The polynomial method in combinatorics

Notice: To solve the problems, you will need to refer to facts in the back of the handout. Once you go to the back of the handout, if you don't remember a definition or theorem, you can solve the problems there to remind you of the ideas. *If you go to the back of the handout and you realize that you are lost — **don't worry** — ask an instructor!*

Vagueries: Polynomials are useful to solve some problems in combinatorics. The fundamental usefulness of polynomials in combinatorics seems to derive from — in part — the following features:

1. (Flexibility) In n variables there are lots of polynomials: the dimension of polynomials of degree D in n variables is roughly D^n.
2. (Rigidity) If a degree D polynomial vanishes at more than D points of a line $l \subset F^n$, then P vanishes on the whole line l.

1. The basic lemmas of the polynomial method

1.1 Exercise. Recall the definition of a field F. Give a few examples of some finite fields.Give three examples of a field with infinitely many elements. Is $\mathbb{Z}/6\mathbb{Z}$ with the usual addition and multiplication a field?

1.2 Definition. Let F be a field. We denote

$$\mathrm{Poly}(F^n) \coloneqq \left\{ \sum_{j=1}^{m} a_j x_1^{k_1^j} \cdots x_n^{k_n^j}, \quad a_1, \cdots, a_m \in F, \quad k_1^1, \cdots, k_n^m \in \mathbb{N} \right\}.$$

In words, $\mathrm{Poly}(F^n)$ is the collection of polynomials in n variables with coefficients in the field F.

1.3 Exercise. Write down 5 different polynomials in $\mathrm{Poly}((\mathbb{Z}/2\mathbb{Z})^3)$.

1.4 Exercise. Give an example of a field F and a polynomial $P \in \mathrm{Poly}(F)$ so that

$$P(a) = 0 \quad \text{for all } a \in F.$$

Must it be the case that p is the zero polynomial? If not give an example of a polynomial $P \in \text{Poly}(F)$ so that $P(a) = 0$ for all a but $P \neq 0$ as a polynomial.

Hint: this exercise is not so easy... there is a big difference if the field has finitely many or infinitely many elements.

1.5 Exercise. Prove that $\text{Poly}(F^n)$ is a vector space. Recall the definition of **vector space dimension**. What is the dimension of this vector space?

1.6 Definition. Let $P \in \text{Poly}(F^n)$ be a polynomial of the form

$$P(x) = ax_1^{k_1} \cdots x_n^{k_n}.$$

We call such a polynomial a **monomial** and we define its **degree** to be the following number

$$\deg(P) \coloneqq k_1 + \cdots + k_n.$$

If P is not a monomial but the sum of monomials we define

$$\deg(P) \coloneqq \text{ the degree of its largest monomial term.}$$

1.7 Exercise. State the degree of the following polynomials

1. $xy^2 + y^2 \in \text{Poly}(\mathbb{R}^2)$,
2. $xyz + z^2x^2 \in \text{Poly}(\mathbb{Z}^3)$,
3. $4x^{100} + x \in \text{Poly}(\mathbb{Z}/2\mathbb{Z})$ (careful).

1.8 Definition. Let F be a field. We denote

$$\text{Poly}_D(F^n) \coloneqq \{P \in \text{Poly}(F^n) : \deg(P) \leq D\}.$$

In words, these are the polynomials with degree at most D.

1.9 Exercise. How many elements does $\text{Poly}_3((\mathbb{Z}/2\mathbb{Z})^2)$ have?

1.10 Exercise. Prove that $\text{Poly}_D(F^n)$ is a vector space. What is the dimension of this vector space?

1.11 Exercise (Hint # 1)**.** Can you write down a simple basis for $\text{Poly}_D(F^n)$? First, write down a basis for $\text{Poly}_3(F)$. Then try to write down a basis for $\text{Poly}_3(F^2)$.

1.12 Exercise (Hint # 2, Stars and Bars)**.** Once you have a good basis, you need to count how many elements you have written down! It may help to remember how to answer the following combinatorial question: how many ways are there to distribute n cookies to m students (students can get 0 cookies)? It is useful to think of placing dividers between the cookies to assign them to the students.

1.13 Exercise (Hint # 3)**.** When trying to count the basis you constructed, what are the cookies and what are the students? Can you use this to count the dimension of $\mathrm{Poly}_D(F^n)$? You may have to be a bit clever since $\mathrm{Poly}_D(F^n)$ includes polynomials with degree less than D (not just those with degree equal to D).

1.14 Exercise. Prove that

$$\dim(\mathrm{Poly}_D(F^n)) \geq \frac{D^n}{n!}.$$

1.1. Small polynomials can vanish on big sets

This exercise asks you to prove one of the key inputs of the polynomial method.

1.15 Exercise. Let $S \subset F^n$ be a **finite** set of points. Suppose that

$$\dim(\mathrm{Poly}_D(F^n)) > |S|$$

prove that there exists a **non-zero** polynomial P with degree at most D so that

$$p(s) = 0 \quad \text{for all } s \in S.$$

Hint: this requires a very fundamental theorem from Linear Algebra. Which one is it?

1.16 Exercise. Let $S \subset F^n$ be a finite set. Prove there is a **non-zero** polynomial P which vanishes on S and whose total degree D satisfies

$$D \leq n|S|^{\frac{1}{n}}.$$

Idea: in larger dimensions, we can find a low degree polynomial which vanishes on many points.

1.2. Towards the vanishing lemma

The following fact is called the division algorithm.

1.17 Exercise. If $P(x) \in \mathrm{Poly}_D(F)$ is a polynomial in one variable and $x_1 \in F$, then we can uniquely write $P(x)$ in the form

$$P(x) = (x - x_1)P_1(x) + r$$

where $P_1(x) \in \mathrm{Poly}_{D-1}(F)$ and $r \in F$.

Hint: use induction.

1.18 Exercise. Find $P_1(x), r$ so that

$$x^3 - 3x - 2 = (x - 2)P_1(x) + r.$$

Assume we are working in $\mathrm{Poly}(\mathbb{R})$.

1.19 Exercise. If $P(x) \in \mathrm{Poly}_D(F)$ is a polynomial over a field F and $P(x_1) = 0$ for some $x_1 \in F$, then we can write $P(x)$ in the form:

$$P(x) = (x - x_1)P_1(x)$$

for some polynomial $P_1 \in \mathrm{Poly}_{D-1}(F)$.

1.3. The vanishing lemma

1.20 Exercise. Let $P \in \mathrm{Poly}_D(F)$, suppose that P vanishes on $D + 1$ points in F then P must vanish on all of F.

Hint: use induction again.

1.21 Exercise (The vanishing lemma)**.** Let $P \in \mathrm{Poly}_D(F^n)$, suppose that P vanishes on $D + 1$ points of a line $l \subset F^n$ then P must vanish on all of l.

1.22 Exercise. Let W be a subspace of functions $F^n \to F$ which satisfies the following: whenever $f \in W$ vanishes at $D+1$ points on a line, then f vanishes at every point on the line.

Show that if $|F| > D + 1$ then $\dim W \leq (D + 1)^n$.

It is an open problem to understand the maximum possible $\dim W$ as a function of n and D.

2. Two "famous" problems and the polynomial method

In March 2008 Zeev Dvir posted a solution to the finite field Kakeya set problem. This problem had been open for many years and was considered to be very difficult. Previous results on the finite field Kakeya set problem were not very good, and the Kakeya set problem in infinite fields was known to be notoriously hard (only this past year it has been proved in a 100 page paper in dimension 3). It therefore came as a huge shock that Dvir's paper was only 5 pages long. The proof was completely straightforward and used only elementary properties of polynomials. In this section we will rehearse the proof.

2.1 Notation. F_q refers to the finite field with q elements. There is not always a finite field with q elements. In your head, you may proceed acting like q is a prime and $F_q = \mathbb{Z}/q\mathbb{Z}$ (check this is a field!).

2.2 Exercise. How many points does F_q^n contain?

2.3 Exercise. Draw F_5^2. Draw 3 different lines in F_5^2 which are not parallel.

2.4 Exercise. Draw F_3^2. How many different directions are there in F_3^2? Draw a line in each direction in F_3^2, does your collection of lines cover all the points of F_3^2?

2.5 Question (Open ended)**.** In your opinion, what does it mean that a subset $A \subset F_q^n$ contains "much fewer points" than all of F_q^n?

The following lemma will be an important input for the two problems we are going to solve in this section.

2.6 Exercise. Let $P \in \mathrm{Poly}_{q-1}(F_q^n)$. If $P(x) = 0$ for every $x \in F_q^n$ prove that P is the zero polynomial. Prove also that this conclusion does not need to be true if $P \in \mathrm{Poly}_q(F_q^n)$.

Hint: use the vanishing lemma and induction.

2.1. The finite field Nikodym set problem

A set $N \subset F_q^n$ is called a **Nikodym set** if, for each point $x \in F_q^n$, there is a line $L(x)$ so that $L(x) - \{x\} \subset N$.

2.7 Question (Main question)**.** A trivial example of a Nikodym set is $N = F_q^n$. Can one find a significantly smaller Nikodym set?

2.8 Exercise. Draw a Nikodym set for F_3^2 which is not the whole set.

2.9 Exercise. Draw a Nikodym set for F_5^2 which is not the whole set and which has strictly fewer points than the set you drew for the previous exercise.

2.10 Exercise (The finite field Nikodym Theorem (Dvir 2009))**.** Prove that any Nikodym set $N \subset F_q^n$ contains at least $c_n q^n$ elements. Here, c_n means any constant which may depend on n but does not depend on q. If you like you may prove this with the specific choice $c_n = (10n)^{-n}$.

Hints:

1. Proceed by contradiction, you will need Exercise 1.16.
2. The number $c_n = (10n)^{-n}$ is small enough that $n((10n)^{-n}q^n)^{\frac{1}{n}} = \frac{1}{10} \cdot q$ is strictly less than q. Why is this useful? You may want to look at some of the previous exercises.

2.2. The finite field Kakeya set problem

A set $K \subset F_q^n$ is called a **Kakeya set** if it contains a line in each direction. In other words, if for every $a \in F_q^n - \{0\}$ there is a vector $b \in F_q^n$ so that

$$\{at + b : t \in F_q\} \subset K.$$

2.11 Question (Main question)**.** A trivial example of a Kakeya set is $K = F_q^n$. Can one find a significantly smaller Kakeya set?

2.12 Exercise. Draw a Kakeya set in F_3^2 which is not the whole set.

2.13 Exercise. Draw a Kakeya set in F_7^2 which is not the whole set.

2.14 Exercise (The finite field Kakeya Theorem (Dvir 2009))**.** Prove that any Kakeya set $K \subset F_q^n$ contains at least $c_n q^n$ elements. If you like you may prove this with the specific choice $c_n = (10n)^{-n}$.

Hint: proceed like the Nikodym set problem ... but you will have to be more clever.

If you like you may prove this with the specific choice $c_n = (10n)^{-n}$.

3. Encore: questions to test understanding

3.1 Exercise. Suppose that $S = \{x_0, \cdots, x_D\}$ is a set of $D+1$ distinct points in the field F. Define the evaluation map

$$E_S : \mathrm{Poly}_D(F) \to F^{D+1}$$

by

$$E_S(Q) = (Q(x_0), \cdots, Q(x_D)).$$

Prove that E_S is an isomorphism.

Hint: you will need a) to remember the definition of isomorphism (ask if you don't know) b) to use a very important theorem of linear algebra c) prove that E_S is surjective (or use the following exercise).

3.2 Exercise. Let

$$x^0, \cdots, x^D \in \mathrm{Poly}_D(F)$$

be a basis for $\mathrm{Poly}_D(F)$. Using this basis write E_S as a matrix.

Can you find a formula for the determinant of E_S?

3.3 Exercise (The Schwartz-Zippel lemma I)**.** Let $P \in \mathrm{Poly}_D(F^n)$ be a nonzero polynomial, prove that

$$\#\{\text{zeros of } P\} \leq |F|^{n-1} D.$$

3.4 Exercise (The Schwartz-Zippel lemma II)**.** Suppose that $A_1, \cdots, A_n \subset F$ are finite subsets with $|A_1| = \cdots |A_n| = N$. Suppose that $P \in \mathrm{Poly}_D(F^n)$ is a non-zero polynomial. Prove that the number of zeros in the set $A_1 \times \cdots \times A_n$ is at most DN^{n-1}.

What does this tell you about the number of zeros in comparison to the size of $|F|$?

3.5 Exercise. Let $P(x, y)$ and $Q(x)$ be polynomials such that $P(x, Q(x))$ is the zero polynomial. Then there exists a polynomial $S(x, y)$ such that:

$$P(x, y) = (y - Q(x))S(x, y) \tag{3.1}$$

3.6 Exercise. We consider a collection of curves $\Gamma_{\mathbf{a}} \subset \mathbb{F}_q^n$ parametrized by $\mathbf{a} \in \mathbb{F}_q^{n-1}$. For each $\mathbf{a} \in \mathbb{F}_q^{n-1}$ and $1 \leq j \leq n-1$, suppose that $Q_{\mathbf{a},j} \in \mathrm{Poly}_d(\mathbb{F}_q)$

is a polynomial in one variable of degree at most d. Let $\Gamma_{\mathbf{a}}$ be defined as the graph:

$$\Gamma_{\mathbf{a}} := \{(Q_{\mathbf{a},1}(t), Q_{\mathbf{a},2}(t), \ldots, Q_{\mathbf{a},n-1}(t), t) \in \mathbb{F}_q^n \mid t \in \mathbb{F}_q\}.$$

Suppose also that $(Q_{\mathbf{a},1}(0), \ldots, Q_{\mathbf{a},n-1}(0)) = \mathbf{a}$, so that $(\mathbf{a}, 0) \in \Gamma_{\mathbf{a}}$.

Prove that there is a constant $c(d, n) > 0$ so that the size of the union of all such curves is bounded below by:

$$\left| \bigcup_{\mathbf{a} \in \mathbb{F}_q^{n-1}} \Gamma_{\mathbf{a}} \right| \geq c(d, n) q^n$$

3.7 Exercise (Hard)**.** Let q be an odd prime. Can you give a general (i.e. works for all $n \in \mathbb{N}$) construction of a Kakeya set in F_q^n that with the property that

$$\frac{|K_n|}{|F_q^n|} \to 0 \quad \text{as } n \text{ becomes very large}$$

where K_n denotes the Kakeya set constructed in F_q^n.

4. Appendix: Linear Algebra

4.1. The definition of a vector space

4.1 Definition (Vector Space (Informal))**.** A **vector space** V over a field F is a set whose elements (called **vectors**) can be added together and multiplied ("scaled") by elements of F (called **scalars**), satisfying several axioms (closure, associativity, commutativity, existence of identity/inverse elements for both operations, and distributivity).

4.2 Definition (Vector Space)**.** A set V is a **vector space** over a field F (whose elements are called **scalars**) if there are two operations:

1. **Vector Addition ($\mathbf{v} + \mathbf{w}$):** Takes two vectors $\mathbf{v}, \mathbf{w} \in V$ and returns a unique vector in V.
2. **Scalar Multiplication ($c\mathbf{v}$):** Takes a scalar $c \in F$ and a vector $\mathbf{v} \in V$ and returns a unique vector in V.

These operations must satisfy the following ten axioms:

Axioms for Vector Addition (V1-V5). Vector addition must make $(V, +)$ an Abelian group:

V1. Closure: For all $\mathbf{u}, \mathbf{v} \in V$, $\mathbf{u} + \mathbf{v} \in V$.

V2. Commutativity: For all $\mathbf{u}, \mathbf{v} \in V$, $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.

V3. Associativity: For all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.

V4. Zero Vector: There exists a **zero vector** $\mathbf{0} \in V$ such that $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for all $\mathbf{v} \in V$.

V5. Additive Inverse: For every $\mathbf{v} \in V$, there exists a vector $-\mathbf{v} \in V$ such that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$.

Axioms for Scalar Multiplication (V6-V10). Scalar multiplication links the field F to the vector space V:

V6. Closure: For all $c \in F$ and $\mathbf{v} \in V$, $c\mathbf{v} \in V$.

V7. Distributivity (Vector Sum): For all $c \in F$ and $\mathbf{u}, \mathbf{v} \in V$, $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$.

V8. Distributivity (Scalar Sum): For all $c, d \in F$ and $\mathbf{v} \in V$, $(c + d)\mathbf{v} = c\mathbf{v} + d\mathbf{v}$.

V9. Associativity: For all $c, d \in F$ and $\mathbf{v} \in V$, $(cd)\mathbf{v} = c(d\mathbf{v})$.

V10. Unit Identity: For the multiplicative identity $1 \in F$, $1\mathbf{v} = \mathbf{v}$ for all $\mathbf{v} \in V$.

4.2. More definitions about vectors

4.3 Definition (Linear Combination and Linear Independence)**.**

1. A **linear combination** of vectors $v_1, \ldots, v_k \in V$ is an expression of the form $c_1 v_1 + \cdots + c_k v_k$, where $c_1, \ldots, c_k \in F$ are scalars.
2. A set of vectors $\{v_1, \ldots, v_k\}$ is **linearly independent** if the only way to express the zero vector as a linear combination of these vectors is the trivial way, i.e., $c_1 v_1 + \cdots + c_k v_k = \mathbf{0}$ implies $c_1 = c_2 = \cdots = c_k = 0$.

4.4 Exercise. Prove that

$$\text{Poly}(\mathbb{R}) = \text{polynomials with coefficients in } \mathbb{R}$$

is a vector space over $\mathbb{R}$. Show that x and x^2 are linearly independent as elements of this vector space.

4.5 Definition (Basis)**.** A **basis** $\mathcal{B}$ for a vector space V over a field F is a subset of V that satisfies two conditions:

1. The set $\mathcal{B}$ is **linearly independent**.
2. Every vector in V can be written as a linear combination of vectors in $\mathcal{B}$ (i.e., $\mathcal{B}$ **spans** V).

4.6 Example. Is $1, x$ a basis for $\text{Poly}(\mathbb{R})$? Is $x + 2, x + 3$ a basis for $\text{Poly}_1(\mathbb{R})$?

4.7 Definition (Dimension)**.** Let V be a vector space over a field F. The **dimension** of V is defined as the size of $\mathcal{B}$ where $\mathcal{B}$ is **any** basis of V.

4.8 Remark. It is not obvious that the dimension of a vector space V only depends on the vector space! You can try to prove this for yourself if you like. In the meantime we content ourselves with the following exercise.

4.9 Exercise. Find three different bases for $\mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}$ and check that they all give the same dimension.

4.10 Exercise. Find 5 different bases for $\text{Poly}_3(\mathbb{C})$ and check that they all give the same dimension.

4.11 Example (Examples of Vector Spaces and Dimensions).

1. Coordinate Space $\mathbb{R}^n$ (or F^n): The set of all n-tuples $\mathbf{x} = (x_1, \ldots, x_n)$ with $x_i \in \mathbb{R}$ (or F). The **standard basis** is $\{e_1, \ldots, e_n\}$, where e_i is the vector with 1 in the i-th position and 0 elsewhere.

$$\mathrm{Dim}(\mathbb{R}^n) = n.$$

2. Space of Low-Degree Polynomials ($\mathrm{Poly}_D(\mathbb{R})$): The set of polynomials $a_0 + a_1 x + \cdots + a_D x^D$ with real coefficients a_i. The basis is the set of monomials $\{1, x, x^2, \ldots, x^D\}$.

$$\mathrm{Dim}(\mathrm{Poly}_D(\mathbb{R})) = D + 1.$$

4.3. (Perhaps) the most important theorem in linear algebra

4.12 Definition. A **linear transformation** $T : V \to W$ where V, W are vector spaces over F is a function so that

$$T(av + bw) = aT(v) + bT(w)$$

for every $a, b \in F$, $v, w \in V$.

4.13 Exercise. If T is a linear map what is $T(0)$?

4.14 Exercise. Write down all the linear maps from $(\mathbb{Z}/2\mathbb{Z})^3$ to $(\mathbb{Z}/2\mathbb{Z})^2$.

4.15 Exercise. Let $T : V \to W$ be a linear transformation, show that

1. $\ker T \coloneqq \{v \in V : T(v) = 0\}$ is a vector space,
2. $\mathrm{Im}\, T \coloneqq \{T(v) : v \in V\}$ is a vector space.

4.16 Theorem (Rank-Nullity Theorem)**.** *Let V and W be finite-dimensional vector spaces over a field F, and let $T : V \to W$ be a linear transformation. Then*

$$\mathit{Dim}(V) = \mathit{Dim}(\ker(T)) + \mathit{Dim}(\mathrm{Im}(T)).$$

4.4. The important version of the Rank-Nullity Theorem i.e. linear algebra pigeonhole principle

4.17 Exercise. Prove the following corollary of the Rank-Nullity Theorem.

4.18 Theorem. *Let $T : V \to W$ be a linear map. If*

$$\dim V > \dim W$$

then there is $v \in V$ so that $T(v) = 0$ and $v \neq 0$.

4.19 Exercise. Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear map so that

$$T(1,0,0) = (2,3), \quad T(0,1,0) = (3,4), \quad T(0,0,1) = (12,3).$$

Show that there is a nonzero vector v in $\mathbb{R}^3$ so that $T(v) = 0$. Can you find what v is constructively?

4.20 Exercise. Let $T : \mathbb{R}^8 \to \mathbb{R}^7$ be the linear map so that

$$\begin{aligned}
T(1,0,0,0,0,0,0,0) &= (1,3,4,5,23242,1,2),\\
T(0,1,3,0,0,0,0,0) &= (1,3,4,1,12,1,4),\\
T(0,0,1,0,0,0,0,0) &= (1,3,4,43,23242,1,2),\\
T(0,0,0,2,2,0,0,0) &= (1,3,123,5,77,1,2),\\
T(0,0,0,0,1,0,0,0) &= (1,3,4,23,43,1,2),\\
T(9,0,0,0,0,1,0,0) &= (77,3,4,5,23242,1,2),\\
T(0,0,2,0,0,0,3,0) &= (1,77,4,5,23242,1,0),\\
T(1,0,0,1,0,0,0,6) &= (1,3,4,85,23242,98,2).
\end{aligned}$$

Show that there is a nonzero vector v in $\mathbb{R}^8$ so that $T(v) = 0$. Can you find what v is constructively?