BOOK REVIEWS

Polynomial methods in combinatorics, by Larry Guth, University Lecture Series, Vol. 64, American Mathematical Society, Providence, RI, 2016, ix+273 pp., ISBN 978-1-4704-2890-7, softcover, US \$48.00

Polynomials $P(x_1, x_2, \ldots, x_d)$ of one or more variables $x_1, \ldots, x_d$ and the algebraic varieties $\{P_1(x_1, \ldots, x_d) = \cdots = P_k(x_1, \ldots, x_d) = 0\}$ that they cut out are of course fundamental objects in algebra in general and in algebraic geometry in particular. But it has gradually been realised over time that they are also fundamentally important in other areas of mathematics and theoretical computer science, such as combinatorial incidence geometry, harmonic analysis, differential geometry, and error correcting codes. One particularly striking manifestation of this phenomenon has been the dramatic successes in recent years of the *polynomial method* in combinatorial geometry, which has been used to solve (or nearly solve) some major open probelms in the subject that did not, on first glance, seem at all related to polynomials.

Roughly speaking, combinatorial geometry is the study of configurations of finitely many geometric objects (such as points, lines, planes, or circles) in some standard geometry (e.g., the Euclidean plane $\mathbf{R}^2$, a higher-dimensional Euclidean space $\mathbf{R}^n$, or a vector space k^n over a more general field k). One is often interested in *extremal* questions, in which one tries to maximise or minimise some combinatorial quantity involving these configurations subject to various constraints. There are many questions in this subject; we will just mention two of these, which are also extensively discussed in the book under review.

(1) *Finite field Kakeya problem.* Suppose one is given a subset E of a finite-dimensional vector space $\mathbb{F}_q^d$ over a finite field $\mathbb{F}_q$ of q elements. Suppose also that E is a *finite field Kakeya set*, which means that it contains a line in every direction (i.e., for every non-zero $v \in \mathbb{F}_q^d$, there exists a line $\{x + tv : t \in k\}$ that is contained in E). For a given choice of k and q, what is the minimum cardinality $|E|$ of E?

(2) *Erdős distinct distances problem.* Suppose one is given a set P of n points in the Euclidean plane $\mathbf{R}^2$. For a given choice of n, what is the minimum number of distinct distances that are formed between the points in P, that is to say what is the minimum cardinality of the set $\{|p_1 - p_2| : p_1, p_2 \in P, p_1 \neq p_2\}$?

Finding exact answers to these questions is probably hopeless, except when most of the parameters n, m, q that are involved are small. Thus, attention has focused instead on asymptotic regimes, in which one or more of these parameters is large or goes to infinity.

We first discuss the finite field Kakeya problem. This is considered a "toy" problem for the *Kakeya conjecture* in geometric measure theory, which studies analogues of finite field Kakeya sets in Euclidean space, and is in turn related to several other important open problems in harmonic analysis, PDE, and number theory; see, e.g., [6] (or Chapter 15 of the book under review) for a survey. The finite field Kakeya problem is ostensibly about lines, and for several years partial progress was made by exploiting simple geometric facts about lines (such as the fact that two points are incident to at most one line, or that any three lines that are incident to each other at different points are necessarily coplanar), or by exploiting the arithmetic structure of lines (for instance, using the fact that the average of any two points on a line also is on the line, at least if the characteristic is odd). Both the Kakeya conjecture and its finite field analogue were considered quite difficult; it was thus a shock when Dvir [1] applied the polynomial method to almost completely settle the finite field Kakeya problem by showing that the minimum cardinality of a Kakeya set is at least $c_d q^d$ for some constant c_n depending only on d (for instance one can take $c_d = \frac{1}{d!}$). The proof is so short that we will be able to sketch it later in this review.

Dvir's method relied quite heavily on the finite field geometry of the problem, and it was initially believed that the polynomial method was not applicable to combinatorial problems in Euclidean geometries. A breakthrough was achieved by Guth [3], who obtained some progress on a variant of the Kakeya problem in Euclidean spaces by introducing an algebraic topology variant of the polynomial method, which was powered by a polynomial version (first established by Stone and Tukey [5]) of the *ham sandwich theorem*. By combining this method with some additional tools from algebraic geometry, Guth and Katz [4] were able to almost fully resolve the distinct distances problem of Erdős mentioned above. Namely, they showed that the minimum number of distinct distances between n points was at least $\frac{cn}{\log n}$ for some absolute constant $c > 0$. (Erdős [2] had previously established an upper bound of $\frac{Cn}{\sqrt{\log n}}$ for an absolute constant C; thus the problem is resolved up to a factor of about $O(\sqrt{\log n})$.)

Informally, the polynomial method is based on somehow combining the following two general principles:

- An arbitrary configuration of geometric objects can be efficiently "captured" by an algebraic variety of controlled "complexity".
- Algebraic varieties of controlled complexity interact with each other in only a limited number of ways.

These principles manifest themselves in a particularly simple way in the case of Dvir's bound on the finite field Kakeya problem. We begin with a discussion of the first principle in this context. We all know that given any two points (x_1, y_1), (x_2, y_2) in a plane, there is a line that passes through them; algebraically, this can be seen because the equation $ax + by + c = 0$ of a line involves three coefficients a, b, c in a linear fashion, and so one can use linear algebra to find non-trivial coefficients a, b, c obeying the two equations $ax_1 + by_1 + c = 0$ and $ax_2 + by_2 + c = 0$. A similar argument shows that given any five points $(x_1, y_1), \ldots, (x_5, y_5)$ in the

plane, one can find a conic section $ax^2 + bxy + cy^2 + dx + ey + f = 0$ that passes through them. More generally, we have

Lemma 1. *Let E be a subset of a finite field vector space $\mathbb{F}_q^d$. Then there exists a non-zero polynomial $P \in \mathbb{F}_q[x_1, \ldots, x_d]$ in d indeterminates that has degree at most $C_d|E|^{1/d}$ and that vanishes identically on E, where C_d depends only on n.*

Proof. If C_d is chosen large enough, then the space of polynomials in $\mathbb{F}_q[x_1, \ldots, x_d]$ of degree at most $C_d|E|^{1/d}$ can be computed to be a vector space (over $\mathbb{F}_q$) of dimension strictly greater than $|E|$. On the other hand, the requirement that such a polynomial vanishes on E imposes precisely $|E|$ linear constraints. Thus the space of polynomials of degree at most $C_d|E|^{1/d}$ and vanishing on E has positive dimension, and the claim follows. $\square$

Now we discuss the second principle. The simplest manifestation of this principle arises when considering the zeroes $\{x \in \mathbb{F}_q : P(x) = 0\}$ of a polynomial $P \in \mathbb{F}_q[x]$ of one variable of degree at most d. Either P is identically zero (in which case the zero set is all of $\mathbb{F}_q$, or else there are at most d zeroes. In other words, the number of zeroes can be less than or equal to d or equal to q, but it is prohibited from ranging strictly between d and q. Geometrically, this is a limitation on how the graph $\{(x, P(x)) : x \in \mathbb{F}_q\}$ may interact with the line $\{(x, 0) : x \in \mathbb{F}_q\}$. In a similar spirit, we have

Lemma 2. *Let $P \in \mathbb{F}_q[x_1, \ldots, x_n]$ be a polynomial of degree D for some $d < q$, let P_D be the polynomial formed from the monomials in P of degree exactly D, and let $\{x + tv : t \in \mathbb{F}_q\}$ be a line in $\mathbb{F}_q^d$. Then either P is not identically vanishing on this line or else $P_D(v) = 0$.*

Proof. If $P_D(v) \neq 0$, then the one-dimensional polynomial $t \mapsto P(x + tv)$ has degree exactly D and thus has at most D roots. As $D < q$, we conclude that this polynomial does not vanish identically on $\mathbb{F}_q$, and the claim follows. $\square$

Now we can prove Dvir's result. Suppose for contradiction that there existed a finite field Kakeya set $E \subset \mathbb{F}_q^d$ of cardinality less than $c_d q^d$ for some sufficiently small $c_d > 0$. By Lemma 1, we may then find a non-zero polynomial $P \in \mathbb{F}_q[x_1, \ldots, x_d]$ of some degree $D < q$ that vanishes on E; in particular, it vanishes on every line $\{x + tv : t \in \mathbb{F}_q\}$ contained in E. Using Lemma 2 and the hypothesis that E is a finite field Kakeya set, we conclude that P_D vanishes for every $v \in \mathbb{F}_q^d$, which is not possible since P_D is a non-zero polynomial of degree strictly less than q. The claim follows.

Lemma 1 is available in any field, not just the finite fields $\mathbb{F}_q$. However, when working with sets E of points in a Euclidean space $\mathbf{R}^d$, this lemma is not always useful, because the polynomial P that it provides is of too high a degree to be usable in applications. The key insight of Guth [3] mentioned previously is that one can also exploit the *topological* structure of $\mathbf{R}^d$ to find a lower degree polynomial P that may not pass through all the points in E any more, but instead *partitions* them in a very uniform fashion. A well-known instance of such a partitioning result is the ham sandwich theorem, which asserts for instance that if U_1, U_2, U_3 are bounded open subsets of $\mathbf{R}^3$, then there exist a plane (that is to say, the zero set of a linear polynomial) that bisects each of the three sets U_1, U_2, U_3 in volume. Setting U_1, U_2, U_3 to be small neighbourhoods of finite sets of points P_1, P_2, P_3 and applying a limiting argument, we conclude that for any finite sets of points $P_1, P_2, P_3 \subset \mathbf{R}^3$,

there exists a plane which bisects each of the P_i in the sense that the two open half-spaces on either side of the plane contain an equal number of elements of P_i. By using a polynomial version of this ham sandwich theorem due to Stone and Tukey [5], Guth observed the following variant of Lemma 1, which turns out to be a very useful manifestation of the first principle mentioned above:

Lemma 3 ([3]). *Let E be a finite subset of $\mathbf{R}^d$, and let D be a natural number. Then there exists a non-zero polynomial $P \in \mathbf{R}[x_1, \dots, x_d]$ in d indeterminates of degree at most D, such that the complement $\mathbf{R}^d \backslash \{x \in \mathbf{R}^d : P(x) = 0\}$ of the zero set of P has at most $C_d D^d$ connected components, each of which contains at most $C_d |E|/D^d$ elements of E, where C_d depends only on d.*

One can think of Lemma 1 as corresponding to an endpoint case of Lemma 3 when the parameter D is chosen to be a little bit larger than $|E|^{1/d}$; however the ability to set the parameter D to be significantly smaller than $|E|^{1/d}$ makes this lemma significantly more flexible than Lemma 1 in applications. Lemma 3 sets up a divide-and-conquer strategy in which an arbitrary set of points E is subdivided into a subset $E \cap \{x \in \mathbf{R}^d : P(x) = 0\}$ that is contained in a relatively low degree hypersurface, together with a collection of smaller sets that are contained in *cells* bounded by this hypersurface. This leads to an important special case of the polynomial method known as the *polynomial partitioning method*, which was used for instance by Guth and Katz [4] in the above-mentioned work on the distinct distances problem of Erdős.

In many applications one needs more advanced manifestations of the second principle than what is provided by Lemma 2. Many such manifestations are supplied by classical theorems of algebraic geometry. For instance, we have the basic theorem of Bézout, which asserts that if one is given two irreducible algebraic curves γ_1, γ_2 in the plane of degrees d_1, d_2, respectively, then the two curves either coincide identically or else intersect in at most $d_1 d_2$ points; note that this generalises Lemma 2. This simple result, valid in any field, is already very useful in applications; however one also has need of more sophisticated results of this type from algebraic geometry. Here is one such. Given a smooth surface S in $\mathbf{R}^3$, we say that a point x on S is a *flecnode* if there is a line $\{x + tv : t \in \mathbf{R}\}$ passing through x which is tangent to S to third order (or equivalently, there is a smooth curve $t \mapsto \gamma(t)$ in S which has a Taylor expansion $\gamma(t) = x + tv + O(t^4)$ for t near zero). For instance, if S is a ruled surface (a union of straight lines), then every point of S is clearly a flecnode. A remarkable theorem of Monge, Cayley, and Salmon asserts a converse, at least in the context of algebraic surfaces:

Lemma 4. *Let $P \in \mathbf{R}[x_1, x_2, x_3]$ be an irreducible polynomial, and let S be the surface $S := \{x : P(x) = 0\}$. Then either S is ruled or else the set of flecnodes of S are contained in a finite union of algebraic curves.*

Informally, this lemma asserts that if "enough" points in S are flecnodal, then S must in fact be a ruled surface. There is a variant of this lemma that asserts (roughly speaking) that if "enough" points of S are doubly flecnodal (in that they have two flecnodal lines passing through them), then S must be doubly ruled; this is a particularly useful result because the doubly ruled surfaces in $\mathbf{R}^3$ have been completely classified (they are all quadric surfaces), and it plays a decisive role in the work of Guth and Katz on the Erdős distinct distances problem. See Chapter

13 of the book under review for an in-depth discussion of these results and their proofs.

The book under review is a highly accessible and readable introduction to this circle of ideas; the author has made particular effort to patiently introduce and explain all the key concepts and ideas in a natural, enjoyable, and almost completely self-contained fashion, requiring little more than an undergraduate mathematics background in most cases. As the title suggests, the core topic of the book is the polynomial method in combinatorics; however, significant space is also devoted to non-polynomial approaches to the same type of combinatorial problems (such as non-polynomial approaches to the Kakeya problem), as well as polynomial methods applied to other fields of mathematics than combinatorics, such as differential geometry or number theory. As such, a reader of this text will not just learn about a single method applied to a single field of mathematics, but will also learn about the context of both the method and the field. I myself found the book very enjoyable and rewarding to read, and certainly recommend it to students who are interested in either the polynomial method or in combinatorial incidence geometry.

References

[1] Z. Dvir, *On the size of Kakeya sets in finite fields*, J. Amer. Math. Soc. **22** (2009), no. 4, 1093–1097, DOI 10.1090/S0894-0347-08-00607-3. MR2525780

[2] P. Erdős, *On sets of distances of n points*, Amer. Math. Monthly **53** (1946), 248–250, DOI 10.2307/2305092. MR0015796

[3] L. Guth, *The endpoint case of the Bennett-Carbery-Tao multilinear Kakeya conjecture*, Acta Math. **205** (2010), no. 2, 263–286, DOI 10.1007/s11511-010-0055-6. MR2746348

[4] L. Guth and N. H. Katz, *On the Erdős distinct distances problem in the plane*, Ann. of Math. (2) **181** (2015), no. 1, 155–190, DOI 10.4007/annals.2015.181.1.2. MR3272924

[5] A. H. Stone and J. W. Tukey, *Generalized "sandwich" theorems*, Duke Math. J. **9** (1942), 356–359. MR0007036

[6] T. Tao, *From rotating needles to stability of waves: emerging connections between combinatorics, analysis, and PDE*, Notices Amer. Math. Soc. **48** (2001), no. 3, 294–303. MR1820041

TERENCE TAO

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA, LOS ANGELES
E-mail address: tao@math.ucla.edu