INTRODUCTION

In the last five years, several challenging problems in combinatorics have been solved in an unexpected way using polynomials. This new approach is called the polynomial method, and the goal of these notes is to study and explore it.

The polynomial method has roots in some algorithms about polynomials developed in coding theory in the 80's and 90's. Ideas from these algorithms were then applied to mathematical problems that aren't obviously related to polynomials. Some problems that seemed very hard can now be solved in a couple pages with this new perspective.

The following problem is in the background of the polynomial method. Consider a field $\mathbb{F}$ and a finite set of points $S \subset \mathbb{F}^n$.

Problem: Find a non-zero polynomial P that vanishes on S with degree as small as possible.

For example, consider the points $(j, 2^j) \in \mathbb{R}^2$ with $j = 1, ..., 10^6$. What is the lowest degree of a (non-zero) polynomial that vanishes on all of these points? Let me try to find some polynomials that vanish on these points beginning with simple ones. One polynomial that vanishes on all these points is $(x-1)...(x-10^6)$. It has degree 10^6. Another is $(y-2)(y-4)...$ which also has degree 10^6. I can be a little cleverer by choosing a linear polynomial L_1 that vanishes on the first two points, then a second linear polynomial that vanishes on the next two points, etc. The product of these linear factors has degree only 500,000.

We can get a better perspective on the problem by thinking about the general situation of a finite set of points $x_1, ..., x_s \in \mathbb{F}^n$. Let $V(d)$ be the vector space of polynomials of degree $\leq d$ in n variables over $\mathbb{F}$. Let E be the evalution map defined by

$$E(P) := (E(x_1), ..., E(x_s)) .$$

The map E is linear, and the kernel of E is exactly the set of polynomials of degree $\leq d$ that vanish on the given points. Our original problem reduces to linear algebra. Using linear algebra, we can draw two important corollaries.

Corollary 0.1. *There is a polynomial-time algorithm to find a minimal degree polynomial that vanishes on a given finite set.*

(The running time is polynomial in the number of points s, and also in the dimension n. We will usually have a fixed n and consider $s \to \infty$.)

Corollary 0.2. *If $dim V(d) > s$, then there is a non-zero polynomial P of degree $\leq d$ that vanishes on the given finite set.*

The dimension of $V(d)$ is $\binom{d+n}{n} \geq d^n/n!$. For n fixed and d large, $d^n/n!$ is a good approximation of the dimension. Therefore, we get the following corollary.

Corollary 0.3. *For any set of s points in $\mathbb{F}^n$, there is a non-zero polynomial that vanishes on the set with degree $\leq ns^{1/n}$.*

Returning to our example, we see that there is a polynomial vanishing on the million points $\{(j, 2^j) | j = 1, ..., 10^6\}$ with degree ≤ 2000. This polynomial is much more efficient than the examples I came up with above. It's extremely messy and it would be very difficult to write down explicitly, but for simple abstract reasons we know that it exists.

Here is one moral of this discussion. Suppose that we are trying to find a polynomial with some special properties. One approach is to try to write down the polynomial and find a clever formula. But this discussion gives another approach - proving that such a polynomial exists by a dimension-counting argument. Sometimes this approach is more effective than any polynomial that I could craft.

A bare outline of the polynomial method goes as follows.

(1) Begin with a problem about some points in a vector space.
(2) Find or consider a polynomial that vanishes at these points with degree as small as possible.
(3) Use the polynomial to attack the problem.

After this general discussion, let's mention some of the applications of this method. I'm going to mention four applications that we'll study in this course.

1. Algorithms in coding theory

Suppose $\mathbb{F}$ is a finite field with q elements and $P : \mathbb{F} \to \mathbb{F}$ is a polynomial of low degree, $degP \leq q^{1/2}$. In the coding theory scenario, we could imagine that this polynomial is a piece of data that we want to send over an unreliable channel. In transmission, the data gets corrupted, and the other side receives a function $F : \mathbb{F} \to \mathbb{F}$. Let's suppose that a slim majority of the data is correct: in other words $F(x) = P(x)$ for at least $(51/100)q$ values of x. Is it possible to recover P from F? If so, can we do it efficiently?

As long as q is sufficiently large, it is possible to recover P from F in theory because of a fundamental property of polynomials.

Lemma 1.1. *If $P : \mathbb{F} \to \mathbb{F}$ has degree $\leq d$ and vanishes at more than d points, then P is the zero polynomial.*

(This simple lemma will have a lot of applications in our course.)

Corollary 1.2. *If $q > 10^4$, for any function $F : \mathbb{F} \to \mathbb{F}$, there is at most one polynomial P of degree $\leq q^{1/2}$ so that $F(x) = P(x)$ for at least $(51/100)q$ values of x.*

Proof. Suppose that P_1 and P_2 are such polynomials. Since $P_1 = F$ 51 % of the time and $P_2 = F$ 51 % of the time, it follows that $P_1(x) = P_2(x)$ for $\geq (2/100)q$ values of x. So $P_1 - P_2$ is a polynomial of degree $\leq q^{1/2}$ with at least $(2/100)q$ zeroes. If q is big enough $(2/100)q > q^{1/2}$ and so $P_1 - P_2 = 0$. □

In theory, we can recover P from F by trying all the polynomials of degree $\leq q^{1/2}$ until we find the one that agrees with F 51 % of the time. But this algorithm is very inefficient. Berlekamp and Welch found an efficient algorithm to recover P from F.

Theorem 1.3. *(Berlekamp-Welch, 1986) There is a polynomial time algorithm to recover P from F.*

Berlekamp and Welch consider the graph of P and the graph of F. The graph of P is a nice algebraic curve in $\mathbb{F}^2$. The graph of F contains a lot of points from the graph of P, together with some error points. We are given the graph of F. We don't know which points lie in the graph of P and which are errors. In this cloud of points, we are hoping to find a hidden algebraic structure - the graph of P. The main idea of Berlekamp and Welch is to consider a lowest degree polynomial $R(x, y)$ that vanishes on the graph of F in $\mathbb{F}^2$. (In fact, they consider the lowest degree polynomial of the special form $R(x, y) = R_0(x) + yR_1(x)$.) As we discussed above, it's possible to find this polynomial R in polynomial time. Then it turns out that the zero set of R is exactly the graph of P together with a vertical line through each error. In other words, for each $e \in \mathbb{F}$ where $F(e) \neq P(e)$, the graph of R contains the line $x = e$. With the help of R we can immediately tell which values of F agree with P and which were corrupted. After that, it's straightforward to recover P.

2. The finite field Nikodym conjecture

The next problem that we consider originates in geometry and analysis.

A set N in the cube $[0, 1]^n \subset \mathbb{R}^n$ is called a Nikodym set if, for each point $x \in [0, 1]^n$, there is a line $L(x)$ so that

- The point x lies in $L(x)$.
- Except for x, $L(x) \cap [0, 1]^n$ lies in N.

For example, if I remove the line $y = 1/2$ from the square $[0, 1]^2$, the result is a Nikodym set. If I remove a circle, then it isn't. In the 1920's, Nikodym proved the following counterintuitive result:

Theorem 2.1. *There are Nikodym sets of measure zero in each dimension $n \geq 2$.*

The sets Nikodym constructed have full Hausdorff dimension, as do all known constructions. This suggests the following conjecture:

Conjecture 2.2. *Every Nikodym set $N \subset [0,1]^n$ has Hausdorff dimension n.*

This conjecture turns out to be related to many deep problems in analysis, and it has come to play an important role. (There is also a more famous cousin, the Kakeya conjecture). Although they may look rather arbitrary at first, the Nikodym and Kakeya conjectures underlie a variety of important and natural problems in Fourier analysis, PDE, and number theory. A lot of effort has gone into studying the problem, and we are still far from resolving it. Faced with the difficult problem, mathematicians have looked at cousin problems and toy problems that might give some insight. For example, Tom Wolff formulated a finite field version.

Let $\mathbb{F}$ be a finite field with q elements. A set $N \subset \mathbb{F}^n$ is called a Nikodym set if, for each point $x \in \mathbb{F}^n$, there is an affine line $L(x)$ so that

- The point x lies in $L(x)$.
- Except for x, the line $L(x)$ lies in N.

The analogue of the Nikodym problem is to ask how many points there must be in a Nikodym set. The finite field Nikodym conjecture says that every Nikodym set must have at least $c_n q^n$ points. For a while, the two problems seemed about equally hard. About five years ago, Dvir proved the finite field Nikodym conjecture. The proof was only a page long and it shocked a lot of mathematicians in the area. The proof uses the polynomial method, somewhat in the spirit of the Berlekamp-Welch algorithm.

Here is a sketch of the proof. Suppose that N is a small Nikodym set, with only $(2n)^{-n} q^n$ elements. By dimension counting, we can then find a non-zero polynomial P that vanishes on N with degree at most $(q/2)$. Fix a point $x \in \mathbb{F}^n$, and consider the line $L(x)$. By the definition of a Nikodym set, at least $q-1$ points of $L(x)$ lie in N. Therefore, P must vanish on $q-1$ points of $L(x)$. Since the degree of P is $< q-1$, P must vanish on the whole line $L(x)$, in particular $P(x) = 0$. Now x was arbitrary so $P(x) = 0$ at every point $x \in \mathbb{F}^n$. Given that P vanishes at every point and that the degree of P is $< q$, it's not hard to show that P is the zero polynomial, giving a contradiction.

Filling in all details of the proof takes two more short paragraphs, and we'll do it later. Previously, people tried hard to prove the result without this polynomial trick, and it seems to be extremely difficult. The situation raises a lot of questions. Do polynomials really play such an important role in this problem? If so, why? What does the method have to do with the problem? We'll come back to these kinds of questions a number of times throughout the notes.

People tried to adapt the polynomial method to attack the original Nikodym conjecture, but there are serious difficulties. The polynomial method hasn't yet led

to any significant progress on Nikodym-type problems in Euclidean space. But it has had a lot of success in combinatorial problems involving finitely many lines in Euclidean space.

3. The distinct distance problem

The polynomial method has led to solutions for several challenging problems in extremal combinatorics, as well as giving new proofs and perspectives for some important known results. We will study most of these new proofs. The result that we will spend the most time tackling is an estimate for the distinct distance problem in the plane.

Suppose $P \subset \mathbb{R}^2$ is a finite set with N elements. We let $d(P)$ denote the set of non-zero distances between elements of P:

$$d(P) := \{|p-q|\}_{p,q \in P, p \neq q}.$$

(We are using the standard Euclidean distance on $\mathbb{R}^2$.)

Let's consider some examples.

(1) N generic points in the plane gives $|d(P)| = \binom{N}{2} \sim N^2$.
(2) N evenly spaced points along a line gives $|d(P)| = N - 1$.
(3) N points arranged in a $\sqrt{N} \times \sqrt{N}$ square grid gives $|d(P)| \sim N(\log N)^{-1/2}$.

In the 1940's, Erdős raised the question how small the distance set $d(P)$ could possibly be. He worked out the example of the square grid, and he conjectured that the square grid is minimal up to constant factors: in other words, any set of N points should have $|d(P)| \geq cN(\log N)^{-1/2}$. A number of people have proven lower bounds for the distinct distance problem using different techniques. Before the polynomial method, the best lower bound proved that the number of distinct distances is $\gtrsim N^{.864}$. The book *The Erdős Distance Problem*, by Garibaldi, Iosevich, and Singer, describes various approaches to the problem. Using the polynomial method, Nets Katz and I proved the following theorem.

Theorem 3.1. *(G.-Katz, 2010) For any set of N points in the plane, the number of distinct distances is at least $cN(\log N)^{-1}$.*

This proof is more difficult than the proof of finite field Kakeya or joints. It will take us 40-60 pages all in all. There is a new ingredient coming from topology and a new ingredient coming from ruled surfaces in algebraic geometry. Nevertheless, the proof is pretty elementary, and I hope it will be accessible to a broad range of readers.

4. Number Theory

The polynomial method is also connected with work in number theory from the early 20th century. In particular, there was an important breakthrough by Thue in the study of diophantine equations. Thue was able to prove that many polynomial equations have only finitely many integer solutions. Here are a couple of examples.

A. The polynomial $y^3 - 2x^3 = 1$ has only finitely many integer solutions.

B. The polynomial $y^4+6x^2y^2+7x^3y+101x^4$ has only finitely many integer solutions.

These are just special cases of Thue's general theorem.

Theorem 4.1. *(Thue 1908) If $P(x, y)$ is an irreducible homogeneous polynomial of degree ≥ 3, and A is an integer, then the equation $P(x, y) = A$ has only finitely many integer solutions.*

Before Thue, people usually studied single equations or small families of equations. Thue's theorem was much more general. Looking at irreducible polynomials is not that big a restriction, because one can study reducible polynomials by considering the factors. Being homogeneous is a restriction, but Siegel was able to generalize Thue's work to develop a systematic theory of diophantine equations in two variables.

Thue's argument involved some 'auxiliary polynomials'. In order to study a particular diophantine equation, like equation A above, Thue needed an infinite sequence of 'auxiliary polynomials' with special properties. Thue tried to construct these polynomials explicitly. He was able to do it in some examples like equation A, but he couldn't do it for many other equations. Then he realized that the auxiliary polynomials needed to exist for every equation because of a simple counting argument like the one at the beginning of this section. This was probably the most important idea in Thue's breakthrough.

Reviewing Thue's work at the 1974 ICM, Schmidt described it as follows: "The idea of asserting the existence of certain polynomials rather than explicitly constructing them is the essential new idea in Thue's work. As Siegel [1970] points out, a study of Thue's papers reveals that Thue at first tried hard to construct the polynomials explicity (and he could actually do so [for equations of the form $y^d - Bx^d = A$])."

5. Goals of the course

The polynomial method gives several strikingly short applications. The first goal of the course is to study these. I'll emphasize three examples: the Berlekamp-Welch algorithm, the finite field Nikodym and Kakeya problems, and the joints problem.

These short proofs are hard to appreciate without context. The next goal is to learn the context of these results. In particular, we will learn about incidence geometry: combinatorial estimates about how lines and other basic geometric objects intersect each other.

The third goal of the course is to prove the estimate about the distinct distance problem.

The fourth goal is to explore connections between the polynomial method and different parts of mathematics. We will see some connections involving computer science, algebraic geometry, topology, harmonic analysis, and number theory.

The fifth goal is to mull over some philosophical questions related to the polynomial method. For example, what is special about polynomials? Why are polynomials involved in these problems?

18.S997 The Polynomial Method
Fall 2012