart 27 Proposition 8.3] for the polynomial \(f(x, y) = xy(x + y) |\).

The exceptional role played by the additive and multiplicative forms suggests that groups play a special role in this type of theorem. We will touch on this in Section 2, but see Elekes and Szabó 19 Subsection 1.2] for further discussion. Other (older) expositions of Theorem 1.1 can be found in Elekes’s magnificent survey 14, and in Matoušek’s book 32 notes to Section 4.1].

1.2 Extensions

It was observed in 33 that a bound like (3) should also hold when \(A |\) and \(B |\) have different sizes, and this was proved in a weak sense. In 39 such an “unbalanced” form of Theorem 1.1 was proved: If \(f |\) is not additive or multiplicative, and \(A, B \subset \mathbb{R} |\), then

\[
|f(A \times B)| = \Omega_d \left( \min \left\{ |A|^{2/3} |B|^{2/3}, |A|^2, |B|^2 \right\} \right). \tag{6}
\]

Another way in which Theorem 1.1 can be extended is by replacing polynomials by rational functions. This was indeed done in 16, but not in 39. Here the exceptions include the same forms (4) and (5) with \(g, h, k |\) rational functions, but surprisingly, a third special form shows up here, namely

\[
f(x, y) = g \left( \frac{k(x) + l(y)}{1 - k(x)l(y)} \right) \tag{7}
\]

with rational functions \(g, k, l |\). It was pointed out in 7 that over \(\mathbb{C} |\) this can be seen as a multiplicative form, because \(g(\overline{k(x)} + \overline{l(y)})/(1 - k(x)l(y)) = G(K(x)L(y)) |\) if we set \(G(z) = (z - 1)/(iz + 1)) |\), \(K(x) = (1 + ik(x))/(1 - ik(x)) |\), and \(L(y) = (1 + il(y))/(1 - il(y)) |\) (and if we do some tedious computation).

It remains an open problem to improve the bound \(|f(A \times B)| = \omega(n)|\) of 10 for rational functions \(f |\). For one special case, the rational function \(f(x, y) = (x - y)^2/(1 + y^2) |\), the bound \(|f(A \times A)| = \Omega(|A|^{4/3}) |\) was proved in 35. This was done in the context of the distinct distance problem for distances between points and lines; \(f |\) gives the distance between the point \((x, 0) |\) and the line spanned by the points \((y, 0) |\) and \((0, 1) |\).

Problem 1.2. Prove Theorem 1.1 for rational functions \(f \in \mathbb{R}[x, y] |\) that are not additive, multiplicative, or of the form (7).

Yet another way to extend Theorem 1.1 is from \(\mathbb{R} \) to \(\mathbb{C} |\). Most of the proof in 39 extends easily to \(\mathbb{C} |\), with the exception of the incidence bound used (see Subsection 1.4), which can be replaced by the bound over \(\mathbb{C} |\) proved in 53 (see Theorem 1.6); the details are written down in 62. Thus Theorem 1.1 holds also over \(\mathbb{C} |\).

\[^{3}\text{The result in 27 Proposition 8.3 has the same proof setup as 41, but it appears to be somewhat isolated; it makes no reference to 16, and in turn is not referred to in 41 and 39. This may be because 27 primarily concerns expansion bounds over finite fields. The arXiv publication date of 41 is half a year before that of 27.}\]
The exponent $4/3$ in Theorem 1.1 is most likely not optimal. Elekes [12] in fact conjectured that the bound in Theorem 1.1 can be improved as far as $\Omega(n^{2-\varepsilon})$, but no exponent better than $4/3$ has been established for any polynomial. Elekes [12] noted that for $f(x,y) = x^2 + xy + y^2$ (and many other polynomials) and $A = B = \{1, \ldots, n\}$ we have $|f(A \times B)| = \Theta(n^2/\sqrt{\log n})$ (see [17], Chapter 6) for details, so perhaps the bound in Theorem 1.1 can even be improved to $\Omega(n^{2-\varepsilon}/\log n)$.

Let us combine the above extensions to conjecture the ultimate Elekes-Rónyai-type theorem (although see Section 4 for further variants).

**Conjecture 1.3.** Let $f \in \mathbb{C}(x,y)$ be a rational function of degree $d$ that is not additive or multiplicative. Then for all $A,B \subset \mathbb{C}$ with $|A| = |B| = n$ we have

$$|f(A \times B)| = \Omega_{d,\varepsilon} \left(n^{2-\varepsilon}\right).$$

### 1.3 Applications

**Sum-product bounds.** For a first consequence of Theorem 1.1, we return to the sum-product problem mentioned at the start of Subsection 1.1. The following generalization of the bound (1) was proved by Shen [50]: If $A \subset \mathbb{R}$ and $f \in \mathbb{R}[x,y]$ is a polynomial of degree $d$ that is not of the form $g(\ell(x,y))$, with $g$ a univariate polynomial and $\ell$ a linear bivariate polynomial, then

$$\max\{|A+A|, |f(A \times A)|\} = \Omega_d \left(|A|^{5/4}\right).$$

For many polynomials, Theorem 1.1 improves this bound, and it also shows that in those cases one does not need to consider $|A+A|$ to conclude that $|f(A \times A)|$ is large.

On the other hand, there are many polynomials that have the special form of Theorem 1.1 but that do not have the special form of Shen. Even in those cases, it may be possible to obtain a bound on $|f(A \times A)|$ independent of $|A+A|$; for instance, Elekes, Nathanson, and Ruzsa [15] prove $|f(A \times A)| = \Omega_d(|A|^{5/4})$ for $f(x,y) = x + y^2$ (and many similar functions). Note that for this bound it is crucial that the Cartesian product is of the form $A \times A$ rather than $A \times B$. It may be that such a bound holds for any $f$ that is not of the form $g(h(x) + h(y))$ or $g(h(x) \cdot h(y))$; this question does not seem to have been studied.

**Distances between lines.** As a corollary of their result, Elekes and Rónyai [16] made progress on the following problem of Purdy (see [5], Section 5.5): Given two lines with $n$ points each, what is the minimum number of distances occurring between the two point sets? This problem is a simpler variant of the distinct distances problem of Erdős [21], which asks for the minimum number of distinct distances determined by a point set in the plane. Erdős’s problem was almost completely solved by Guth and Katz [24] using new algebraic methods.

In Purdy’s problem there are two exceptional situations, when the two lines are parallel or orthogonal. Indeed, if on two parallel lines one places two arithmetic progressions of size $n$ with the same common difference, then the number of distinct distances is linear in $n$. On two orthogonal lines, say the $x$-axis and the $y$-axis, one can take the sets $\{(\sqrt{i},0) : 1 \leq i \leq n\}$ and $\{(0,\sqrt{j}) : 1 \leq j \leq n\}$ to get a linear number of distances. The following theorem states that for all other pairs of lines there are considerably more.

---

2The maximum of the degrees of the numerator and denominator, assuming that these do not have a common factor.
The distinct directions problem asks for the minimum number of distinct directions determined by a non-collinear point set in the plane. It is superficially similar to the distinct distances problem, in the sense that it asks for the minimum number of distinct values of a function of pairs of points in the plane. However, it was solved exactly by Ungar [58] in 1982: Any non-collinear set \( P \) in \( \mathbb{R}^2 \) determines at least \( |P| - 1 \) distinct directions.

This leaves the more difficult structural question: What is the structure of sets that determine few distinct directions? Let us write \( S(P) \) for the set of directions (or slopes) determined by \( P \subset \mathbb{R}^2 \). Elekes [13] conjectured that if \( |S(P)| = O(|P|) \), then \( P \) must have many points on a conic; even in the weakest form, where “many” is six, this is unknown. Elekes [13] showed that a (very) restricted version of this conjecture follows from Theorem 1.1. If \( P \) lies on the graph of a polynomial of degree at most \( d \) and has \( |S(P)| = O_d(|P|) \), then the polynomial must be linear or quadratic. See [14, Subsection 3.3] for a detailed discussion. We state here the improvement of this result from [39], obtained as a consequence of Theorem 1.1. In Subsection 3.4 we will discuss the same question for points sets on arbitrary algebraic curves.

---

3The chronology is somewhat confusing here. The paper [16] was published in 2000, and [12] in 1999. However, [12] refers back to [16] and makes it clear that [12] is an improvement on a special case of [16].

4A preprint version of [24] became available in 2010.
Corollary 1.5. Let $P$ be a finite point set that is contained in the graph $y = g(x)$ of a polynomial $g \in \mathbb{R}[x]$ of degree $d \geq 3$. Then the number of distinct directions determined by $P$ satisfies

$$|S(P)| = \Omega_d \left( |P|^{4/3} \right).$$

Proof sketch. The direction determined by the points $(s, g(s)), (t, g(t))$ is given by the polynomial

$$f(s, t) = \frac{g(s) - g(t)}{s - t}.$$ 

It is not hard to verify that this polynomial is not additive or multiplicative (except when $g$ is linear or quadratic), so Theorem 1.1 gives the stated bound.

When $g$ has degree less than three, the number of directions can be linear. Take for instance the parabola $y = x^2$ and the point set $\{(i, i^2) : 1 \leq i \leq n\}$. Then the determined directions are $(i^2 - j^2)/(i - j) = i + j$, so there are $O(n)$ distinct directions. It is not hard to see that a similar example can be constructed for any other quadratic $g$.

1.4 About the proof of Theorem 1.1

We now discuss the proof of Theorem 1.1. We certainly do not give a full account of the proof in [39], but we introduce the setup without too much technical detail. The proof of the weaker bound in [16] used related techniques but was inherently different. The proof setup in [39] originated in [44], and seems to have been discovered independently (and somewhat later) in [27, Proposition 8.3]; a glimpse of this setup can also be seen in the (earlier) proof of [19, Theorem 27]. We use the word “setup” to refer to the overall counting scheme of [39], which is only the surface layer of the proof. The real achievement in [39] was the treatment of high-multiplicity curves, which we discuss only briefly at the end of this subsection.

The main tool for obtaining the bound in Theorem 1.1 is incidence theory, specifically an incidence bound for points and curves of Pach and Sharir [34]. This incidence bound is a generalization of the classical theorem of Szemerédi and Trotter [55] that bounds incidences between points and lines. The following version is particularly convenient for the applications in this survey. It assumes that the point set is a Cartesian product, which allows for a significantly simpler proof over $\mathbb{R}$, and makes it easier to prove an analogue over $\mathbb{C}$ (where the corresponding bound has not yet been established without extra assumptions; see [48]). This theorem was proved by Solymosi and De Zeeuw [53], although the real case was probably folklore.

We write $|I(P, C)|$ for the set of incidences between the points $P$ and the curves $C$, i.e., the set of pairs $(p, C) \in P \times C$ such that $p \in C$. We give a quick sketch of the proof, to show that proving this tool does not require any heavy machinery (at least over $\mathbb{R}$).

Theorem 1.6. Let $P = A \times B$ be a Cartesian product in $\mathbb{R}^2$ or $\mathbb{C}^2$, and let $C$ be a set of algebraic curves of degree at most $d$ in the same plane. Assume that any two points of $P$ are contained in at most $M$ curves of $C$. Then

$$|I(P, C)| = O_{d, M} \left( |P|^{2/3} |C|^{2/3} + |P| + |C| \right).$$

5The proof in [39] did not directly use the bound from [34], but rather adapted the proof from [34] to the incidence situation in [39].
Proof sketch in \( \mathbb{R}^2 \). Let us assume that \(|A| = |B| = n\) and \(C = n^2\) (which roughly holds in most of the statements in this survey). We can partition \( \mathbb{R}^2 \) using \( O(r) \) horizontal and vertical lines, in such a way that each of the \( O(r^2) \) resulting rectangles contains roughly \( O(n^2/r^2) \) points of \( A \times B \). Moreover, we can ensure that the partitioning lines do not contain any points of \( A \times B \), and are not contained in any of the curves in \( C \). We observe that an algebraic curve of degree at most \( d \) intersects \( O_d(r) \) rectangles, since it can intersect a partitioning line in at most \( d \) points.

We split the incidences as follows: \( I_1 \) is the set of incidences \((p, C)\) such that \( p \) is the only incidence on \( C \) in the rectangle that \( p \) lies in, and \( I_2 \) is the remaining set of incidences. Since each curve hits \( O_d(r) \) rectangles, we have \(|I_1| = O_d(rn^2)\). On the other hand, if a curve has an incidence from \( I_2 \) in a certain rectangle, then it has at least one additional incidence in that same rectangle. By the assumption of the theorem, any two points are together contained in at most \( M \) curves. Thus the \( O(n^2/r^2) \) points in a rectangle are involved in at most \( O_M(n^3/r^4) \) incidences from \( I_2 \), which altogether gives \(|I_2| = O_M(n^4/r^2)\). Choosing \( r = n^{2/3} \) optimizes 

\[
|I(\mathcal{P}, \mathcal{C})| = |I_1| + |I_2| = O_{d,M}(rn^2 + n^4/r^2) = O\left(n^{8/3}\right),
\]

which is the stated bound when \(|\mathcal{P}| = |\mathcal{C}| = n^2\).

The proof of Theorem \([1, 1]\) is based on an upper bound for the size of the following set of quadruples:

\[
Q = \{(a, b, a', b') \in A \times B \times A \times B : f(a, b) = f(a', b')\}.
\]

Given such an upper bound, the Cauchy-Schwarz inequality gives a lower bound on the size of the image set \( f(A \times B) \), using the following calculation:

\[
|Q| = \sum_{c \in f(A \times B)} |f^{-1}(c)|^2 \geq \frac{1}{|f(A \times B)|} \left( \sum_{c \in f(A \times B)} |f^{-1}(c)| \right)^2 = \frac{n^4}{|f(A \times B)|}. \tag{9}
\]

Specifically, when \( f \) is not additive or multiplicative we obtain the upper bound \(|Q| = O_d(n^{8/3})\), and then (9) implies \(|f(A \times B)| = \Omega_d(n^{1/3})\). A similar application of Cauchy-Schwarz played a central role in [24].

To obtain an upper bound on \(|Q|\), we define a set of curves and a set of points based on the given polynomial \( f \) and the given sets \( A, B \), and then we apply Theorem \([1, 6]\). For each \((a, a') \in A \times A\), define

\[
C_{aa'} = \{(x, y) \in \mathbb{R}^2 : f(a, x) = f(a', y)\}.
\]

This is an algebraic curve of degree at most \( d \) (the degree of \( f \)). Note that for \((b, b') \in B \times B\), we have \((b, b') \in C_{aa'}\) if and only if \((a, b, a', b') \in Q\). Thus, if we set

\[
\mathcal{P} = B \times B \quad \text{and} \quad \mathcal{C} = \{C_{aa'} : (a, a') \in A \times A\},
\]

then \(|I(\mathcal{P}, \mathcal{C})| = |Q|\).

If we could apply Theorem \([1, 6]\) to \( \mathcal{P} \) and \( \mathcal{C} \), then we would immediately get the desired bound \(|Q| = O_d(n^{8/3})\). However, \( \mathcal{P} \) and \( \mathcal{C} \) need not satisfy the degrees-of-freedom condition of Theorem \([1, 6]\) that two points are contained in a bounded number of curves. This is to be expected, since the bound should fail when \( f \) is additive or multiplicative. In fact,
even when \( f \) is not additive or multiplicative, the degrees-of-freedom condition may be violated. However, it was shown in [39] that when \( f \) is not additive or multiplicative, the condition is only violated in a weak sense. Specifically, one can remove negligible subsets of the points and curves so that the remainder does satisfy the condition.

A key insight in the proof is that the curves \( C_{aa'} \) satisfy a kind of duality. Indeed, we can define “dual curves” of the form \( C_{bb'}^{*} = \{(s, t) \in \mathbb{R}^2 : f(s, b) = f(t, b')\} \), so that the point \((a, a')\) lies on the dual curve \( C_{bb'}^{*}\), if and only if the point \((b, b')\) lies on the curve \( C_{aa'} \). Thus, to check the degrees-of-freedom condition of Theorem 1.6 that many curves have high multiplicity, then there must be polynomials \( g, h, k \) such that \( h(a) + k(x) - h(a') - h(y) = 0 \), which corresponds to a component that can have high multiplicity (and the same happens for \( f(x, y) = g(h(x)\cdot k(y)) \)).

The key challenge in the proof of Theorem 1.1 is to obtain the converse, i.e., to show that when many curves have high multiplicity, then there must be polynomials \( g, h, k \) that explain the multiplicity in one of the ways above. In [39] this is done by algebraically prying out \( g, h, k \) from specific coefficients of the polynomial \( f \). For instance, roughly speaking, when the curves \( C_{aa'} \) have many common components and the coefficient of the leading term of \( f(a, x) - f(a', y) \) is not constant as a polynomial in \( a \), then this polynomial turns out to be the \( h \) in the multiplicative form \( f(x, y) = g(h(x)k(y)) \). On the other hand, if only the constant term of \( f(a, x) - f(a', y) \) depends on \( a \) and \( a' \), then this leads to the polynomial \( h \) in the additive form \( f(x, y) = g(h(x) + k(y)) \).

2 The Elekes-Szabó problem

2.1 Intersecting varieties with Cartesian products

To introduce a generalization of the Elekes-Rónyai problem due to Elekes and Szabó, we take a step back and approach from a different direction; after a while we will see what the connection between the problems is. In this section we work primarily over \( \mathbb{C} \), which is the most natural setting for the Elekes-Szabó problem and the relevant proofs.

Let us consider the Schwartz-Zippel lemma (see [30] for the curious history of this lemma). The simplest non-trivial case is the following bound on the intersection of a curve with a Cartesian product\(^6\). If \( F \in \mathbb{C}[x, y] \) is a polynomial of degree \( d \) and \( A, B \subset \mathbb{C} \)

\( B' \) is easily bounded by Bézout's inequality, unless the curves \( C_{bb'} \) have many common components, with high multiplicity, then \( f \) must be additive or multiplicative. By symmetry, we may as well consider this question for the original curves \( C_{bb'} \).

Let us see what happens when \( f \) is additive or multiplicative. First consider the case where \( f(x, y) = h(x) + k(y) \). Then \( C_{aa'} \) is defined by \( k(x) - k(y) = h(a') - h(a) \). Thus \( C_{aa'} \) is the same curve whenever \( h(a_1') - h(a_1) = h(a_2') - h(a_2) \); this means that as many as \( \Theta(|A|) \) pairs \((a, a')\) may define the same curve \( C_{aa'} \). Similarly, when \( f(x, y) = h(x)k(y) \), then \( C_{aa'} \) is defined by \( k(x) = (h(a')/h(a)) \), and again we can have high multiplicity. Finally, when \( f(x, y) = g(h(x) + k(y)) \), then \( f(a, x) - f(a', y) \) has the factor \( h(a) + k(x) - h(a') - h(y) \), which corresponds to a component that can have high multiplicity (and the same happens for \( f(x, y) = g(h(x) \cdot k(y)) \)).

\(^6\)The word “grid” is often used in this context, but may lead to confusion with integer grids.
are finite sets of size $n$, then\footnote{We write $Z(F)$ for the zero set of a polynomial $F$, i.e., the set of points at which $F$ vanishes.}

$$|Z(F) \cap (A \times B)| = O_d(n).$$

(10)

This statement is “tight” in the sense that, for any fixed polynomial, there are sets $A, B$ for which the bound is best possible. Indeed, we can arbitrarily choose $n$ points on $Z(F)$, let $A$ be the projection of this set to the $x$-axis, and let $B$ be the projection to the $y$-axis; then $A \times B$ shares at least $n$ points with $Z(F)$.

Now let us consider the tightness for the next case of the Schwartz-Zippel lemma (see Subsection 4.11 for the general statement), which says that for $F \in \mathbb{C}[x, y, z]$ and $A, B, C \subset \mathbb{C}$ of size $n$ we have

$$|Z(F) \cap (A \times B \times C)| = O_d(n^2).$$

(11)

It is not so clear if this bound is tight, since the trick used above to show that (10) is tight does not work here. The best we could try is to choose $A, B$ of size $n$, take $n^3$ points on $Z(F)$ above $A \times B$, and then project to the $z$-axis to get $C$; but the resulting $C$ is likely to have many more than $n$ points.

Nevertheless, for certain special polynomials the bound in (11) is tight. Take for instance $F = x + y - z$ and $A = B = C = \{1, \ldots, n\}$; then $|Z(F) \cap (A \times B \times C)| = \Theta(n^2)$. Of course, one can construct similar examples for any polynomial of the form $F = f(g(x) + h(y) + k(z))$ with $f, g, h, k$ univariate polynomials. In analogy with Theorem 1.1, one might guess that these are the only special polynomials, but this is not quite true. It turns out that the right class of special polynomials consists of those of the form $F = f(g(x) + h(y) + k(z))$ with $f, g, h, k$ analytic functions that are defined in a local way.

**Theorem 2.1** (Elekes-Szabó, Raz-Sharir-De Zeeuw). Let $F \in \mathbb{C}[x, y, z]$ be an irreducible polynomial of degree $d$ with with each of $F_x, F_y, F_z$ not identically zero. Then one of the following holds.

(i) For all $A, B, C \subset \mathbb{C}$ with $|A| = |B| = |C| = n$ we have

$$|Z(F) \cap (A \times B \times C)| = O_d(n^{11/6}).$$

(ii) There exists a one-dimensional subvariety $Z_0 \subset Z(F)$, such that every $v \in Z(F) \setminus Z_0$ has an open neighborhood $D_1 \times D_2 \times D_3$ and analytic functions $\varphi_i : D_i \to \mathbb{C}$, such that for every $(x, y, z) \in D_1 \times D_2 \times D_3$ we have

$$(x, y, z) \in Z(F) \ if \ and \ only \ if \ \varphi_1(x) + \varphi_2(y) + \varphi_3(z) = 0.$$  

Qualitatively, this result was proved by Elekes and Szabó\footnote{The publication year of [19] is 2012, but an essentially complete version of the paper existed much earlier; Elekes [14] referred to the result in 2002.} who proved that $|Z(F) \cap (A \times B \times C)| = O(n^{2-\eta_d})$ for a constant $\eta_d$ depending only on the degree $d$ of $F$ unless $F$ has the special form described in Theorem 2.1(ii) below. Using the new proof setup in [14, 39], this bound was improved to $O(n^{11/6})$ by Raz, Sharir, and De Zeeuw\footnote{Earlier versions of [19] claimed a bound of the form $O(n^{2-\eta})$ for an absolute $\eta > 0$, and this was restated in [14]. But the published version states the theorem with $\eta_d$ depending on $d$.}. An exposition of the algebraic geometry underlying the proof in [19] was written by Wang\footnote{[60].}.
The statement also holds if $C$ is replaced by $\mathbb{R}$ (and the analytic $\varphi_1 : D \to \mathbb{C}$ are replaced by real-analytic $\varphi_1 : D \to \mathbb{R}$). We can also allow $A, B, C$ to have different sizes. This does not affect the description in condition $(ii)$, and the bound in condition $(i)$ becomes

$$|Z(F) \cap (A \times B \times C)| = O_d \left( |A|^{2/3} |B|^{2/3} |C|^{1/2} + |A||C|^{1/2} + |B||C|^{1/2} + |C| \right);$$

of course the same bound holds for any permutation of $A, B, C$. We can again conjecture that the bound can be significantly improved, perhaps as far as $O_d(n^{1+\epsilon})$ (in the balanced case), or even $O(n \sqrt{\log n})$. No better lower bound is known than $\Omega(n \sqrt{\log n})$, for instance for $F = x^2 + xy + y^2 - z$ (which comes directly from the polynomial $f = x^2 + xy + y^2$ that provides the best known upper bound for Theorem 1.1 as mentioned in Subsection 1.2).

It is not likely that condition $(ii)$ can be replaced by a purely polynomial condition, i.e., without mentioning analytic functions. This can be seen from the fact that the group law on an elliptic curve gives constructions for which the bound in $(i)$ does not hold, while on the other hand, it is well-known that parametrizing elliptic curves as in $(ii)$ requires analytic functions. We will give more details on the connection with elliptic curves at the end of Subsection 3.3.

In [19], condition $(ii)$ of Theorem 2.1 is formulated in a somewhat stronger “global” form, although for all applications in this survey, the formulation in Theorem 2.1 seems to be more convenient. Specifically, the local functions $\varphi_i$ can be replaced by analytic multi-functions from $\mathbb{C}$ to a one-dimensional connected algebraic group $G$, so that $Z(F)$ is the image of the variety $\{(x, y, z) \in G^3 : x \oplus y \oplus z = e\};$ we refer to [19, 60] for definitions.

### 2.2 A derivative test for special $F$

In applications, it may not be easy to determine whether a given polynomial $F$ satisfies condition $(ii)$. For relatively simple polynomials, we have the following derivative condition. It is mentioned in [16 Subsection 1.1] and stated in [19 Lemma 33]; in [16] the sufficiency of the condition is ascribed to Jarai, although no proof is provided in [16] or [19]. We give a short sketch of the proof; a detailed proof can be found in [37].

**Lemma 2.2.** Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a twice-differentiable function with $f_y \neq 0$. There exist differentiable functions $\psi, \varphi_1, \varphi_2 : \mathbb{R} \to \mathbb{R}$ such that

$$f(x, y) = \psi(\varphi_1(x) + \varphi_2(y))$$

if and only if

$$\frac{\partial^2 (\log |f_x/f_y|)}{\partial x \partial y} = 0.$$

(14)

The same holds for analytic $f : \mathbb{C}^2 \to \mathbb{C}$ with $\psi, \varphi_1, \varphi_2 : \mathbb{C} \to \mathbb{C}$ analytic.

**Proof sketch.** If $f(x, y) = \psi(\varphi_1(x) + \varphi_2(y))$, then $f_x/f_y = \varphi'_1(x)/\varphi'_2(y)$, which gives

$$\log |f_x/f_y| = \log |\varphi'_1(x)| - \log |\varphi'_2(y)|.$$

Differentiating with respect to $x$ and $y$ gives $0$.

Conversely, if $\partial^2 (\log |f_x/f_y|)/\partial x \partial y = 0$, integrating gives $\log |f_x/f_y| = g_1(x) - g_2(y)$. Then, setting $\varphi_1(x) = \int e^{g_1(y)} dx$ and $\varphi_2(y) = \int e^{g_2(y)} dy$, we have $f_x/f_y = \varphi'_1(x)/\varphi'_2(y)$. We express $f$ in terms of the new variables $u = \varphi_1(x) + \varphi_2(y)$ and $v = \varphi_1(x) - \varphi_2(y)$, so that the chain rule gives $f_x = \varphi'_1(x)(f_u + f_v)$ and $f_y = \varphi'_2(y)(f_u - f_v)$. Combining these equations gives $0 = \frac{f_x}{\varphi'_1(x)} - \frac{f_y}{\varphi'_2(y)} = 2f_u$. Thus $f$ depends only on the variable $u$, which means that we can write it as $f(x, y) = \psi(\varphi_1(x) + \varphi_2(y))^2$. 

\[\square\]
To apply this lemma to a polynomial \( F(x, y, z) \), we need to locally write the implicit surface \( Z(F) \) as an explicit surface \( z = f(x, y) \), for an analytic function \( f \). Then the expression \( \varphi_1(x) + \varphi_2(y) + \varphi_3(z) = 0 \) in condition (ii) of Theorem 2.1 is equivalent to \( f(x, y) = \varphi_3^{-1}(\varphi_1(x) + \varphi_2(y)) \). In theory, such an \( f \) exists by the implicit function theorem, but in practice we can only calculate \( f \) when \( F \) has low degree (in one of the variables).

### 2.3 Applications

**Expanding polynomials.** Given \( f \in \mathbb{C}[x, y] \), we can set \( F(x, y, z) = f(x, y) - z \) and apply Theorem 2.1 with \(|A| = |B| = n \) and \( C = f(A \times B) \). If condition (i) applies, we get from the unbalanced bound (12) that

\[
n^2 = |Z(f(x, y) - z) \cap (A \times B \times C)| = O_d(n^{4/3}|C|^{1/2}),
\]

so \( |f(A \times B)| = |C| = \Omega(n^{4/3}) \). Otherwise, condition (ii) tells us that locally we have

\[
f(x, y) = \psi(\varphi_1(x) + \varphi_2(y)), \tag{15}
\]

with \( \psi, \varphi_1, \varphi_2 \) analytic functions.

Note that the multiplicative form of \( f \) also falls under (15), since we can (locally) write

\[
g(h(x) \cdot k(y)) = (g \circ \log^{-1})(\log |h(x)| + \log |k(y)|)
\]

with all functions analytic. We thus have almost deduced Theorem 2.1 except that the special form of \( f \) is local, and we do not know that \( \varphi_1, \varphi_2, \psi \) are polynomials. It would be interesting to find a way to deduce the full Theorem 2.1 from this local analytic form. Something close to that is done by Tao in [57, Theorem 41] using arguments from complex analysis.

**Distances from three points.** The following result follows from Theorem 2.1.

**Theorem 2.3** (Elekes-Szabó, Sharir-Solymosi). Given three non-collinear points \( p_1, p_2, p_3 \) and a point set \( P \) in \( \mathbb{R}^2 \), there are \( \Omega(|P|^{6/11}) \) distinct distances from \( p_1, p_2, p_3 \) to \( P \).

**Proof sketch.** Let \( D \) denote the set of squared distances between \( p_1, p_2, p_3 \) and the points in \( P \). A point \( q \in P \) determines three squared distances to \( p_1, p_2, p_3 \), given by

\[
a = (x_q - x_{p_1})^2 + (y_q - y_{p_1})^2, \quad b = (x_q - x_{p_2})^2 + (y_q - y_{p_2})^2, \quad c = (x_q - x_{p_3})^2 + (y_q - y_{p_3})^2.
\]

The variables \( x_q \) and \( y_q \) can be eliminated from these equations to yield a quadratic equation \( F(a, b, c) = 0 \) with coefficients depending on \( p_1, p_2, p_3 \) (in [19] \( F \) can be seen written out). By construction, for each point \( q \in P \), the corresponding squared distances \( a, b, c \) belong to \( D \). The resulting triples \( (a, b, c) \) are all distinct, so \( F \) vanishes at \(|P|\) triples of \( D \times D \times D \).

The polynomial \( F(x, y, z) \) turns out to be quadratic in \( z \), so we can locally express it as \( z = f(x, y) \). Then we can apply Lemma 2.2 to \( f \) to see that, if \( p_1, p_2, p_3 \) are not collinear, then \( f \) does not have the form in (15), which implies that \( F \) does not satisfy property (ii) of Theorem 2.1. Then property (i) gives \( |P| = O(|D|^{11/6}) \), or \(|D| = \Omega(|P|^{6/11}) \). When \( p_1, p_2, p_3 \) are collinear, \( F \) becomes a linear polynomial, so it does satisfy property (ii). 

11
This problem was introduced by Elekes [9], who showed that if \(p_1, p_2, p_3\) are collinear (and equally spaced), then one can place \(P\) so that there are only \(O(|P|^{1/2})\) distances from \(p_1, p_2, p_3\) to \(P\). Elekes and Szabó [19] proved Theorem 2.3 with the weaker bound \(\Omega(|P|^{1/2} + \eta)\) for some small absolute constant \(\eta > 0\), as a consequence of their version of Theorem 2.1 (although they formulated the result in terms of triple points of circles; see the next application). Sharir and Solymosi [46] used the setup of [44] and ad hoc arguments to improve this \(\eta\) to \(1/22\); their work preceded [41] and was the first extension of [44] that does not follow from [39].

Theorem 2.5 provides curious new information on Erdős’s distinct distances problem (see [5]), which asks for the minimum number of distinct distances determined by a point set in \(\mathbb{R}^2\). Erdős conjectured that this minimum is \(\Theta(n/\sqrt{\log n})\), and this was almost matched by Guth and Katz [24,10] who established \(\Omega(n/\log n)\). Theorem 2.3 suggests that something stronger is true: There are many distances that occur just from three fixed non-collinear points. If, as conjectured, the bound in Theorem 2.1 can be improved from \(O(n^{11/6})\) to \(O(n^{1+\varepsilon})\), or even \(O(n^{\sqrt{\log n}})\), then Erdős’s conjectured bound would already hold if one only considers distances from three non-collinear points in the point set (note that if the entire point set is collinear, there are \(\Omega(n)\) distances from any given point).

While waiting for improvements in the bound of Theorem 2.3, we could instead consider distances from more than three points. Given \(k\) points in a suitable non-degenerate configuration and a point set \(P\) in \(\mathbb{R}^2\), the number of distances from the \(k\) points to \(P\) should be \(\Omega(|P|^{1/2} + \alpha_k)\), where we would expect \(\alpha_k\) to grow with \(k\). Let us pose the first unknown step as a problem.

**Problem 2.4.** Let \(P \subset \mathbb{R}^2\), and consider four points in \(\mathbb{R}^2\) such that no three are collinear (or such that some stronger condition holds). Then the number of distinct distances from the four points to \(P\) is \(\Omega(|P|^{1/2} + \alpha)\), with \(\alpha \geq 1/22\).

**Triple points of circle families.** Elekes and Szabó [19] formulated the problem of Theorem 2.3 in a different way. They considered three points \(p_1, p_2, p_3\) in \(\mathbb{R}^2\) and three families of \(n\) concentric circles centered at the three points, and they looked for an upper bound on the number of triple points of these families, i.e., points covered by one circle from each family. Theorem 2.3 states that if \(p_1, p_2, p_3\) are not collinear, then the number of triple points is \(O(n^{11/6})\); a construction in [9] shows that if \(p_1, p_2, p_3\) are collinear, there can be as many as \(\Omega(n^2)\) triple points.

One can ask the same question for any three one-dimensional families of circles, or even more general curves. Such statements were studied by Elekes, Simonovits, and Szabó [17], with a special interest in the case of concurrent unit circles, i.e., unit circles passing through a fixed point. This case was improved by Raz, Sharir, and Solymosi [40], again using the setup of [44], and ad hoc analytic arguments. By the arguments from [17], the improvement also follows from Theorem 2.1.

**Theorem 2.5** (Elekes-Simonovits-Szabó, Raz-Sharir-Solymosi). Three families of \(n\) concurrent unit circles (concurrent at three distinct points) determine \(O(n^{11/6})\) triple points.

This result shows an interesting distinction between lines and unit circles, because for three families of concurrent lines it is possible to determine \(\Omega(n^2)\) triple points. We can

---

\[10\] The new algebraic methods introduced in [24] indirectly led to the improvement of [44] in Theorem 1.4, and thus to many of the recent results in this survey.
for instance take the horizontal and vertical lines of an \( n \times n \) integer grid, and \( n \) lines at a 45° angle that cover \( \Omega(n^2) \) points of the grid.

It is natural to extend the question to ask for \( k \)-fold points of \( k \) families of concurrent unit circles (or other curves). This is largely unexplored when \( k \) is small. For large \( k < \sqrt{n} \) and an arbitrary set of \( n \) unit circles, it follows from the incidence bound of Pach and Sharir \cite{34} (the general case of Theorem 1.6) that there are at most \( O(n^2/k^3) \) points where at least \( k \) circles meet.

### 2.4 About the proof of Theorem 2.1

Let us briefly discuss the proof of Theorem 2.1, although again we will not go into too much detail. The proof is based on that of Theorem 1.1 as described in Subsection 1.4, but now we have to deal with \( F(x, y, z) = 0 \) instead of \( f(x, y) = 0 \).

The first challenge is that we can no longer define the quadruples and curves using the equation \( f(a, b) = f(a', b') \). Instead, we define the quadruples by

\[
Q = \{(a, b, a', b') \in A \times B \times A \times B : \exists c \in C \text{ such that } F(a, b, c) = F(a', b', c) = 0\}.
\]

Using the Cauchy-Schwarz inequality we get

\[
|Z(F) \cap (A \times B \times C)| = \sum_{c \in C} |\{(a, b) \in A \times B : F(a, b, c) = 0\}| \\
\leq |C|^{1/2} \left( \sum_{c \in C} |\{(a, b) \in A \times B : F(a, b, c) = 0\}|^2 \right)^{1/2} \\
= O_d \left( |C|^{1/2} |Q|^{1/2} \right).
\]

In the last step we use the fact that for \( (a, b) \in A \times B \), there are at most \( d \) values of \( c \in C \) for which \( F(a, b, c) = 0 \) (unless \( F(x, y, z) \) contains a vertical line, but this happens at most \( O_d(1) \) times). Again the goal is to obtain the upper bound \( |Q| = O_d(n^{8/3}) \) using an incidence bound, which will result in \( |Z(F) \cap (A \times B \times C)| = O_d(n^{1/2}(n^{8/3})^{1/2}) = O_d(n^{11/6}) \).

The set \( Q \) can be viewed as the projection of a fiber product. A fiber product\footnote{This is a special case of a more general object from category theory; what we call a fiber product here is sometimes called a \textit{set-theoretic} fiber product, or also a \textit{relative product}.} of a set with itself has the form \( X \times \varphi X = \{(x, x') \in X \times X : \varphi(x) = \varphi(x')\} \) for some function \( \varphi : X \to Y \). This type of product is useful for counting, because the Cauchy-Schwarz inequality gives \( |X| \leq |X \times \varphi X|^{1/2} |Y|^{1/2} \). In the calculation above, we have \( X = |Z(F) \cap (A \times B \times C)|, Y = C, \) and \( \varphi(a, b, c) = c \). Then \( Q \) is the projection of \( X \times \varphi X \) to the coordinates \((a, b, a', b')\), and has essentially the same size as \( X \times \varphi X \). See \cite{7,57} for similar uses of fiber products, as well as further discussion of the technique.

The step in which we project from the fiber product to \( Q \) is necessary to make the next step work; specifically, we need \( Q \) to lie on a codimension one subvariety of \( C^4 \), in order to be able to define the algebraic curves that we apply the incidence bound to. Unfortunately, this projection brings in the problem of quantifier elimination. The set \( Q \) lies on the set

\[
\{(x, y, x', y') \in C^4 : \exists z \in C \text{ such that } F(x, y, z) = F(x', y', z) = 0\},
\]

which is not quite a variety, but only a constructible set (see \cite{41} for details and references). We can eliminate the quantifier in the sense that there is a variety \( Z(G) \subset C^4 \) that contains
the constructible set, and it differs only in a lower-dimensional set. However, we have little
grip on $G$ other than that its degree is bounded in terms of that of $F$. Note that in the
proof of Theorem 1.1 in Subsection 1.4 we had $F = f(x, y) - z$, so that we could easily
eliminate $z$ to get $f(x, y) = f(x', y')$.

To obtain $|Q| = O_d(n^{8/3})$, we define curves as in Subsection 1.4 but this becomes
more complicated due to the quantifier elimination. We set

$$C_{aa'} = \{(x, y) \in \mathbb{C}^2 : \exists z \in \mathbb{C} \text{ such that } F(a, x, z) = F(a', y, z) = 0\}.$$  

This set is a one-dimensional constructible set, i.e., an algebraic curve with finitely many
points removed. This leads to many technical complications, but we can basically still
apply an incidence bound to the points and curves

$$\mathcal{P} = B \times B, \quad \mathcal{C} = \{C_{aa'} : (a, a') \in A \times A\},$$

and we essentially have $|Q| = |I(\mathcal{P}, \mathcal{C})|$. As in Subsection 1.4 we can use Theorem 1.6
to obtain $|I(\mathcal{P}, \mathcal{C})| = O_d(n^{8/3})$, unless the curves badly violate the degrees-of-freedom
condition. The hardest part of the proof is then to connect the failure of the degrees-of-
freedom condition to the special form (ii) in Theorem 2.1.

3 Elekes-Rónyai problems on curves

In this section we discuss some variants of the Elekes-Rónyai and Elekes-Szabó problems
for point sets contained in algebraic curves. We still work with Cartesian products of “one-
dimensional” finite sets, but instead of finite subsets of $\mathbb{R}$ or $\mathbb{C}$, we take finite subsets of
algebraic curves. We start with the first known instance of an Elekes-Rónyai problem on
curves, where the function is the Euclidean distance. After that we discuss more general
polynomial functions on curves, and finally we look at Elekes-Szabó problems on curves.

3.1 Distances on curves

We have already seen one instance of the Elekes-Rónyai problem for distances on curves
in Theorem 1.4, which concerned distances between two point sets on two lines. A natural
generalization is to consider distances between two point sets on two algebraic curves in $\mathbb{R}^2$.
More precisely, given algebraic curves $C_1, C_2 \subset \mathbb{R}^2$ and finite point sets $P_1 \subset C_1, P_2 \subset C_2$
with $|P_1| = |P_2| = n$, can we get a superlinear lower bound on $|D(P_1 \times P_2)|$?\footnote{As in Subsection 1.3 $D(p, q) = (p_x - q_x)^2 + (p_y - q_y)^2$ is the squared Euclidean distance function.}

Observe that there are pairs of curves for which we cannot expect a superlinear lower
bound on $|D(P_1 \times P_2)|$, as we saw in Subsection 1.3 for parallel lines and orthogonal
lines. There is one more construction involving curves other than lines. If we take two
concentric circles with equally spaced points, then there is also only a linear number of
distances between the two point sets. It was proved by Pach and De Zeeuw\footnote{Pach and De Zeeuw claim that these three constructions are the only exceptions to a superlinear lower bound; the proof used (once again) the setup in \cite{41}, together with ad hoc arguments.} that these
three constructions are the only exceptions to a superlinear lower bound; the proof used
(once again) the setup in \cite{41}, together with ad hoc arguments.

**Theorem 3.1 (Pach-De Zeeuw).** Let $C_1, C_2 \subset \mathbb{R}^2$ be irreducible algebraic curves of degree
at most $d$. For finite subsets $P_1 \subset C_1, P_2 \subset C_2$ of size $n$ we have

$$|D(P_1 \times P_2)| = \Omega_d(n^{4/3}),$$

unless the curves are parallel lines, orthogonal lines, or concentric circles.
The proof in [35] allows $C_1$ and $C_2$ to be the same curve, which leads to a statement for distances on a single curve that is interesting in its own right. A version of this corollary was proved earlier by Charalambides [8], but with a weaker bound $\Omega_d(n^{5/4})$ (the same exponent as in [12], coming from the same counting scheme). The proof in [8] relied on an interesting connection with graph rigidity.

**Corollary 3.2** (Charalambides, Pach-De Zeeuw). Let $C \subset \mathbb{R}^2$ be an irreducible algebraic curve of degree $d$. For a finite subset $P \subset C$ we have

$$|D(P \times P)| = \Omega_d(n^{4/3}),$$

unless $C$ is a line or a circle.

Let us discuss some of the new issues involved in the proofs of these results. In the case of lines, a bound could be deduced from the Elekes-Rónyai theorem (or proved directly), unless the lines are parallel or orthogonal. For general algebraic curves, this does not seem possible. If the curves happen to be parametrized by polynomials, then plugging that parametrization into $D(p,q)$ would give a polynomial in two variables, and we could apply Theorem 1.1 (although it requires some work to translate the exceptional form of the polynomial to exceptional curves). This was done in [39] to prove that if $C$ is polynomially parametrizable, then the bound in Corollary 3.2 holds, unless $C$ is a line (a circle is not polynomially parametrizable).

If the curves are not polynomially parametrizable, then this approach will not work. The curve $y^2 = x^3 + 1$, for instance, has no parametrization with polynomials or rational functions. We could locally write $y = \sqrt{x^3 + 1}$ and plug that into $D(p,q)$, but this gives an algebraic function in two variables; moreover, most curves do not even have such an explicit solution by radicals, and the best we can do is to use the implicit function theorem to locally write $y$ as an analytic function of $x$. This suggests that Theorem 2.1 may provide a larger framework for Theorem 3.1. We will see in Subsection 3.3 how one can manipulate a question about curves to fit it into the framework of Theorem 2.1. Even then, extracting the exceptional forms of the curves from property (ii) is not straightforward; for distances on curves this was done by Raz and Sharir [35], who used considerations from graph rigidity (similar to those in [8]) to deduce Theorem 3.1 from Theorem 2.1.

A natural way to extend Theorem 3.1 or Corollary 3.2 is to consider curves in higher dimensions. This was done for Corollary 3.2 by Charalambides [8] with the bound $\Omega_d(n^{5/4})$, using the counting scheme of [12], together with various tools from analysis. He determined that in $\mathbb{R}^3$ the exceptional curves are again lines and circles, while in $\mathbb{R}^4$ and above, the class of exceptional curves consists of so-called algebraic helices; see [8, 36] for definitions and exact statements. The bound in $\mathbb{R}^D$ was improved to $\Omega(n^{4/3})$ in [39] for the case of polynomially parametrizable curves using Theorem 1.1, and finally for all curves except algebraic helices by Raz [36], using both Theorem 2.1 and the analysis of Charalambides [8]. Bronner, Sharir, and Sheffer [6] considered variants involving curves in $\mathbb{R}^D$ that are not necessarily algebraic.

Another variant of Theorem 3.1 was studied by Sheffer, Zahl, and De Zeeuw [49]: Suppose that $P_1$ is contained in a curve $C$, and $P_2$ is an arbitrary set in $\mathbb{R}^2$. In general it is difficult to obtain bounds in this situation, but [49] showed that if $C$ is a line or a circle, then the number of distances determined by $P_1 \cup P_2$ is reasonably large. This result was used to show that a point set that determines $o(n)$ distinct distances cannot have too many points on a line or a circle, which is a small step towards the conjecture of Erdős that a set of $n$ points with $o(n)$ distinct distances must resemble an integer grid.
3.2 Other polynomials on curves

We can ask the same questions for any polynomial function $\mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, considered as a function $C_1 \times C_2 \to \mathbb{R}$. For clarity, we focus on the case $C_1 = C_2$. We also switch from $\mathbb{R}$ to $\mathbb{C}$, because some of the statements become more natural over $\mathbb{C}$.

Charalambides [8] also considered the function $A(p, q) = p_x q_y - p_y q_x$, which (in $\mathbb{R}^2$) gives twice the signed area of the triangle spanned by $p$, $q$, and the origin. He proved that, for $P$ contained in an irreducible algebraic curve $C$ in $\mathbb{R}^2$, we have $|A(P \times P)| = \Omega_d(|P|^{5/4})$, unless $C$ is a line, an ellipse centered at the origin, or a hyperbola centered at the origin. This result was generalized by Valculescu and De Zeeuw [59], to any bilinear form over $\mathbb{C}$, with a broader class of exceptional curves. The condition on the curve is tight in the sense that for any excluded curve there is a bilinear form that can take a linear number of values on that curve. To the authors, it was surprising that such curves can have large degree, whereas previous evidence (namely [8] and [20]) suggested that the exceptional curves in such problems have degree at most three.

**Theorem 3.3.** Let $C$ be an irreducible algebraic curve in $\mathbb{C}^2$ of degree $d$, and consider the bilinear form $B_A(p, q) = p^T A q$ for a nonsingular $2 \times 2$ matrix $A$. For $P \subset C$ we have

$$|B_A(P \times P)| = \Omega_d \left( |P|^{4/3} \right),$$

unless $C$ is a line, or linearly equivalent to a curve of the form $x^k = y^\ell$, with $k, \ell \in \mathbb{Z} \setminus \{0\}$.

The description of the exceptional curves in Theorem 3.3 is very succinct but requires some clarification. Lines through the origin are linearly equivalent to a curve of the form $x^k = y^\ell$, but other lines are not, which is why they are listed separately. When $k$ or $\ell$ is negative, one obtains a more natural polynomial equation after multiplying by an appropriate monomial. Thus hyperbola-like curves of the form $x^k y^\ell = 1$ with coprime $k, \ell \geq 1$ are included, since they can also be defined by $x^k = y^{-\ell}$. Ellipses centered at the origin are also included, since these are linearly equivalent to the unit circle $(x - iy)(x + iy) = 1$, which is linearly equivalent to $xy = 1$. Thus all the exceptional curves of Charalambides are special. Note that these curves are all rational, i.e., they have a parametrization by rational functions.

The reason that these curves are exceptional in the proof of Theorem 3.3 is that they have infinitely many linear automorphisms. Here an automorphism of a curve $C$ is a map $T : \mathbb{C}^2 \to \mathbb{C}^2$ such that $T(C) = C$, and it is linear if it is a linear transformation (and similarly one can define affine, projective, or rational automorphisms). It was proved in [59] that the algebraic curves with infinitely many linear automorphisms are exactly those excluded in Theorem 3.3. From the proof of Theorem 3.3 in [59], it appears that if one considers more general polynomial functions instead of $B_A$, the exceptional curves will be those that have infinitely many rational automorphisms. By a theorem of Hurwitz (see for instance [20], Exercise IV.2.5)], a nonsingular curve with infinitely many rational automorphisms must have genus zero or one, or in other words, it must be rational or elliptic.

If we consider the statement of Theorem 3.3 for more general polynomials, then we also encounter exceptional polynomials that can take a linear number of values on any curve. Already for the bilinear forms $B_A$, this occurs when $A$ is a singular matrix; in that case we can write $B_A(p, q) = L_1(p) \cdot L_2(q)$ with linear polynomials $L_1, L_2$, which

---

13We say that two curves are linearly equivalent if there is a linear transformation $(x, y) \mapsto (ax + by, cx + dy)$ that gives a bijection between the point sets of the curves.
is reminiscent of the multiplicative form in Theorem 1.1. More generally, for functions of the form $G(H(p) + K(q))$ or $G(H(p) \cdot K(q))$ there are exceptional constructions. We arrive at the following conjecture (where again we could conjecture a larger exponent).

**Conjecture 3.4.** Let $C \subset \mathbb{C}^2$ be an algebraic curve of degree at most $d$ and $F : C \times C \to \mathbb{C}$ a polynomial of degree at most $d$. Then for any $P \subset C$ we have

$$|F(P \times P)| = \Omega_d(|P|^{4/3}),$$

unless $F(p, q) = G(H(p) + K(q))$ or $F(p, q) = G(H(p) \cdot K(q))$, or unless $C$ is rational.

### 3.3 Elekes-Szabó problems on curves

An Elekes-Szabó theorem on curves would take the following form. Let $C_1, C_2, C_3 \subset \mathbb{C}^2$ be algebraic curves of degree at most $d$, and let $G \in \mathbb{C}[x, y, s, t, u, v]$ be a polynomial of degree at most $d$. Then for point sets $P_1 \subset C_1, P_2 \subset C_2, P_3 \subset C_3$ of size $n$, we would want to bound $|Z(G) \cap (P_1 \times P_2 \times P_3)|$. We would expect exceptions for certain $G$, of a form related to that in Theorem 2.1(ii), and we would expect exceptions for certain curves, including low-degree curves and those in Conjecture 3.4. We won’t state a full conjecture here, but we will discuss some of the instances that have been considered.

One possible choice of polynomial is

$$G(x, y, s, t, u, v) = \frac{1}{2} \begin{vmatrix} x & s & u \\ y & t & v \\ 1 & 1 & 1 \end{vmatrix} - 1,$$

for which $G(x, y, s, t, u, v) = 0$ if and only if the triangle determined by the points $(x, y), (s, t), (u, v)$ has unit area. The problem of determining the maximum number of unit area triangles determined by $n$ points in $\mathbb{R}^2$ is ascribed to Oppenheim in [22]. The best known lower bound is $\Omega(n^2 \log \log n)$, due to Erdős and Purdy [22], and the best known lower bound is $O(n^{20/9})$, due to Raz and Sharir [37]. It is easy to show that for $n$ points on a curve, or three sets of $n$ points on three curves, the bound $O(n^2)$ holds. The following problem was suggested by Solymosi and Sharir [10].

**Problem 3.5.** Given $n$ points on an algebraic curve $C$ in $\mathbb{R}^2$, prove a bound better than $O(n^2)$ on the number of unit area triangles, or show that there are point sets on $C$ with $\Omega(n^2)$ unit area triangles.

The only case of this problem that has been studied is the one where $C$ is the union of three distinct lines. Surprisingly, Raz and Sharir [37] showed that on any three distinct lines there are non-trivial constructions that determine $\Omega(n^2)$ unit area triangles. They discovered this using the derivative criterion in Lemma 2.2.

Another choice of $G$ (which we won’t try to write out) is the polynomial such that $G(x, y, s, t, u, v) = 0$ if the points $(x, y), (s, t), (u, v)$ lie on a common unit circle, or in other words, the circle determined by the three points has radius equal to one. Since the radius of the circle determined by three points can be written as a rational function, multiplying out denominators gives a polynomial $G$ with the property above. Raz, Sharir, and Solymosi showed that for three distinct unit circles in $\mathbb{R}^2$ with three finite point sets of size $n$, the number of determined unit circles is $O(n^{11/6})$ (this statement is equivalent to Theorem 2.5). One can generalize this problem as follows.
Problem 3.6. Given three algebraic curves \( C_1, C_2, C_3 \subset \mathbb{R}^2 \) containing three sets of \( n \) points, prove a bound better than \( O(n^2) \) on the number of unit circles containing one point from each of the three sets. Are there curves for which \( \Omega(n^2) \) unit circles is possible?

Another problem of this form is to bound collinear triples on curves. More is known on this problem, and we will discuss it in detail in the next subsection. Note that it can be seen as a degenerate case of the unit area triangle problem, where instead of unit area we ask for zero area.

Let us see how the Elekes-Szabó problem on curves is related to the Elekes-Szabó problem in \( \mathbb{C}^3 \). We can think of the three curves \( C_1, C_2, C_3 \subset \mathbb{C}^2 \) as a Cartesian product \( C_1 \times C_2 \times C_3 \subset \mathbb{C}^6 \). We can choose a generic projection \( \varphi : \mathbb{C}^2 \to \mathbb{C} \), in such a way that the three-fold product \( \pi = (\varphi \times \varphi \times \varphi) : \mathbb{C}^6 \to \mathbb{C}^3 \) maps a product \( P_1 \times P_2 \times P_3 \subset C_1 \times C_2 \times C_3 \) with \( |P_1| = |P_2| = |P_3| = n \) to a product \( \varphi(P_1) \times \varphi(P_2) \times \varphi(P_3) \subset \mathbb{C}^3 \) with \( \varphi(P_1) = \varphi(P_2) = \varphi(P_3) = n \).

The variety \( X = Z(F) \cap (C_1 \times C_2 \times C_3) \) is two-dimensional (unless \( G \) happens to vanish on \( C_1 \times C_2 \times C_3 \), in which case the problem is trivial). Again by choosing \( \varphi \) generically, we get that \( \pi(X) \subset \mathbb{C}^3 \) is also a two-dimensional variety, and thus can be written as \( X = Z(F) \). Now we have

\[
|Z(G) \cap (P_1 \times P_2 \times P_3)| \leq |Z(F) \cap (\varphi(P_1) \times \varphi(P_2) \times \varphi(P_3))|,
\]

so if the upper bound of Theorem 2.1(i) applies to \( F \), then we also have that upper bound for \( G \) on \( C_1 \times C_2 \times C_3 \). Otherwise, \( F \) satisfies property (ii) of Theorem 2.1.

Unfortunately, it is not clear how to transfer back property (ii) for \( F \) to the original setting. Note that \( F \) not only encodes information about \( G \), but also about the curves \( C_1, C_2, C_3 \). Thus property (ii) for \( F \) should imply that either \( G \) has an exceptional form, or the curves \( C_1, C_2, C_3 \) have an exceptional form. This was made to work for the problem of collinear triples (see Subsection 3.4), but it remains difficult to do this in general.

The projection from \( \mathbb{C}^6 \) to \( \mathbb{C}^3 \) above lets us connect Theorem 2.1 to the group law on elliptic curves (irreducible nonsingular algebraic curves of degree three). We refer to [51] for an introduction to the group structure of an elliptic curve, but we summarize it here as follows: On any elliptic curve \( E \) there are an operation \( \oplus \) and an identity element \( O \) that turn the point set \( E \) into a group, with the special property that \( P, Q, R \in E \) are collinear if and only if \( P \oplus Q \oplus R = O \).

This group structure on elliptic curves allows us to construct finite point sets with many collinear triples. If we take a finite subgroup \( H \) of size \( n \) of an elliptic curve \( E \) (which exists for any \( n \)), then \( |Z(G) \cap (H \times H \times H)| = \Omega(n^2) \), where \( G \) is the polynomial that represents collinearity. This follows from the fact that for any two distinct elements \( P, Q \in H \), the line spanned by \( P \) and \( Q \) intersects \( E \) in \( R = \cap(P \oplus Q) \) (the group element \( R \) such that \( (P \oplus Q) \oplus R = O \)), which must also be in the subgroup \( H \). We could have \( R = P \) or \( R = Q \), but there are only \( O(n) \) pairs that satisfy \( P \oplus P \oplus Q = O \), so we get \( \Omega(n^2) \) collinear triples of distinct points.

Applying the generic projection \( \pi : \mathbb{C}^6 \to \mathbb{C}^3 \), we get a polynomial \( F \in \mathbb{C}[x, y, z] \) such that \( |Z(F) \cap (\varphi(H) \times \varphi(H) \times \varphi(H))| = \Omega(n^2) \), which implies that \( G \) satisfies property (ii) of Theorem 2.1. It follows that, locally, there are \( \varphi_i : \mathbb{C} \to \mathbb{C} \) such that \( P, Q, R \in E \) are collinear if and only if \( \varphi_1(\varphi(P)) + \varphi_2(\varphi(Q)) + \varphi_3(\varphi(R)) = 0 \). This is why it appears that the local analytic nature of property (ii) is necessary; it would be a big surprise if such maps existed globally, or could be described by polynomial or rational functions.
3.4 Collinear triples on curves

The one instance of an Elekes-Szabó problem on curves that has been studied in some detail (and is not an Elekes-Rónyai problem) is the problem where $F$ represents collinearity. This question was first considered by Elekes and Szabó [20]; they proved a weaker bound in $\mathbb{R}^2$ using the main bound from [19], which was then improved in [41] to the following statement.

**Theorem 3.7** (Elekes-Szabó, Raz-Sharir-De Zeeuw). Let $C$ be an irreducible algebraic curve in $\mathbb{C}^2$ of degree $d$ and let $P \subset C$ be a finite set. Then $P$ determines $O_d(n^{1/6})$ proper collinear triples, unless $C$ is a line or a cubic curve.

As we saw at the end of Subsection 3.3, elliptic curves must be exceptions to this statement, because their group structure gives constructions with a quadratic number of collinear triples. In fact, on any cubic curve (including the union of a conic and a line, or a union of three lines) there is a “quasi-group law”; see Green and Tao [25, Proposition 7.3]. This law does not quite give the whole point set a group structure (and there may not be an identity element), but comes close enough to allow for a construction with $\Omega(n^2)$ collinear lines. These constructions are worked out in [20].

Green and Tao [25] proved the Dirac-Motzkin conjecture for large point sets, which states that any non-collinear point set $P \subset \mathbb{R}^2$ of size $n$ determines at least $n/2$ ordinary lines (lines containing exactly two points of $P$). As a by-product, they also solved Sylvester’s “orchard problem” for large $n$, which asks for the maximum number of lines with at least three points (triple lines) from a set of $n$ points. It is easy to see that this number is at most $\frac{1}{3} \binom{n}{3}$, and Sylvester noted (in 1868; see [25]) that there are constructions on elliptic curves with almost exactly this number of triple lines. Green and Tao showed that any set with at least $\frac{2}{5} n^2 - O(n)$ triple lines must have most of its points on a cubic curve. Elekes made the bolder conjecture that any set with $\Omega(n^2)$ collinear triples has many points on a cubic; to underline our ignorance he suggested to show that ten of the points are on a cubic (which is one more than the trivial number). The conjecture was stated in [20] but already present in disguise in [14].

**Conjecture 3.8** (Elekes). If $P \subset \mathbb{R}^2$ determines $\Omega(n^2)$ collinear triples, then at least ten points of $P$ lie on a cubic.

An interesting application of Theorem 3.7 pointed out in [20], is to the problem of distinct directions mentioned in Subsection 1.3. Theorem 2.5 allows us to generalize Corollary 1.5 to any algebraic curve. The connection with triple lines is that if two pairs of points determine lines in the same direction, then these lines intersect the line at infinity in the same point. Thus a point set on a curve $C$ that has few directions determines few such points at infinity, or conversely, adding these points at infinity gives a point set with many collinear triples. An unbalanced form of Theorem 3.7 then gives the following lower bound on the number of directions. There is an exception when $C$ together with the line at infinity is a cubic curve, which means that $C$ is a conic.

**Corollary 3.9.** Let $C$ be an irreducible algebraic curve in $\mathbb{C}^2$ of degree $d$ and let $P \subset C$ be a finite set. Then $P$ determines $\Omega_d(n^{4/3})$ distinct directions, unless $C$ is a conic.
4 Other variants

4.1 Longer one-dimensional products

One way to extend Theorems 1.1 and 2.1 is to consider “longer” Cartesian products, i.e., products with more factors. The reference point for such bounds is the Schwartz-Zippel lemma, which we now state in full generality. This version was proved by Lang and Weil \[29, \text{Lemma } 1\]; see also Tao \[56\] for a proof sketch (both focus on finite fields, but their proofs work also over \(\mathbb{C}\)). The bound in fact holds over any field, and can be modified to allow factors \(A_i\) of different sizes.

**Theorem 4.1.** Let \(X \subset \mathbb{C}^D\) be a variety. Then for finite sets \(A_1, \ldots, A_D \subset \mathbb{C}\) of size \(n\) we have

\[
|X \cap (A_1 \times \cdots \times A_D)| = O_{D, \deg(X)}(n^{\dim(X)}).
\]

Theorem 2.1 told us that in the case with \(D = 3\) and \(\dim(X) = 2\), this bound can be improved on, unless the defining polynomial of \(X\) is special. Analogously, one would expect that the bound in Theorem 4.1 can be improved on, unless the variety is of some special type. The only other case that has so far been studied is \(D = 4\) and \(\dim(X) = 3\), for which Raz, Sharir, and De Zeeuw \[42\] proved the following.

**Theorem 4.2** (Raz-Sharir-De Zeeuw). Let \(F \in \mathbb{C}[x,y,s,t]\) be an irreducible polynomial of degree \(d\) with each of \(F_x, F_y, F_s, F_t\) not identically zero. Then one of the following holds.

(i) For all \(A, B, C, D \subset \mathbb{C}\) of size \(n\) we have

\[
|Z(F) \cap (A \times B \times C \times D)| = O_d(n^{8/3}).
\]

(ii) There exists a one-dimensional subvariety \(Z_0 \subset Z(F)\), such that every \(v \in Z(F)\setminus Z_0\) has an open neighborhood \(D_1 \times D_2 \times D_3 \times D_4\) and analytic functions \(\varphi_i : D_i \to \mathbb{C}\), such that for every \((x, y, s, t) \in D_1 \times D_2 \times D_3 \times D_4\) we have

\[
(x, y, s, t) \in Z(F) \quad \text{if and only if} \quad \varphi_1(x) + \varphi_2(y) + \varphi_3(s) + \varphi_4(t) = 0.
\]

The proof of Theorem 4.2 mostly uses the same techniques as that of Theorem 2.1 and the setup turns out to be simpler. The reason is that (because four is an even number) we can define curves by

\[
C_{cd} = \{(x, y) : F(x, y, c, d) = 0\},
\]

which avoids quantifier elimination and thus many technical complications. The theorem is in fact closely related to the incidence bound in Theorem 1.6. In most applications of that bound, the curves are defined as in \(16\), for some given polynomial \(F\) (for instance, the proof of Theorem 1.1 in Subsection 1.4 uses \(F = f(s, x) - f(t, y)\)). Theorem 4.2 thus gives the same bound as Theorem 1.6, but it replaces the combinatorial condition (two points being contained in a bounded number of curves) by an algebraic condition (that \(F\) is not of the form in \((ii)\)). Much like condition \((ii)\) in Theorem 2.1, condition \((ii)\) in Theorem 4.2 can often be checked using derivatives.

\[14\] This brings into question if “Schwartz-Zippel” is the right name, but it has become standard in combinatorics.
A consequence of Theorem 4.2 is an expansion bound for three-variable polynomials $f \in \mathbb{C}[x, y, z]$ on sets $A, B, C \subset \mathbb{C}$ with $|A| = |B| = |C| = n$, namely
\[
|f(A \times B \times C)| = \Omega(n^{3/2}),
\]
unless $f$ is in a local sense of the form $\psi(\varphi_1(x) + \varphi_2(y) + \varphi_3(s))$ with analytic $\psi, \varphi_1, \varphi_2, \varphi_3$. A weaker bound for this question was proved in [43], with the more precise condition that $f$ is not of the form $g(h(x) + k(y) + l(z))$ or $g(h(x) \cdot k(y) \cdot l(z))$ with $g, h, k, l$ polynomials. It should be possible to use techniques similar to those in [39] to prove the stronger bound with the precise condition.

We make the following conjecture for the general case.

**Conjecture 4.3.** There is a constant $c > 0$ such that one of the following holds for any irreducible $F \in \mathbb{C}[x_1, \ldots, x_D]$ of degree $d$ with each of $F_{x_1}, \ldots, F_{x_D}$ not identically zero.

(i) For all $A_1, \ldots, A_D \subset \mathbb{C}$ of size $n$ we have
\[
|Z(F) \cap (A_1 \times \cdots \times A_D)| = O_d(n^{D-1-c}).
\]

(ii) In a local sense we have
\[
(x_1, \ldots, x_D) \in Z(F) \quad \text{if and only if} \quad \sum_{i=1}^{D} \varphi_i(x_i) = 0.
\]

A largely unexplored question is how the bound of Theorem 4.2 can be improved for a variety $X \subset \mathbb{C}^D$ with $\dim(X) < D - 1$. When $\dim(X) = 1$, the bound $O(n)$ is tight: We can place $n$ points on $X$ and project in the coordinate directions to get $A_i$ such that $|X \cap (A_1 \times \cdots \times A_D)| = n$. Thus the first open case is $D = 4$, $\dim(X) = 2$.

**Problem 4.4.** Determine for which two-dimensional varieties $X \subset \mathbb{C}^4$ the Schwartz-Zippel-type bound $|X \cap (A_1 \times A_2 \times A_3 \times A_4)| = O_{\deg(X)}(n^2)$ is tight.

### 4.2 Two-dimensional products

A different way of extending Theorems 1.1 and 2.1 would be to consider Cartesian products of “two-dimensional” sets instead of “one-dimensional” sets, i.e., finite subsets of $\mathbb{C}^2$ instead of finite subsets of $\mathbb{C}$. For instance, the two-dimensional analogue of the Elekes-Rónyai problem would be: Given a map $F : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}^2$, defined by $F(x, y, s, t) = (F_1(x, y, s, t), F_2(x, y, s, t))$ for polynomials $F_1, F_2 \in \mathbb{C}[x, y, s, t]$, and given finite sets $P, Q \subset \mathbb{C}^2$, can we obtain a non-trivial lower bound on $|F(P \times Q)|$?

The study of such questions was initiated by Nassajian Mojarrad et al. [33]. Even analogues of the Schwartz-Zippel lemma become more complicated: The easiest case would be a bound of the form $|X \cap (P \times Q)|$ for a variety $X \subset \mathbb{C}^4$ and finite sets $P, Q \subset \mathbb{C}^2$, but there are varieties for which no non-trivial bound holds. Let us call a polynomial $F \in \mathbb{C}[x, y, s, t]$ Cartesian if it can be written as
\[
F(x, y, s, t) = G(x, y)H(x, y, s, t) + K(s, t)L(x, y, s, t),
\]
with $G \in \mathbb{C}[x, y] \setminus \mathbb{C}$, $K \in \mathbb{C}[s, t] \setminus \mathbb{C}$, and $H, L \in \mathbb{C}[x, y, s, t]$. We say that a variety $X \subset \mathbb{C}^4$ is Cartesian if every polynomial vanishing on $X$ is Cartesian with the same $G, K$. For a Cartesian variety $X$, we can have $|X \cap (P \times Q)| = |P||Q|$, since we can take $P \subset Z(G)$ and $Q \subset Z(K)$ to get $P \times Q \subset X$.

It is proved in [33] that if $X$ is not Cartesian, then one can obtain non-trivial upper bounds on $|X \cap (P \times Q)|$. As a consequence, we obtain the following lower bound on the number of distinct values of a polynomial map $F = (F_1, F_2) : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}^2$.\[
\]
Theorem 4.5 (Nassajian Mojarrad et al.). Let $F_1, F_2 \in \mathbb{C}[x, y, s, t]$ be polynomials of degree $d$ and $F = (F_1, F_2)$. Then for $P, Q \subset \mathbb{C}^2$ with $|P| = |Q| = n$ we have

$$|F(P \times Q)| = \Omega_d(n),$$

unless $F_1 = \varphi_1 \circ \psi$ and $F_2 = \varphi_2 \circ \psi$ for a nonlinear polynomial $\psi$, or $F_1$ and $F_2$ are both Cartesian with the same $G, K$.

This theorem provides a starting point for the study of Elekes-Rónyai problems for two-dimensional polynomial maps. To compare with the one-dimensional case, for any $G, K$ Cartesian with the same $F$ degree (other than those already excluded in Theorem 4.5) the bound $\Omega(n)$ can be improved on. The bound of Theorem 4.5 is tight for certain polynomial maps. For instance, if we take the vector addition map $F = (x + s, y + t)$ and we take $P$ to be any arithmetic progression on a line, then we have $|F(P \times P)| = O(n)$; note that $x + s$ and $y + t$ are both Cartesian, but not with the same $G, K$. We can do something similar for any map of the form $F = (f_1(x, s), f_2(y, t))$, where $f_1, f_2$ are additive or multiplicative in the sense of Theorem 1.1. Just like for the Elekes-Rónyai problem, composing these maps with other polynomials gives more exceptions: If we write $(x, y) \oplus (s, t) = (x + s, y + t)$, and we set $F = \psi(\varphi_1(x, y) \oplus \varphi_2(s, t))$ with reasonable maps $\psi, \varphi_1, \varphi_2 : \mathbb{C}^2 \to \mathbb{C}^2$, then again the image of $F$ can have linear size.

**Problem 4.6.** Determine for which polynomial (or even rational) maps $F : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}^2$ there exists an $\alpha > 0$ such that

$$|F(P \times Q)| = \Omega_d(n^{1+\alpha})$$

for all $P, Q \subset \mathbb{C}^2$ of size $n$, perhaps with further restrictions on $P$ and $Q$.

Superlinear bounds like in Problem 1.6 are known for only a few functions (which happen to be rational maps), and only over $\mathbb{R}$. Beck’s “two extremities” theorem [2, Theorem 3.1] can be phrased in terms of the rational map $L : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$ that maps a pair of points to the line spanned by the pair (for instance represented by the point in $\mathbb{R}^2$ whose coordinates are the slope and intercept of the line). Beck’s theorem says that $|L(P \times P)| = \Omega_c(|P|^2)$, unless $P$ has at least $c|P|$ points on a line. Raz and Sharir [37] work with the map that sends a pair of points to the line that consists of all the points that span a unit area triangle (of a fixed orientation) with the pair. They prove a superlinear bound on the number of distinct values of this map, for point sets with not too many points on a line (this statement is not made explicit in the paper, but is implicit in the proof). Finally, Lund, Sheffer, and De Zeeuw [31] study the rational map $B : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$ that sends a pair of points to the line that is their perpendicular bisector. They prove that $|B(P \times P)| = \Omega_{M, \varepsilon}(|P|^{8/5-\varepsilon})$ if $P \subset \mathbb{R}^2$ has at most $M$ points on a line or circle.

Elekes and Szabó [19] proved a more general theorem in this vein (their “Main Theorem” [19, Theorem 27]). It is difficult even to state this theorem precisely, so we give only a rough description. Let $Y$ be a $D$-dimensional irreducible variety over $\mathbb{C}$ of degree at most $d$, and let $X \subset Y \times Y \times Y$ be an irreducible $2D$-dimensional subvariety of degree at most $d$ with surjective and generically finite projections onto any two of the three factors.
Let $A \subset Y$ be a finite set, which is in very general position in the following sense: For any proper subvariety $Z \subset Y$ of degree at most $M$ we have $|A \cap Z| \leq N$. Then

$$|X \cap (A \times A \times A)| = O_{d,D,M,N}(|A|^{2-\eta})$$

with $\eta > 0$ depending on $d, D, M, N$, unless $X$ is in some specific way related to an algebraic group.

The results of [2, 37, 31] mentioned above fit into this framework: Given a reasonable map $F : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}^2$, its graph $X \subset \mathbb{C}^2 \times \mathbb{C}^2 \times \mathbb{C}^2$ is a variety that satisfies the conditions of the Main Theorem of Elekes and Szabó. Note that [2, 37, 31] not only provide explicit values of $\eta$ in certain special cases, but they also make the “very general position” condition more precise in those cases. Indeed, they replace the condition that $P$ avoids any algebraic curve (a proper subvariety of $\mathbb{C}^2$) of bounded degree, with the condition that $P$ avoids only lines in the case of [2, 37], or lines and circles in [31].

### 4.3 Expanding polynomials over other fields

Let us finish by briefly discussing Elekes-Rónyai-type questions over a finite field $\mathbb{F}_q$ and the field $\mathbb{Q}$ of rational numbers. We focus on statements of the form “$|f(A \times A)| = \Omega(|A|^{1+c})$ for a polynomial $f$ that is not of a special form”; there has been much work on conditional expansion bounds (especially over finite fields), of the form “if $|A + A|$ is small, then $|f(A \times A)|$ is large”, but we will not discuss these here.

Over finite fields, the question is considerably harder, because there is not yet any finite field equivalent of an incidence bound for algebraic curves like Theorem 1.1. For simplicity, let us restrict to a prime field $\mathbb{F}_p$, and let us ignore small $p$. There is a dichotomy between small subsets of $\mathbb{F}_p$, where very few incidence bounds are known, and large subsets of $\mathbb{F}_p$, for which various techniques can be used to obtain incidence bounds.

For instance, Bourgain [3] used the incidence bound of Bourgain, Katz, and Tao [4] for small subsets of $\mathbb{F}_p$ to prove the following expansion bound. For $f(x,y) = x^2 + xy$ and $A \subset \mathbb{F}_p$ with $|A| < p^{c'}$, we have $|f(A \times A)| > |A|^{1+c}$, with $c > 0$ depending on $c'$, $< 1$. This bound has been improved slightly, and generalized to some other polynomials, but the range of polynomials for which such bounds are known remains limited; see Aksoy Yazici et al. [1] for some recent developments.

For large subsets of finite fields (think of $|A| > p^{7/8}$), more comprehensive bounds have been proved. Bukh and Tsimerman [7] proved that $|f(A \times A)| = \Omega_d(|A|^{1+c})$ for $|A| > p^{7/8+c'}$, with $c > 0$ depending on $c'$, if $f \in \mathbb{F}_p[x,y]$ is monic in each variable, not of the form $p(q(x,y))$ with $\deg p \geq 2$, and not of the form $g(x) + h(y)$ or $g(x)h(y)$. The exceptional cases here are close to those in Theorem 1.1. Tao [57] then proved that $|f(A \times A)| = \Omega_d(p)$ for $|A| > p^{15/16}$, unless $f \in \mathbb{F}_p[x,y]$ is additive or multiplicative, matching the condition in Theorem 1.1. The proofs in [7, 57] both used fiber products and Cauchy-Schwarz (somewhat like in Subsection 2.4), and both relied on the bound of Lang and Weil [29, Theorem 1] for points on varieties over finite fields. Some of the ideas in [57] played a role in the proof of Theorem 2.1 in [11]. In [7, Section 9], an Elekes-Szabó-type statement over finite fields is conjectured.

One can also ask the Elekes-Rónyai question over $\mathbb{Q}$. Of course, Theorem 1.1 gives a bound there, but it is not clear that every exceptional polynomial from Theorem 1.1 is also exceptional over $\mathbb{Q}$. For instance, although $f(x,y) = x^2 + y^2$ is additive, the construction from Subsection 1.1 would require us to choose $A$ so that $A^2$ is an arithmetic progression, which is not possible in $\mathbb{Q}$. Solymosi made the following conjecture.
Conjecture 4.7. Let \( f \in \mathbb{Q}[x, y] \) be a polynomial of degree \( d \). For \( A \subset \mathbb{Q} \) we have
\[
|f(A \times A)| = \Omega_d(|A|^{1+c}),
\]
unless \( f(x, y) = g(ax + by) \) or \( f(x, y) = g((x + a)^\alpha(y + b)^\beta) \) for \( a, b \in \mathbb{Q} \) and positive integers \( \alpha, \beta \).

References

[1] E. Aksoy Yazici, B. Murphy, M. Rudnev, and I. Shkredov, Growth estimates in positive characteristic via collisions, arXiv:1512.06613, 2015.

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

[3] J. Bourgain, More on the sum-product phenomenon in prime fields and its applications, International Journal of Number Theory 1, 1–32, 2005.

[4] J. Bourgain, N. Katz, and T. Tao, A sum-product estimate in finite fields, and applications, Geometric And Functional Analysis 14, 27–57, 2004.

[5] P. Brass, W. Moser, and J. Pach, Research Problems in Discrete Geometry, Springer, 2005.

[6] A. Bronner, A. Sheffer, and M. Sharir, Distinct distances on a line and a curve, manuscript, 2016.

[7] B. Bukh and J. Tsimerman, Sum-product estimates for rational functions, Proceedings of the London Mathematical Society 104, 1–26, 2012.

[8] M. Charalambides, Distinct distances on curves via rigidity, Discrete & Computational Geometry 51, 666–701, 2014.

[9] G. Elekes, Circle grids and bipartite graphs of distances, Combinatorica 15, 167–174, 1995.

[10] G. Elekes, On the number of sums and products, Acta Arithmetica 81, 365–367, 1997.

[11] G. Elekes, A combinatorial problem on polynomials, Discrete & Computational Geometry 19, 383–389, 1998.

[12] G. Elekes, A note on the number of distinct distances, Periodica Mathematica Hungarica 38, 173–177, 1999.

[13] G. Elekes, On Linear Combinatorics III. Few Directions and Distorted Lattices, Combinatorica 1, 43–53, 1999.

[14] G. Elekes, SUMS versus PRODUCTS in Number Theory, Algebra and Erdős Geometry, Paul Erdős and his Mathematics II, Bolyai Society Mathematical Studies 11, 241–290, 2002.

[15] G. Elekes, M.B. Nathanson, and I.Z. Ruzsa, Convexity and sumsets, Journal of Number Theory 83, 194–201, 1999.
[16] G. Elekes and L. Rónyai, A combinatorial problem on polynomials and rational functions, Journal of Combinatorial Theory, Series A 89, 1–20, 2000.

[17] G. Elekes, M. Simonovits, and E. Szabó, A combinatorial distinction between unit circles and straight lines: How many coincidences can they have?, Combinatorics, Probability, and Computing 18, 691–705, 2009.

[18] G. Elekes and M. Sharir, Incidences in Three Dimensions and Distinct Distances in the Plane, Combinatorics, Probability and Computing 20, 571–608, 2011.

[19] G. Elekes and E. Szabó, How to find groups? (And how to use them in Erdős geometry?), Combinatorica 32, 537–571, 2012.

[20] G. Elekes and E. Szabó, On triple lines and cubic curves: The Orchard Problem revisited, arXiv:1302.5777, 2013.

[21] P. Erdős, On sets of distances of n points, American Mathematical Monthly 53, 248–250, 1946.

[22] P. Erdős and G. Purdy, Some extremal problems in geometry, Journal of Combinatorial Theory 10, 246–252, 1971.

[23] P. Erdős and E. Szemerédi, On sums and products of integers, in: Studies in Pure Mathematics, Birkhäuser, 213–218, 1983.

[24] L. Guth and N.H. Katz, On the Erdős distinct distances problem in the plane, Annals of Mathematics 181, 155–190, 2015.

[25] B. Green and T. Tao, On Sets Defining Few Ordinary Lines, Discrete & Computational Geometry 50, 409–468, 2013.

[26] R. Hartshorne, Algebraic Geometry, Springer, 1977.

[27] N. Hegyvári and F. Hennecart, Conditional expanding bounds for two-variable functions over prime fields, European Journal of Combinatorics 34, 1365–1382, 2013.

[28] S. Konyagin and I. Shkredov, On sum sets of sets, having small product set, arXiv:1503.05771, 2015.

[29] S. Lang and A. Weil, Number of points of varieties in finite fields, American Journal of Mathematics 76, 819–827, 1954.

[30] R.J. Lipton, blog post, available at http://rjlipton.wordpress.com/2009/11/30/the-curious-history-of-the-schwartz-zippel-lemma/, 2009.

[31] B. Lund, A. Sheffer, and F. de Zeeuw, Bisector energy and few distinct distances, 31st International Symposium on Computational Geometry (SoCG 2015), 537–552, 2015. Also in arXiv:1411.6868.

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

[33] H. Nassajian Mojarrad, T. Pham, C. Valculescu, and F. de Zeeuw, Schwartz-Zippel bounds for two-dimensional products, arXiv:1507.08181, 2015.
[34] J. Pach and M. Sharir, *On the number of incidences between points and curves*, Combinatorics, Probability and Computing 7, 121–127, 1998.

[35] J. Pach and F. de Zeeuw, *Distinct distances on algebraic curves in the plane*, to appear in Combinatorics, Probability and Computing. Also in [arXiv:1308.0177](http://arxiv.org/abs/1308.0177) and Proceedings of the thirtieth annual symposium on Computational geometry, 549–557, 2014.

[36] O.E. Raz, *A note on distinct distance subsets*, [arXiv:1603.00740](http://arxiv.org/abs/1603.00740), 2016.

[37] O.E. Raz and M. Sharir, *The number of unit-area triangles in the plane: Theme and variations*, [arXiv:1501.00379](http://arxiv.org/abs/1501.00379), 2015.

[38] O.E. Raz and M. Sharir, *Rigidity of complete bipartite graphs in the plane*, manuscript, 2016.

[39] O.E. Raz, M. Sharir, and J. Solymosi, *Polynomials vanishing on grids: The Elekes-Rónyai problem revisited*, to appear in American Journal of Mathematics. Also in [arXiv:1401.7419](http://arxiv.org/abs/1401.7419) and Proceedings of the thirtieth annual symposium on Computational geometry, 251–260, 2014.

[40] O.E. Raz, M. Sharir, and J. Solymosi, *On triple intersections of three families of unit circles*, Discrete & Computational Geometry 54, 930–953, 2015. Also in: Proceedings of the thirtieth annual symposium on Computational geometry, 198–205, 2014.

[41] O.E. Raz, M. Sharir, and F. de Zeeuw, *Polynomials vanishing on Cartesian products: The Elekes-Szabó Theorem revisited*, to appear in Duke Mathematical Journal. Also in [arXiv:1504.05012](http://arxiv.org/abs/1504.05012) and 31st International Symposium on Computational Geometry (SoCG 2015), 522–536, 2015.

[42] O.E. Raz, M. Sharir, and F. de Zeeuw, *The Elekes-Szabó Theorem in four dimensions*, manuscript, 2016.

[43] R. Schwartz, J. Solymosi, F. de Zeeuw, *Extensions of a result of Elekes and Rónyai*, Journal of Combinatorial Theory, Series A 120, 1695–1713, 2013.

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

[45] M. Sharir, S. Smorodinsky, C. Valculescu, and F. de Zeeuw, *Distinct distances between points and lines*, [arXiv:1512.09006](http://arxiv.org/abs/1512.09006), 2015.

[46] M. Sharir and J. Solymosi, *Distinct distances from three points*, to appear in Combinatorics, Probability and Computing. Also in [arXiv:1308.0814](http://arxiv.org/abs/1308.0814).

[47] A. Sheffer, *The polynomial method*, lecture notes from a course at Caltech, available at [http://www.math.caltech.edu/~2014-15/3term/ma191c-sec2/](http://www.math.caltech.edu/~2014-15/3term/ma191c-sec2/) 2015.

[48] A. Sheffer and J. Zahl, *Point-curve incidences in the complex plane*, [arXiv:1502.07003](http://arxiv.org/abs/1502.07003), 2015.

[49] A. Sheffer, J. Zahl, and F. de Zeeuw, *Few distinct distances implies no heavy lines or circles*, to appear in Combinatorica. Also in [arXiv:1308.5620](http://arxiv.org/abs/1308.5620), 2013.
[50] C.-Y. Shen, *Algebraic methods in sum-product phenomena*, Israel Journal of Mathematics **188**, 123–130, 2012.

[51] J.H. Silverman and J. Tate, *Rational points on elliptic curves*, Springer, 1992.

[52] J. Solymosi, *Bounding multiplicative energy by the sumset*, Advances in Mathematics **222**, 402–408, 2009.

[53] J. Solymosi and F. de Zeeuw, *Incidence bounds for complex algebraic curves on Cartesian products*, [arXiv:1502.05304](https://arxiv.org/abs/1502.05304), 2015.

[54] J. Spencer, E. Szemerédi, and W.T. Trotter, *Unit distances in the Euclidean plane*, in: Graph Theory and Combinatorics, Academic Press, 293–303, 1984.

[55] E. Szemerédi and W.T. Trotter, *Extremal problems in discrete geometry*, Combinatorica **3**, 381–392, 1983.

[56] T. Tao, blog post, available at [http://terrytao.wordpress.com/2012/08/31/the-lang-weil-bound/](http://terrytao.wordpress.com/2012/08/31/the-lang-weil-bound/), 2012.

[57] T. Tao, *Expanding polynomials over finite fields of large characteristic, and a regularity lemma for definable sets*, Contributions to Discrete Mathematics **10**, 22–98, 2015.

[58] P. Ungar, 2N *Noncollinear points determine at least 2N directions*, Journal of Combinatorial Theory, Series A **33**, 343–347, 1982.

[59] C. Valculescu and F. de Zeeuw, *Distinct values of bilinear forms on algebraic curves*, to appear in Contributions to Discrete Mathematics, 2015. Also in [arXiv:1403.3867](https://arxiv.org/abs/1403.3867).

[60] H. Wang, *Exposition of Elekes Szabo paper*, [arXiv:1512.04993](https://arxiv.org/abs/1512.04993), 2015.

[61] J. Zahl, *A Szemerédi-Trotter type theorem in \( \mathbb{R}^4 \)*, Discrete & Computational Geometry **54**, 513–572, 2015.

[62] F. de Zeeuw, *A course in algebraic combinatorial geometry*, lecture notes available from the author’s website [http://dcg.epfl.ch/page-84876-en.html](http://dcg.epfl.ch/page-84876-en.html), 2015.