Lectures in combinatorics

Asaf Ferber [*]

August 26, 2019

Contents

[*] Department of Mathematics, MIT. Email: ferbera@mit.edu.

Disclaimer: I've started writing these notes when I was teaching Algebraic methods in Combinatorics at MIT. The notes are based on many resources that I found online. In particular, I used notes of Dan Spielman from Yale, Jeff Kahn from Rutgers, Benny Sudakov from ETH, Zurich, László Babai from University of Chicago, and from some chapters of a book by Richard Stanley from MIT and more. Apologies to those who I forgot to mention even though that I used their notes. This was the first time that I was giving such a class and I was preparing it on a weekly base, so the reader may find the structure of the topics being far from optimal. It is not recommended to follow these notes in the given order (or in any). If you plan on using these notes in a class that you are giving, don't hesitate to email me and I'll send you the TEX file so you could edit to your convenience (and of course, I would be very grateful to get a more organized version of it in return after your semester ends).

My plan is to keep extending/polishing the notes on a regular base until it will become a book where each chapter is a series of few independent lectures on some topic in advanced combinatorics. I'll do my best to cover as many different topics as possible.

1 Some linear algebra and other useful lemmas

The following theorem will be frequently used throughout the course and we will not prove it here. We refer to it as the Spectral Theorem in these lectures:

Theorem 1.1. *1. The eigenvalues of a graph G are always real (as its adjacency matrix is a real values, symmetric matrix).*

2. A_G is diagonalizable.

3. There is an orthonormal basis of eigenvectors.

The following formula seems important enough to memorize as it is going to be used in almost every proof.

Lemma 1.2. $x^t A_G x = \sum_{i=1}^n x_i \sum_{j=1}^n a_{ij} x_j = \sum_{ij \in E(G)} x_i x_j$. *In particular, if x_S is a characteristic vector of a subset $S \subseteq V(G)$, then $x_S^t A_G x = 2e(S)$.*

The following min-max theorem helps in estimating the eigenvalues of symmetric, real-valued matrices.

Theorem 1.3 (Min-max)**.** *Suppose A is symmetric, real valued matrix and let $\lambda_1 \geq \ldots \geq \lambda_n$ be its eigenvalues. Then, for all i we have*

$$\lambda_i = \max_{dim(F)=i} \min_{x \in F, x \neq 0} \frac{x^t A x}{x^t x} = \min_{dim(F)=n-i+1} \max_{x \in F, x \neq 0} \frac{x^t A x}{x^t x}.$$

Remark 1.4. *F runs over all subspaces of $\mathbb{R}^n$ of the appropriate dimension.*

Note that from the above theorem, in particular we have that for all x,

$$x^t A x \geq \lambda_n x^t x.$$

Indeed, there is exactly one F of $dim(F) = n$, and therefore

$$\lambda_n = \min_{x \neq 0} \frac{x^t A x}{x^t x}.$$

Moreover, we have

$$x^tAx \leq \lambda_1 x^tx$$

by using

$$\lambda_1 = \min_{dim(F)=n} \max_{x\in F, x\neq 0} \frac{x^tAx}{x^tx} = \max_{x\neq 0} \frac{x^tAx}{x^tx}.$$

In particular, this shows that

$$\lambda_1 \geq 2m/n = d(G)$$

(just take $x = 1$).

The next lemma is not from linear algebra, but you may find it as a simple lemma which is useful in various calculations (maybe even PSET?).

Lemma 1.5. *suppose $\alpha_1, \ldots, \alpha_r$ and $\beta_1, \ldots, \beta_s$ are non-zero complex numbers such that for all positive integers ℓ we have*

$$\alpha_1^\ell + \ldots + \alpha_r^\ell = \beta_1^\ell + \ldots + \beta_s^\ell.$$

Then, $r = s$ and $\alpha_i = \beta_i$ for all i, up to a permutation.

Proof. We use *generating functions* for it. Multiply the above equation by x^ℓ and summing over all $\ell \geq 1$, we obtain

$$\sum \frac{\alpha_i x}{1-\alpha_i x} = \sum \frac{\beta_i x}{1-\beta_i x}.$$

Multiply both equations by $1 - \gamma x$ and let $x \to 1/\gamma$. We obtain that LHS is the number of α_i which equal γ and RHS is same just for β. Therefore, we obtain the desired. □

The following lemma helps to simplify some calculations:

Lemma 1.6. *Let G be a triangle free graph on n vertices with m edges. Then,*

$$\sum_{x\in V(G)} d(x)^2 = \sum_{xy\in E(G)} (d(x)+d(y)) \leq mn.$$

Proof. Note that $\sum_{xy}(d(x) + d(y))$ contributes $d(x)$ times $d(x)$ for all x (every edge touching x is counted $d(x)$ times from x point of view). Therefore, we obtain the first equality. On the other hand, note that each edge is being counted $d(x) + d(y)$ times, which is at most n (there are no common neighbors to $xy \in E(G)$ as otherwise it results in a triangle). This completes the proof. □

The following theorem is version of Perron-Frobenius theorem from linear algebra, tailored for our purposes.

Theorem 1.7 (Perron-Frobenius)**.** *Let A be the adjacency matrix of a connected graph which is not an isolated vertex. If ρ is the maximum absolute value of the eigenvalues of B, then $\rho > 0$, and there is an eigenvalue equal to ρ. Moreover, there is an eigenvector for ρ all of whose entries are positive.*

2 Power of linear independence – some classical examples

The following few examples serve as examples for how simple and basic ideas from algebra yield simple solutions to some problems in combinatorics, for which a combinatorial proof is either very hard or even unknown.

2.1 Even and Odd towns

The first problem we describe is the Even and Odd town problem. That is:

Problem 2.1. *Suppose that there are n people in some town. We wish to form* clubs *where each club is of even size and all the intersections are even. How many possible ways exist?*

Apparently, with a small change, the answer becomes completely different.

Problem 2.2. *Same question, just this time each club is of ODD size. How many possible ways exist?*

2.2 The 2-distances problem

The second problem is regarding the number of points one can arrange in the plane with some restrictions on the number of possible pairwise distances. Formally, we can describe it as follows: Let $a_1, \dots, a_t$ be points on the plane such that all pairwise distances are the same. Then it is quite clear that $t \leq n + 1$. Now, what if we allow two distances?

Theorem 2.3. *Let $m(n)$ be the maximum number of such points. Then,*

$$n(n+1)/2 \leq m(n) \leq (n+1)(n+4)/2.$$

Proof. Lower bound can be obtained by considering e_{ij} for all $i < j$. For the upper bound, assume the distances are δ_1, δ_2. For each i define $f_i : \mathbb{R}^n \to \mathbb{R}$:

$$f_i(x) := \left(\|x - a_i\|_2^2 - \delta_1^2\right)\left(\|x - a_i\|_2^2 - \delta_2^2\right).$$

Note that $f_i(a_i) \neq 0$ and $f_i(a_j) = 0$ for all $j \neq i$. Therefore, the f_is are linearly independent over the linear space generated by $(\sum x_k^2)^2, (\sum x_k^2)x_j, x_i x_j, x_i, 1$. This gives the upper bound. □

Exercise 2.4. *obtain something similar for s-distance problem*

$$\binom{n+1}{s} \leq m(n,s) \leq \binom{n+s+1}{s}.$$

2.3 Graham-Pollak

The following theorem due to Graham and Pollack (1972) gives a bound on the number of edge-disjoint, complete bipartite graphs needed in order to cover all the edges of K_n.

Theorem 2.5 (Graham and Pollak)**.** *Let $G_1, \dots, G_t$ be edge disjoint, complete bipartite graphs, such that $\cup G_i = K_n$, then $t \geq n - 1$.*

It's an easy exercise to prove that this bound is tight.

Proof. For each complete bipartite graph (X_k, Y_k) we assign an $n \times n$ matrix A_k in which $a_{ij} = 1$ iff $i \in X_k$ and $j \in Y_k$. Clearly $S := \sum A_k$ satisfies $S + S^t = J - I$. We now claim that $r(S) \geq n - 1$. Indeed, otherwise, there exists x with $\sum x_i = 0$ and $Sx = 0$ (as S is of rank at most $n - 2$ we can consider the $n - 1$ equations in the linear system $Sx = 0$ and $\sum x_i = 0$ so there must be a solution). Thus, $S^t = -x$, and so $0 = x^t S^t x = -\|x\|_2^2$. □

2.4 The number of perfect matchings in d-regular graphs

As a last example, we give a simple proof for the fact that every d-regular, bipartite graph with d being an even integer, has an even number of perfect matchings.

Theorem 2.6. *Let G be bipartite graph which is d-regular, for d = even. Then, the number of perfect matchings is even.*

Proof. Let's work over $\mathbb{Z}_2$. Let A_G be the (bipartite) adjacency matrix of G. Note that over $\mathbb{Z}_2$ we have $Per(A_G) = Det(A_G) =$ parity of the number of perfect matchings in G'. Now, as d is even and the sum of all rows equals the all d vector (which is $0(mod\ 2)$, we conclude that $Det(A_G) = 0(mod\ 2)$. □

The following problem is quite embarrassingly open.

Problem 2.7. *What is the proportion of d-regular, bipartite graphs which have an odd number of perfect matchings where d is an odd integer? Some simulations suggests that in case $d = 3$, the answer should be roughly $1/2$. Any ideas?*

3 Some spectral graph theory

3.1 Walks in Graphs

Before we start, let's first analyze the eigenvalues of a simple matrix that will be used quite often, that is the $n \times n$ matrix, denote by J, consisting of all 1 entries.

Lemma 3.1. *Eigenvalues of J are n (with multiplicity 1) and 0 (with multiplicity $n-1$).*

As an immediate corollary, we obtain the eigenvalues of the adjacency matrix of the complete graph on n vertices, $A(K_n)$.

Proposition 3.2. *Eigenvalues of K_n are $n-1$ and -1 (multiplicities 1 and $n-1$, respectively).*

Proof. Note that $A(K_n) = J - I$, and the rest is trivial. □

Another special graph that we want to analyze its eigenvalues if the graph C_n which is a cycle of length n.

Proposition 3.3. *Eigenvalues of C_n are $2, 2cos(2i\pi/n)$, $i = 1, \ldots, n-1$.*

Proof. Let W be the $n \times n$ matrix whose first row 0100000..., and each subsequent row equals the one above it, but shifted to the right by one position. So the second-to-last row is 00000001, and the last row is 10000.... Now, note that W^k is then the permutation matrix whose first row has a single 1, in position $k+1$ and the rest are shifted to the right. Crucially, observe that $A_{C_n} = W + W^{-1}$ which will let us determine the eigenvalues of C_n. Clearly, $W^n = I$ and each ω^ℓ is an eigenvalue (with eigenvector $1, \omega^\ell, \omega^{2\ell}, \ldots$). Now, for u_ℓ being the eigenvector of W with ω^ℓ, note that $A_{C_n} u_\ell = W u_\ell + W^{-1} u_\ell = (\omega^\ell + \omega^{-\ell}) u_\ell = 2cos(2\pi\ell/n)$. □

Exercise 3.4. *Eigenvalues of $K_{n,m}$.*

Exercise 3.5. Petersen graph *PET can be constructed by taking all* 2*-elements subsets of* $\{1, \ldots, 5\}$ *as vertices, and connecting two by an edge if they are disjoint. It is a* 3*-regular graph WHY? Compute the eigenvalues (with multiplicities) of PET.*

In the next lemma we show a connection between the largest eigenvalue of $A(G)$, for any G, and its average/maximal degree.

Lemma 3.6. *The largest eigenvalue* λ_1 *of* G *satisfies:*

$$\delta(G) \leq d(G) \leq \lambda_1 \leq \Delta(G).$$

In particular, if G *is* d*-regular, then* $\lambda_1 = d$.

Proof. Start with $\lambda_1 \leq \Delta$. Let $(x_v)_{v \in V(G)}$ be an eigenvector corresponds to λ_1 and let x_u be an entry with largest absolute value. Then, for $N(u) = \{v \in V(G) \mid vu \in E(G)\}$ we have:

$$\lambda_1 x_u = \sum_{v \in N(u)} x_v.$$

Therefore,

$$|\lambda_1| \cdot |x_u| \leq \sum |x_v| \leq \sum |x_u| \leq \Delta(G) \cdot |x_u|,$$

gives the desired ($\lambda_1 > 0$ as $tr(A) = 0$).

Now we show $d(G) \leq \lambda_1$. Consider $1^t A 1$. On one hand it equals $\sum d_v = 2|E(G)|$. On the other hand, if we take an orthonormal basis $v_1, \ldots, v_n$ of eigenvectors of A we get: $1 = \sum < 1, v_i > v_i$, $Av_i = \lambda_i v_i$, and $\sum < 1, v_i >^2 = \|1\|^2 = n$. Therefore,

$$1^t A 1 = \sum < 1, v_i > 1^t A v_i = \sum \lambda_i c_i^2 \leq \lambda_1 \sum c_i^2 = \lambda_1 n.$$

This gives the desired. □

Now we start connecting what we've seen to the title of this section, namely, walks in graphs.

Definition 3.7. *A* walk *of length* ℓ *in a graph* G *is a sequence* $v_1 \ldots v_{\ell+1}$ *where* $v_i v_{i+1}$ *is an edge (allowing repeated vertices and edges).*

How are walks in G related to $A(G)$? here we try to give an answer.

Theorem 3.8. *Let* G *be a graph an* $A(G)$ *be its adjacency matrix. For all* $\ell \geq 1$*, the* ij*th entry of* $A(G)^\ell$ *is the number of walks of length* ℓ *from* v_i *to* v_j.

Proof. Note that $A(G)^\ell_{ij} = \sum_{i_1, \ldots, i_{\ell-1}} a_{ii_1} a_{i_1 i_2} \ldots a_{i_{\ell-1} j}$. Therefore, it equals to the number of walks of length ℓ from i to j. □

Our goal is to use the above theorem to obtain explicit formula for the number of walks of length ℓ between two specified vertices. The formula will depend on the eigenvalues of $A(G)$.

Theorem 3.9. *Let* G *be a graph,* $A(G)$ *its adjacency matrix and let* $\lambda_1, \ldots, \lambda_n$ *be its eigenvalues. Then, there exist real numbers* $c_1, \ldots, c_n$ *such that for all* $\ell \geq 1$ *and* i, j *we have*

$$A^\ell_{ij} = c_1 \lambda_1^\ell + \ldots + c_n \lambda_n^\ell.$$

In fact, if $U^{-1} A U = D$*, then* $c_k = u_{ik} u_{jk}$.

Proof. Clearly, $U^{-1}A^{\ell}U = D(\lambda_i^{\ell})$. Therefore, $A^{\ell} = UDU^{-1}$ and the ijth entry is just

$$\sum_k u_{ik}\lambda_k^{\ell}u_{jk}$$

as desired. □

In order to obtain some value from the above result, we need to compute the eigenvalues of $A(G)$ and the matrix U.

Definition 3.10. *A* closed walk *is a walk which starts an ends at the same vertex.*

Note that the number of closed walks of length ℓ is $trace(A^{\ell})$ WHY?. Recall that for any squared matrix M, $tr(M) = \sum \lambda_i$ WHY? and therefore we have

Theorem 3.11. $\sum \lambda_i^{\ell} =$ *number of closed walks of length* ℓ.

As an immediate corollary we obtain an explicit formula for the number of closed walks of length ℓ in the complete graph.

Corollary 3.12. *The number of closed walks of length ℓ in K_n is $(n-1)^{\ell} + (n-1)(-1)^{\ell}$.*

What about walks which are not closed? that if, what about $i \neq j$? Note that

$$(J-I)^{\ell} = \sum_{k=0} \binom{\ell}{k}(-1)^{\ell-k}J^k.$$

Moreover, $J^k = n^{k-1}J$ and $J^0 = I$. Therefore,

$$(J-I)^{\ell} = \sum_{k=1} \binom{\ell}{k}(-1)^{k}p^{k-1} + (-1)^{\ell}I$$

and this can be easily computed using binomial formula.

All in all, we get

$$(A_G)_{ij}^{\ell} = \frac{1}{n}\left((n-1)^{\ell} - (-1)^{\ell}\right).$$

How can we use some knowledge on certain closed walks in order to upper bound the number of edges in our graph G? The following proposition upper bounds the number of edges possible in a C_4-free graph.

Proposition 3.13. *Suppose G is d-regular on n vertices and contains no C_4. What is the largest possible d?*

Proof. Consider $tr(A^4)$. On one hand it is at least d^4. On other hand it is at most $d^2n + \#C_4 = d^2n$. Therefore, $d \leq \sqrt{n}$. □

Exercise 3.14. *Show that x many C_4s force C_6. (Find the best x that you can).*

It is a well known (and quite easy to prove) fact from Graph Theory that a graph G is bipartite if and only if it contains no cycles of odd length. In the following theorem we give a spectral characterization of bipartite graphs.

Theorem 3.15. *A graph G is bipartite if and only if its spectrum is symmetric (that is, if λ is an eigenvalue, then also $-\lambda$, and with same multiplicity).*

Proof. Suppose G is bipartite with parts S and T of sizes s and t. That is,

$$A = \begin{pmatrix} 0_{ss} & B \\ B^t & 0_{tt} \end{pmatrix}.$$

If λ is an eigenvector, then $(\lambda v, \lambda u) = \lambda(v, u) = A(v, u)^t = (Bu, B^t v)$. So $Bu = \lambda v$ and $B^t v = \lambda u$. Then, $-\lambda$ is also eigenvalue with eigenvector $(v, -u)$.

Conversely, suppose the spectrum is symmetric. Then, for all k odd we have

$$\sum \lambda_i^k = 0.$$

Therefore, there are no closed walks of odd length, and also no cycles of odd length. This is equivalent to being bipartite. □

3.2 Independence number of a graph

Let $\alpha(G)$ denote the *independence number* of G. That is, the size of the largest independent set (a set that induces no edges) in G. Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ be its eigenvalues (note that $|\lambda_n|$ is not necessarily the smallest..), and we prove the following:

Theorem 3.16 (Hoffman's bound)**.** *Suppose G is d-regular. Then,*

$$\alpha(G) \leq \frac{n}{1 - \frac{d}{\lambda_n}}.$$

Proof. Let S be an independent set and let x_S be its characteristic vector. As S is independent, we obtain

$$x_S^t A_G x_S = 0.$$

Moreover, as λ_n is the minimal eigenvalue, we have $A - \lambda I$ has nonnegative eigenvalues. Also, $A1 = d1$ so

$$(A - \lambda I)1 = (d - \lambda)1 = \frac{d - \lambda}{n} J1.$$

Define a new matrix

$$M = A - \lambda_n I - \frac{d - \lambda}{n} J.$$

We just saw that $M1 = 0$ so 1 is a null-vector. Let v be an eigenvector orthogonal to 1 with eigenvalue μ. So $Jv = 0$. Then,

$$\mu v = Mv = (A - \lambda_n I)v,$$

hence μ is ev of $A - \lambda_n I$ so it is nonnegative.

Clearly, this way we can show that all ev of M are nonnegative. In particular, this implies that

$$x_S^t M x_S \geq 0.$$

On the other hand,

$$0 \leq x_S^t M x_s = x_S^t A x_S - \lambda x_S^t x_S - \frac{d - \lambda}{n} x_S^t J x_S = -\lambda |S| - \frac{d - \lambda}{n} |S|^2,$$

which gives

$$|S| \leq -\lambda \frac{n}{d-\lambda} = \frac{n}{1-\frac{d}{\lambda}}.$$

□

3.3 Friends and politicians

Imagine a group of n people, where every two distinct guys have exactly one common friend (assume that friendship is a symmetric relation...). It turns out, that there must be a guy (politician...) who is a friend with everybody else. The following beautiful proof of this theorem was discovered by Erdős, Renyí and Sós in the 60's, and is based on a linear algebra trick.

Theorem 3.17. *G is such that every two vertices have exactly one common neighbor. Then, there is a vertex which is adjacent to all other vertices.*

Proof. Suppose that G is a counter example, and we will try to get a contradiction in two steps.

1. First, we claim that G is regular. Let us start by showing that every two nonadjacent vertices have the same degree. Take u, v such that $uv \notin E(G)$. Let w be the common neighbor of u, v, and u' the common neighbor of of u and w, and v' the common neighbor of v and w (useful to draw a picture!). Now, each of the remaining neighbors of u is adjacent to exactly one neighbor of v, which can't be w or v' (as otherwise you'll find a C_4). Therefore, we have a bijection between $N(u) \setminus \{w, v'\}$ and $N(v) \setminus \{w, v'\}$, by mapping each vertex of the former into its unique neighbor in the latter (it's only a $1-1$ function, but by symmetry we obtain the other part). All in all we have $d(u) = d(v) = k$ (for some k). To conclude the regularity, we show that for any two vertices u, v, there exists a sequence $uv_1, \dots, v_s v$ (for some s), where every two consecutive vertices are non-adjacent in G. In other words, we want to show that the complement of G is a connected graph. Suppose it's not connected. In particular, one can partition the vertices into two sets $V(G) = A \cup B$ where all the pairs $ab \in A \times B$ are edges of G. It cannot be that $|A| = 1$ or $|B| = 1$ (as otherwise we are done with the theorem), and every other case will give us a C_4.

   Before we proceed into the next step, observe, crucially, that

   $$n = k^2 - k + 1.$$

   Indeed, since each of the k neighbors of a given v has $k-2$ neighbors which are not in the neighborhood, there are $k(k-2)$ non-neighbors of v. All in all $n = 1 + k + k(k-2) = k^2 - k + 1$ (v plus its neighborhood plus the non-neighbors).

2. Now, observe that $k \geq 3$. Consider A_G. Any row sums to k and every two rows have exactly one common column with both entries 1. A moment's thought now reveals that

   $$A^2 = (k-1)I + J.$$

   WHY? Therefore, *Spectrum*(A^2) : $k - 1 + n = k^2, k - 1$ (where that latter appears with multiplicity $n-1$). Therefore, *Spectrum*(A) is k (multiplicity 1) and $\pm\sqrt{k-1}$. Suppose r are

$+$ and s are $-$. As the trace equals 0 we obtain $k + r\sqrt{k-1} - s\sqrt{k-1} = 0$, and in particular $s \neq r$ and

$$\sqrt{k-1} = \frac{k}{s-r}.$$

Therefore, $\sqrt{k-1}$ is rational and must be an integer! (think about this old argument, it's a nice and simple riddle). All in all, let $t = \sqrt{k-1}$ (which is an integer!), we can rewrite the above equation as

$$t(s-r) = t^2 + 1,$$

and therefore t divides $t^2 + 1$ and therefore $t = 1$, leading to $k = 2$, which is a contradiction.

This completes the proof. □

Conjecture 3.18 (Kotzig's Conjecture). *Let $\ell > 2$. Then there are no finite graphs with the property that between any two vertices there is precisely one path of length ℓ.*

This is known up to $\ell = 33$. Any idea??

3.4 Turán's theorem

Given a graph H, one can ask the following natural question:

Question 3.19. *What is the maximum number of edges a graph G on n vertices can have without having a copy of H?*

Clearly, if $n \geq |V(H)|$ and G has $\binom{n}{2}$ edges, then G contains H. The question is whether one can get a non-trivial upper bound on this number, which from now on we denote by $ex(n, H)$, and refer to it as the *extremal number* of H.

As an example, consider the case where $H = C_3$ (that is, H is a triangle). It is obvious that $ex(n, H) \geq \frac{n^2}{4}$ (when n is even). Indeed, take the complete bipartite graph with parts of sizes exactly $n/2$. It contains $n^2/4$ edges an no odd cycles. To show that it is tight (that is, $e(G) \geq n^2/4 + 1$ implies existence of triangles) using a graph theoretical arguments, is left as an easy exercise (this is called Mantel's theorem).

Before we discuss the more general case, where $H = K_k$ for any $k \geq 3$, let us start with a warmup by giving a spectral proof for Mantel's theorem (again, there are much simpler ways to prove it).

Theorem 3.20 (Mantel's Theorem). *Let G be a triangle free graph on n vertices. Then, G contains at most $\lfloor n^2/4 \rfloor$ edges. Moreover, equality holds if and only if $G = K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$.*

Proof. Let $A(G)$ be the adjacency matrix of G, let $\lambda_1 \geq \ldots \geq \lambda_n$ be its eigenvalues, and let v be the eigenvector corresponding to λ_1. Note that for all $u \in V(G)$ we have

$$\lambda_1 v_u = \sum_{w \in N(u)} v_w. \tag{1}$$

Let x denote the vertex with maximum $|v_x|$, and WLOG we can assume that $v_x = 1$ (if there are multiple such vertices, then just pick one arbitrarily). Clearly, the above equality becomes

$$\lambda_1 = \sum_{y \in N(x)} v_y, \tag{2}$$

and note that this implies that $\lambda_1 \leq d(x)$. Now, by multiplying both sides of (2) by λ_1 and applying (1), we obtain

$$\lambda_1^2 = \sum_{y \in N(x)} \lambda_1 v_y = \sum_{y \in N(x)} \sum_{z \in N(y)} v_z = \sum_{y \in N(x)} \sum_{z \in N(x) \cap N(y)} v_z + \sum_{y} \sum_{z \in N(y) \setminus N(x)} v_z,$$

which is at most (recall that each vector entry is at most 1!)

$$2e(N(x)) + e(N(x), V(G) \setminus N(x)).$$

Now we can turn to the proof of Mantel's. Note that since G is triangle free we have $e(N(x)) = 0$. Moreover, as $\lambda_1 \geq \frac{2m}{n}$, using the above estimate we obtain

$$\frac{4m^2}{n^2} \leq \lambda_1^2 \leq e(N(x), V(G) \setminus N(x)) \leq \lceil \frac{n}{2} \rceil \lfloor \frac{n}{2} \rfloor.$$

Note that equality can occur only if $e(N(x), V(G) \setminus N(x)) = \lfloor n^2/4 \rfloor$. By rearranging one obtain the desired. □

Now, what happens if instead of taking $H = C_3$ we take H to be a complete graph on k vertices? (for graphs which are not complete the proofs are a bit more complicated..). Here one can also easily guess some example which, intuitively, sounds 'extremal'. That is, suppose that $H = K_{k+1}$ and that n is divisible by k. Take the complete k-partite graph with all parts of sizes exactly n/k. Clearly, such a graph has no copies of K_{k+1} and has

$$\left(\frac{k-1}{k}\right) \binom{n}{2}$$

many edges.

The following theorem due to Paul Turán from 1941 settles this problem and is a cornerstone in extremal graph theory. There are many proofs for this theorem (at least 4 that I'm aware of), and here we give a proof which is based on a spectral approach (most likely it is not the easiest one!).

Theorem 3.21. *If a graph on n vertices and m edges has a clique number ω (in particular, it has no clique of size $\omega + 1$), then*

$$m \leq \frac{1}{2} \left(\frac{\omega - 1}{\omega}\right) n^2.$$

Proof. Let λ_1 be its largest eigenvalue. We've seen that $\lambda_1 \geq d(G) = 2m/n$. So in order to complete the proof we need to show that

$$\lambda_1 \leq \frac{\omega - 1}{\omega} n.$$

Recall that for all vectors x we have

$$x^t A_G x = \sum_{uv \in E(G)} x_u x_v.$$

We first need the following claim:

Claim 3.22. *If $G = K_s$, then for all $x \in \mathbb{R}^s$ we have*

$$x^t A_{K_s} x \leq \frac{s-1}{s} \cdot (1^t x)^2.$$

Now, suppose that G has a complete subgraph $H = K_s$. Then, for any vector x with support H we have (by claim)

$$\frac{x^t A_G x}{(1^t x)^2} = \frac{x^t A_H x}{(1^t x)^2} \leq \frac{s-1}{s}.$$

We claim that this is true for all vectors.

Claim 3.23. *The maximum of* $\frac{x^t A_G x}{(1^t x)^2}$ *over all* x *is attained on some vector* y *with support*(y) = *a complete graph.*

Assuming this claim, let's finish the proof of the theorem. Take a unit eigenvector v of λ_1. Then,

$$\frac{v^t A_G v}{(1^t v)^2} = \frac{v^t \lambda_1 v}{(1^t v)^2} = \lambda_1 \frac{v^t v}{(1^t v)^2} \geq \lambda_1 / n,$$

using Cauchy-Schwarz (indeed, $(1^t v)^2 \leq n(v^t v) = n$). On the other hand, the clique number is ω, and therefore, for some x with *support*$(x) = H$, where H is a complete graph on $s \leq \omega$ vertices, we have

$$\frac{v^t A_G v}{(1^t v)^2} \leq \frac{x^t A_H x}{(1^t x)^2} \leq \frac{\omega - 1}{\omega},$$

as desired.

Now, let's prove the claims.

Proof of Claim 3.22. Note that

$$x^t A_{K_s} x = \sum_{u \neq v} x_u x_v = \sum_u x_u \left(\sum_{v \neq u} x_v \right) = \sum_u x_u \left(1^t x - x_u\right)$$

$$= 1^t x \left(\sum_u x_u \right) - \left(\sum_u x_u^2 \right) = (1^t x)^2 - x^t x. \tag{3}$$

Now, using Cauchy-Schwarz we obtain

$$(1^t x)^2 \leq s \cdot x^t x,$$

which yields

$$x^t x \geq \frac{1}{s}(1^t x)^2.$$

Therefore, we can upper bound (3) by

$$\frac{s-1}{s} \cdot (1^t x)^2$$

as desired. □

Proof of Claim 3.23. Let y be a vector maximizing $\frac{x^t A x}{(1^t x)^2}$, scaled so that $1^t y = 1$. We show that if $y_u, y_v \neq 0$ for some $uv \notin E(G)$, then one can make one of those into 0 without changing $y^t A y$. By repeating this argument, we end up with y supported on a clique and we're done. Sounds like a plan!

Suppose that $a_{uv} = 0$ for $y_u, y_v \neq 0$ (note that they are both positive! WHY?). Then, in $y^t A y$ the only summands corresponding to these entries are

$$a = \sum_{w \neq v} a_{uw} y_u y_w + \sum_{w \neq u} a_{vw} y_v y_w = y_u \sum_{w \neq v} + y_v \sum_{w \neq u}.$$

WLOG we can assume that $\sum_{w\neq v} \geq \sum_{w\neq u}$. Therefore,

$$y^tAy \leq y^tAy - a + (y_u + y_v) \cdot (\sum_{w\neq v} a_{uw}y_w) = z^tAz,$$

if we define $z_u = y_u + y_v$, $z_v = 0$, and $z_w = y_w$ for $w \neq u, v$. Therefore, z also maximizing. □

This completes the proof of the theorem. □

3.5 K_{10} is not a union of edge-disjoint Petersen's graphs

In PSET1 you had to calculate the eigenvalues of Petersen's graph. Assuming you've already done it, let's see how to use it in order to prove the following nice proposition.

Proposition 3.24. *K_{10} cannot be decomposed into 3 Petersen's graphs.*

Proof. As we've already checked, the eigenvalues of K_{10} are $9, -1, -1, \ldots, -1$ and the eigenvalues of PET are $3, 1, 1, 1, 1, 1, -2, -2, -2, -2$. Now, $A_{K_{10}} = J - I$ and assume that $A_G = A + B + C$, where each of the summands is the adjacency matrix of Petrsen's graph (with some permutation applied to it). Let V_A and V_B be the eigenspaces corresponding to the eigenvalue 1 of A and B. They are both orthogonal to the vector 1. Moreover, each of them is of dimension 5 and therefore $dim(V_A \cap V_B) \geq 1$. Therefore, there exists $v \in V_A \cap V_B$. Recall that $1 \cdot v = 0$. So, $Cv = (J - I - A - B)v = -3v$ and hence (-3) is an eigenvalue of C, contradiction. □

3.6 Maxcut and another proof for Hoffman's bound

In this section we show two almost immediate corollaries from the following inequality.

Lemma 3.25. *Suppose that G is d-regular graph with $d = \lambda_1 \geq \ldots \geq \lambda_n$. Then, for $x_1, \ldots, x_n$ we have*

$$\sum_{i<j, ij\in E(G)} (x_i - x_j)^2 \leq (d - \lambda_n) \sum_{i=1}^{n} x_i^2.$$

Furthermore, if $\sum x_i = 0$, then

$$(d - \lambda_2) \sum x_i^2 \leq \sum_{i<j, ij\in E(G)} (x_i - x_j)^2.$$

Proof. Note that

$$\sum_{i<j, ij\in E(G)} (x_i - x_j)^2 = d\sum_i x_i^2 - 2 \sum_{i<j, ij\in E(G)} x_ix_j = d\sum x_i^2 - \sum_{i,j} a_{ij}x_ix_j. \tag{4}$$

A useful observation is that $x^tAx = \sum_{ij\in E(G)} x_ix_j = 2\sum_{i<j, ij\in E(G)} x_ix_j$, and $x^tAx \geq \lambda_n x^tx$. Now, let $x = \sum \alpha_i v_i$ where v_i is some orthonormal basis with $v_1 = \frac{1}{\sqrt{n}}1$ (in particular, $\alpha_i = x \cdot v_i$ for all i). Then,

$$\sum x_i^2 = x \cdot x = \sum \alpha_i^2.$$

Combining the above estimates we obtain

$$\sum_{i<j, ij\in E(G)} (x_i - x_j)^2 = d\sum x_i^2 - x^t Ax \leq d\sum x_i^2 - \lambda_n \sum x_i^2,$$

as desired.

For the second part, note that $\sum x_i = 0$ is equivalent to $\alpha_1 = x^t v_1 = 0$. Now

$$\sum a_{ij}x_i x_j = x^t Ax = x^t \sum_{i=2}^{n} \alpha_i \lambda_i v_i = \sum_{i=2}^{n} \lambda_i \alpha_i^2 \leq \lambda_2 x^t x.$$

Plugging into (4) we obtain the desired. □

First, we show how to derive Hoffman's bound easily from Lemma 3.25.

Corollary 3.26 (Hoffman's bound). $\alpha(G) \leq \frac{1}{1-d/\lambda_n} \cdot n.$

Proof. Let S be an independent set of size s. Define: $x_i = n - s$ is $i \in S$, and $x_i = -s$ otherwise. Note that $e_G(S, S^c) = d|S|$, as S is independent. Moreover, the x_i's have been defined in such a way that

$$\sum_{ij\in E(G)} (x_i - x_j)^2 = n^2 e(S, S^c).$$

All in all, together with the previous lemma, we obtain

$$n^2 ds \leq (d - \lambda_n)\sum x_i^2 = (d - \lambda_n)(s(n-s)^2 + (n-s)s^2) = (d - \lambda_n)sn(n-s),$$

which yields

$$nd \leq (d - \lambda_n)(n - s)$$

which is equivalent to

$$s \leq -\lambda_n n/(d - \lambda_n).$$

□

Definition 3.27. *$e(A, B)$ = the number of edges with one endpoint in A and the other in B. The* max cut *of a graph G is defined as*

$$\max_{A\cap B=\emptyset, A\cup B=V(G)} e(A, B).$$

As a second application, we show how to upper bound the size of the maxcut using the smallest eigenvalue of $A(G)$.

Corollary 3.28. *For any d-regular graph G, the max cut is at most*

$$\frac{n}{4}(d - \lambda_n) = \frac{e(G)}{2} - \frac{n\lambda_n}{4}.$$

Proof. Let $A \cup B = V(G)$ be a partition of $V(G)$ that gives a max cut, and let $|A| = a$. Define $x_i = n - a$ is $i \in A$ and $x_i = -a$ otherwise. As in previous corollary, $\sum (x_i - x_j)^2 = n^2 e(A, B)$. Moreover, as $a(n-a) \leq n^2/4$, we have

$$\sum x_i^2 = a(n-a)^2 + (n-a)a^2 = a(n-a)n \leq \frac{n^3}{4}.$$

Now, using lemma 3.25 we obtain

$$n^2 e(A, B) \leq (d - \lambda_n)\frac{n^3}{4}.$$

□

3.7 Expander mixing lemma

Now we show the power of the second largest eigenvalue (in absolute value). Recall that

$$\lambda(G) = \max\{|\lambda_2|, |\lambda_n|\}.$$

Lemma 3.29 (Expander mixing lemma). *Let G be a d-regular graph on n vertices, and let $\lambda := \lambda(G)$. Then, for all $S, T \subseteq V(G)$ we have*

$$\left| e(S,T) - \frac{d|S||T|}{n} \right| \leq \lambda \sqrt{|S|\left(1 - \frac{|S|}{n}\right)|T|\left(1 - \frac{|T|}{n}\right)}.$$

Remark 3.30. *Edges in the intersection of S and T are being counted twice.*

Proof. Let $S, T \subseteq V(G)$ and let 1_S and 1_T be the corresponding characteristic functions. Expand these vectors in an orthonormal basis of eigenvectors $v_1, \ldots, v_n$ to obtain

$$1_S = \sum \alpha_i v_i, \text{ and } 1_T = \sum \beta_i v_i.$$

Then,

$$e(S,T) = 1_S^t A 1_T = \sum \lambda_i \alpha_i \beta_i.$$

Since $\alpha_1 = 1_S^t \frac{1}{\sqrt{n}} = \frac{|S|}{\sqrt{n}}$, $\beta_1 = \frac{|T|}{\sqrt{n}}$, and $\lambda_1 = d$, we have

$$e(S,T) = \frac{d|S||T|}{n} + \sum_{i=2}^{n} \lambda_i \alpha_i \beta_i.$$

Thus,

$$\left| e(S,T) - \frac{d|S||T|}{n} \right| \leq \lambda \sum_{i=2}^{n} |\alpha_i \beta_i|.$$

By applying Cauchy-Schwarz, we obtain

$$\sum_{i=2}^{n} |\alpha_i \beta_i| \leq \left(\sum_{i=2}^{n} \alpha_i^2\right)^{1/2} \left(\sum_{i=2}^{n} \beta_i^2\right)^{1/2} = \sqrt{\left(|S| - \alpha_1^2\right)\left(|T| - \beta_1^2\right)}$$

$$= \sqrt{|S|\left(1 - \frac{|S|}{n}\right)|T|\left(1 - \frac{|T|}{n}\right)}.$$

Plugging it into the above inequality we obtain the desired. □

Definition 3.31. *A d-regular graph G on n vertices is said to be an (n, d, λ)-graph if $\lambda(G) \leq \lambda$.*

Exercise 3.32. *Show that for an (n, d, λ)-graph G we have*

$$\alpha(G) \leq \lambda n.$$

Question 3.33. *How small can $\lambda(G)$ be?*

Theorem 3.34. *Let G be a d-regular graph on n vertices, where $d \leq (1 - \varepsilon)n$. Then,*

$$\lambda(G) = \Omega(\sqrt{d}).$$

Proof. Note that

$$2e(G) = dn = trA^2 \leq d^2 + (n-1)\lambda^2.$$

Therefore,

$$\lambda^2 \geq \frac{dn - d^2}{n-1} = \Omega(d),$$

as desired. □

Remark 3.35. *Alon and Boppana observed that in fact $\lambda \geq 2\sqrt{d-1} - o(1)$. Therefore, a graph is said to be 'perfect expander' (or* Ramanujan *graph) if $\lambda \leq 2\sqrt{d-1}$. There is a fascinating theory about ramanujan graphs! For example, it is known that for $d = o(\sqrt{n})$, almost every d-regular graph on n vertices satisfies $\lambda(G) \leq 2\sqrt{d-1} + \varepsilon$ (Friedman).*

As an application we present the following neat argument due to Krivelevich and Sudakov.

Definition 3.36. *A graph G is said to be k-*edge-connected *if and only if it cannot be made disconnected by deleting at most $k - 1$ edges. In particular, it is equivalent to: in every cut there are at least k edges WHY?.*

We will also need the following theorem due to Tutte.

Theorem 3.37 (Tutte's Theorem). *A graph G has a perfect matching if and only if for every subset $S \subseteq V(G)$, the number of connected components of odd size in $G[V \setminus S]$ is at most $|S|$.*

Now we are ready to state and prove the theorem that we wanted.

Theorem 3.38. *Let G be an (n, d, λ)-graph with $\lambda \leq d - 2$. Then,*

1. *G is d-edge-connected, and*
2. *G contains a perfect matching.*

Proof. For 1 we wish to show that for all $S \subseteq V(G)$ we have $e(S, S^c) \geq d$. Note that as in every cut either we have $|S| \leq n/2$ or $n - |S| \leq n/2$, and therefore it is enough to consider only sets S of size at most $n/2$. Now, if $|S| \leq d$, then there is nothing to prove WHY?. Therefore, assume that $d < |S| \leq n/2$. Fix such a subset S, and by the expander mixing lemma we obtain

$$e(S, S^c) \geq \frac{|S|(n - |S|)d}{n} - \frac{|S|(n - |S|)\lambda}{n} \geq (d - \lambda)|S|/2 > d$$

as desired.

For part 2, let us act as follows. Let $S \subseteq V(G)$ (note that we may assume $|S| \geq 1$ as we must take n to be even). Let $C_1, \dots, C_t$ be all the connected components of $G[V \setminus S]$. Our goal is to show that the number of indices i for which $|C_i|$ =odd is at most $|S|$, and then we complete the proof by using Tutte's Theorem. In fact, we show something stronger, namely, we show that $t \leq |S|$. To this end, note that as G is d-edge-connected (by 1.), every connected component must send at least d edges to S WHY?. On the other hand, the total number of edges touching S is at most $d|S|$ (as G is d-regular). Combining these two bounds we obtain

$$dt \leq$$

□

3.8 Low-rank approximation and spectral partitioning

In this section we discuss 'low rank approximation' of a matrix and the problem of recovering a 'planted' partitioning in random graphs, using spectral methods. Let us start by defining how would we like to measure distance between matrices. We usually do it either by the *operator norm* $\|A-B\|$ or the Frobenius norm $\|A-B\|_F$, where

$$\|M\| = \max_x \frac{\|Mx\|}{\|x\|} \text{ and } \|M\|_F = \sqrt{\sum_{i,j} M_{ij}^2}.$$

(observe that $\|M\|_F = \sqrt{tr(M^tM)}$.)

Now, recall from linear algebra that if A is symmetric with eigenvalues $\lambda_1 \geq \ldots \geq \lambda_n$ and a corresponding orthonormal basis of column vectors $v_1, \dots, v_n$, then

$$A = \sum \lambda_i v_i v_i^t.$$

Indeed, $U = [v_1, \dots, v_n]$ satisfies

$$U^{-1}AU = D(\lambda_1, \dots, \lambda_n),$$

and therefore

$$A = UDU^t.$$

It follows that

$$A = \sum \lambda_i v_i v_i^t.$$

Using the Min-max theorem, one can show that for every k, the best approximation of A by a rank-k matrix is given by summing the terms $\lambda_i v_i v_i^t$ over the largest k values of λ_i in absolute value, and this holds for both norms. If the difference is small, it explains why the largest k eigenvalues of A should provide a lot of information about A.

Let us illustrate why approximations are useful. Consider the problem of a *planted partition.* That is, suppose that $S \subseteq V(G)$ which is 'planted' in some sense. Our goal is to recover it. Without loss of generality, one can assume that S is the set of the first $|S|$ vertices. Therefore, $A(G)$ can be written as

$$A(G) = \begin{pmatrix} A(S) & 0 \\ 0 & A(V \setminus S) \end{pmatrix} + \begin{pmatrix} 0 & A(S, V \setminus S) \\ A(V \setminus S, S) & 0 \end{pmatrix}.$$

The set S can be discovered from examining the eigenvectors of the left-hand matrix: it has an eigenvector which is positive on all the entries of S and 0 otherwise, and an eigenvector which is positive on $V \setminus S$ and 0 elsewhere. Therefore, if the right-hand matrix is 'small' in some sense, then we expect similar eigenvectors to remain, and therefore the partition is recovered by finding this eigenvector, and partitioning the vertices according to the sign in the corresponding entry.

Let's try to make things a bit more rigorous. The simplest model of this form is the following: partition $[n]$ into two sets of the exact same size, $X \cup Y$. Then, choose probabilities $p > q$, and place edges between vertices according to the following rule: if $uv \subseteq X$ or $uv \subseteq Y$, then add uv with probability p. Otherwise, probability q. All choices are being made independently at random.

The expected number of edges between X and Y is $q|X||Y|$. If p is sufficiently larger than q, then every other partitioning will have more crossing edges. On the other hand, if p is too close to q, then $X \cup Y$ doesn't necessarily have the smallest number of crossing edges. The main question is how to recover the partitioning in an efficient way (that is, without checking all the 2^n possible partitions)? or even a weaker question: can you recover a (say) 2/3 fraction of the partitioning?

Exercise 3.39. *assume that $p = 1/2$ and $q = 1/3$. Can you think about an easy way to recover the partitioning?*

Here we present a general strategy that works for various ranges of p, q, but it will be simple for us just to work with specific values, even if are easy to handle with different methods. Let us consider the case $p = 1/2$ and $q = p - \frac{100}{\sqrt{n}}$. Note that if $q = p - \varepsilon/\sqrt{n}$ for very small ε, then basically there is no chance to recover the partition (do you see why?).

Our main goal is to show that the partition can mostly be recovered from the eigenvector of the second eigenvalue of the adjacency matrix of our graph. The idea was introduced by McSherry in 2001. For more details about recent developments and history of the problem, just google 'Stochastic Block Model' (there is a very nice survey of E. Abbe about the problem). The main idea is to consider the adjacency matrix as a perturbation of one *ideal probability matrix.* Apparently, in this ideal matrix (to be defined bellow) the second eigenvector provides a partitioning into two blocks (according to the sign of the entries). McSherry showed that the difference between the ideal matrix and the actual matrix is 'small', and therefore, using some concentration results, he could show that the second largest eigenvector of the actual matrix is 'more or less' the same like the one of the ideal matrix. Therefore, one can recover most of the partition.

From now on we can assume that $S = \{1, \ldots, n/2\}$, and define a matrix

$$M = \begin{pmatrix} pJ_{n/2} & qJ_{n/2} \\ qJ_{n/2} & pJ_{n/2} \end{pmatrix}.$$

M is our 'ideal probability matrix'. Note that the adjacency matrix of the planted partition graph is obtained according to the probabilities in M (minus diagonal of course). Basically, the algorithm to recover S goes as follows: compute v, the eigenvector of λ_2. Then, set $X = \{x \mid v_x < 0\}$. As we show, whp X is mostly one set of the partition.

Let us first consider the eigenvectors of M. Clearly $\mathbb{1}$ is an eigenvector with eigenvalue $\frac{n}{2}(p+q)$. The second eigenvector of M has two values, one on S and one on S^c. That is, we can take $w_i = \frac{1}{\sqrt{n}}$ if $i \in S$ and $w_i = -\frac{1}{\sqrt{n}}$ otherwise. Clearly, $Mw = \frac{n}{2}(p-q)w$. As M has rank 2, all the other eigenvalues are 0. Now, let us define a matrix

$$B = A(G) + pI.$$

Note that A and B have the same eigenvectors so it doesn't change our analysis, but intuitively, B is a bit 'closer' to M. Our goal is to measure the difference $R = B - M$ and to obtain some useful knowledge out of it.

Note that for xy in the same part, we have $\Pr[R_{xy} = 1 - p] = p$ and $\Pr[R_{xy} = -p] = 1 - p$.

For xy in different components we can replace p by q in the above line. At this point we can already use some concentration bounds to show that $\|R\|$ is small. For example, by a simple application of Talagrand's inequality, Alon, Krivelevich and Vu showed that (in particular) we can assume that $\|R\| \leq 3\sqrt{pn}$ (they have a much more general statement about concentration of all eigenvalues and for a wide range of p). As here we are more interested in the linear algebra part rather than the probability part, then from now on, let's just assume that $\|R\|$ is small.

Let $\alpha_1 \geq \ldots \geq \alpha_n$ be the eigenvalues of B and let $\mu_1 > \mu_2 > 0 = \ldots = \mu_n$ be the eigenvalues of M. It follows from the Min-max theorem (more or less trivially...) that $|\alpha_i - \mu_i| \leq \|R\|$ for all i. In particular, if

$$\|R\| < \frac{n}{4}(p-q),$$

then

$$\frac{n}{4}(p-q) < \alpha_2 < \frac{3n}{4}(p-q).$$

Moreover, as $q > p/3$, we have

$$\alpha_1 > \mu_1 - \|R\| > \frac{n}{2}(p+q) - \frac{n}{4}(p-q) \geq \frac{3n}{4}(p-q).$$

Therefore, we can view α_2 as a perturbation of μ_2. The main question is whether we can see v (the eigenvector of α_2) as a perturbation of w (the one we've already discussed)?

To address this question we are going to use the following theorem that says that v will be close to w, in angle, if the norm of R is significantly less than the distance between μ_2 and the other eigenvalues of M. That is, the eigenvector doesn't move much if it corresponds to a 'well separated' eigenvalue.

Theorem 3.40 (Davis-Kahan)**.** *Let B, M be symmetric matrices. Let $R = M - B$. Let α_i be the eigenvalues of B with eigenvectors v_i, and μ_i of M with corresponding eigenvectors w_i. Let θ_i be the angle between v_i and w_i. Then*

$$\sin 2\theta_i \leq \frac{2\|R\|}{\min_{j \neq i} |\mu_i - \mu_j|}.$$

For simplicity, we'll only prove a bit weaker version of the theorem soon (and also only use the weaker version). That is, we prove it with 2θ replaced by θ.

How to continue? Let

$$\delta = v_2 - w_2.$$

For every vertex i that is misclassified by v_2, we have $|\delta(i)| \geq \frac{1}{\sqrt{n}}$ (otherwise it wouldn't change a sign). So, if v_2 misclassifies k vertices, then

$$\|\delta\| \geq \sqrt{\frac{k}{n}}.$$

As w_2 and v_2 are unit vectors, we may apply the crude inequality

$$\|\delta\| \leq \sqrt{2} \sin \theta_2.$$

(to see this inequality, just expand $\|\delta\|^2 = \|w\|^2 + \|v\|^2 - 2w \cdot v = 2 - 2\cos\theta_2$. The right hand side is clearly upper bounded by $2(1 - \cos^2\theta_2) = 2\sin^2\theta_2$.)

To combine all, recall that $q > p/3$ (so α_2 is perturbation of μ_2) and compute

$$\min_{j \neq 2} |\mu_2 - \mu_j| = \frac{n}{2}(p - q).$$

Recall that $\|R\| \leq 3\sqrt{pn}$ so we find (by the weaker theorem)

$$\sin\theta_2 \leq \frac{2 \cdot 3\sqrt{pn}}{\frac{n}{2}(p-q)} = \frac{12\sqrt{p}}{\sqrt{n}(p-q)}.$$

Therefore, the number of misclassified vertices satisfies

$$\sqrt{\frac{k}{n}} \leq \frac{\sqrt{2} \cdot 12\sqrt{p}}{\sqrt{n}(p-q)},$$

which implies $k \leq \frac{288p}{(p-q)^2}$.

In particular, if p and q are both constants, we expect to misclassify at most constant number of vertices. For our choice of parameters, we get constant fraction of the vertices (and we can control it by weaker restriction on $p - q$).

It thus remain to prove some version of the Davis-Kahan Theorem.

Proof. By considering the matrices $M - \mu_i I$ and $B - \mu_i I$ instead of M and B, we can assume that $\mu_i = 0$. The theorem is vacuous if μ_i has multiplicity more than 1 so we can assume the multiplicity is 1 and that w_i is a unit eigenvector in the nullspace of M. Note that our assumption $\mu = 0$ also gives $|\alpha_i| \leq \|R\|$.

Now, expand v_i in the eigenvector-basis of M as

$$v_i = \sum c_j w_j,$$

where $c_j = w_j^t v_i$. For

$$\delta = \min_{j \neq i} |\mu_j|,$$

we can compute

$$\|Mv_i\|^2 = \sum_j c_j^2 \mu_j^2 \geq \sum_{j \neq i} c_j^2 \delta^2 = \delta^2(1 - c_i^2) = \delta^2 \sin^2\theta_i.$$

On the other hand,

$$\|Mv_i\| \leq \|Bv_i\| + \|Rv_i\| = \alpha_i + \|Rv_i\| \leq 2\|R\|.$$

Therefore, we obtain

$$\sin\theta_i \leq \frac{2\|R\|}{\delta},$$

as desired. □

3.9 Random walks on graphs

Let G be a graph on n vertices, and consider a *random walk* on it. That is, start with some vertex u^0, chosen according to some distribution. Now, in each Step i, where you are at vertex u^i, choose one of its neighbors, u^{i+1}, uniformly at random, and move towards it. The basic question to be asked is the following:

Question 3.41. *What is the probability to be at a given vertex after ℓ steps?*

For simplicity, throughout our discussion, we assume that our graph G is d-regular and simple (both requirements can be dropped, for more details just read Chapter 3, or follow the hints at the end of this section to rewrite the proofs). Consider the matrix $M = \frac{1}{d} \cdot A(G)$, and observe that one can view each entry M_{uv} of M as 'the probability that in the next step our walk will end at v, given that it currently ends at u'. Clearly, M^{ℓ}_{uv} is just the probability that one ends up at vertex v after ℓ steps, assuming that we started at vertex u. Now, suppose that the starting vertex is not specified, but given as a probability distribution. Let $P = (\rho(v_1), \dots, \rho(v_n))$ be the corresponding vector of this distribution. Again, it is quite obvious to see that for $\sigma^{(\ell)} := (PM^{\ell})$, $\sigma^{(\ell)}_v$ is the probability to land at v after ℓ steps, given the probability distribution P. Therefore, as it should be clear by now (see Theorem 3.9), if we can compute the eigenvalues and eigenvectors of M, then we can actually compute all the crucial probabilities.

As a simple example, imagine that one starts with $P = \pi$ being the uniform distribution (that is, all entries are $1/n$). It is quite obvious to see that in this case, for all ℓ we have $(\pi M^{\ell})_v = 1/n$. Indeed, $\pi M = \pi$, so by induction we obtain the desired. We say that π is the *stationary distribution* of the random walk (what if G was not d regular?). The next theorem shows that no matter how do we pick P, the process will always converge to π.

Theorem 3.42. *If G is a connected nonbipartite graph, then $\sigma^{(\ell)} \to \pi$ for every P.*

(can you see why this theorem is not true for bipartite?).

Before proving the theorem, let us discuss the statement a bit. Suppose that we want to sample an object uniformly at random from a large collection of objects. Apparently, this is not as simple as it sounds... One way of doing it, is to construct a connected nonbipartite regular graph on this set, and start a random walk on this graph. By the above theorem, after sufficiently many steps, we get an object which is essentially uniformly distributed. Clearly, in order to make it efficient, it is important to know what 'sufficiently many steps' is, or in other words, what is the rate of convergence to π? this is usually referred to as the *mixing rate*. As it turns out (hopefully this is no longer a surprise for you by now..), this relates to the eigenvalue gap.

Theorem 3.43. *Let G be a connected, non-bipartite, d-regular graph on n vertices, and let $M = \frac{1}{d} \cdot A(G)$. Let $\lambda_1 \geq \ldots \geq \lambda_n$ be the eigenvalues of M and let $\lambda := \max\{|\lambda_2|, |\lambda_n|\}$. Then, for every starting vertex u, any vertex v, and any $\ell \geq 0$, we have*

$$\left| \Pr[u^{\ell} = v] - \frac{1}{n} \right| \leq \lambda^{\ell}.$$

Exercise 3.44. *While reading the proof, try to think how to state and prove the analogous theorem for the non-regular case!*

Proof. As M is symmetric (what if not?), one can write

$$M = \sum \lambda_k v_k v_k^t,$$

where the v_is form an orthonormal basis of eigenvectors. Clearly, one can take $v_1 = \frac{1}{\sqrt{n}} \cdot \mathbf{1}$. It follows that

$$\begin{aligned}\Pr[u^\ell = v] &= (M^\ell)_{uv} = e_u^t M^\ell e_v = \sum \lambda_k^\ell e_u^t (v_k v_k^t) e_v \\ &= \sum_{k=1}^n \lambda_k^\ell v_{ku} v_{kv} = \frac{1}{n} + \sum_{k=2}^n \lambda_k^\ell v_{ku} v_{kv}.\end{aligned}$$

It thus remains to bound the absolute value of the second summand:

$$\left|\sum_{k=2}^n \lambda_k^\ell v_{ku} v_{kv}\right| \le \lambda^\ell \sum_{k=1}^n |v_{ku} v_{kv}| \le \lambda^\ell \left(\sum_{k=1}^n v_{ku}^2\right)^{1/2} \left(\sum_{k=1}^n v_{kv}^2\right)^{1/2} = \lambda^\ell.$$

This completes the proof. □

Clearly, as smaller λ is, the better rate of convergence we obtain!

Another natural problem arises is about the first time 'visiting' a particular vertex. That is, suppose that G is a connected graph and let u, v be two of its vertices. Define $H(u, v)$ to be the expected number of steps needed for a random walk, starting at u, to 'hit' v at the first time. This is called the *hitting time*. Let p_m be the probability that we hit v for the first time after m steps, then

$$H(u, v) = \sum m p_m.$$

Let us consider some easy example. Suppose G is the path of length two uwv. To compute $H(u, v)$ is quite simple: after one step we are at w. Then, with probability $1/2$ we move to v or u. Therefore,

$$H(u, v) = \frac{1}{2} \cdot 2 + \frac{1}{2} \cdot (2 + H(u, v)).$$

Solving it gives $H(u, v) = 4$.

The question we are interested at is how to get such a formula (for more complicated graphs of course..) using linear algebra? Before answering it we need some preparation.

A matrix B is called *nonnegative* if all its entries are nonnegative. We say that B is *irreducible* if it is not the 1×1 matrix $[0]$ and if there is no permutaion matrix such that $PBP^{-1} = \begin{pmatrix} C & D \\ 0 & E \end{pmatrix}$, where C and E are square matrices of size greater than 0. In terms of an adjacency matrix of a graph, $A(G)$ is irreducible if and only if G is connected and is not an isolated vertex (WHY?). The following version of Perron-Frobenius theorem (from linear algebra) will be crucial in order to complete the proof of the theorem that appears bellow (I'll be a bit hand-wavy here, but a detailed explanation appears in Chapter 3 of Stanley's book).

Theorem 3.45 (Perron-Frobenius)**.** *Let B be a nonnegative, irreducible, square matrix. If ρ is the maximum absolute value of the eigenvalues of B, then $\rho > 0$, and there is an eigenvalue equal to ρ. Moreover, there is an eigenvector for ρ all of whose entries are positive.*

Now, let $M = \frac{1}{d}A(G)$. Let $M[v]$ denote M with the row/column corresponding to v deleted. Let $T[v]$ be the column vector of length $n-1$, indexed by $w \neq v$, where in each entry w we have $T[v]_w = 1/d$ if $vw \in E(G)$ and 0 otherwise. The following theorem gives an explicit formula for $H(u,v)$.

Theorem 3.46. *$I_{n-1} - M[v]$ is invertible, and*

$$H(u,v) = \left((I_{n-1} - M[v])^{-2} T[v]\right)_u .$$

Proof. The probability that when we take s steps from u, we never reach v and end up at some vertex w is $(M[v]^s)_{uw}$ WHY?. The probability that once we reach w, the next step is to v, is $\frac{1_{wv \in E(G)}}{d}$. Therefore, by definition of expectation we have

$$H(u,v) = \sum_{w \neq v} \sum_{s \geq 0} (s+1) \frac{1_{wv \in E(G)}}{d} \cdot (M[v]^s)_{uw}.$$

Using the equality

$$\sum (s+1)x^s = (1-x)^{-2}$$

in a quite suspicious way, we obtain

$$H(u,v) = \sum_{w \neq v} (I - M[v])^{-2}_{uw} \frac{1_{wv \in E(G)}}{d} = \left((I - M[v])^{-2} T[v]\right)_u$$

as desired.

Let's convince ourselves that our last move was legal. We had a matrix (say) B and a term of the form $\sum_s (s+1)(B^s)_{uw}$. Note that $\sum_s (s+1)B^s = C$ if and only if $\sum_n (s+1)(B^s)_{uw} = C_{uw}$ for all u, w.

Now, observe that for all m we have

$$(I - B)^2 (I + 2B + 3B^2 + \ldots + mB^{m-1}) = I - (m+1)B^m + mB^{m+1}.$$

Suppose that B is diagonalizable and that all its eigenvalues are smaller than 1 in absolute value. Then, by Theorem 3.9 we obtain

$$(B^m)_{uv} = c_1 \lambda_1^m + \ldots + c_r \lambda_r^m,$$

where the c_i's are some complex numbers independent of m. Therefore, from the above equality, we see that as m tends to infinity, it tends to I, and therefore $(I-B)^{-2}$ tends to $\sum (s+1)B^s$.

It thus remain to show that $M[v]$ is diagonalizable and all its eigenvalues are smaller than 1 in absolute value. Diagonalizability is trivial as $M[v]$ is symmetric. For the eigenvalues, we need some computational trick. Let N be a submatrix of $M[v]$ consisting of a connected component of the graph $G - v$. Clearly (WHY?) $M[v]$ has the same eigenvalues as in the union of all N's. By Perron-Frobenious, there is a positive eigenvector u of the largest eigenvalue in absolute value of N, say call it λ. Now, expand in two ways the expression

$$1^t M[v] u = \lambda \sum u_i = \sum_i \text{ sum of column } i \ u_i,$$

and observe that all summands are at most u_i, and at least one of the summands is strictly smaller than u_i. This completes the proof. □

3.10 The Matrix Tree Theorem

In this section we want to prove Cayley's formula for the number of labeled, spanning trees in the complete graph K_n. In fact, we prove something much stronger. Let G be a graph on n vertices, and let $t(G)$ denote the number of spanning trees in G. Consider the *incidence matrix* B of G. That is, B is an n by $e(G)$ matrix, with $B_{ij} = 1$ if and only if $v_i \in e_j$. EXAMPLE?

Now, replace one of the two 1's in each column by a -1 (arbitrarily), to obtain a matrix C with all column sums 0, and define $M = CC^t$. That is, M is an $n \times n$ symmetric matrix, which is

$$D(d(1), \ldots, d(n)) - A(G).$$

(this is sometimes called the Laplacian matrix of a graph G)

The following theorem is due to Kirchoff and is known as the *matrix-tree theorem.*

Theorem 3.47. *The number of spanning trees in G is*

$$t(G) = \det M_{ii},$$

where M_{ii} is the ith minor of M (the formula holds for all i).

A key ingredient in the proof is the following theorem of Binet and Cauchy, that we will prove later in this section.

Theorem 3.48. *If P is an $r \times s$ matrix and Q is an $s \times r$ matrix with $r \leq s$, then*

$$\det PQ = \sum_Z \det P_Z \det Q_Z,$$

where P_Z is the $r \times r$ submatrix of P with column set Z, and Q_Z is the $r \times r$ submatrix of Q with the corresponding rows Z, and the sum is over all r-subsets Z of $[s]$.

Let us now present the proof of the Matrix Tree Theorem.

Proof of Theorem 3.47. Note that C has at least $n-1$ columns, because G is connected (and therefore has at least $n-1$ edges). This means that we can apply Theorem 3.48 to M_{ii} and get

$$\det M_{ii} = \sum_N \det N \det N^t = \sum (\det N)^2,$$

where N runs over all $(n-1) \times (n-1)$ submatrices of $C \setminus \{$ row i $\}$. The $n-1$ columns of N correspond to a subgraph of G with $n-1$ edges on n vertices. Therefore, it remains to show that

$$\det N = \pm 1 \text{ if these edges span a tree },$$

and

$$\det N = 0 \text{ otherwise }.$$

Suppose the $n-1$ edges do not span a tree. In particular, a graph on n vertices with $n-1$ edges which is not a tree is not connected. Therefore, one can find a connected component that does not contain the vertex i. Clearly, the corresponding rows of the matrix sum to 0 and therefore dependent. This shows that $\det(N) = 0$ in this case.

Now suppose that the columns of N span a tree. Then, there is a vertex $j_1 \neq i$ of degree 1. Let e_1 be the incident edge. Deleting j_1, e_1 we obtain a tree with $n-2$ edges. Again, one can find such a vertex $j_2 \neq i$ and an edge e_2. Continue until $j_1, \ldots, j_{n-1}$ and $e_1, \ldots, e_{n-1}$ with $j_i \in e_i$ are determined. Now, permute the rows and columns to bring j_k into the kth row and e_k into the kth column. Since by construction, $j_k \notin e_\ell$ for all $k < \ell$, we see that the new matrix N' is lowertriangular with all elements on the main diagonal equal to ± 1. Therefore, $\det N' = \pm \det N = \pm 1$. This completes the proof. □

3.11 Binet-Cauchy

Recall that

$$\det M = \sum_{\sigma \in S_n} sign(\sigma) \prod_{i=1}^{n} m_{i\sigma(i)}.$$

Let us now define a graph where the vertices $a_1, \ldots, a_n$ stand for rows and $b_1, \ldots, b_n$ for columns of M. For each pair i, j, draw an arrow from a_i to b_j and assign it with a weight m_{ij}. In terms of graphs we have the following interpretation:

- each entry m_{ij} corresponds to the weight of the unique directed path from a_i to b_j.
- the determinant is the weighted sum over all *vertex-disjoint path systems* from $A = \{a_i\}$ to $B = \{b_i\}$. Such a system P_σ is just $a_i b_{\sigma(i)}$, for all i, and the weight is the product of all weights.

In this language we can rewrite

$$\det M = \sum_{\sigma} sign(\sigma) \omega(P_\sigma).$$

A natural generalization from bipartite to arbitrary graphs was found by Gessel and Viennot. This widely applicable result has a very simple and elegant proof, as we will see bellow.

Before stating the result we need some preparation. Let G be a finite acyclic directed graph. Every edge e carries a weight $\omega(e)$. If P is a directed path from a to b (we include all self loops), then we define the weight of P as

$$\omega(P) = \prod_{e \in P} \omega(e).$$

(this is defined as 1 for loops).

Now, let $A = \{a_1, \ldots, a_n\}$ and $B = \{b_1, \ldots, b_n\}$ be two sets of n vertices, not necessarily disjoint. To A and B we associate the *path matrix* M with

$$m_{ij} = \sum_{P : a_i \to b_j} \omega(P).$$

A *path system* $\mathcal{P}$ from A to B consists of a permutation σ together with n paths $P_i : a_i \to b_{\sigma(i)}$ for all i. We write $sign(\mathcal{P}) = sign(\sigma)$. The *weight* of $\mathcal{P}$ is the product of the path weights

$$\omega(\mathcal{P}) = \prod_{P \in \mathcal{P}} \omega(P).$$

Finally, we say that the path system $\mathcal{P} := \{P_1, \ldots, P_n\}$ is vertex disjoint if the paths P_i's are such.

Lemma 3.49 (Gessel-Viennot)**.** *Let G be a finite weighted acyclic directed graph, $A = \{a_1, \ldots, a_n\}$ and $B = \{b_1, \ldots, b_n\}$ two subsets of vertices, and M be the path matrix from A to B. Then,*

$$\det M = \sum_{\mathcal{P} \text{ vertex-disjoint path system}} sign(\mathcal{P})\omega(\mathcal{P}).$$

Proof. Note that a typical summand of $\det M$ is of the form

$$sign(\sigma)\prod_i m_{i\sigma(i)},$$

which can be written (by definition of M) as

$$sign(\sigma)\prod_{i=1}^{n}\left(\sum_{P_i : a_i \to b_{\sigma(i)}} \omega(P_i)\right).$$

Summing over all σ we get that

$$\det M = \sum_{\mathcal{P}} sign(\mathcal{P})\omega(\mathcal{P}),$$

where $\mathcal{P}$ runs over all path systems from A to B (vertex-disjoint or not). Hence, in order to complete the proof we have to show that

$$\sum_{\mathcal{P}\in\mathcal{N}} sign(\mathcal{P})\omega(\mathcal{P}) = 0,$$

where $\mathcal{N}$ is the set of all path systems which are *not* vertex-disjoint. To this end, we define a bijection $\pi : \mathcal{N} \to \mathcal{N}$ without fixed points such that for $\mathcal{P}$ and $\pi\mathcal{P}$

$$\omega(\pi\mathcal{P}) = \omega(\mathcal{P}) \text{ and } sign(\pi\mathcal{P}) = -sign(\mathcal{P}).$$

This will clearly give us the desired.

Let $\mathcal{P} \in \mathcal{N}$ with paths $P_i : a_i \to b_{\sigma_i}$. By definition, some pair of paths will intersect. Let i_0 be the first index such that P_{i_0} intersect with some path in $\mathcal{P}$. Let x be the first such common vertex on P_{i_0}, and let j_0 be the minimal index such that P_{j_0} has the vertex x in common with P_{i_0}. Now just swap between the subpaths $a_{i_0}P_{i_0}x$ and $a_{j_0}P_{j_0}x$ to obtain new paths P'_{i_0} and P'_{j_0} and define $\pi\mathcal{P}'$ be the path system obtained by replacing P_{i_0} and P_{j_0} by the P'-s. Note that

$$P'_{i_0} : a_{i_0} \to b_{\sigma(j_0)} \text{ and } P_{j_0} : a_{j_0} \to b_{\sigma(i_0)},$$

and therefore $sign(\mathcal{P}') = -sign(\mathcal{P})$. Clearly, we also have $\pi(\pi\mathcal{P}) = \mathcal{P}$ and therefore π is a bijection. Moreover, as both systems contain the exact same edges, their weights are the same. This completes the proof. □

The Gessel-Viennot Lemma can be used to derive basic properties of determinants just by looking at appropriate graphs. For us, it will serve as a tool to prove the Cauchy-Binet formula:

Proof. Let A and B be the vertex sets corresponding to the rows and columns of the matrix P. Similarly, let B and C correspond to the matrix Q. Consider now the 3 levels graph with A =TOP, B =MIDDLE and C =BOTTOM, where all edges directed from top to bottom. The ij-entry m_{ij} of the path matrix M from A to C is precisely $m_{ij} = \sum_k p_{ik}q_{kj}$, thus $M = PQ$.

The vertex-disjoint path systems from A to C in this graph correspond to pairs of systems from A to Z and Z to C, where Z is any r-element subset of B. Therefore, the result follows immediately from Lemma 3.49. □

4 Nearly orthogonal vectors

First, let us introduce the following useful lemma.

Lemma 4.1. *For any symmetric matrix A we have*

$$rank(A) \geq \frac{(trA)^2}{trA^2}.$$

Proof. Let $r = rankA$ and let $\lambda_1, \dots, \lambda_r$ all A's non-zero eigenvalues. Then,

$$trA = \sum \lambda_i,$$

and

$$trA^2 = \sum \lambda_i^2.$$

Now, using Cauchy-Schwarz we conclude that

$$(trA)^2 = \left(\sum \lambda_i\right)^2 \leq r \sum \lambda_i^2 = r \cdot trA^2$$

as desired. □

Now, let us define what a set of nearly orthogonal vectors actually means.

Definition 4.2. *A set X of unit vectors in $\mathbb{R}^d$ is said to be* nearly orthogonal *if for every* 3 *distinct vectors in X there is some pair of vectors which are orthogonal.*

The following theorem is obtained as a nice corollary from Lemma 4.1.

Theorem 4.3. *[Rosenfeld] Let X be a set of nearly orthogonal vectors in $\mathbb{R}^d$. Then,*

$$|X| \leq 2d.$$

Before proving it, we also need to recall Parseval's inequality:

Lemma 4.4. *if X is a set of orthogonal, unit vectors, then for all v we have*

$$\sum (x \cdot v)^2 \leq v \cdot v.$$

Proof. Extend X into an orthonormal basis B of $\mathbb{R}^d$. Now, take v and write it as

$$\sum_{x \in B} (x \cdot v)x.$$

Observe that

$$v \cdot v = \sum_{x \in B} (x \cdot v)^2 \geq \sum_{x \in X} (x \cdot v)^2.$$

This completes the proof. □

Finally, we are ready to prove Theorem 4.3.

Proof. Let $X = \{v_1, \dots, v_n\}$ be a set of nearly orthogonal vectors in $\mathbb{R}^d$ and let A the Gram matrix of these vectors (that is, $A_{ij} = v_i \cdot v_j$). Note that if we take M to be the matrix consisting of all v_i's as its column vectors, then $A = M^T M$, and in particular we have

$$rank(A) \le rank(M) \le d.$$

Therefore, it will be enough for us to show that $rank(A) \ge n/2$. To this end, observe that

$$trA = \sum v_i \cdot v_i = n,$$

and

$$trA^2 = \sum_i \sum_j (v_j \cdot v_i)^2.$$

Now we need to use the nearly orthogonal property: if i, j, k are distinct numbers, then at least one pair of v is orthogonal. In particular, for all i we have that the set of v_j's for which $v_i \cdot v_j \neq 0$ is a set of orthogonal unit vectors! Therefore, fixing an i, by Parseval's inequality we have

$$\sum_{j \neq i : v_j \cdot v_i \neq 0} (v_j \cdot v_i)^2 \le v_i \cdot v_i = 1.$$

Thus, for all i we have

$$\sum_{j=1}^{n} (v_j \cdot v_i)^2 \le 1 + v_i \cdot v_i = 2$$

and hence

$$trA^2 \le 2n.$$

Applying Lemma 4.1 to A we conclude that

$$rankA \ge \frac{(trA)^2}{trA^2} \ge n^2/2n = n/2.$$

This completes the proof. □

5 The Sperner Property

In this section we discuss some extremal problems related (maybe indirectly) to chains/antichains in posets. Let us first refresh our memory about what a poset is.

Definition 5.1. *A* poset *is a finite set, also denoted by P, together with a binary relation $\le$ which is: reflexive, antisymmetric, and transitive.*

For example, consider all subsets of $[n]$ with the relation $\subseteq$. If P consists of all subsets of $[n]$, then we call it a *boolean algebra* of *rank n*, and denote it by B_n.

To present small posets visually, one can draw their *Hasse diagram.* Roughly speaking, we draw all the elements such that the smaller ones bellow large ones, and we draw an edge between two *consecutive* elements. For example, draw the Hasse diagram of B_3.

Two posets P, Q are *isomorphic* if there is a bijection between them that preserves the binary relation.

A *chain* C in a poset is a totally ordered subset of P. If C has $n+1$ elements, we say it is of length n (like paths in graphs). We say that a poset is *graded of rank* n if every maximal chain is of length n. For example, B_n is such. A chain C is said to be *saturated* if every two consecutive elements in the chain are consecutive in P. If P is graded of rank n, then an element $x \in P$ is said to be of *rank* j, if the length of the largest saturated chain ending at x is j. If x is of rank j we set $\rho(x) = j$. For example, the rank of every element x in B_n is exactly its size. Clearly, one can partition P into $n+1$ 'levels', $P_0, \dots, P_n$ according to the rank of its elements. Moreover, note that every maximal chain is touching each P_j in exactly 1 element. Let us define $p_j = |P_j|$, and define the *rank generating function*

$$F(P,q) = \sum_{i=0}^{n} p_i q^i = \sum_{x \in P} q^{\rho(x)}.$$

For example, note that $F(B_n, q) = (1+q)^n$ WHY?

A graded poset of rank n is said to be *rank symmetric* if $p_i = p_{n-i}$ for all i, and *rank unimodal* if for some j we have $p_0 \leq \ldots \leq p_j \geq p_{j+1} \geq \ldots \geq p_n$. If P is both rank symmetric and rank unimodal, then we clearly have that $j = m$ if $n = 2m$ or $n = 2m+1$ and in the latter we also have $p_m = p_{m+1}$. We also say that the sequence (p_i) itself or the rank generating function is symmetric or unimodal, depends on P. A subset A of P is called an *antichain* if no two elements in A are comparable. For example, all the level sets are also antichains. The problem we are considering in this section is about finding/copmuting the size of a largest antichain.

Let's focus at the beginning at B_n. In this case, the problem of finding the largest antichain is equivalent to the problem of finding the largest family of subsets of $[n]$ such that no set is contained in the other. An intuitive guess should be the the level set $P_{n/2}$ is also a maximal antichain, which gives a lower bound of $\binom{n}{n/2}$. The question is how to show that there are no larger antichains? The main theorem that we want to present is due to E. Sperner from 1927 and is known as *Sperner's Theorem*. We give three proofs of this theorem, two are tailored for B_n and another one, based on linear algebra, that can be applied in a more general setting. Before stating the theorem we need the following definition:

Definition 5.2. *Let P be a graded poset of rank n. We say that P has the* Sperner property *if the maximum size of an antichain equals to the largest size of a level set.*

5.1 Sperner's theorem

Now we are ready to state Sperner's theorem.

Theorem 5.3 (Sperner's theorem)**.** *B_n has the Sperner's property.*

Note that it doesn't prove uniqueness of a maximal antichain!

The first proof was obtained in 1966 by David Lubell

Proof 1. Given a subset $X \subseteq [n]$ and a permutation $\pi \in S_n$, we say that π contains X as an initial segment if $\{\pi(1), \dots, \pi(|X|)\} = X$. Now, let A be an antichain and take a permutation $\pi \in S_n$ uniformly at random. For every $X \in A$, let E_X be the event 'π contains X as an initial segment'. Clearly, as A is an antichain, we have that for all $X \neq X' \in A$ the events E_X and $E_{X'}$ are disjoint! Therefore,

$$\sum_{X \in A} \Pr[E_X] \leq 1.$$

Now, fix some $X \in A$ of size $|X| = x$. Clearly,

$$\Pr[E_X] = \frac{x!(n-x)!}{n!} = \frac{1}{\binom{n}{x}}.$$

As $\binom{n}{x} \geq \binom{n}{n/2}$ holds for all x, we conclude that

$$\frac{|A|}{\binom{n}{n/2}} \leq \sum_{X \in A} \Pr[E_X] \leq 1.$$

Rearranging the above inequality gives the desired bound. □

Proof 2. Fix an integer $0 \leq k \leq n$. Consider a bipartite graph with parts $A = \binom{n}{k}$ and $B = \binom{n}{k-1}$, where $a \in A$ and $b \in B$ are adjacent if and only if $b \subseteq a$. Note that each vertex in A has degree k and each vertex in B has degree $n - k + 1$. Therefore, for a fixed subset $X \subseteq B$, we have

$$|X|(n-k+1) = e(X, N(X)) \leq |N(X)|k.$$

Now, if $k \leq \frac{n+1}{2}$, we obtain that $|N(X)| \geq |X|$. Therefore, there exists a matching of size $|B|$ for all $k = 1, \ldots, n/2$. Similarly, one can show that if $k > \frac{n+1}{2}$, then there is a matching of size $|A|$. Next, fix such a matching for each k, and observe that their union consist of saturated, disjoint, chains, covering all the subsets of $[n]$. Observe that in every obtained chain, there must be an element (unique...) from the the level set $n/2$ WHY?. Since each antichain can intersect every chain at most once, we obtain the desired. □

5.2 Application – the Erdős-Littlewood-Offord inequality

Before diving into the more complicated (but much more general though) proof of Sperner's theorem, let us give two, similar in nature, applications.

In 1938, Littlewood and Offord, in considering the distribution of zeroes in random polynomials, raised the following question. Suppose that $a_1, \ldots, a_n$ are given, real numbers, with absolute value at least 1. How many sums of the form $\sum_i \varepsilon_i a_i$ having $\varepsilon_i \in \{-1, 1\}$ can lie within an open unit interval? They proved that this number is at most $\frac{c \log n}{\sqrt{n}} 2^n$ for some fixed constant c. Later on, Erdős found an elegant way to obtain an optimal bound using Sperner's theorem. This result is now known as the Erdős-Littlewood-Offord inequality and has tons of applications and extensions.

Theorem 5.4 (Erdős, 1945)**.** *Let $a_1, \ldots, a_n$ be real numbers of absolute value at least one. For all open unit intervals I, there are at most $\binom{n}{\lfloor n/2 \rfloor}$ vectors $(\varepsilon_i)_{i=1}^n \in \{-1, 1\}^n$ such that $\sum \varepsilon_i a_i \in I$.*

Proof. Note that by changing signs of the a_is we do not change the distribution, and therefore we are allowed to assume that they are all positive. Now, fix an open, unit interval I, and let $S_I = \{(\varepsilon_i)_{i=1}^n \in \{\pm 1\}^n : \sum_i \varepsilon_i a_i \in I\}$. For each vector $\varepsilon = (\varepsilon_i) \in S$, let $A_\varepsilon \subseteq [n]$ be the set of all indices i for which $\varepsilon_i = 1$, and let $\mathcal{A} := \{A_\varepsilon : \varepsilon \in S_I\}$. In order to complete the proof, it is enough to claim that $\mathcal{A}$ is an antichain, and then to apply Sperner's theorem. Indeed, suppose that there are $\varepsilon \neq \varepsilon'$ with $A_\varepsilon \subset A_{\varepsilon'}$. As all the a_is have absolute value at least 1, it follows that

$$\left| \sum \varepsilon_i a_i - \sum \varepsilon'_i a_i \right| \geq 1$$

and therefore they cannot both lie in I. This completes the proof. □

Note that the above theorem is best possible as the sequence $a_i = 1$ for all i shows. Clearly, the number of vectors ε can be large only if there are many cancelations. That is, intuitively, it means that the sequence a_i has some 'nice' additive properties. What if, for example, we enforce all the a_i's to be distinct integers? can we do better? The following beautiful argument is due to Erdős and Moser, and was later improved by Sarkozy and Szemeredi, and was also proven in full generality by Halasz, using Fourier analysis.

Theorem 5.5 (Erdős-Moser)**.** *Suppose that all the a_i's are distinct integers. Then, the number of vectors $\varepsilon \in \{0,1\}^n$ for which $\sum_i \varepsilon_i a_i = m$ is $O(\log^{3/2} n/n^{3/2})$.*

Proof. I'll leave it to you as an exercise to prove that by switching to $\{0,1\}$ instead of $\{\pm 1\}$, and assuming that all a_i's are positive don't change the conclusion. So from now on, we assume the above and reenumerate in such a way that $a_1 \leq a_2 \leq \ldots \leq a_n$. Let $m \in \mathbb{N}$ be any integer, and let

$$B_m = \{I \subseteq [n] : \sum_i a_i = m\}.$$

We wish to show that B_m is small. To this end, we need the following two claims.

Claim 5.6. *Suppose that there are $i_1 < i_2 < \ldots < i_t$ for which*

$$2a_{i_j} \leq a_{i_{j+1}}$$

for all $1 \leq j \leq t-1$. Then,

$$|B_m| \leq 2^{n-t}.$$

Proof. Give any assignment to the ε_i's for $i \neq i_j$. In order to make the full sum be equal m, there is a unique assignment on the indices i_j. This gives the desired. □

Claim 5.7. *Suppose that $b_1, \ldots, b_s$ is a sum-free subset (that is, no partial sum gives an element b_i). Then, the number of solutions to $\sum \varepsilon_i b_i = m'$ is at most $\frac{10 \cdot 2^s}{s^{3/2}}$.*

Proof. Let $B_{m'}$ defined as before with respect to the sequence b_i. One can assume that every $I \in B_{m'}$ is of size at least $s/4$, as otherwise there are much less than $1/s^2$ such solutions. Now, for each $I \in B_{m'}$ let S_I be its 1-*shadow*. That is, S_I consists of all subsets $J \subseteq I$ of size $|I|-1$. Observe that for $I \neq I'$ we have $S_I \cap S_{I'} = \emptyset$. Indeed, as $\sum_{i \in I} a_i = \sum_{i \in I'} a_i = m$, if we delete only one element from each, this element can be recovered in a unique way.

Now to the key observation: The set $\mathcal{S} := \cup S_I$ is a Sperner family. Indeed, suppose that there are $J \subset I$ and $J' \subset I'$ with $J \subset J'$. Then, by definition we have

$$\sum_{j \in J} b_j + \sum_{j \in I' \setminus J} b_j = \sum_{j \in J} b_j + b^*,$$

where b^* is the unique element in $I \setminus J$. In particular, we obtain $\sum_{j \in I' \setminus J} b_j = b^*$ which contradicts the sum-free assumption. To complete the proof, we need to upper bound the size of $B_{m'}$. To this end, let us first observe that one can make the assumption that all the sets in $B_{m'}$ are of size at least $s/10$, as otherwise we get a much better bound WHY?. Moreover, as the S_I's are disjoint, we clearly have that every S_I consists of at least $s/10$ sets and $|mathcalS| \geq |B_{m'}|s/10$. Moreover, as $\mathcal{S}$ is a Sperner family, we obtain that

$$|B_{m'}|s/10 \leq \binom{s}{s/2},$$

yielding

$$|B_{m'}| \leq 10 \cdot 2^s/s^{3/2}.$$

This completes the proof. □

Finally, let us show how to deduce the proof of the theorem from the claims. Let us consider the disjoint intervals $I_k := [2^{k-1}, 2^k)$, $k = 1, \ldots \infty$. Clearly, $\mathbb{N} = \cup I_k$ and they are all disjoint. Moreover, crucially observe that every interval I_k is a sum-free set WHY?. Therefore, if the sequence a_i intersect at least (say) $2 \log n$ distinct intervals, then by Claim 5.6 we are done. So, we can assume that it touches at most $2 \log n$ intervals, and therefore, there must be I_k for which $s := |\{a_1, \ldots, a_n\} \cap I_k| \geq \frac{n}{2 \log n}$. By claim 5.7, we conclude that for every assignment on the complement of this set (2^{n-s} such assignments), there are at most $10 \cdot 2^s/s^{3/2}$ completions for a solution. This completes the proof. □

5.3 Sperner's theorem – Stanley's proof

The main question we want to deal with now is: which combinatorial condition guarantees that certain graded posets P have the Sperner property? Two sections ago we showed that the boolean algebra has it, but what if we work with other posets?

One natural property is similar to the approach we used in Proof 2. That is, the existence of matchings between any two consecutive level sets. More formally, define an *ordered matching* from P_i to P_{i+1} to be a one-to-one function $f : P_i \to P_{i+1}$ satisfying $x < f(x)$ for all $x \in P_i$. Clearly, if such f exists then $|P_i| \leq |P_{i+1}|$. Similarly, one can define an ordered matching from P_i to P_{i-1} (here we want $f(x) < x$ for all x). The proof of the following proposition is easy (see the previous proof) and is left as an exercise.

Proposition 5.8. *Let P be a graded poset of rank n. Suppose that there exists an integer j and ordered matchings*

$$P_0 \to P_1 \to \ldots \to P_j \leftarrow P_{j-1} \leftarrow \ldots \leftarrow P_n.$$

Then P is rank unimodal and Sperner.

Now we want to add some linear algebra into the discussion. Note that working in B_n is quite easy as it is a quite simple poset and we know everything about it. What if we replace it by a general and more abstract one? We clearly won't have the same luxury of using Hall's theorem in an easy way, so we need a new idea. For any finite set S, let $\mathbb{R}S$ denote the real vector space consisting of all formal linear combinations (with real coefficients) of elements of S. Thus, S is a basis for it. The next lemma is the linear-algebra ingredient that we need in order to prove the assumptions of Proposition 5.8.

Lemma 5.9. *Suppose there are linear transformation $U : \mathbb{R}P_i \to \mathbb{R}P_{i+1}$ satisfying*

- *U is one-to-one, and*
- *for all $x \in P_i$, $U(x)$ is a linear combination of elements $y \in P_{i+1}$ with $x < y$ (we say that U is an* order raising operator*).*

Then, there exists an order-matching $f : P_i \to P_{i+1}$.

Similarly, suppose that there exists linear transformation $U : \mathbb{R}P_i \to \mathbb{R}P_{i+1}$ satisfying:

- *U is onto, and*

- *U is an order raising operator.*

Then, there exists an order-matching $f : P_{i+1} \to P_i$

Proof. Suppose $U : \mathbb{R}P_i \to \mathbb{R}P_{i+1}$ is a one-to-one order raising operator. Let $[U]$ denote the matrix representing U with respect to the bases P_i of $\mathbb{R}P_i$ and P_{i+1} of $\mathbb{R}P_{i+1}$. Thus, the rows are indexed by the elements $\{y_i\} = P_{i+1}$ and the columns by $\{x_j\} = P_i$. Since U is one-to-one we have $rank[U] = p_i$ and therefore there are p_i linearly independent rows. By relabeling if necessary, we may assume that the first p_i (out of p_{i+1}) rows are independent. Let A be the $p_i \times p_i$ submatrix consisting of these rows. Since the rows of A are linearly independent, we have

$$\det(A) = \sum_{\pi \in S_{p_i}} sign(\pi) a_{1\pi(1)} \dots a_{p_i\pi(p_i)} \neq 0.$$

Therefore, we can pick a $\pi \in S_{p_i}$ with $a_{1\pi(1)} \dots a_{p_i\pi(p_i)} \neq 0$, and observe that since U is an order-raising operator, we have that $y_k > x_{\pi(k)}$ for all k. Indeed, consider $e_{\pi(k)}$ (the coordinates vector representing $x_{\pi(k)}$ in the corresponding basis), $Ue_{\pi(k)}$ has a y_k term in one of its coordinates. Hence, the map $f : P_i \to P_{i+1}$ defined by $f(x_k) = y_{\pi^{-1}(k)}$ is an ordered matching as desired. The second part of the lemma is similar so we omit it. This completes the proof. □

Finally, we want to apply Proposition 5.8 and Lemma 5.9 to the boolean algebra B_n in order to conclude Sperner's theorem.

To this end, we need to find a linear transformation $U_i : \mathbb{R}(B_n)_i \to \mathbb{R}(B_n)_{i+1}$ for all $0 \leq i < n$, and then prove it has the desired properties. We can define U_i in the most natural way as:

$$U_i(x) = \sum_{y \in (B_n)_{i+1}, y > x} y.$$

By definition U_i is order-raising operator and we need to show it is a one-to-one for $i < n/2$ and onto for $i \geq n/2$. In order to do so, let us introduce a 'dual' operator $D_i : \mathbb{R}(B_n)_i \to \mathbb{R}(B_n)_{i-1}$ as follows:

$$D_i(y) = \sum_{x \in (B_n)_{i-1}, x < y} x.$$

Let $[U_i]$ denote the matrix of U_i with respect to the bases $(B_n)_i$ and $(B_n)_{i+1}$, and similarly let $[D_i]$ denote the matrix od D_i with respect to the bases $(B_n)_i$ and $(B_n)_{i-1}$. Observe that

$$[D_{i+1}] = [U_i]^t.$$

Let us set $U_n = 0$ and $D_0 = 0$. The following lemma states the property that we need from B_n in order to make everything work.

Lemma 5.10. *Let $0 \leq i < n$. Then*

$$D_{i+1}U_i - U_{i-1}D_i = (n - 2i)I_i.$$

Proof. Let $x \in (B_n)_i$, and observe that

$$D_{i+1}U_i(x) = \sum_{|y|=i+1, x \leq y} \sum_{|z|=i, z \leq y} z.$$

Note that if $x, z \in (B_n)_i$ satisfy $|x \cap z| < i-1$, then there is no $y \in (B_n)_{i+1}$ with $x \cup z \subseteq y$, so the corresponding coefficients of such zs are 0. If $|x \cap z| = i-1$, then there is a unique y containing them both, namely, $y = x \cup z$. Finally, if $x = z$, then there are $n - i$ options to choose y. All in all,

$$D_{i+1}U_i(x) = (n-i)x + \sum_{|z \cap x| = i-1} z.$$

Similarly, one can show

$$U_{i-1}D_i(x) = ix + \sum_{|z \cap x| = i-1} z,$$

and we obtain the desired. □

Theorem 5.11. *U_i is one-to-one if $i < n/2$, and onto otherwise.*

Proof. As we observed before, $[D_i] = [U_{i-1}]^t$. Moreover, for every matrix A, we know that A^tA is a positive semidefinite matrix, and therefore it has only real, non-negative eigenvalues. By Lemma 5.10 we have

$$D_{i+1}U_i = U_{i-1}D_i + (n-2i)I.$$

Thus, the eigenvalues of the LHS are shifted eigenvalues of RHS. By assumption we have $n - 2i > 0$ so all the eigenvalues are strictly positive! Therefore, we obtain that U_i is one-to-one. The case $i \geq n/2$ is left as an exercise. □

5.4 Group actions on the Boolean Algebras

Suppose that X is an n-element set and that G is a group. We say that G *acts* on the set X if for every element π of G we associate a permutation π of X, such that for all $x \in X$ and $\pi, \sigma \in G$ we have

$$\pi(\sigma(x)) = (\pi\sigma)(x).$$

This gives us a homomorphism $\varphi : G \to S_X$.

Example 5.12. *Let a real number α act on the xy-plane by rotating counter clockwise around the origin by an angle of α radians.*

Recall the notion of an *orbit* of a group G on a set X. Namely, we say that $x, y \in X$ are G-*equivalent* if $\pi(x) = y$ for some $\pi \in G$. This is clearly an equivalence relation WHY? and each equivalence class is called an orbit. The orbits partition X and are disjoint. The orbit containing x is denoted by Gx. The set of all orbits is denoted X/G.

Let us consider now the case where X is the boolean algebra B_n. An *automorphism* of a poset P is an isomorphism $\varphi : P \to P$. The set of all automorphisms forms a group, denoted by $Aut(P)$ and called the *automorphism group* of P, under the operation of composition of functions. Note that any permutation of $[n]$ acts on B_n as follows: $\pi\{i_1, \ldots, i_k\} = \{\pi(i_1), \ldots, \pi(i_k)\}$. This action is clearly an automorphism. In particular, any subgroup G of S_n acts on B_n like above.

We now define the class of posets which will be of interest to us here. Let G be a subgroup of S_n, and define the *quotient poset* B_n/G as follows: the elements are the orbits of G. If o' and o'' are orbits, then define $o' \leq o''$ if there exist $x \in o'$ and $y \in o''$ such that $x \leq y$ in B_n. Check that this is indeed a partial order.

Proposition 5.13. *The quotient poset B_n/G defined above is graded of rank n and rank-symmetric.*

Proof. Graded of rank n is easy. We show rank-symmetric. Observe that the rank of each orbit is just the same as the rank of each of its elements in B_n (see that?). Therefore, the number of elements in the ith level set of B_n/G is just the number of orbits $o' \in (B_n)_i/G$. If $x \in B_n$, let $\bar{x}$ denote its complement $([n] \setminus x)$. Then $\{x_1, \ldots, x_j\}$ is an orbit of i-element subsets if and only if $\{\bar{x}_1, \ldots, \bar{x}_j\}$ is an orbit of $n-i$-element subsets. Therefore we obtain the symmetry property. □

Let $\pi \in S_n$. We associate π with a linear transformation

$$\pi : \mathbb{R}(B_n)_i \to \mathbb{R}(B_n)_i$$

defined as

$$\pi\left(\sum_{x \in (B_n)_i} c_x x\right) = \sum c_x \pi(x).$$

Clearly, this defines an action of S_n on the vector space $\mathbb{R}(B_n)_i$. The matrix of π with respect to the basis $(B_n)_i$ is just a permutation matrix. We will be interested in elements of $\mathbb{R}(B_n)_i$ which are fixed by every element of a subgroup G of S_n. The set of all such elements is denoted $\mathbb{R}(B_n)_i^G$ and is consisting of all $v \in \mathbb{R}(B_n)_i$ with $\pi(v) = v$ for all $\pi \in G$.

Lemma 5.14. *A basis for $\mathbb{R}(B_n)_i^G$ consists of all elements*

$$v_{o'} := \sum_{x \in o'} x,$$

where $o' \in (B_n)_i/G$.

Proof. First note that if o' is an orbit and $x \in o'$, then by definition we have $\pi(x) \in o'$ for all $\pi \in G$. Since π permutes the elements of $(B_n)_i$, it follows that π permutes the elements of o'. Thus, $\pi(v_{o'}) = v_{o'}$ and $v_{o'} \in \mathbb{R}(B_n)_i^G$. Moreover, all the $v_{o'}$s are linearly independent as each element $x \in (B_n)_i$ appears with a non-zero coefficient in exactly one of them.

It thus remains to show that they span $\mathbb{R}(B_n)_i^G$. That is, we want to show that every $v = \sum_x c_x x \in \mathbb{R}(B_n)_i^G$ can be written as a linear combination of the $v_{o'}$s. Given $x \in (B_n)_i$, let $G_x = \{\pi \mid \pi(x) = x\}$ be the *stabilizer* of x. Recall that $\pi(x) = \sigma(x)$ if and only if $\pi G_x = \sigma G_x$. It follows that in the multiset $\{\pi(x) \mid \pi \in G\}$, every element y in the orbit Gx appears $|G_x|$ times, and no other element appears. Therefore,

$$\sum_{\pi \in G} \pi(x) = |G_x| \cdot v_{Gx}.$$

Now, apply π to v and sum on all $\pi \in G$. Since $\pi(v) = v$, we obtain

$$|G| \cdot v = \sum_{\pi \in G} \pi(v) = \sum_{\pi \in G}\left(\sum_{x \in (B_n)_i} c_x \pi(x)\right) = \sum_{x \in (B_n)_i} c_x \left(\sum_{\pi \in G} \pi(x)\right) = \sum_{x \in (B_n)_i} c_x |G_x| v_{Gx}.$$

This completes the proof. □

Now we analyze the affect of applying the order-raising operator U_i to an element $v \in \mathbb{R}(B_n)_i^G$.

Lemma 5.15. *If $v \in \mathbb{R}(B_n)_i^G$ then $U_i(v) \in \mathbb{R}(B_n)_{i+1}^G$.*

Proof. Note that since $\pi \in G$ is an automorphism of B_n, we have $x < y$ iff $\pi(x) < \pi(y)$. Therefore, if $v \in (B_n)_i$ we obtain

$$\pi(U_i(v)) = U_i(\pi(v)),$$

which equals $U_i(v)$ for all $v \in \mathbb{R}(B_n)_i^G$. Therefore, $U_i(v) \in \mathbb{R}(B_n)_{i+1}^G$, as desired. □

Now we are ready to state and prove the main result on the Sperner's property, and this is basically the main tool to obtain all the more complicated results about general posets which satisfy the Sperner's property.

Theorem 5.16. *Let G be a subgroup of S_n. Then, B_n/G is graded of rank n, rank-symmetric, rank-unimodal, and Sperner.*

Proof. Let $P = B_n/G$. We've already seen in Proposition 5.13 that P is graded of rank n and rank-symmetric. We want to define order-raising operators $\hat{U}_i : \mathbb{R}P_i \to \mathbb{R}P_{i+1}$ and order-lowering operators $\hat{D}_i : \mathbb{R}P_i \to \mathbb{R}P_{i-1}$. Let us first consider just $\hat{U}_i$. The idea is to identify the basis elements $v_{o'}$ of $\mathbb{R}B_n^G$ with the basis element o' of $\mathbb{R}P$ and to let

$$\hat{U}_i : \mathbb{R}P_i \to \mathbb{R}P_{i+1}$$

correspond to the usual order raising operator

$$U_i : \mathbb{R}(B_n)_i \to \mathbb{R}(B_n)_{i+1}.$$

That is, for the order-raising operator as defined in the previous section, suppose that

$$U_i(v_{o'}) = \sum_{o'' \in (B_n)_{i+1}/G} c_{o',o''} v_{o''}$$

(observe that by Lemma 5.15 $U_i(v_{o'})$ indeed has this form, as $v_{o'} \in (B_n)_i^G$ and $U_i(v_{o'}) \in (B_n)_{i+1}^G$, and the $v_{o''}$ form a basis for this vector subspace). Now, define the linear operator

$$\hat{U}_i(o') = \sum_{o'' \in (B_n)_{i+1}/G} c_{o',o''} o''.$$

We claim the $\hat{U}_i$ is order-raising operator. That is, we need to show that if $c_{o',o''} \neq 0$, then $o' < o''$ in B_n/G. Since

$$v_{o''} = \sum_{x'' \in o''} x'',$$

the only way $c_{o',o''} \neq 0$, by definition of U_i, is to some $x'' \in o''$ to satisfy $x'' > x'$ for some $x' \in o'$. But this is the definition of $o'' > o'$, as we wanted to show.

Finally, to complete the proof we need to show that $\hat{U}_i$ is one-to-one for $i < n/2$ and $\hat{D}_i$ is one-to-one order-lowering for $i \geq n/2$. As the latter can be handled similarly to the $\hat{U}_i$'s, we omit the proof. We've already seen in the previous section that U_i is one-to-one for $i < n/2$. Thus, the restriction of U_i to the subspace $\mathbb{R}(B_n)_i^G$ is one-to-one. Note that U_i and $\hat{U}_i$ are the exact same transformations (in terms of their representative matrix), and therefore $\hat{U}_i$ is one-to-one as well. This completes the proof. □

An application: let $n = \binom{m}{2}$ and let $M = \{1, \dots, m\}$. Set $X = \binom{M}{2}$. Think about X as the set of all possible edges of a graph on vertex set M. Let B_X be the boolean algebra on X, then $x \in B_X$ is a collection of edges. Define a subgroup G of S_X as follows: G consists of all permutation which are obtained by permuting vertices. That is, if $\pi \in S_m$, then define $\hat{\pi}\{i, j\} = \{\pi(i), \pi(j)\}$. Thus, G is isomorphic to S_m. Now, observe that two elements $x, y \in B_X$ are in the same orbit iff they are isomorphoc graphs. Therefore, the elements of B_X/G are the isomorphism classes of simple graphs on m vertices. In particular, $|B_X/G|$ is the number of non-isomorphic such graphs, and $|(B_X/G)_i|$ is the number of non-isomorphic graphs with exactly i edges. In B_X we have $x \leq y$ iff x is a subgraph of y. This immediately gives us the following theorem:

Theorem 5.17. (*a*) *Fix $m \geq 1$. Let p_i be the number of non-isomorphic graphs on m vertices and exactly i edges. Then, the sequence p_i is symmetric and unimodal.*

(*b*) *Let T be a collection of simple graphs with m vertices such that no element of T is isomorphic to a spanning subgraph of another element of T. Then $|T|$ is maximized by taking T to consists of all nonisomorphic simple graphs with $\lfloor \frac{1}{2}\binom{m}{2} \rfloor$ edges.*

6 Combinatorial nullstellensatz

6.1 Chevalley-Warning

In this section we discuss a classical theorem which is based on a similar idea like the nullstellensatz one that will appear in the next section. Throughout the section we consider $q = p^k$, where p is prime and $k \in \mathbb{N}$.

Theorem 6.1 (The Chevalley-Warning Theorem)**.** *Let $f_1, \dots, f_t \in \mathbb{F}_q[x_1, \dots, x_n]$. If $\sum deg(f_i) < n$, then*

$$\left|\{x \in \mathbb{F}_q^n : f_i(x) = 0 \textit{ for all } i\}\right| = 0(mod\ q).$$

This theorem was proved by Warning in 1935, extending the following result due to Chevalley:

Theorem 6.2 (Chevalley's Theorem)**.** *Same assumptions as before. If the f_is have a common zero, then they have at least two.*

The theorems are quite easy to understand in case that all the f_is are linear functions WHY?

Let us give a quick application of Theorem 6.1 before proving it. Define

$$s(p, n) := \min\{m : \text{ for every } a^1, \dots, a^m \in \mathbb{Z}_p^n \ \exists \emptyset \neq I \subseteq [m] \text{ s.t. } \sum a^i = 0(mod\ p)\}.$$

In other words, $s(p, n)$ is the minimal m such that any $n \times m$ matrix in $\mathbb{Z}_p$, M, satisfy $Mx = 0$ for some 0/1 vector.

It is easy to check that $s(p, n) \geq (p - 1)n + 1$ as you can just take $p - 1$ copies of some basis. We prove the following theorem

Theorem 6.3 (Olson)**.** *For any prime p and any n, $s(p, n) = (p - 1)n + 1$.*

In general the *Davenport constant* of a finite abelian group G, denoted $s(G)$, is the least m such that for any $a^1, \dots, a^m \in G$, there's a nonempty $I \subseteq [m]$ for which $\sum_{i \in I} a^i = 0$. Thus $s(p, n) = s(\mathbb{Z}_p^n)$. Olson and D. Kruyswijk independently (and by different methods) determined the Davenport constants for all p-groups. In general, it is not even known whether the above theorem is true if p is not a prime. That is, is it true that $s(\mathbb{Z}_k^n) = (k - 1)n + 1$ for every k and n?

Proof. Let $m = (p-1)n + 1$. Suppose $a^1, \dots, a^m \in \mathbb{Z}_p^n$, and for each j let $f_j \in \mathbb{Z}_p[y_1, \dots, y_m]$ be

$$f_j(y) = \sum_i a_j^i y_i^{p-1}.$$

Note that $\sum deg(f_j) = (p-1)n < m$ and $f_j(0) = 0$ for all j. Therefore, by Theorem 6.1 we have that there exists $y \neq 0$ which is a common zero of the f_j's. Then, we can take as I the support of y and observe that $x^{p-1} = 1$ for all $x \neq 0$ in $\mathbb{Z}_p$. This completes the proof. □

Now let's prove the theorem. We say that $g \in \mathbb{F}_q[x_1, \dots, x_n]$ is *reduced* if $deg_i(g) \leq q-1$ for every $i \in [n]$ ($deg_i(g)$ is the degree of x_i in g). Then for $f \in \mathbb{F}_q[x_1, \dots, x_n]$, the reduced polynomial corresponding to f is the reduced polynomial $\bar{f}$ obtained from f by iterating until no longer possible: replace some term x_i^ℓ with $\ell \geq q$ by $x_i^{\ell-q+1}$. Observe that f and $\bar{f}$ both agree on $\mathbb{F}_q^n$. Note that in finite fields, $f(x) = 0$ for all x doesn't imply all coefficients are 0! (for example, consider $x^q - x$). But the reduced polynomials imply that (the proof is obtained by a simple induction on the number of variables):

Lemma 6.4. *If g is reduced, then*

$$g(x) = 0 \forall x \in \mathbb{F}_q^n \text{ if and only if } g = 0.$$

Now we are ready to prove the theorem:

Proof. Let Z be the set of common zeros of the f_i's. It is easy to check that each of the following is a polynomial representing the function 1_Z:

$$f(x) = \prod_{j=1}^{t} (1 - f_j(x)^{q-1}),$$

$$h(x) = \sum_{a \in Z} \prod_{i=1}^{n} (1 - (x_i - a_i)^{q-1}).$$

Note that h is reduced, and therefore $\bar{f} = h$ and

$$deg(h) = deg\bar{f} \leq \deg(f) \leq (q-1) \sum deg(f_j) < (q-1)n.$$

BUT then the leading term in h, $|Z| \prod_{i=1}^{n} (-x_i)^{q-1}$ must vanish! that is, we must have $|Z| = 0(mod\ p)$. This completes the proof. □

6.2 Combinatorial nullstellensatz

In this section we are going to focus on the 'theory of zeros'. In general, our tools are the following two theorems that we will prove later.

Theorem 6.5. *Let F be an arbitrary field, and let $f(x_1, \dots, x_n)$ be a polynomial in $F[x_1, \dots, x_n]$. Let $S_1, \dots, S_n$ be nonempty subsets of F and define $g_i(x) = \prod_{s \in S_i}(x_i - s)$. If f vanishes over all the common zeros of $g_1, \dots, g_n$ (that is, $f(s_1, \dots, s_n) = 0$ for all $s_i \in S_i$), then there are polynomials $h_1, \dots, h_n \in F[x_1, \dots, x_n]$ satisfying $deg(h_i) \leq deg(f) - deg(g_i)$ so that*

$$f = \sum h_i g_i.$$

Moreover, if $f, g_1, \dots, g_n$ lie in $R[x_1, \dots, x_n]$ for some subring R of F then there are polynomials as above with $h_i \in R[x_1, \dots, x_n]$.

As a corollary of the above theorem we obtain the following:

Theorem 6.6. *Let F be an arbitrary field, and let $f(x_1, \ldots, x_n) \in F[x_1, \ldots, x_n]$. Suppose that $deg(f) = \sum t_i$, where each t_i is a nonnegative integer, and suppose the coefficient of $\prod_{i=1}^{n} x^{t_i}$ in f is nonzero. Then, if $S_1, \ldots, S_n$ are subsets of F with $|S_i| > t_i$, there are $s_i \in S_i$ for which*

$$f(s_1, \ldots, s_n) \neq 0.$$

These two theorems are known as Combinatorial Nullstellensatz and where introduced by Alon (based on the so called Hilbert Nullstellensatz Theorem).

Before proving the above theorems, let us start with some simple applications. The first application gives us a short proof for a famous theorem by Cauchy and Davenport, and it has numerous of applications in Additive Number Theory.

Theorem 6.7. *If p is a prime, and A, B are two nonempty subsets of $\mathbb{Z}_p$, then*

$$|A + B| \geq \min\{p, |A| + |B| - 1\}.$$

Cauchy proved this theorem in 1813 and applied it to give a new proof to a lemma of Lagrange which asserts that any integer is a sum of four squares. Davenport, having other applications in mind, rediscovered it in 1835. Here we give a very short proof for this theorem due to Alon, Nathanson and Ruzsa.

Proof. If $|A| + |B| > p$ then the result is trivial, as for all $g \in \mathbb{Z}_p$, the set $g - B$ intersects A. Assume therefore that $|A| + |B| \leq p$ and suppose that the result is false. That is, assume that $|A + B| \leq |A| + |B| - 2$. Let C be a subset of $\mathbb{Z}_p$ satisfying $A + B \subseteq C$ and $|C| = |A| + |B| - 2$. Define

$$f(x, y) = \prod_{c \in C} (x + y - c).$$

By definition, we have that

$$f(a, b) = 0 \text{ for all } a \in A \text{ and } b \in B.$$

Let $t_1 = |A| - 1$, $t_2 = |B| - 1$ and note that the coefficient of $x^{t_1} y^{t_2}$ is $\binom{|A|+|B|-2}{|A|-1}$ which is nonzero in $\mathbb{Z}_p$, since $|A| + |B| - 2 < p$. Therefore, by applying Theorem 6.6 with $S_1 = A$ and $S_2 = B$ we obtain a contradiction. □

As a second application we prove the Erdős-Ginzburg-Ziv theorem from 1961. One of the first exercises in pigeonhole principle tells us that of $a_1, \ldots, a_n$ is any sequence of integers (not necessarily distinct), then there is a non-empty subset summing to 0 mod n. A natural question that one can ask is about the *size* of this subset. This is the aim of the next theorem.

Theorem 6.8 (Erdős-Ginzburg-Ziv)**.** *For any $a_1, \ldots, a_{2n-1} \in \mathbb{Z}_n$, there is an $I \subseteq [2n - 1]$ with $\sum_{i \in I} a_i = 0(mod\ n)$ and $|I| = n$.*

Exercise 6.9. *Show that $2n - 2$ is not enough.*

It turns out that it is enough to prove this theorem when n is a prime (do you see how is it related to one of the problems in PSET3?). So, we only prove it for this case.

Before proving the theorem, we need the following immediate corollary from the Cauchy-Davenport theorem which can be easily obtained by induction.

Corollary 6.10. *For any t, a prime p and nonempty $A_1, \ldots, A_t \subseteq \mathbb{Z}_p$,*

$$|A_1 + \ldots + A_t| \geq \min\{\sum |A_i| - t + 1, p\}.$$

Now we are ready to prove the theorem.

Proof. Let $A_i = \{a_{2i-1}, a_{2i}\}$ for all $1 \leq i \leq n-1$, and $A_n = \{a_{2n-1}\}$. Then, the above corollary gives us

$$|\sum A_i| \geq \min\{\sum |A_i| - n + 1, n\} = n.$$

That is,

$$\sum A_i = \mathbb{Z}_n,$$

so in particular we have $0 \in \sum A_i$ and we are done. Clearly, the above argument is wrong! as it can be that some of the A_i contain two elements which are the same... In order to fix the proof, we need the following:

Claim 6.11. *If no $a \in \mathbb{Z}_n$ appears more than n times in the sequence $a_1, \ldots, a_{2n-1}$, then we can reindex in such a way that all the A_i are of size 2 (except of the obvious A_n).*

Proof. If no a_i appears twice, then we are done. Otherwise, delete one appearance of an a_i that appears at least twice and label it a_{2n-1}. To complete the argument – we are left with $2n-2$ numbers (with multiplicities) and consider a graph with those a_is as vertices and edges between two a_i's which are non-equal. This graph has minimum degree at least $n-1$ (by assumption), and therefore it is quite straightforward to find a perfect matching in it. □

In order to complete the proof, just observe that if some a_i appears n times then it is trivial to find such an index set I. □

The next application is in graph theory:

Theorem 6.12 (Alon-Friedland-Kalai). *For any prime p, any loopless graph G with average degree bigger than $2p-2$ and maximum degree at most $2p-1$ contains a p-regular subgraph.*

Proof. Let $(a_{v,e})_{v \in V(G), e \in E(G)}$ be the incidence matrix of G defined by $a_{v,e} = 1$ if $v \in e$ and 0 otherwise. Associate each edge e with a variable x_e and consider the polynomial

$$F = \prod_{v \in V(G)} \left[1 - (\sum_{e \in E(G)} a_{v,e} x_e)^{p-1} \right] - \prod_{e \in E(G)} (1 - x_e),$$

over $\mathbb{Z}_p$. Note that $deg(F) = |E|$ since the degree of the first product is $(p-1)|V(G)| < |E|$. Moreover, the coefficient of $\prod_{e \in E(G)} x_e$ in F is $(-1)^{|E|+1} \neq 0$. Therefore, there are values $x_e \in \{0, 1\}$ with $F(x) \neq 0$. Note that the vector x is not the zero vector! Now, since $\sum_{e \in E} a_{v,e} x_e = 0(mod\ p)$ for every v (otherwise F will vanish), it follows that the subgraph consisting of all edges $e \in E$ with $x_e = 1$ is p regular. □

We will see many more applications later on in this class.

6.3 Proof of the main theorems

Before proving Theorem 6.5 we need the following lemma.

Lemma 6.13. *Let $P(x_1, \ldots, x_n)$ be a polynomial in n variables over an arbitrary field F. Suppose that the degree of P as a polynomial in x_i is at most t_i, and let $S_i \subset F$ be a set of $t_i + 1$ distinct elements of F. If $P(x_1, \ldots, x_n) = 0$ for all n-tuples $(x_1, \ldots, x_n) \in S_1 \times \ldots \times S_n$, then $P = 0$.*

Proof. We apply induction on n. For $n = 1$, the lemma is simple. Suppose it holds for $n - 1$ and let's prove it for n. Given P and sets S_i as in the lemma, let us write

$$P = \sum_{i=0}^{t_n} P_i(x_1, \ldots, x_{n-1}) x_n^i.$$

For each such fixed $(n-1)$-tuple $(x_1, \ldots, x_{n-1}) \in S_1 \times \ldots \times S_{n-1}$, the polynomial in x_n obtained from P by substituting the values of $x_1, \ldots, x_{n-1}$ vanishes for all $x_n \in S_n$, and therefore is 0. Thus, $P_i = 0$ for all such $(n-1)$-tuples. All in all, by induction, $P_i = 0$ for all i. □

Now we can prove Theorem 6.5.

Proof. Let $t_i = |S_i| - 1$ for all i. By assumption, for every $(x_1, \ldots, x_n) \in S_1 \times \ldots \times S_n$ we have

$$f(x_1, \ldots, x_n) = 0.$$

For each i let

$$g_i(x) = \prod_{s \in S_i} (x_i - s) = x_i^{t_i+1} - \sum_{j}^{t_i} g_{ij} x_i^j.$$

Observe that if $x_i \in S_i$, then $g_i(x_i) = 0$. That is, $x_i^{t_{i+1}} = \sum g_{ij} x_i^j$.

Let $\bar{f}$ be the polynomial obtained by writing f as a linear combination of monomials and replacing, repeatedly, each occurrence of $x_i^{f_i}$, where $f_i > t_i$, by a linear combination of smaller powers of x_i as above. This gives a polynomial of degree at most t_i in x_i, for each i. Clearly, $\bar{f}$ is obtained from f by subtracting from it a product of the form $h_i g_i$, where the degree of each h_i does not exceed $deg(f) - deg(g_i)$. Moreover, as $\bar{f}(x) = f(x)$ for all $x \in S_1 \times \ldots \times S_n$, by Lemma 6.13 we have $\bar{f} = 0$. This implies that $f = \sum h_i g_i$ as desired. □

Now we can prove Theorem 6.6.

Proof. By deleting elements if necessary, we may assume that $|S_i| = t_i + 1$ for all i. Suppose the result is false and define $g_i(x_i)$ as before. By Theorem 6.5 there are polynomials $h_1, \ldots, h_n \in F[x_1, \ldots, x_n]$ satisfying $deg(h_j) \leq \sum t_i - deg(g_j)$ so that

$$f = \sum h_i g_i.$$

By the assumption, the coefficient of $\prod x_i^{t_i}$ in f is nonzero, and so is the coefficient of this monomial in the right hand side. BUT, the degree of $h_i g_i = h_i \prod_{s \in S_i}(x_i - s)$ is at most $deg(f)$, and if there are monomials of degree $deg(f)$ in it, they are divisible by $x_i^{t_i+1}$! this shows that the coefficient of $\prod x_i^{t_i}$ in RHS must be 0, which gives us a contradiction. □

6.4 More applications of the combinatorial nullstellensatz

6.4.1 Latin transversals

Suppose that M is a $k \times k$ matrix with entries from a set $\mathcal{S}$ of symbols. A *Latin transversal* of M is a set of k cells, no two sharing a row/column/symbol. EXAMPLE. That is, a latin transversal is a permutation $\pi \in S_k$ with all entries $m_{i\pi(i)}$ being distinct.

We are interested in the case that M is a submatrix of the addition table of some abelian group G. That is, we fix $A, B \subset G$ and consider the submatrix $(m_{a,b})_{a\in A, b\in B}$ where $m_{a,b} = a + b$.

Conjecture 6.14 (H. Snevily, 1999)**.** *If G is abelian group of odd order, then any square submatrix of the addition table of G has a Latin transversal.*

An equivalent way to say it: given $A, B \subset G$ of sizes $|A| = |B| = k$, there are orderings $A = \{a_1, \dots, a_k\}$ and $B = \{b_1, \dots, b_k\}$ for which $a_1 + b_1, \dots, a_k + b_k$ are all distinct. Observe that this is false if $|G|$ is even! the easiest way to see it is is considering $\mathbb{Z}_2$, but one can also easily find examples for larger sizes.

Snevily's conjecture is still open but has been proven for the cyclic groups $\mathbb{Z}_n$. It was proved in two steps

Theorem 6.15 (Alon 2000)**.** *True for $\mathbb{Z}_p$ where p is prime.*

Theorem 6.16 (Dasgupta, Károlyi, Serra and Szegedy, 2000)**.** *True for all odd n.*

As the latter relies on the former we need to prove them both. A related conjecture which is worth mentioning is the following:

Conjecture 6.17. *If k is odd and no symbol appears more than once in any row/column of M, then M has a Latin transversal.*

We now prove the following theorem, which is slightly stronger than Theorem 6.15

Theorem 6.18. *Suppose p is prime and $k < p$. Then for any (not necessarily distinct) $a_1, \dots, a_k \in \mathbb{Z}_p$ and $B \subseteq \mathbb{Z}_p$ of size k, there is an ordering $b_1, \dots, b_k$ of the elements of B for which $a_1+b_1, \dots, a_k+b_k$ are all distinct.*

Note that this is false for $k = p$. Indeed, one can take $a_1 = \ldots = a_{p-1} = 0$ and $a_p \neq 0$. It is also false for any p which is not prime (WHY?). On the other hand, it implies Theorem 6.15 as the case $k = p$ is easy to verify.

Proof. Let us define the following polynomial:

$$f(x) = \prod_{i<j}(x_i - x_j) \prod_{i<j}(x_i + a_i - x_j - a_j) \in \mathbb{Z}_p[x_1, \dots, x_n].$$

Note that an ordering as in the theorem is the same as an $x \in B^k$ with $f(x) \neq 0$. Note that $deg(f) = 2\binom{k}{2} = k(k-1)$ and therefore we need to prove that the coefficient of $\prod x_i^{k-1}$ is not 0. Now, note that the coefficient is also the same as the coefficient of $\prod x_i^{k-1}$ in $\prod(x_i - x_j)^2 = (\det M)^2$, where M is the Vandermonde matrix $V(x_1, \dots, x_k)$. RECALL WHAT IS IT?

We have

$$\det M = \sum_{\sigma \in S_k} sign(\sigma) \prod_i x_i^{\sigma(i)-1},$$

and therefore, the coefficient of $\prod x_i^{k-1}$ in $(\det M)^2$ is

$$\sum_{\sigma} sign(\sigma) sign(\sigma')$$

where σ' is given by

$$\sigma'(i) = k + 1 - \sigma(i).$$

The key observation is that $x = sign(\sigma)sign(\sigma')$ is the same for all σ. WHY?

Therefore, the coefficient is $k!x$ which is not 0 mod p. This completes the proof. □

Now we prove the more general Theorem 6.16

Proof. The new idea here is to write the group multiplicatively. That is, we write C_n instead of $\mathbb{Z}_n$ and embed it in the multiplicative group of an appropriate field (and then we can use the Nullstellensatz). For the field we choose $\mathbb{F}_q$ with

$$q = 2^{\varphi(n)},$$

where φ is Euler's totient function. As you'll see soon, the choice of characteristic 2 is crucial. The exponent is chosen to guarantee that C_n is indeed a subgroup of $\mathbb{F}_q^\times$: since $2 \in \mathbb{Z}_n^\times$ (since n is odd) and $|\mathbb{Z}_n^\times| = \varphi(n)$, we have $q = 2^{\varphi(n)} = 1$ in $\mathbb{Z}_n$. That is, $n \mid (q-1)$. So, since the multiplicative group of any finite field is cyclic (IF YOU DON'T SEE IT, IT IS A NICE EXERCISE!), we have

$$C_n \leq C_{q-1} \cong \mathbb{F}_q^\times.$$

So, we are given distinct $a_1, \ldots, a_k \in C_n \leq \mathbb{F}_q^\times$ and $B \subset C_n$ of size k, and should find an ordering $b_1, \ldots, b_k$ for which the products $a_i b_i$ are all distinct. Define

$$f(x) = \prod (x_i - x_j) \prod (a_i x_i - a_j x_j) \in \mathbb{F}_q[x].$$

As before, we're looking for an $x \in B^k$ with $f(x) \neq 0$. The degree of f is $k(k-1)$ (same as in the previous proof), so existence of the desired x will follow from the nullstellensatz theorem if we can show that the coefficient of $\prod x_i^{k-1}$ is non-zero. Here, the a_i's do play a role! note that we can write

$$f(x) = V(x_1, \ldots, x_k) V(a_1 x_1, \ldots, a_k x_k).$$

Therefore,

$$f(x) = \left(\sum_{\sigma} sign(\sigma) \prod x_i^{\sigma(i)-1} \right) \left(\sum_{\tau} sign(\tau) \prod (a_i x_i)^{\tau(i)-1} \right).$$

The desired coefficient in this expression is

$$\sum sign(\sigma) sign(\sigma') \prod a_i^{\sigma'(i)-1} = \pm \mathrm{per}(V(a_1, \ldots, a_k)),$$

where σ' is as in the previous proof. In order to complete the proof we need to show that $\mathrm{per}(V(a_1, \ldots, a_k)) \neq 0$. Here we use the fact that the chracteristic is 2 and the x_i's are distinct! that is, $per = det \neq 0$. This completes the proof. □

6.4.2 Graceful labeling of certain trees

7 Working over the reals – integer rounding

7.1 Shannon Capacity

Let's start with few definitions. For a graph G, define its *independence number*, $\alpha(G)$, to be the size of the largest independent set in G. For two graphs H, G, the *product* of G and H, denoted $G \times H$ is the graph with vertex set $V(G) \times V(H)$ and

$$(x,y) \sim (x',y') \Leftrightarrow \left(x = x' \text{ or } x \sim x'\right) \text{ and } \left(y = y' \text{ or } y \sim y'\right).$$

(where (x,y) and (x',y') are distinct vertices of course.) More generally, for graphs $G_1, \ldots, G_n$, the product $G_1 \times \ldots \times G_n$ is defined as a graph on

$$V(G_1) \times \ldots \times V(G_n)$$

with

$$x \sim y \Leftrightarrow (x_i = y_i \text{ or } x_i \sim y_i \text{ for every } i)\,.$$

Note that $G \times H \times K = (G \times H) \times K = G \times (H \times K)$ (very simple exercise).

We are interested in $\alpha(G^n)$, where $G^n = G \times G \times \ldots \times G$ (n times). Let us first observe that we trivially have $\alpha(G \times H) \geq \alpha(G)\alpha(H)$. (Indeed, if I is an ind. set of G and J of H, then $I \times J$ is an ind. set of $G \times H$). As one can easily check, for (very) small graphs the lower bound is actually tight. The smallest graph for which it's not tight for is C_5, which is convenient to think about as $\mathbb{Z}_5$ with edges of the form $(i, i+1)$ (where addition is being made modulo 5). It is quite simple to show that $\alpha(C_5) = 2$ but $\alpha(C_5^2) \geq 5$ (check for example the set $\{(0,0),(1,2),(2,4),(3,1),(4,3)\}$), so the above inequality is a strict one for C_5.

The following quantity was introduced by Shannon in 1956:

Definition 7.1. *The* Shannon capacity *of G is*

$$\theta(G) = \sup_n \alpha(G^n)^{1/n}.$$

Let us give some motivation for the above definition. Shannon was interested in error-free communication over noisy channels. That is, imagine you have alphabet, denote by V, in which some pairs of letters are indistinguishable. A message is just a string of letters from V, and the question is, what is the largest size of a set of messages of length n, any two of which are distinguishable. That is, you want the largest collection of strings of length n such that for all $x \neq y$ we have that x_i is distinguishable from y_i for some i.

Now, define a graph G on V with $x \sim y$ if and only if x and y are indistinguishable. Then, clearly, one want to find independent sets of G^n as large as possible. The Shannon capacity then measures the *rate* of transmission (which is defined as $\log \theta(G^n)$), and this quantity can be thought of the number of bits sent per unit of time.

Before stating the bounds we wish to prove, let us show that the sup is actually the lim.

Lemma 7.2. *For any G, $\lim_{n\to\infty} \alpha(G^n)^{1/n}$ exists and is at least $\alpha(G^n)^{1/n}$ for every n.*

To prove the lemma we need the following form of Fekete's Lemma (the proof is a simple exercise):

Lemma 7.3. *(Fekete's Lemma) If $f : \mathbb{N} \to \mathbb{N}$ satisfies $f(m+n) \geq f(m) + f(n)$ for all m, n. Then $\lim \frac{f(n)}{n}$ exists (it may be ∞) and is at least $\frac{f(k)}{k}$ for any particular n.*

In order to prove Lemma 7.2, simply apply Fekete's Lemma on $b_n = \log \alpha(G^n)$ and use the inequality above. To conclude, the lemma shows that any lower bound on $\alpha(G^n)$ for a particular n implies a lower bound $\alpha(G^n)^{1/n}$ (which is at least $\alpha(G)$) on $\theta(G)$. For example, we have that $\alpha(C_5^2) \geq 5$ and therefore

$$\theta(C_5^2) \geq \sqrt{5}.$$

Our main goal in this section is to show that this is tight, but to do so we still need some preparation.

Let $\mathcal{C}(G)$ denote the set of all cliques in G. A *clique cover* of G is a collection of cliques whose union is $V(G)$, and the *clique cover number*, which here we denote by $\rho(G)$, is the smallest size of a clique cover. This is trivially an upper bound on the independence number as distinct vertices in an independent set must belong to different cliques.

Let us show that

$$\alpha(G \times H) \leq \rho(G)\alpha(H).$$

Indeed, suppose that I is independent set of $G \times H$ and K is a clique of G. Then, if $x \neq x' \in K$ and $(x, y), (x', y') \in I$, then y and y' are nonadjacent. Therefore, for each such I and a clique cover $\mathcal{K}$ of G we have

$$|I| \leq \sum_{x \in V(G)} |I \cap (\{x\} \times V(H))| \leq \sum_{K \in \mathcal{K}} \sum_{x \in K} |I \cap (\{x\} \times V(H))|$$

where the latter equals

$$\sum_{K \in \mathcal{K}} |I \cap (K \times V(H))| \leq |\mathcal{K}|\alpha(H).$$

The key point here is that

$$\theta(G) \leq \rho(G).$$

Indeed, from the above bound we get $\alpha(G^n) \leq \rho(G)\alpha(G^{n-1}) \leq \rho(G)^n$ and we are done.

This is already enough to determine the capacities of all graph with $\alpha(G) = \rho(G)$ (this includes for example, complete graphs, bipartite graphs and more). The first graph for which this argument is not helpful is C_5 so we need new ideas.

One possible idea is to define a *fractional clique cover* of G. This is just a function $t : \mathcal{C}(G) \to \mathbb{R}^+$ satisfying

$$\sum_{x \in K} t(K) \geq 1, \ \forall x \in V(G).$$

Note that a clique cover is also a fractional cover with $t(K) \in \{0, 1\}$. Therefore, $\rho(G)$ is lower bounded by the *fractional clique cover number* which is

$$\rho^*(G) = \min\{\sum t(K) : t \text{ a fractional clique cover of } G\}.$$

For example, one can easily show that

$$\rho^*(C_{2k+1}) = k + 1/2.$$

(you have it in PSET4). Do you see what is the clique cover number of C_{2k+1}?

Proposition 7.4. *For any G and H,* *1.* $\alpha(G \times H) \leq \rho^*(G)\alpha(H)$.

2. $\rho^*(G \times H) = \rho^*(G)\rho^*(H)$.

We won't prove 2. as is required to define duality in linear prgramming but it is a relatively simple exercise.

Proof. We only prove 1. For any independent set I in $G \times H$ and a fractional clique cover t of G,

$$|I| = \sum_{x \in V(G)} |I \cap (\{x\} \times V(H))| \leq \sum_{x \in V(G)} \sum_{x \in K} t(K)|I \cap (\{x\} \times V(H))|$$

which equals

$$\sum_K t(K)|I \cap (K \times V(H))| \leq \sum_K t(K)\alpha(H).$$

This completes the proof. □

Clearly the above gives

$$\theta(G) \leq \rho^*(G).$$

In particular, this improves what we know for C_5 to

$$\sqrt{5} \leq \theta(C_5) \leq 5/2,$$

which was the best known bound for this problem for more than 20 years until

Theorem 7.5. *(Lovász 1979)* $\theta(C_5) = \sqrt{5}$.

The problem is still widely open for $\theta(C_{2k+1})$ for all $k \geq 3$. For proving the theorem, we work with the *tensor* product of $u \in \mathbb{R}^m$ and $w \in \mathbb{R}^n$, defined as

$$u \otimes w = (u_i w_j : i \in [m], j \in [n]).$$

(note that this is a vector of length mn). The following easy property is crucial to us

$$\langle u \otimes w, u' \otimes w' \rangle = \langle u, u' \rangle \langle w, w' \rangle.$$

Definition 7.6. *An (orthonormal) representation of a graph G with vertex set V is a list* $(u_x : x \in V)$ *of unit vectors in some space* $\mathbb{R}^a$ *satisfying*

$$\langle x, y \rangle = 0 \text{ for all distinct, non adjacent } x, y \in V.$$

The value *of the representation is*

$$\min_c \max_{x \in V} \langle c, u_x \rangle^{-2},$$

where c runs over all unit vectors in $\mathbb{R}^a$*.*

It is not hard to see in the definition above that there is a unique c achieving the minimum. This is called the *handle* of the representation. Finally, the Lovász *theta function* of G, $\vartheta(G)$, is the minimum value of a representation, and an *optimal* representation is one with value $\vartheta(G)$.

Now we are ready to prove the theorem:

Proof. Let us start with the following claim that will make the above definition look a bit more natural:

Claim 7.7. *For any G we have $\alpha(G) \leq \vartheta(G)$.*

Proof. Let I be an indepedent set of G. Then, for any representation $(u_x : x \in V(G))$ and unit vector c, by orthonormality we obtain:

$$1 = \|c\|^2 \geq \sum_{x \in I} \langle c, u_x \rangle^2 \geq |I| \min_x \langle c, u_x \rangle^2.$$

□

Next, we show:

Claim 7.8. *For all G and H we have $\vartheta(G \times H) \leq \vartheta(G)\vartheta(H)$*

Proof. Suppose (u_x) and (v_y) are optimal representations of G and H, respectively, with handles c and d. It is immediate to see that $(u_x \otimes v_y : (x, y) \in V(G \times H))$ is a representation of $G \times H$ and $c \otimes d$ is a unit vector. Moreover, for all $(x, y) \in V(G \times H)$ we have

$$\langle c \otimes d, u_x \otimes v_y \rangle^{-2} = \langle c, u_x \rangle^{-2} \langle d, v_y \rangle^{-2} \leq \vartheta(G)\vartheta(H).$$

Therefore, $\vartheta(G)\vartheta(H)$ is an upper bound on the value of this representation and we obtain the desired. □

Note that combining both claims give us

$$\theta(G) \leq \vartheta(G).$$

Therefore, in order to complete the proof, it is enough to show that $\vartheta(C_5) \leq \sqrt{5}$. To this end, take an umbrella with unit handle c and unit ribs $u_1, \ldots, u_5$. Open it until u_i and u_{i+2} are all orthogonal, and by the spherical cosine theorem one can hsow that $\langle c, u_i \rangle = 5^{-1/4}$. This completes the 'proof'. □

Lovász result was the beginning of what is now an area. Here we give one of several ways of defining $\vartheta(G)$ and use (either in the notes or in a pset) this alternate definition to determine ϑ for some graphs.

Theorem 7.9. *For G on vertex set $[n]$,*

$$\vartheta(G) = \min \lambda_1(A),$$

where the minimum is over all real symmetric $n \times n$ matrices A with $a_{ij} = 1$ if $i = j$ or $i \nsim j$.

Proof. First, we show that the minimum in the theorem, say λ, is an upper bound. Let A be a matrix as in the theorem with $\lambda_1(A) = \lambda$. Then, $\lambda I - A$ is positive semidefinite, so there are vectors $x_1, \ldots, x_n$ (in some space) with

$$x_i \cdot x_j = \lambda \delta_{ij} - a_{ij}$$

(this follows from the fact that every real values, symmetric, positive semidefinite matrix is a gram matrix of some n vectors – exercise).

Let c be a unit vector orthogonal to all the x_i's (there is no constraint on the space in which the x_i's and c lie and therefore the existence of such a c is not an issue), and set

$$u_i = \lambda^{-1/2}(c + x_i).$$

Then,

$$u_i \cdot u_j = \lambda^{-1}(1 + x_i \cdot x_j)$$

which equals 1 if $i = j$ and 0 if $i \nsim j$. In particular, (u_i) is a representation, and therefore, as $c \cdot u_i = \lambda^{-1/2}$, $\vartheta(G)$ is at most λ.

Now we wish to show that $\lambda \leq \vartheta(G)$. To this end we more or less reverse the above construction. Let (u_i) be a representation of G with handle c, and set

$$x_i = \frac{u_i}{c \cdot u_i} - c.$$

Then,

$$x_i \cdot x_j = (c \cdot u_i)^{-2} - 1 \leq \vartheta - 1$$

if $i = j$ and -1 if $i \nsim j$. Now, let A be given by

$$a_{ij} = 1 \text{ if } i = j$$

and

$$a_{ij} = -x_i \cdot x_j \text{ otherwise.}$$

Then A satisfies the conditions in the theorem. So,

$$\lambda \leq \lambda_1(A).$$

On the other hand, observe that

$$\vartheta I - A \text{ is positive semidefinite .}$$

Indeed, note that $(\vartheta I - A)_{ij}$ is at least $x_i \cdot x_j$ if $i = j$ and equals $x_i \cdot x_j$ otherwise. And since the Gram matrix is positive semidefinite and we only added positive values on the diagonal, then it follows. All in all, we get

$$\lambda \leq \lambda_1(A) \leq \vartheta$$

as desired. This completes the proof. □

Note that once we know $\lambda = \vartheta$, the construction in the first part of the argument shows that there is an optimal representation (u_i) with handle c such that $(c \cdot u_i)^{-2} = \lambda$ for every i.

A graph G is *vertex transitive* if $Aut(G)$ acts transitively on $V(G)$. *Edge transitivity* is defined similarly.

Theorem 7.10. *If G is d-regular and $d = \lambda_1 \geq \ldots \geq \lambda_n$, then,*

$$\vartheta(G) \leq \frac{-n\lambda_n}{d - \lambda_n}.$$

Equality holds whenever G is edge-transitive.

Does it look familiar?

Proof. Let $A = J - \gamma A_G$ with the value γ to be chosen later. Then, A satisfies the condition in Theorem 7.9, so by this theorem, its largest eigenvalue is an upper bound on $\vartheta(G)$. We would like to choose γ to minimize the upper bound. The eigenvalues of A are $n - \gamma\lambda_1$ and $-\gamma\lambda_i$, for all $i \geq 2$. The maximum eigenvalue of A is the maximum of $n - \gamma\lambda_1$ and $-\gamma\lambda_n$, and is minimized at $\gamma = \frac{n}{d-\lambda_n}$, where both of these are qual to the RHS of the inequality stated in the theorem. This proves the first part of the theorem.

For the second part, we show that

Claim 7.11. *For any G on vertex set $[n]$,*

$$\vartheta(G) = \min \lambda_1(A),$$

where the minimum is over all matrices as defined in Theorem 7.9 and which satisfy

$$a_{\sigma(i)\sigma(j)} = a_{ij}$$

for all $i, j \in [n]$ and $\sigma \in Aut(G)$.

Note that the second part of the theorem now follows immediately. Indeed, for an edge transitive graph G, any A satisfying the condition is of the form $J - \gamma A_G$, and we have already seen that the minimum of $\lambda_1(A)$ for such matrices is equal the right hand side of the theorem. We therefore need to prove the claim.

Proof. For $\sigma \in Aut(G)$, let P be the corresponding permutation matrix. Then

$$(P^{-1}MP)_{\sigma(i)\sigma(j)} = M_{ij}$$

for any $M \in M_n(\mathbb{R})$. Let $\Gamma = \{P_\sigma : \sigma \in Aut(G)\}$. Let A be a matrix that achieves the minimum in Theorem 7.9 (that is, $\lambda_1(A) = \vartheta$) and

$$B = |\Gamma|^{-1} \sum_{P \in \Gamma} P^{-1}AP.$$

Clearly B is symmetric and satisfies the assumptions in the claim. Indeed, for $\sigma \in Aut(G)$ and $i, j \in [n]$,

$$B_{ij} = (P_\sigma^{-1}BP_\sigma)_{\sigma(i)\sigma(j)} = |\Gamma|^{-1} \sum_{P \in \Gamma} (P_\sigma^{-1}P^{-1}APP_\sigma)_{\sigma(i)\sigma(j)}$$

which equals

$$= |\Gamma^{-1}| \sum_{P \in \Gamma} (P^{-1}AP)_{\sigma(i)\sigma(j)} = B_{\sigma(i)\sigma(j)}.$$

Therefore, the claim follows from

$$\lambda_1(B) \leq \vartheta.$$

Note that this is just a consequence from the minmax theorem that we have and the fact that all $P_\sigma^{-1}AP_\sigma$ have the same spectrum. That is,

$$\lambda_1(A) = \max_{\|x\|=1} x^t Bx$$

which equals

$$= \max x^t(|\Gamma|^{-1} \sum_{P \in \Gamma} P^{-1}AP)x \leq |\Gamma|^{-1} \max x^t P^{-1}APx = \vartheta.$$

This completes the proof. □

□

Application: cycles of odd length (all are edge-transitive so it follows immediately after calculating the eigenvalues).

Application: Peterson graph. Here the upper bound is also the independence number so it is tight.

7.2 Discrepency and Beck-Fiala

Let's continue with more examples of attacking combinatorial problems by looking at related real problems.

Recall that a hypergraph H is just a collection of subsets of vertices (not necessarily of size 2). If all hyperedges are of size k, then H is called k-uniform.

Definition 7.12. *The* discrepancy *of H is*

$$disc(H) = \min_{V=L \cup R} \max_{E \in E(H)} ||L \cap E| - |R \cap E|| .$$

WRITE IT IN TERMS OF ±1 ASSIGMNMENTS TO VERTICES.

To understand the definition it may be useful to consider few simple examples:

- Imagine that n points in the plane are given. For every line ℓ which is parallel to one of the axis, let E_ℓ be the hyperedge consisting of all the given points which lie on ℓ. This gives a hypergraph H. Show that $\text{disc}(H) \leq 1$.
- for a given connected graph G, let H be the hypergraph on vertex set $E(G)$ be the vertex stars of G. Show that $disc(H) = 1$ if some vertex of G has odd degree. $disc(H) = 0$ if all degrees are even and $|E(G)|$ is even. 2 if all degrees are even and $|E(G)|$ is odd.

Let's give an alternative definition for the discrepancy that will sometimes be easier to work with and which leads to some nice generalizations of the problem. The *discrepancy* of a real matrix M with columns indexed by a set V is

$$\text{disc}(M) = \min_{\varepsilon \in \{\pm 1\}^V} \|M\varepsilon\|_\infty.$$

Note that if M is the incidence matrix of H then the above definition is exactly $disc(H)$. As the problem of computing the discrepancy of a specific H is quite hard, it make sense to restrict ourselves to certain families. Aa natural thing to do is to consider hypergraphs with some bounds on their maximal degree. Having said that, for a positive integer t, let

$$f(t) = \sup\{disc(H) : \Delta(H) \leq 2\}.$$

Is it clear that $f(t) < \infty$?? Can you compute $f(t)$? what about $f(3)$?

Theorem 7.13. *(Beck-Fiala 1981) For every positive t we have $f(t) \leq 2t - 1$.*

The actual conjecture though is that $f(t) = O(\sqrt{t})$.

Before proving the above theorem, let us mention one of the most famous conjectures in this area

Conjecture 7.14 (Komlós)**.** *There exists a consitant $K > 0$ such that every matrix M with all columns of Euclidean length at most 1 has discrepency at most K.*

MAYBE WORTH MENTIONING THE 6 S.D THEOREM OF SPENCER.

Proof. Let M be the incidence matrix of H. We want to show that there exists $\varepsilon \in \{\pm 1\}^n$ with

$$\|M\varepsilon\|_\infty \leq 2t - 1.$$

We obtain ε via a sequence of approximations in $[-1, 1]^n$. We start with

$$\varepsilon^0 = 0,$$

which clearly gives $\|M\varepsilon\| = 0$. We try to reach a corner of the cube without making too much damage. We do this a step at a time. At the end of step i we will have ε^i and define

$$S^i = \{j \in [n] : \varepsilon(j)^i \in (-1, 1)\},$$

$$T^i = [n] \setminus S^i,$$

and

$$H^i = \{E \in E(H) : |E \cap S^i| > t\}.$$

It will be convenient to think about S^i as the set of *live* variables, and H^i is the set of live edges. Define M^i to be the submatrix of M indexed by $H^i \times S^i$, and we wish to maintain the following properties:

1. $\varepsilon^i \equiv \varepsilon^{i-1}$ on T^{i-1};
2. $T^i \supset T^{i-1}$;
3. $E \in H^i$ implies $\sum_{j \in E} \varepsilon^i_j = 0$.

2. means that we are actually making progress and 3. means that as long as an edge is alive, it is perfectly balanced.

Assume we can maintain the above properties, we stop when there are no more live variables. That is, we stop at stage $k = \min\{i : T^i = [n]\}$, and set $\varepsilon = \varepsilon^k$.

Claim 7.15. *ε is the desired.*

Intuitively it's obvious. An edge is perfectly balanced until it is no longer alive. At this point it has at most t live variables and changing the values of these variables can affect by less than ±$2t$. Making it formal is also easy and is left as an exercise.

It thus remains to show that we can indeed maintain Properties $1 - 3$. That is, given ε^{i-1} and the associated $S^{i-1} \neq \emptyset$, T^{i-1} and H^{i-1} satisfying $1 - 3$ (with i replaced by $i - 1$), we wish to show that there is an ε^i that also has these properties.

The main point (which is just a relatively trivial observation) is that

$$|H^{i-1}| < |S^{i-1}|.$$

Indeed, this follows immediately since M^{i-1} has more columns than rows (column sums are at most t by the degree assumption, and row sums are greater than t by definition of H^i). It follows that there is some $y^i \in \mathbb{R}^n \setminus \{0\}$ with

$$\mathit{support}(y^i) \subseteq S^{i-1}$$

and

$$\sum_{j \in E} y^i_j = 0 \ \forall E \in H^{i-1}.$$

(that is, y is in the orthogonal complement of the row space).

Take α to be the smallest positive number for which there is some $j \in S^{i-1}$ with

$$\varepsilon^{i-1}_j + \alpha y^i_j \in \{\pm 1\}$$

and set

$$\varepsilon^i = \varepsilon^{i-1} + \alpha y^i.$$

Intuitively, we are choosing a direction that does not involve coordinates in T^{i-1} and, starting from ε^{i-1}, follow y^i until we hit some face of $[-1,1]^n$ that was not among the facets containing ε^{i-1}.

We now need to check that ε^i satisfies the properties. Note that 1 and 2 are trivial. For 3, note that for all $E \in H^{i-1}$,

$$\sum_{j \in E} \varepsilon^i_j = \sum_{j \in E} \varepsilon^{i-1} + \alpha \sum_{j \in E} y^i_j.$$

This is clearly 0. Indeed, the second sum is 0 by the choice of y, and the first is 0 because $E \in H^{i-2}$ and induction. □

7.3 A vector balancing problem

Given any norm, one can define its *unit ball*

$$B_1 = \{x \in \mathbb{R}^d : \|x\| \leq 1\}.$$

Proposition 7.16. *The unit ball of any norm is a compact, convex set with the origin in its interior.*

The following proposition can be see as the 'reverse' implication.

Proposition 7.17. *Let $K \subseteq \mathbb{R}^d$ be compact, convex set with the origin in its interior. Define*

$$\|x\| = \min\{\alpha > 0 : \alpha^{-1} x \in K\},$$

and $\|0\| = 0$. Then $\|\cdot\|$ is a norm.

The only interesting part of the above proposition is the triangle inequality and is left as an exercise.

A set K of special interest to us is the *regular simplex* centered at the origin of $\mathbb{R}^d$. For simplicity of representation, we can add a dimension and take

$$K = \{x \in \mathbb{R}^{d+1} : x_i \geq 0, \sum_i x_i = 1\}.$$

This is a d-dimensional simplex contained in the affine hyperplane $\sum x_i = 1$. It has a center $c = \frac{1}{d+1}(1, 1, \ldots, 1)$ which plays the role of 0, and vertices $e_0, \ldots, e_d$. Note that the corresponding norm is quite 'asymmetric'. For example, the vector $v_i = e_i - c$ has norm 1 (WHY?) ad its negative has norm d WHY?

Before stating the theorem we would like to prove in this section, let us start with the following:

Theorem 7.18 (Steinitz 1914)**.** *For any d, there is a c_d such that: for any norm on $\mathbb{R}^d$ and $v_1, \ldots, v_n$ with $\|v_i\| \leq 1$ and $\sum v_i = 0$ there exists a permutation $\sigma \in S_n$ for which*

$$\|\sum_{i=1}^{t} v_{\sigma(i)}\| \leq c_d \ \forall t \in [n].$$

That is, one can reorder the vectors such that all initial sums are small (depends on d, not on the number of vectors). Note that if we don't assume that they sum to 0 so clearly it cannot be true WHY?. Knowing this theorem, what is the next most natural question to ask? Hint: what is the minimal possible c_d? Grinberg and Sevastyanov found in 1979 the precise answer:

Theorem 7.19. *The above theorem is true with $c_d = d$.*

To show that it is best possible take the norm with unit ball a regular simplex K. Indeed, let $n = d + 1$ and take $v_1, \ldots, v_n$ to be the vertices of K, then no matter how we choose σ we have

$$\|\sum_{i=1}^{n-1} v_{\sigma(i)}\| = \| - v_{\sigma(n)}\| = d.$$

We thus only need to prove the upper bound. As you will see shortly, the proof is based on the same principle as in the Beck-Fiala theorem so maybe it is about time to write the hidden argument in a more formal way.

A *linear constraint* on the real variables $x_1, \ldots, x_n$ is an inequality of one of the forms

$$a \cdot x \leq b, \ a \cdot x \geq b, \ a \cdot x = b,$$

where $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$. Note that the last two can be covered by the first ($\leq$) so general set of constraints can be put in the form

$$a_i x \leq b_i, \ i \in [m],$$

or equivalently

$$Ax \leq b.$$

The *rank* of the set of constraints indexed by $I \subseteq [m]$ is the rank of A_I (the submatrix with the corresponding rows).

A constraint $a \cdot x \leq b$ is *saturated* if it holds with equality at x. The following lemma, which is getting close to the basic ideas of linear programming, says that if the system $Ax \leq b$ can be satisfied, then it can be satisfied with 'many' constraints saturated.

Lemma 7.20. *If there exists x for which $Ax \leq b$, then there is such x for which the rank of the saturated constraints is the rank of A.*

The proof is based on the same idea as the proof of the Beck-Fiala theorem so we won't do it here (think about it as an exercise).

Now we turn into the proof of Theorem 7.19.

Proof. It will be convenient to start with the full set $V_n = \{v_1, \dots, v_n\}$ and remove one element at a time, producing a sequence

$$V_n \supset V_{n-1} \supset \dots \supset V_d.$$

We then take $v_{\sigma(i)}$ to be the unique vector in $V_i \backslash V_{i-1}$ so that $V_i = \{v_{\sigma(1)}, \dots, v_{\sigma(i)}\}$ and our condition is

$$\| \sum_{v \in V_t} v \| \leq d, \ \forall t.$$

Note that we don't care what happens beyond V_d as the norms are upper bounded by 1. The key idea is to find the 'correct' hypothesis that enables us to continue.

Lemma 7.21. *There are $V_n, \dots, V_d$ and $\lambda_i : V_i \to [0,1]$ for $i = n, \dots, d$ satisfying*

$$\sum_{v \in V_i} \lambda_i(v) v = 0$$

and

$$\sum_{v \in V_i} \lambda_i(v) = i - d.$$

Note that the above lemma gives us the desired. Indeed,

$$\| \sum_{v \in V_t} v \| = \| \sum_{v \in V_t} (v - \lambda_t(v) v) \| \leq \sum_{v \in V_t} (1 - \lambda_t(v)) \|v\| = d.$$

It thus remains to prove the lemma.

Proof. We begin with $V_n = V$ and $\lambda_n(v) = \frac{n-d}{n}$ which clearly satisfies the conditions.

Given V_i and λ_i (with $i > d$), we wish to show that we can extend to V_{i-1} and λ_{i-1}. To this end set

$$\lambda^*_{i-1} = \frac{i-1-d}{i-d} \lambda_i.$$

Then, $\lambda^*_{i-1} : V_i \to [0,1]$ satisfies

$$\sum_{v \in V_i} \lambda^*_{i-1}(v) v = 0$$

and

$$\sum_{v \in V_i} \lambda^*_{i-1}(v) = i - 1 - d.$$

Therefore, we just need to 'shrink' the domain. That is, it is enough to find $\lambda^{**}_{i-1} : V_i \to [0,1]$ satisfying the above conditions with $\lambda^{**}_{i-1}(w) = 0$ for some $w \in V_i$. We then let $V_{i-1} = V_i \setminus \{w\}$ and take λ_{i-1} to be the restrictions of λ^{**}_{i-1} to V_{i-1}.

To do so, we wish to use Lemma 7.20. Working in $\mathbb{R}^{V_i}$, we are given a solution, λ^*_{i-1} of the system

$$\sum_{v\in V_i} x(v)v = 0;$$

$$\sum_{v\in V_i} x(v) = i-1-d;$$

$$0 \leq x(v) \leq 1, \ \forall v \in V_i.$$

This is a system with rank $i = |V_i|$ (because of the constraints $x(v) \geq 0$ which are linearly independent). Therefore, by Lemma 7.20 there is some solution, λ^{**}_{i-1}, saturating at least i linearly independent constraints. Note that there are at most $d+1$ linearly independent constraints in the first two conditions, therefore, λ^{**}_{i-1} must saturate at least $i-1-d$ of the constraints in the third condition. This already gives as a $w \in V_i$ with $\lambda^{**}_{i-1}(w) = 0$. Indeed, assume otherwise, then there are $i-1-d$ w's with $\lambda^{**}_{i-1}(w) = 1$. Therefore we already have $\sum x(w) = i-1-d$ so the sum of the rest is 0 (in particular, as all of them non negative, there is one vanishing term). □

This completes the proof. □

7.4 Baranayai's Theorem

As was promised in class, in this section we (finally) going to show that the complete k-uniform hypergraph admits a 1-factorization, up to some trivial divisibility conditions. This result was actually the original inspiration for the Beck-Fiala theorem. We will make use of the following consequence of Lemma 7.20.

Lemma 7.22. *For any real $m \times n$ matrix M with integer row and column sums, there is an integer $m \times n$ matrix M' having the same row and column sums as M and satisfying:*

$$|m_{ij} - m'_{ij}| < 1, \ \forall i,j.$$

Proof. We prove it by induction on $m+n$. We're looking for an integer solution for the system

$$\sum_j x_{ij} = \sum_j m_{ij} \ \forall i;$$

$$\sum_i x_{ij} = \sum_i m_{ij} \ \forall j;$$

$$\lfloor m_{ij} \rfloor \leq x_{ij} \leq \lceil m_{ij} \rceil \ \forall i,j.$$

The individual constraints show that the rank of the system is mn and the m_{ij} themselves provide a real solution. Therefore, by Lemma 7.20, there is a solution, say x, where the rank of the saturated constraints is mn. Since the rank of the first two conditions of constraints is $m+n-1$ (DO YOU SEE WHY?), x must saturate at least $mn-m-n+1$ linearly independent constraints from the third condition. That is, at most $m+n-1$ of the x_{ij}'s are not integers. Now, note that we can finish it by induction if all the x_{ij}'s in some *line* (that is, same row or column) are integers (then just ignore this line and apply induction). On the other hand, clearly no line can have exactly one noninteger entry, so we may assume that each line has at least two. But this says that at least $m+n$ of the x_{ij}'s are not integers, which is a contradiction. □

Let us restate the problem that we are interested at. Is it true that whenever $k \mid n$, then there is a partition of H_n^k into perfect matchings? Note that numerically it makes sense as

$$\frac{\binom{n}{k}}{n/k} = \binom{n-1}{k-1}.$$

Theorem 7.23. *[Baranayai, 1973] yes!*

To begin, consider a more general problem. Given n and nonnegative integers $k_1, \ldots, k_r$, set H_i to be the complete k_i-uniform hypergraph on vertex set $[n]$.

Question 7.24. *For which $k_1, \ldots, k_r$ and n is there a partition of $\cup H_i$ into perfect matchings?*

Note that the above union is a multiset union (in case that there are k_i's which are the same).

Suppose we do have such a partition into perfect matchings $\mathcal{F}_1, \ldots, \mathcal{F}_\ell$. For each i, j set

$$|H_i \cap \mathcal{F}_j| = \alpha_{ij}.$$

Then, the α_{ij}'s satisfy

$$\sum_j \alpha_{ij} = \binom{n}{k_i} \ \forall i;$$

and

$$\sum_i \alpha_{ij} k_i = n \ \forall j.$$

So, a reasonable definition of 'numerical feasibility' might be the existence of α_{ij} satisfying these two conditions. Apparently, this is enough!

Theorem 7.25 (Baranayai 1973)**.** *For any n and $k_1, \ldots, k_r$, if there are nonnegative integers α_{ij} satisfying the above conditions, then there is a partition as in the question (with intersections α_{ij} as defined above).*

Easy exercise: assuming that and prove Theorem 7.23.

Proof. We proceed by induction on n (base case $n = 0$ is trivial). A useful trick that we're going to use is the following (essentially, dividing the edges of each H_i into those that contain n and those which don't): For $i = 1, \ldots, r$ define

$$H_i' = \{A \setminus \{n\} : n \in A \in H_i\},$$

and

$$H_i'' = \{A \in H_i : n \notin A\}.$$

We also set $k_i' = k_i - 1$ and $k_i'' = k_i$, so that H_i' is a copy of $\binom{[n-1]}{k_i'}$ and H_i'' is a copy of $\binom{[n-1]}{k_i''}$. We thus have a new instance of the problem with n replaced by $n-1$ and $H_1, \ldots, H_r$ by $H_1', \ldots, H_r', H_1'', \ldots, H_r''$. Observe that a solution to the original problem is the same thing as a partition of $\bigcup_i H_i' \cup \bigcup H_i''$ into perfect matchings of $[n-1]$, each containing exactly one element of $\cup H_i'$ (so we can add n to this unique element).

With a bit more details, we should find α'_{ij} and α''_{ij} satisfying the appropriate modification of the 'numerical feasible' definition we had so that finding $\mathcal{F}^*_j$'s with

$$|H'_i \cap \mathcal{F}^*_j| = \alpha'_{ij}$$

and

$$|H''_i \cap \mathcal{F}^*_j| = \alpha''_{ij}$$

is equivalent to finding $\mathcal{F}_j$'s with

$$|H_i \cap \mathcal{F}_j| = \alpha_{ij}.$$

Formally, we have

Claim 7.26. *To prove the theorem it is enough to show that there are $\alpha'_{ij} \in \{0,1\}$ satisfying*

$$\sum_j \alpha'_{ij} = \binom{n-1}{k_i - 1} \ \forall i;$$

$$\sum_i \alpha'_{ij} = 1 \ \forall j;$$

and

$$\alpha_{ij} = 0 \Rightarrow \alpha'_{ij} = 0 \ \forall i, j.$$

Let's prove the claim.

Proof. Given α_{ij} as described, set

$$\alpha''_{ij} = \alpha_{ij} - \alpha'_{ij}.$$

Then,

$$\alpha''_{ij} \geq 0$$

WHY? and we have

$$\sum_j \alpha''_{ij} = \sum_j \alpha_{ij} - \sum_j \alpha'_{ij} = \binom{n}{k_i} - \binom{n-1}{k_i - 1} = \binom{n-1}{k''_i}$$

for all i, and

$$\sum_i \alpha'_{ij} k'_i + \sum_i \alpha''_{ij} k''_i = \sum_i \alpha_{ij} k_i - \sum_i \alpha'_{ij} = n - 1 \ \forall j.$$

The existence of the desired $\mathcal{F}^*_j$'s now follows by induction as the α' and α'' satisfy the condition of the theorem with $n-1$.

Observe that according to the assumption of the claim, we do have the desired property that each $\mathcal{F}^*_j$ contains exactly one set from $\cup H'_i$, and therefore by adding the element n back to it, one obtain the desired $\mathcal{F}_j$'s. □

Finally, we need to prove that it is possible to find α'_{ij}'s as in the claim. For every i, j let us define

$$x_{ij} = \frac{k_i}{n} \alpha_{ij}.$$

Then,

$$\sum_j x_{ij} = \frac{k_i}{n}\sum_j \alpha_{ij} = \frac{k_i}{n}\binom{n}{k_i} = \binom{n-1}{k_i-1};$$

$$\sum_i x_{ij} = \sum_i \alpha_{ij}k_i/n = 1;$$

$$\alpha_{ij} = 0 \Rightarrow x_{ij} = 0;$$

and

$$x_{ij} \in [0,1].$$

The existence of α'_{ij} now just follows by Lemma 7.22.

□

8 A brief introduction to the polynomial method

Let $\mathbb{F}$ be a field. Let $P_D(\mathbb{F}^n)$ be the space of polynomials in n variables over $\mathbb{F}$ with all monomials of degree at most D. Observe that $P_D(\mathbb{F}^n)$ is a vector space over $\mathbb{F}$ (and a sub vector space of $\mathbb{F}[x_1,\dots,x_n]$). Suppose that $S \subseteq \mathbb{F}^n$ is a finite set. We would like to know if there is a non zero polynomial $P \in P_D$ that vanishes over S. As you can guess, we can do it using some dimensional arguments.

Proposition 8.1. *If $dim(P_D) > |S|$, then there is a non-zero polynomial $P \in P_D$ which vanishes on S.*

Proof. Suppose that $S = \{t_1,\dots,t_{|S|}\}$ and define $f : P_D \to \mathbb{F}^{|S|}$ as follows:

$$f(P) = (P(t_1),\dots,P(t_{|S|})).$$

Observe that f is a linear map. Note that the kernel of f consists of all $P \in P_D$ for which P vanishes on S. In order to complete the proof, recall from linear algebra that for a linear transformation $L : V \to U$ we have

$$dim(V) = dim(Ker(L)) + dim(Im(L)),$$

which in our case translates to

$$dim(P_D) \le dim(Ker(f)) + |S|.$$

Finally, as $dim(P_D) > |S|$ we obtain that $Ker(f) \neq 0$. This completes the proof. □

A natural question now to ask is: what is the dimension of P_D? Note that a basis for P_D is given by all monomial of the form

$$\prod_i x_i^{t_i},$$

with $\sum_i t_i \le D$. In particular, by counting solutions to

$$\sum_{i=1}^{n} t_i \le D$$

we obtain

Lemma 8.2. *The dimension of P_D is $\binom{D+n}{n}$. In particular, we have*

$$dim(P_D) \geq \frac{D^n}{n!}.$$

As an immediate corollary from the previous two lemmas we obtain

Corollary 8.3. *Is $S \subseteq \mathbb{F}^n$ is of size $|S| < \binom{D+n}{n}$, then there exists a non-zero $P \in P_D$ which vanishes on S.*

Or, a little bit less sharp but easier to work with, we have

Corollary 8.4. *For any finite set $S \subseteq \mathbb{F}^n$, there exists a non-zero polynomial P that vanishes on S with degree at most $n|S|^{1/n}$.*

Proof. Choose $D = n|S|^{1/n}$ and observe that

$$\binom{D+n}{n} > |S|.$$

Now apply the previous corollary. □

The following lemma plays a central role in what we will see today.

Lemma 8.5. *If $P \in P_D(\mathbb{F})$ (that is, a polynomial of one variable) and if P vanishes at $D+1$ points, then P is the 0 polynomial.*

A *line* ℓ in $\mathbb{F}^n$ is a 1-dimensional affine subspace.

Lemma 8.6 (The vanishing lemma)**.** *If $P \in P_D(\mathbb{F}^n)$ and P vanishes over $D+1$ points on some line ℓ, then P vanishes at every point of ℓ.*

Proof. We can write $\ell(t) = at + b$ for some vectors $a, b \in \mathbb{F}^n$ with $a \neq 0$. Then, define

$$Q(t) = P(\ell(t)).$$

Clearly, Q is a polynomial of degree at most D with only one variable that vanishes on $D+1$ points, so by previous lemma it is the zero polynomial. □

8.1 The finite-field Nikodym problem

Let $\mathbb{F}_q$ be a finite field with q elements. A set $N \subset \mathbb{F}_q^n$ is called a *Nikodym set* if, for every points $x \in \mathbb{F}_q^n$ there is a line ℓ_x containing x so that $\ell_x \setminus \{x\} \subseteq N$. A trivial example for such a set is $\mathbb{F}_q^n$. Can one find a significantly smaller Nikodym set?

Theorem 8.7 (Dvir, 2009)**.** *Any Nikodym set is of size at least $c_n q^n$, with $c_n = (10n)^{-n}$.*

Proof. Suppose that N is a Nikodym set with $|N| < c_n q^n$. By Corollary 8.4 can find a non-zero polynomial P that vanishes on N with degree bounded by

$$2n|N|^{1/n} < q - 1.$$

Next, we claim that P vanishes at any point of $\mathbb{F}_q^n$ which is clearly absurd (the degree is smaller than q so it has at most $q-1$ 0's. Don't get confused with the fact that $x^q - x$ is 0 on $\mathbb{F}_q$, we are talking about a 'reduced' polynomial). Let x be any point in $\mathbb{F}_q^n$. By definition, there exists a line ℓ_x containing x for which $\ell_x \subset N$. The polynomial P vanishes on $\ell_x \setminus x$, so it vanishes on at least $q-1$ points of ℓ_x. Since $deg(P) < p-1$, it means that P vanishes on ℓ_x, so in particular, $P(x) = 0$. □

8.2 The finite field Kakeya problem

A set $K \subseteq \mathbb{F}_q^n$ is called a *Kakeya set* if it contains a line in every direction. In other words, for every vector $a \in \mathbb{F}_q^n \setminus \{0\}$ there is a vector b so that the line $at + b$ is fully contained in K. Again, a trivial example for a Kakeya set is the whole space, and we should ask for how small does a Kakeya set can be?

Theorem 8.8 (Dvir, 2009). *A Kakeya set has at least $c_n q^n$ many elements, where $c_n = (10n)^{-n}$.*

Proof. Suppose K is a Kakeya set with less than $c_n q^n$ elements. Therefore, there exists a non-zero polynomial $P \in P_D(\mathbb{F}_q^n)$ that vanishes on K, for $D \leq n|K|^{1/n} < q$. Write P as a sum of two polynomial $P = Q + S$ where Q is the set of all monomial of max degree in P (say D). Now, let a be any non-zero vector in $\mathbb{F}_q^n$ and choose b in such a way that the line $at + b$ is contained in K. Consider the polynomial in one variable $R(t) = P(at + b)$. This polynomial vanishes on the line but has degree smaller than q. Therefore, R is the 0 polynomial. That is, all coefficients of R are 0. BUT, the coefficient in R of t^d is exactly $Q(a)$. So we see that Q vanishes for all $a \in \mathbb{F}_q^n \setminus \{0\}$. Since Q is homogeneous of degree $D \geq 1$, Q also vanishes at 0 and therefore is the 0 polynomial. This gives a contradiction. □

8.3 The joints problem

Let $\mathcal{L}$ be a collection of lines in $\mathbb{R}^3$. A *joint* of $\mathcal{L}$ is a point which lies in three non-coplanar lines of $\mathcal{L}$. The joint problem asks what is the maximal number of joints that can be formed from L lines?

This problem was posed in the 90's by Chazelle, Edelsbrunner, Guibas, Pollack, Seidel, Sharir and Snoeyink. They proved that the number of joints formed by L lines is at most (around) $L^{7/4}$.

Let's look at few examples.

- Consider an $S \times S \times S$ grid of points, and let $\mathcal{L}$ be the set of all axis-parallel lines that intersect the grid. The number of lines in $\mathcal{L}$ is $3|S|^2 = L$. Each point in the grid is a joint for $\mathcal{L}$, so there are S^3 joints. Therefore, the number of joints is roughly $L^{3/2}$.
- Consider the edges of a tetrahedron. A tetrahedron has six edges and for vertices. Each vertex lies in three non-coplanar edges, and so tetrahedron gives a set of six lines with four joints. This example can be generalized to large numbers of lines in the following way: Let $S \geq 3$ be some parameter. Consider S planes in $\mathbb{R}^3$ in a general position (that is, every two intersect in a line and every three intersect in a point). Let $\mathcal{L}$ be the set of all these lines. The number of lines in $\mathcal{L}$ is $L = \binom{S}{2}$. Note that as each three planes intersect in a point, every such point is a joint for $\mathcal{L}$. Therefore, $\mathcal{L}$ has $\binom{S}{3}$ joints. If we take $S = 4$, we recover the tetrahedron case. Note that in this example we still have around $L^{3/2}$ joints, but the constant is better. This is actually the best known example!

Theorem 8.9 (Guth and Katz, 2010). *Any L line in the space determine at most $10L^{3/2}$ joints.*

To prove the theorem we need the following main lemma

Lemma 8.10 (Main Lemma). *If $\mathcal{L}$ is a set of lines in $\mathbb{R}^3$ that determines J joints, then one of the lines contains at most $3J^{1/3}$ joints.*

Before proving the lemma, let's see how to use it in order to prove the theorem. Let $J(L)$ be the maximum number of joints that can be formed by L lines. If $\mathcal{L}$ is a set of L lines, then the main lemma tells us that one of the lines contains at most $J(L)^{1/3}$ many joints. The number of joints not on this line is at most $J(L-1)$, and therefore we obtain a recurrence relation:

$$J(L) \leq J(L-1) + 3J(L)^{1/3}.$$

By iterating we obtain

$$J(L) \leq L \cdot 3J(L)^{1/3},$$

which gives us

$$J(L) \leq 3^{3/2}L^{3/2} \leq 10L^{3/2}$$

as desired. Now we can prove the main lemma.

Proof. Let P be a non-zero polynomial that vanishes at every joint of $\mathcal{L}$ and with degree as small as possible. By Corollary 8.4 its degree is at most $3J^{1/3}$. Now, suppose that every line contains more than $3J^{1/3}$ joints. By the vanishing lemma, P must vanish on every line of $\mathcal{L}$. Next, we study the gradient of P at each joint of $\mathcal{L}$.

Claim 8.11. *If x is a joint of $\mathcal{L}$, and if a smooth function $F : \mathbb{R}^3 \to \mathbb{R}$ vanishes on the lines of $\mathcal{L}$, then ∇F vanishes at x.*

Proof. Note that x lies in three non-coplanar lines of $\mathcal{L}$. Let v_1, v_2, v_3 be tangent vectors for these three lines. The directional derivative of F in the direction v_i must vanish at x. So we have $\nabla F(x) \cdot v_i = 0$ for each i. Since the v_i are a basis of $\mathbb{R}^3$, we have $\nabla F(x) = 0$. □

By Claim, we see that the derivatives of P vanish at each joint. The derivatives have smaller degree than P. Since P was a minimal degree non-zero polynomial that vanishes at each joint, each derivative must be identically 0. Then P must be constant. Since P is non-zero, it gives us a contradiction. □

8.4 The capset problem

A *capset* in the vector space $\mathbb{Z}_3^n$ over $\mathbb{Z}_3$ is a collection of vectors where no three of them lie on the same line. That is, for no x, y, z there exists r for which $\{x, y, z\} = \{x, x+r, x+2r\}$ and $r \neq 0$. A basic problem in combinatorics is to determine the size of the largest capset.

Trivially, if A is a capset, then $|A| \leq 3^n$. Using Fourier methods (a beautiful proof by the way), Meshulam (1995) obtained the bound $|A| = O(3^n/n)$. In 2012 this bound was improved to $O(3^n/n^{1+c})$ by Bateman and Katz (JAMS, 28 pages long paper). In a quite recent paper, Ellenberg and Gijswijt, based on a week earlier result by Croot, Lev and Pach, obtained an exponential improvement $|A| = O(2.756^n)$, using a version of the polynomial method. There is also a construction of Edel that gives a lower bound around $(2.2174)^n$.

The proof is very short (the paper is around 3 pages long) and the goal of this section is to present its proof, based on the very elegant exposition of Terry Tao from his blog. The starting point is the following: if A is a capset, then

$$\delta_0(x+y+z) = \sum_{a \in A} \delta_a(x)\delta_a(y)\delta_a(z)$$

for all $x, y, z \in A^3$. Indeed, $x + y + z = 0$ exactly when x, y, z form a line or are all equal and the former is ruled our by the assumption on A.

The basic idea is to bound the 'rank' (in some form) of the above equality. Intuitively, imagine that one can write a matrix with Ω as its row and columns, and suppose that we are restricting our attention to a principal submatrix of the form $A \times A$. If this sub matrix is invertible, then clearly the rank of the all matrix is at least as large as the size of A. On the other hand, if we can obtain a non-trivial upper bound on the rank of the full matrix, then we clearly obtain an upper bound on $|A|$.

A function $F : A \times A \to \mathbb{F}$ is of *rank one* if it is non-zero and of the form $F(x, y) = f(x)g(y)$ for some $f, g : A \to \mathbb{F}$. The *rank* of a general function is the minimum number of rank one functions needed to represent F as a linear combination of. More generally, if $k \geq 2$, we define the rank of a function $F : A^k \to \mathbb{F}$ as the least number of rank one functions of the form

$$F(x_1, \dots, x_k) = f(x_i)g(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_k)$$

that are needed to represent F as a linear combination. For example, if $k = 3$, the tank one functions take the form $f(x)g(y, z)$, $f(y)g(x, z)$, and $f(z)g(x, y)$.

It is standard from linear algebra that diagonal matrices are of full rank (assuming that all diagonal elements are non-zero). We would like to extend it to higher dimensions.

Lemma 8.12. *Let $k \geq 2$, let A be a finite set, let $\mathbb{F}$ be a field, and for each $a \in A$, let $c_a \in \mathbb{F}$ be a coefficient. Then, the rank of the function $\sum_{a \in A} c_a \delta_a(x_1) \dots \delta_a(x_k)$ is equal to the number of non-zero c_a's.*

Proof. The proof goes by induction on k, where the case $k = 2$ is standard. Suppose now that $k > 2$ and that we already know it for $k - 1$. Since every summand on the right hand side is of rank one, it is clear that the rank is at most the number of non-zero coefficients. Therefore, WLOG we can assume that all the c_a's are not zero (if few of them are, then just delete them and the corresponding elements from A). Assume towards a contradiction that the rank is at most $|A| - 1$. Then, we can write

$$\sum_{a \in A} c_a \delta_a(x_1) \dots \delta_a(x_k) = \sum_{i=1}^{k} \sum_{\alpha \in I_i} f_{i,\alpha}(x_i) g_{i,\alpha}(x \setminus \{x_i\}) \tag{5}$$

for some sets $I_1, \dots, I_k$ of cardinalities adding up to at most $|A| - 1$.

Consider the space of functions $h : A \to \mathbb{F}$ that are orthogonal to all the $f_{k,\alpha}$, $\alpha \in I_k$ in the sense that

$$\sum_{x \in A} f_{k,\alpha}(x)h(x) = 0$$

for all $\alpha \in I_k$. This is a vector space with dimension $d \geq |A| - |I_k|$. A basis for this space generates a $d \times |A|$ coordinate matrix of full rank, which implies that there is at least one non-singular $d \times d$ minor. This implies that there exists a function $h : A \to \mathbb{F}$ which is non-vanishing on some subset A' of A of cardinality at least $|A| - |I_k|$ WHY?

If we multiply (5) by $h(x_k)$ and sum in x_k, we conclude that

$$\sum_{a \in A} c_a h(a) \delta_a(x_1) \dots \delta_a(x_{k-1}) = \sum_{i=1}^{k-1} \sum_{\alpha \in I_i} f_{i,\alpha}(x_i) \bar{g}_{i,\alpha}(x - \{x_i, x_k\})$$

where $\bar{g}_{i,\alpha} = \sum_{x_k \in A} g_{i,\alpha} h(x_k)$. The right hand side has rank at most $|A| - 1 - |I_k|$, since all the summands are rank 1 functions (we assumed that the rank is at most $|A| - 1$ and deleted $|I_k|$ terms). On the other hand, by induction we know that the left hand side has rank at least $|A| - |I_k|$, which gives us the desired contradiction. This completes the proof. □

To finish off the proof we need the following (surprisingly simple!) lemma.

Lemma 8.13. *Over $\mathbb{Z}_3^n$, the function $\delta_0(x+y+z)$ has rank at most $(3-c)^n$.*

Proof. Using the identity $\delta_0 = 1 - x^2$ for $x \in \mathbb{Z}_3$, we have

$$\delta_0(x+y+z) = \prod_{i=1}^{n} \left(1 - (x_i + y_i + z_i)^2\right).$$

The right hand side is a polynomial of degree $2n$ in x, y, z, which is a linear combination of monomials of the form $(\prod x_k^{i_k})(\prod y_k^{j_k})(\prod z_k^{\ell_k})$, with all $i_k, j_k, \ell_k \in \{0,1,2\}$ and the sum $i_1 + \ldots + i_n + j_1 + \ldots + j_n + \ell_1 + \ldots + \ell_n \leq 2n$. In particular, by pigeonhole principle, at least one of the partial sums $\sum i_s$, $\sum j_s$, or $\sum \ell_s$ is at most $2n/3$.

Consider the contribution of the monomials for which $\sum i_s \leq 2n/3$. For each such possibility, one can regroup all the y, z terms together to form a rank 1 function of the form $\prod x_k^{i_k} g(y,z)$. To complete the proof, note that if we focus on the i_s for example. Then for a random choice one expects to have sum n. But, we only consider sums which are at most $2n/3$ so one can obtain an exponential improvement by using (some version of) Chernoff's bounds. □

Let's try to do sunflowers of size 3. A sunflower of size 3, is three vectors $x, y, z \in \{0,1\}^n$ for which in every coordinate either all agree or there exists exactly one 1. Therefore, if one has a sunflower free set, there must be a coordinate i where $x_i + y_i + z_i = 2$.

9 Cayley graphs

Let us first give a short backgroung necessary for this section. Let G be a finite group. The elements of a subset S of G are called *generators* of G, and S is called a *generating set* of G, if every element of G can be expressed as a finite product of elements of S. Usually, we denote the identity group element by e and the operation as a multiplication (unless stated otherwise). A subset $S \subseteq G$ is called *symmetric* if $s \in S$ implies that $s^{-1} \in S$.

Example: S_n is generated by by the transpositions $\{(1,2),(2,3),\ldots,(n-1,n)\}$. WHY?

Let $S \subseteq G$ be an identity free and symmetric subset of a finite group G. The *Cayley graph* $\Gamma = Cay(G,S)$ is a graph with vertices correspond to the elements of G, and the edges correspond to multiplication on the right. That is, $E(\Gamma) = \{\{g, gs\} : g \in G, s \in S\}$. Note that the identity free assumption excludes self loops and the symmetric assumption makes Γ an undirected graph.

A permutation σ of the vertex set of a graph Γ is called an *automorphism* if $\{u,v\} \in E(\Gamma)$ if and only if $\{\sigma(u).\sigma(v) \in E(\Gamma)\}$. A graph is called *vertex-transitive* if for any two vertices u and v, there exists an automorphism σ with $\sigma(u) = v$. Clearly, any vertex-transitive graph must be regular, and it is very simple to build examples of regular graphs which are not vertex-transitive CAN YOU SEE ONE?.

A graph is *edge-transitive* if for any two edges $xy, uv \in E(\Gamma)$, there exists an automorphism σ for which $\{\sigma(x)\sigma(y)\} = \{u,v\}$.

Exercise: try to think about a vertex-transitive graph which is not edge-transitive and about an edge-transitive graph which is not vertex-transitive.

Exercise: Petersen graph is both!

Proposition 9.1. *Let S be a set of generators for a group G. The Cayley graph $\Gamma = Cay(G, S)$ has the following properties:*

1. *Γ is a connected, regular graph of degree equal to $|S|$.*

2. *Γ is vertex transitive.*

Proof. As S is a generating set of G, it is clear that Γ is connected. Moreover, as the neighborhood of every vertex v is of the form $\{\{v, vs\} : s \in S\}$ it is clearly $|S|$-regular.

To prove vertex-transitivity, let us fix any element $g \in G$. Define a permutation σ_g of G as follows: $\sigma_g(x) = gx$. This is an automorphism as every edges $\{v, vs\}$ is being mapped to $\{gv, gvs\}$ which is also an edge of Γ. Moreover, for every $h \in G$ there exists a unique y for which $\sigma_g(y) = gy = h$, and we are done. □

Proposition 9.2. *Not every vertex-transitive graph is a Cayley graph.*

Proof. The simplest example is the Petersen graph which is a vertex-transitive graph but not a Cayley graph. To see this, recall that it has order 10, it is 3-regular, and it has a diameter 2. To convince ourselves that this is a valid counter example, one should consider all pairs (G, S) where G is a group of order 10 and the sized of S is 3. There are only two nonisomorphic groups of order 10 and by some case analysis one can convince himself that it is impossible to obtain the Petersen graph. □

Exercise: Convince yourself that for a Cayley graph, the diameter is the maximum (over $g \in G$) of the length of a shortes expression for g as a product of generators.

Let us now give some examples of Cayley graphs:

- The complete graph K_n is a Cayley graph on the additive group $\mathbb{Z}_n$ with $S = \mathbb{Z}_n \setminus \{0\}$.

- The *circulant* is the Cayley graph $Cay(\mathbb{Z}_n, S)$, where S is an arbitrary generating set. Note that if $S = \{1, -1\}$ then it is C_n.

- The *n-dimensional hypercube* Q_n is a Cayley graph. Indeed, take $G = \mathbb{Z}_2^n$ and S to be all the vectors consisting of exactly one 1.

9.1 Hamiltonicity of Cayley graphs

In this section we discuss some sufficient conditions for a Cayley graph to contain a Hamiltonian cycle. In general, testing whether a graph is Hamiltonian is an NP-hard problem, so the search for 'nice' sufficient conditions is natural.

One easy example is to show that Q_n contains a Hamiltonian cycle for all n. This can be proven by induction on n, noting that one can write Q_n as a disjoint union of two copies of Q_{n-1} plus a specific perfect matching between them (for example take all the vertices that start with a 0 to be one copy and all the vertices that start with a 1 to be the other copy.

A famous conjecture of Lovász from 1970 states:

Conjecture 9.3. *Every connected, vertex-transitive graph with more than two vertices has a Hamiltonian path.*

A weaker conjecture is also widely open:

Conjecture 9.4. *Every connected Cayley graph on a finite group has a Hamiltonian cycle.*

There is no consensus what on the correctness of the above conjectures. In particular, Babai conjectured in 1996 that:

Conjecture 9.5. *For some $\varepsilon > 0$, there exist infinitely many connected, vertex-transitive graphs (also Cayley graphs) Γ without cycles of length at least $(1-\varepsilon)|V(\Gamma)|$.*

For *abelian* groups though, it was proved by Marusic in 1983 that:

Theorem 9.6. *A Cayley graph of an abelian group with at least three vertices contains a Hamiltonian cycle.*

The proof is quite simple and is left as an exercise.

Another result which is worth mentioning is one obtained in 2009 by Pak and Radoicic and which we are going to prove shortly:

Theorem 9.7. *Every finite group with at least 3 elements has a generating set of size at most $\log_2 |G|$ such that the corresponding Cayley graph has a Hamiltonian cycle.*

The bound on the size is obtained by $G = \mathbb{Z}_2^n$ for which the size of its smallest generating set is exactly $\log_2 |G|$. The following conjecture is stronger than the above theorem (if $\varepsilon \leq 1$):

Conjecture 9.8. *There exists $\varepsilon > 0$, such that for every finite group G and every $k \geq \varepsilon \log_2 |G|$, the probability that the Cayley graph Γ, obtained by picking a random set of generators S of size k contains a Hamiltonian cycle, tends to 1 as $|G|$ goes to infinity.*

The best known bound is due to Krivelevich and Sudakov, where they proved the conjecture for $k \geq \varepsilon \log^5 |G|$.

Now we turn into the proof of Theorem 9.7. For this we need some preparation. Suppose that G is a group of size n and let $H \subset G$ be a subgroup of G. For $g \in G$ the sets

$$gH := \{gh : h \in H\} \text{ and } Hg := \{hg : h \in H\}$$

are left and right cosets of H in G. A subgroup H is called *normal* (and we denote it by $H \lhd G$) if the sets of left and right cosets of this subgroup are the same for any $g \in G$. A *simple group* is a nontrivial group which only normal subgroups are the trivial group and the group itself. A *factor group G/H* of a group G with a normal subgroup H is the set of all cosets of H such that $(aH)(bH) = (ab)H$. A *composition series* of a group G is a series such that

$$e = H_0 \lhd H_1 \lhd \ldots \lhd H_t = G,$$

where each H_i is a maximal normal subgroup of H_{i+1} for all i. Equivalently, each factor group H_{i+1}/H_i is simple. The factor groups are called *composition factors.*

Let us start with the following lemma:

Lemma 9.9. *Let G be a finite group generated by two elements α, β such that $(\alpha\beta)^2 = 1$. Then the Cayley graph $\Gamma = Cay(G, \{\alpha, \beta\})$ has a Hamiltonian cycle.*

Proof. For every $z \in G$ and every $X \subset G$, denote

$$\vartheta_z(X) = \{g \in G - X : g = xz, x \in X\}.$$

Let $H =< \beta >$, $X_1 = H$, and construct a Hamiltonian cycle in Γ by induction. At step i we obtain a cycle which spans a set $X_i \subset G$ for which $\vartheta_\beta(X_i) = \vartheta_{\beta^{-1}}(X_i) = 0$. We also assume by induction that the restriction of Γ to X_i contains a Hamilton cycle C_i which contains only labels β and α^{-1}.

The base of the induction is obvious. Namely, $\vartheta_\beta(X_1) = \vartheta_{\beta^{-1}}(X_1) = 0$ and X_1 contains a Hamiltonian cycle using only β and α^{-1} (only β actually...). For the induction step, consider $y = x\alpha \in \vartheta_\alpha(X_i) - X_i$ or $y = x\alpha^{-1} \in \vartheta_{\alpha^{-1}}(X_i) - X_i$. If such a y does not exists, then we are done. WLOG assume that $y = x\alpha$ (not exactly WLOG, but the second case is similar). Note that the edge oriented towards $x \in X_i$ in C_i cannot have label α^{-1} (as otherwise it is $\{y, x\}$ but $y \notin X_i$), and also not the labels α or β^{-1} (by assumption). Therefore, this edge has label β, and $\{x\beta^{-1}, x\} \in C_i$. Now, consider a cycle R on yH with labels β on all edges, and observe that

$$x \to x\alpha = y \to x\alpha\beta = y\beta \to x\beta^{-1} = x\alpha\beta\alpha \to x$$

is a square which connects R and C_i. Formally, let

$$C_{i+1} = C_i \cup R + \{x, y\} + \{y\beta, x\beta^{-1}\} - \{x\beta^{-1}, x\} - \{y, y\beta\}$$

and observe that C_i connects C_i to R into a Hamiltonian cycle of $X_{i+1} = X_i \cup yH$. Let it inherit the orientation from C_i and check that it satisfies the conditions with respect to the orientation. □

We also need the following lemma:

Lemma 9.10. *Let G be a finite group and let $H \triangleleft G$ be a normal subgroup. Suppose $S = S_1 \cup S_2$ is a generating set of G such that $S_1 \subset H_1$ generate H, and the projection S_2' of S_2 onto G/H generates G/H. Suppose both $\Gamma_1 = Cay(H, S_1)$ and $\Gamma_2 = Cay(G/H, S_2')$ contain Hamiltonian paths. Then, $\Gamma = Cay(G, S)$ also contains a Hamiltonian path.*

Proof. Let $k = |G/H|$ and let $g_1 = e \in G$. Consider a Hamiltonian path in the Cayley graph Γ_2:

$$H := Hg_1 \to Hg_2 \to \ldots \to Hg_k.$$

Now proceed by induction. Fix a Hamiltonian path in the coset Hg_1 so that $1 \in G$ is its starting point. Suppose h_1g_1 is its endpoint. Add an edge $\{h_1g_1, h_1g_2\} \in \Gamma$ (such an edge exists by the assumption that Γ_2 is hamiltonian). Suppose h_2g_2 is its endpoint, and repeat until getting a path that ends at h_kg_k. This completes the proof. □

Let $\ell(G)$ be the number of composition factors of G. Let $r(G)$ and $m(G)$ be the number of abelian and non-abelian composition factors, respectively. Clearly, $\ell(G) = r(G) + m(G)$.

Theorem 9.11. *Let G be a finite group and let $r(G)$ and $m(G)$ as above. Then, there exists a generating set S of G with $|S| \leq r(G) + 2m(G)$ such that the corresponding Cayley graph $\Gamma = Cay(G, S)$ contains a Hamiltonian path.*

Proof. It is well known (wasn't to me if you wonder..) that every non-abelian finite simple group can be generated by two elements, one of which is an involution (that is $x^2 = 1$). Therefore, Lemma 9.9 applies (with $S = \{\alpha^{-1}, \alpha\beta\}$) and gives a generating set of size 2 for which the corresponding Cayley graph is Hamiltonian. If the group is cyclic though (that is, the abelian case), then it trivially has such a cycle.

Now we want to use Lemma 9.10. Observe that (using the notation from the lemma) any generating set S_2' of G/H can be lifted to $S_2 \subset G$, so that $S = S_1 \cup S_2$ is a generating set of G. Therefore, if H and G/H have generating sets of sizes k_1 and k_2, respectively, so that the corresponding Cayley graphs contain Hamiltonian paths, then G contains such a generating set of size $k_1 + k_2$. To complete the proof, fix any composition series of a finite group G. By Lemma 9.10 we can construct a generating set of size $r(G) + 2m(G)$, so that the corresponding Cayley graph has a Hamiltonian path. This completes the proof. □

Now we are ready to prove Theorem 9.7.

Proof. Fix a composition of G. Let r, m as before, and let $K_1, \dots, K_r, L_1, \dots, L_m$ be the abelian and non-abelian composition factors of G, respectively. Since the smallest simple non-abelian group has order 60 (CHECK IN WIKIPEDIA IF YOU DON'T REMEMBER IT), then $|L_j| \geq 60 > 4$ for all $j \in [m]$. Therefore, we have

$$2^{r+2m} = 2^r 4^m \leq \prod_{i=1}^{r} |K_i| \prod_{j=1}^{m} |L_i| = |G|.$$

Therefore, $r + 2m \leq \log_2 |G|$, and equality actually holds only if $G = \mathbb{Z}_2^n$. Note that in the latter it's easy to prove existence of a Hamilton cycle by induction, and in any other case one can add one more generator in order to close the Hamiltonian path into a cycle. This completes the proof. □

9.2 Eigenvalues of Cayley graphs

In this section, how not, we are going to consider the problem of finding the eigenvalues of a Cayley graph from a finite abelian group. In this case, as we will see, the eigenvalues are the *characters* of the group, and the eigenvalues have a very simple description. Before stating the results we are interested at, let us give a brief reminder of characters.

Definition 9.12. *A* character *of a finite abelian group G is a group homomorphism $\chi : G \to \mathbb{C}$.*

As we usually write groups multiplicatively, we have $\chi(gh) = \chi(g)\chi(h)$, and $\chi(1) = 1$.

Exercise 9.13. *It might be useful to know the following few facts:*

- *If χ and ϕ are characters, then so $\bar{\chi}$, $\chi \cdot \phi$ and $\chi \cdot \bar{\phi}$.*
- *If G is a finite group and χ is a character for G, then $|\chi(g)| = 1$ for all $g \in G$. (In particular, one can view χ : as a function from G to the unit sphere of $\mathbb{C}$.)*
- *If χ is a character of a finite group G and χ is not the identically 1 function, then $\sum_{g \in G} \chi(g) = 0$. (Note that one can view $\chi(g)$ as a random variable, where g is being chosen uniformly at random from G. Then, we basically show that $\mathbb{E}\chi = \frac{1}{|G|} \sum_{g \in G} \chi(g) = 0$.)*

- *The set of all characters of a finite group G is orthonormal (the inner product is defined as $<\chi,\phi>=\frac{1}{|G|}\sum_{g\in G}\chi(g)\phi(g)$.)*

  Indeed, $\sum_{g\in G}\chi(g)\bar{\chi(g)}=\sum_{g\in G}|\chi(g)|=|G|$. Moreover, $\chi\cdot\bar{\phi}$ is a character which is not identically 1*, and therefore, $<\chi,\phi>=\mathbb{E}[\chi\cdot\bar{\phi}]=0$*

Let us also prove the following useful-to-know lemma.

Lemma 9.14. *Let G be a finite cyclic group of size n with a chosen generator g. Then, there are exactly n characters of G, each determined by sending g to a different nth root of unity.*

Proof. Since g generates G, a character is determined by its value on g, which is clearly root of unity. Therefore, there are at most n characters (could also obtain it from the fact that they are linearly independent). As clearly all characters $\chi_j(g)=e^{2j\pi i/n}$ are distinct, we are done. □

Next we see how to determine the eigenvalues/eigenvectors all Cayley graphs of any finite abelian group.

Lemma 9.15. *Let G be a finite abelian group, $\chi : G\to\mathbb{C}$ be a character of G, $S\subseteq G$ be a symmetric set. Let M be the normalized adjacency matrix of the Cayley graph $\Gamma=Cay(G,S)$. Consider the vector $x\in\mathbb{C}^G$ such that $x_a=\chi(a)$. Then, x is an eigenvector of G with eigenvalue*

$$\frac{1}{|S|}\sum_{s\in S}\chi(s).$$

Proof. Consider the a-th entry of Mx:

$$(Mx)_a=\sum_b M_{a,b}x_b=\frac{1}{|S|}\sum_{b:b\cdot a^{-1}\in S}\chi(b)=\frac{1}{|S|}\sum_{s\in S}\chi(as)=x_a\cdot\frac{1}{|S|}\sum_{s\in S}\chi(s).$$

The eigenvalues are of the form

$$\frac{1}{|S|}\sum_{s\in S}\chi(s)$$

where χ is a character. This completes the proof. □

Observe that

- Every character is an eigenvector;
- The characters are linearly independent (as functions from G to $\mathbb{C}$, or equivalently, as vectors in $\mathbb{C}^G$);
- There are as many characters as group elements, and so many characters as nodes in the corresponding Cayley graphs.

To conclude, we have a 'recipe' how to find all of the eigenvalues/eigenvectors.
Let's see two examples, one we already know (hopefully...) and the other is new to us:

- The cycle: The cycle C_n is a Cayley graph of $\mathbb{Z}_n$ with $S=\{-1,1\}$. Recall that every $m\in\{0,\ldots,n-1\}$ has a character $\chi_m(x)=e^{2\pi imx/n}$. This means that for every m we have the eigenvalue

$$\lambda_m=\frac{1}{2}e^{2\pi im/n}+\frac{1}{2}e^{-2\pi im/n}=cos(2\pi m/n).$$

- The hypercube: For every $r \in \mathbb{Z}_2^n$ we have a character

  $$\chi_r : \mathbb{Z}_2^n \to \{-1, 1\},$$

  defined as

  $$\chi_r(x) = (-1)^{\sum_i r_i x_i}.$$

  Let S be the set of the generators $e_1, \ldots, e_n$ (CHECK THAT THE ABOVE ARE INDEED CHARACTERS!). This means that, for every $r \in \mathbb{Z}_2^n$, the hypercube has the eigenvalue

  $$\frac{1}{n}\sum_j \chi_r(e_j) = \frac{1}{n}\sum(-1)^{r_j} = \frac{1}{n}(n-2s),$$

  where s is the number of 1's in r.

  Corresponding to $r = 0$ we have the eigenvalue 1. For each of the n vectors with exactly one 1, we have the eigenvalue $2/n$, and more. In particular, the multiplicity of $n - 2s$ is $\binom{n}{2s}$.

9.3 A random set of generators

To continue the fun, let's also add some probability into the picture. Our goal here is to show that if we choose the set of generators at random, for some constant multiplication of the dimension, then we obtain a graph which is a 'good' approximation of the complete graph (in a spectral way). Let us set, say, $k = 100d$ (where d is the dimension. It will be convenient here to denote $n = 2^d$ so we can compare the bound to those of K_n). For a vector $b \in \{0,1\}^d$ which is not all 0 and for g which is chosen uniformly at random from $\{0,1\}^d$, $b^t g$ mod 2 is a uniformly distributed number in $\{0,1\}$, and so

$$(-1)^{b^t g}$$

is uniformly distributed in $\{\pm 1\}$. So, pick $g_1, \ldots, g_k$ independently at random from $\{0,1\}^d$, the eigenvalues corresponding to the eigenvector χ_b is

$$\lambda_b = \sum_i (-1)^{b^t g_i}.$$

This is a sum of independent, uniformly chosen ±1 random variables. Therefore, we know it is concentrated around 0! To determine how concentrated it is, we can use Chernoff's bounds.

10 Some extremal graph theory

Recall that the *extremal* number of a graph H, denoted by $\mathrm{ex}(n, H)$ is the maximum number of edges in a graph G on n vertices that contains no copies of H. For example, as we've already seen before, $\mathrm{ex}(n, C_3) = \frac{n^2}{4}$ (Mantel's theorem). In general, Erdős and Stone proved that for every graph H with chromatic number $\chi(H) = k$ we have

$$\mathrm{ex}(n, H) = \left(1 - \frac{1}{k-1}\right)\binom{n}{2} + o(n^2).$$

Note that even though the above bound captures most of the graphs H, it doesn't give a clue on the actual extremal number of bipartite graph H (that is, when $k = 2$). Let us prove the following theorem due to Kővari, Sós and Turán:

Theorem 10.1. *For any natural numbers s and t with $s \leq t$, there exists a constant c such that*

$$ex(n, K_{s,t}) \leq cn^{2-1/s}.$$

Proof. Let G be a graph on n vertices with at least $cn^{2-1/s}$ edges, and we wish to show that G must contain a copy of $K_{s,t}$. Assume not, and let us count the number of pairs (v, S) consisting of $v \in V(G)$ and $S \subseteq N(v)$ of size exactly s. The number of such pairs is

$$\sum_v \binom{d(v)}{s} \geq n\binom{\frac{1}{n}\sum d(v)}{s} \geq n\binom{2cn^{1-1/s}}{s} \geq \frac{c^s n^s}{s!}.$$

On the other hand, every subset of size S is being counted at most $t-1$ times (otherwise one can find a copy of $K_{s,t}$) so we have

$$\frac{c^s n^s}{s!} \leq \binom{n}{s}(t-1),$$

and this is a contradiction in case that c is a sufficiently large constant (depending on s and t). □

For example, if we take $K_{2,2}$ (which is just a cycle of length four), the above theorem tells that $\mathrm{ex}(n, K_{2,2}) = O(n^{3/2})$. Let us show that this is tight (up to a constat factor). To this end, let us build a graph on $n = q^2$ vertices (where q is some arbitrarily large prime) which is $K_{2,2}$-free and contains $\Omega(n^{3/2})$ edges. The vertex set consists of all vectors in $\mathbb{F}_q^2$, and we put an edge $(x, y) \sim (s, t)$ if and only if $s + yt + x = 0$ (note that this is a symmetrical relation). There is a simple intuitive way to think about why this relation works, but we will not discuss it here. First, let's count the number of edges in the graph we've just defined. Fix an (x, y), The number of solutions to $s + yt + x = 0$ is exactly q (choose t arbitrarily, and there is a unique s to satisfy this equation). That is, the graph is q regular and therefore there are exactly $\frac{q^3}{2} = \frac{1}{2}n^{3/2}$ many edges (among them there might be few self loops, then we can just delete them as there are not too many such). Finally, let us convince ourselves that there are no copies of $K_{2,2}$. Suppose there are. In particular, there are $(x, y), (x', y')$ that have at least two common neighbors. We distinguish between two cases and show that in each of them this is impossible. First, if $y \neq y'$, then

$$s + yt = -x$$

and

$$s + y't = -x'$$

has more than one solution. As the rank of the coefficient matrix is 2 (because $y \neq y'$), it cannot be. Second, assume $y = y'$ and $x \neq x'$. Then we are looking for s, t such that

$$yt + s = -x$$

and

$$yt + s = -x'$$

which is clearly absurd. This completes the argument.

A nice corollary which is obtained by the special structure of our graph is the following:

Theorem 10.2. *For sufficiently large n we have that one can decompose the edges of K_n into $(1 + o(1))\sqrt{n}$ edge-disjoint C_4-free graphs.*

Proof. Suppose that $n = q^2$, where q is some prime. If not, then there exists a prime $\sqrt{n} \le q \le (1+o(1))\sqrt{n}$, choose such and let $n' := q^2 = (1+o(1))n$. For all $i \in [q]$, define $F_i(x,y,s,t) = x+yt+s+i$ and define G_i on vertex set $\mathbb{Z}_q^2$ with edge set consists of all pairs $(x,y) \sim (s,t)$ if and only if $F_i(x,y,s,t) = 0$. Clearly the G_i's are edge disjoint, and by a similar argument like before, each of them is C_4 free. This completes the proof. □

11 Introduction to discrete Fourier analysis

Suppose $f : G \to \mathbb{C}$, where G is any set. Then, $\mathbb{E}_x f(x)$ means the *average* over all x. That is,

$$\mathbb{E}_x f(x) = \frac{1}{|G|} \sum_{x \in G} f(x).$$

Moreover, for $f : G^k \to \mathbb{C}$ we write

$$\mathbb{E}_{x_1,\dots,x_k} f(\bar{x}) = \frac{1}{|G|^k} \sum_{\bar{x} \in G^k} f(\bar{x}).$$

Another easy thing that we will use a lot so it is nice to recall: $e^{2\pi i x} = \cos(2\pi x) + i\sin(2\pi x)$. Since both cos and sin are 2π-periodic, it follows that $e^{2\pi i x} = e^{2\pi i (x+k)}$ for any integer k. In particular, if we take $x = \frac{y}{n}$, then all the distinct values $e^{2\pi i x}$ are obtained by values of x of the form y/n, where $0 \le y \le n-1$.

11.1 Working in $\mathbb{Z}_n$

Let us start by defining all of our notation over $\mathbb{Z}_n$, and later we repeat the whole thing by doing the same over any abelian group G (the reason for that, as will be clear later, is that there is a relatively simple isomorphism between $\mathbb{Z}_n$ and $\widehat{\mathbb{Z}}_n$ – the set of all *characters* of $\mathbb{Z}_n$ – a notion which will be introduced bellow). Suppose $f : \mathbb{Z}_n \to \mathbb{C}$. We define its *discrete Fourier transform* $\widehat{f} : \mathbb{Z}_n \to \mathbb{C}$ by the formula

$$\widehat{f}(r) = \mathbb{E}_x f(x)\omega^{-rx},$$

where $\omega = \exp(2\pi i k/n)$ is any (but fixed) primitive nth root of unity. For those which are familiar with Fourier transform, note that it is more correct to think about $\widehat{f}$ as a function from $\widehat{\mathbb{Z}}_n$ instead of from $\mathbb{Z}_n$. The main difference which is important for us between $\mathbb{Z}_n$ and $\widehat{\mathbb{Z}}_n$ (except of being completely different sets, but will turn out to be isomorphic as groups...) is in the measure we put on these sets. For $\mathbb{Z}_n$ we use the uniform probability measure (that is, an averaging), and for the latter we use the counting measure (summing). This is illustrated in the following few definitions/reminders (here you can assume that $\mathbb{Z}_n = \widehat{\mathbb{Z}}_n$ as it won't matter for the discussion): suppose $f, g : \mathbb{Z}_n \to \mathbb{C}$.

a. $\langle f, g\rangle = \mathbb{E}_x f(x)\overline{g(x)}$. (Inner product)

b. $\|f\|_p = \left(\mathbb{E}_x |f(x)|^p\right)^{1/p}$. ($p$-norm)

c. $(f * g)(x) = \mathbb{E}_{y+z=x} f(y)g(z)$. (convolution)

The corresponding definitions for functions $\widehat{f}, \widehat{g} : \widehat{\mathbb{Z}}_n \to \mathbb{C}$ are:

(a) $\langle \widehat{f}, \widehat{g}\rangle = \sum_x \widehat{f}(x)\overline{\widehat{g}(x)}$. (inner product) (*b*) $\|\widehat{f}\|_p = \left(\sum_x |\hat{f}(x)|^p\right)^{1/p}$. ($p$-norm)

(*c*) $(\widehat{f} * \widehat{g})(x) = \sum_{y+z=x} \widehat{f}(y)\widehat{g}(z)$. (convolution)

The following few rules are going to be our bread and butter:

1. $\langle f, g\rangle = \langle \hat{f}, \hat{g}\rangle$ (Parseval's identity)
2. $\|f\|_2 = \|\hat{f}\|_2$ (also Parseval's)
3. $f(x) = \sum_r \hat{f}(r)\omega^{rx}$ (the inversion formula)
4. $\widehat{f * g}(r) = \hat{f}(r)\hat{g}(r)$ (the convolution identity)
5. If a is invertible mod n and $g(x) = f(ax)$ for all $x \in \mathbb{Z}_n$, then $\hat{g}(r) = \hat{f}(a^{-1}r)$ for every r (the dilation rule).

The proofs are easy exercises so we will only sketch them:

Proof. 1. Note that

$$\langle \hat{f}, \hat{g}\rangle = \sum_r \hat{f}(r)\overline{\hat{g}(r)} = \sum_r \left(\mathbb{E}_x f(x)\omega^{-rx}\right)\left(\mathbb{E}_y \overline{g(y)}\omega^{ry}\right)$$

which equals

$$\mathbb{E}_x\mathbb{E}_y f(x)\overline{g(y)}\sum_r \omega^{-rx}\omega^{ry}.$$

Note that $\sum_r \omega^{-rx}\omega^{ry}$ is either n (if $x = y$) or 0 otherwise. Therefore, we obtain that the above equals

$$\mathbb{E}_x f(x)\overline{g(x)} = \langle f, g\rangle,$$

as desired.

2. Trivial from 1.

3. The set ω^{rx} forms an orthonormal basis to $\mathbb{C}^{\mathbb{Z}_n}$.

4. Note that

$$\widehat{f * g}(r) = \mathbb{E}_x\left[(f * g)(x)\omega^{-rx}\right] = \mathbb{E}_x\mathbb{E}_{y+z=x} f(y)g(z)\omega^{-rx} = \mathbb{E}_{y,z} f(y)g(z)\omega^{-r(y+z)} = \widehat{f}(r)\widehat{g}(r).$$

5. Note that

$$\widehat{g}(r) = \mathbb{E}_x g(x)\omega^{-rx} = \mathbb{E}_x f(ax)\omega^{-rx}$$

which, by a change of variable is

$$\mathbb{E}_x f(x)\omega^{-ra^{-1}x} = \widehat{f}(a^{-1}r).$$

This completes the argument.

□

We often want to use characteristic functions of subsets A of $\mathbb{Z}_n$. That is, for a subset A we define $A : \mathbb{Z}_n \to \mathbb{C}$ as follows:

$$A(x) = 0 \text{ if } x \notin A, \text{ and } A(x) = 1 \text{ otherwise.}$$

Given a subset A, we let $\alpha = |A|/n$ be its *density.* The following three observations are going to be useful:

- $\widehat{A}(0) = \alpha$. Indeed, $\widehat{A}(0) = \mathbb{E}_x A(x)\omega^{-0} = \alpha$.
- $\sum_r |\widehat{A}(r)|^2 = \alpha$. To see this, note that by Parseval's we have $\sum_r |\widehat{A}(r)|^2 = \mathbb{E}_x A(x)\overline{A(x)} = \alpha$.
- $\widehat{A}(-r) = \overline{\widehat{A}(r)}$. This is a trivial exercise.

As an illustrative example, we show how to approach the following theorem by Roth (here over $\mathbb{Z}_n$, but the proof can be easily adopted to $\mathbb{Z}$):

Theorem 11.1 (Roth's Theorem for finite fields)**.** *Let $\alpha > 0$ and n be a sufficiently large prime. Then, every subset $A \subseteq \mathbb{Z}_n$ of density at least α contains an arithmetic progression of length* 3.

We won't prove this theorem in full here but only show why Fourier's analysis is useful in it. Later we provide a full proof using slightly different arguments as was recently obtained by Croot and Sisask. Before diving into the details, let us define the function $A_2(x)$ as $A(x/2)$ (we assume here that 2 is invertible mod n, and hence we take n to be an odd integer). The key observation is that the number of arithmetic progressions in A can be expressed in terms of convolutions, inner products and dilations. Indeed, as a triple (x, z, y) of distinct numbers forms an arithmetic progression if and only if $x + y = 2z$, the expected number of arithmetic progressions in A can be written as

$$\mathbb{E}_{x+y=2z} A(x)A(y)A(z) = \mathbb{E}_{x+y=z} A(x)A(y)A(z/2).$$

Note that this equals (by Parseval's inequality)

$$\mathbb{E}_z (A * A)(z)A_2(z) = \langle A * A, A_2 \rangle = \langle \widehat{A * A}, \widehat{A}_2 \rangle = \langle \widehat{A}^2, \widehat{A}_2 \rangle.$$

By definition we obtain that the latter equals:

$$\sum_r \widehat{A}(r)^2 \overline{\widehat{A}_2(r)} = \sum_r \widehat{A}(r)^2 \overline{\widehat{A}(2r)} = \sum_r \widehat{A}(r)^2 \widehat{A}(-2r).$$

To make use of the above calculations, let us write the last expression (by isolating the $r = 0$ term from the rest) as

$$\alpha^3 + \sum_{r \neq 0} \widehat{A}(r)^2 \widehat{A}(-2r),$$

and obtain that

$$\mathbb{E}_{x+y=2z} A(x)A(y)A(z) = \alpha^3 + \sum_{r \neq 0} \widehat{A}(r)^2 \widehat{A}(-2r).$$

Intuitively, α^3 is exactly the probability that a fixed 3-term arithmetic progression will be part of A if the set A is being chosen *at random* (among all sets of density α). Therefore, the ugly sum will be a measure for the 'pseudo-randomness' of the set A. Our main goal is to show that the sum is not 'too large'– this will clearly imply the desired. To obtain a useful upper bound note that

$$|\sum_{r\neq 0}\widehat{A}(r)^2\widehat{A}(-2r)| \leq \max_{r\neq 0}|\widehat{A}(r)|\sum_{r\neq 0}|\widehat{A}(r)||\widehat{A}(-2r)| \leq \alpha\max_{r\neq 0}|\widehat{A}(r)|,$$

where the last inequality is obtained by Cauchy-Schwarz.

In particular we obtain that

$$\mathbb{E}_{x+y=2z}A(x)A(y)A(z) \geq \alpha^3 - \alpha\max_{r\neq 0}|\widehat{A}(r)|.$$

To complete the proof, the idea is to show that if there is a large Fourier coefficient, then there exists an arithmetic progression $P \subseteq [n]$ for which the density of A restricted to P is strictly larger than the density of A in $\mathbb{Z}_n$. Then, by letting P playing the role of $[n]$, one can iterate until either we find a 3-AP in A or we get that A has density 1 on some arithmetic progression P of length larger than 3, where we are done as well. In later sections we try to give a bit more formal discussion that should help with developing a bit of intuition.

11.2 General groups – first step – characters

In the previous section we illustrated how to work with Fourier's over $\mathbb{Z}_n$. Here we extend the above arguments to general abelian groups.

Let's start by defining the super useful notion of a *character*. Suppose G is a finite abelian group of order n, written additively. A *character* of G is a homomorphism $\chi : G \to \mathbb{C}^\times$ of G to the multiplicative group of nonzero complex numbers. That is,

$$\chi(a+b) = \chi(a)\chi(b), \text{ for all } a, b \in G.$$

Clearly,

$$\chi(a)^n = \chi(na) = \chi(0) = 1 \text{ for all } a \in G,$$

and therefore

$$\chi : G \to S^1,$$

where S^1 is the 1-dimensional unit sphere. In particular we have

$$\chi(-a) = \chi(a)^{-1} = \overline{\chi(a)}.$$

We define the *principal character* as

$$\chi_0(a) = 1 \text{ for all } a \in G.$$

Proposition 11.2. *For any non-principal character of G we have*

$$\sum_{a\in G}\chi(a) = 0.$$

Proof. Indeed, let $b \in G$ be such that $\chi(b) \neq 1$. Then,

$$\sum_a \chi(a) = \sum_a \chi(a+b) = \sum_a \chi(a)\chi(b) = \chi(b)\sum_a \chi(a).$$

This completes the proof. □

As an immediate corollary we obtain the first orthogonality relation for characters:

Corollary 11.3 (First orthogonality relation)**.** *Let χ and ψ be any two characters of G. Then*

$$\sum_{a\in G} \chi(a)\overline{\psi(a)} = 0 \textit{ if } \chi \neq \psi,$$

and

$$\sum_{a\in G} \chi(a)\overline{\psi(a)} = n \textit{ otherwise }.$$

Proof. The case $\chi = \psi$ is trivial. If $\chi \neq \psi$, then $\psi\overline{\chi}$ is also a character and is non-principal. Therefore, by the previous proposition we are done. □

Let $\widehat{G}$ be the set of all characters of G. It is quite obvious that $\widehat{G}$ forms an abelian group under the operation

$$(\chi\psi)(a) = \chi(a)\psi(a).$$

This group is called the *dual* group of G. To justify why we could identify $\widehat{\mathbb{Z}}_n$ with $\mathbb{Z}_n$ in the previous section, let us prove the following:

Proposition 11.4. *Let ω be a primitive nth root of unity. Then the map $\chi_j : \mathbb{Z}_n \to \mathbb{C}^\times$ defined by*

$$\chi_j(a) = \omega^{ja}$$

is a character for all $j \in \mathbb{Z}_n$. Moreover:

(a) $\chi_j = \chi_k$ *if and only if* $j = k$,

(b) $\chi_j = \chi_1^j$,

(c) $\widehat{\mathbb{Z}}_n = \{\chi_0, \ldots, \chi_{n-1}\}$,

(d) $\widehat{\mathbb{Z}}_n \cong \mathbb{Z}_n$.

Proof. Easy exercise. □

As a first step towards generalizing this notion to general abelian groups we need the following:

Proposition 11.5. *If G is a* direct sum*: $G = H_1 \oplus H_2$, and φ_i is a character of H_i, $i \in \{1, 2\}$, then the function $\chi = \varphi_1 \oplus \varphi_2$, define by*

$$\chi(h_1, h_2) = \varphi_1(h_1)\varphi_2(h_2)$$

is a character of G. Moreover, all characters of G are of this form. Therefore, $\widehat{G} \cong \widehat{H}_1 \oplus \widehat{H}_2$.

Proof. Exercise. □

Corollary 11.6. *For any finite abelian group G we have $G \cong \widehat{G}$.*

Proof. As G is a finite abelian group, we know that $G \cong \mathbb{Z}_{n_1} \oplus \mathbb{Z}_{n_2} \oplus \ldots \oplus \mathbb{Z}_{n_k}$ for some n_i's. Therefore, by the previous propositions we obtain the desired. □

It is worth remarking that there is no natural isomorphism between G and $\widehat{G}$. Even for cyclic groups, the isomorphism depends on the choice of ω. The isomorphism $G \cong \widehat{\widehat{G}}$ however is natural:

Corollary 11.7. *G can be identified with $\widehat{\widehat{G}}$ in the following natural way: for $a \in G$, define $\bar{a} \in \widehat{\widehat{G}}$ by*

$$\bar{a}(\chi) = \chi(a) \text{ for all } \chi \in \widehat{\widehat{G}}.$$

Proof. Exercise. □

Now, let's consider $\mathbb{C}^G$. That is, the linear space consisting of all functions from G to $\mathbb{C}$. As in the previous section, we introduce an inner product

$$\langle f, g \rangle = \mathbb{E}_x f(x)\overline{g(x)}.$$

It is quite obvious (and we've already seen it) that $\widehat{G}$ forms an orthonormal basis for $\mathbb{C}^G$. Let $\widehat{G} = \{\chi_0, \dots, \chi_{n-1}\}$. The $n \times n$ matrix C defined as

$$C_{ij} = (\chi_i(a_j))$$

is called the *character table* of G.

Corollary 11.8. *The matrix $A = \frac{1}{\sqrt{n}}C$ is unitary.*

Proof. Indeed, $(CC^*)_{ij} = \sum_k \chi_i(a_k)\overline{\chi_j(a_k)} = 0$ if $i \neq j$ and n if $i = j$. □

Note that as $AA^* = I$ we also have $A^*A = I$ which immediately gives the following corollary:

Corollary 11.9 (Second orthogonality relation)**.** *Let $a, b \in G$. Then*

$$\sum_{\chi \in \widehat{G}} \chi(a)\overline{\chi(b)} = 0 \text{ if } a \neq b$$

and equals n if $a = b$.

Proof. A different proof than just showing $A^*A = I$ can be obtained by stating the problem for the abelian group $\widehat{\widehat{G}}$ instead of G. □

In particular we have that

Corollary 11.10. *For any non-zero $a \in G$ we have*

$$\sum_{\chi \in \widehat{G}} \chi(a) = 0.$$

11.3 General abelian groups – second step

We have already seen that any function $f \in \mathbb{C}^G$ can be written as

$$f = \sum_{\chi \in \widehat{G}} c_\chi \chi,$$

where $c_\chi = \langle f, \chi \rangle = \mathbb{E}_x f(x)\overline{\chi(x)}$. The coefficients c_χ are called the *Fourier coefficients* of f.

The function $\widehat{f} : \widehat{G} \to \mathbb{C}$ defined by

$$\widehat{f}(\chi) = c_\chi = \frac{1}{n} \sum_{x \in G} f(x)\overline{\chi(x)}$$

is called the *Fourier transform* of f. This transformation is easily inverted:

$$f = \sum_{\chi \in \widehat{G}} c_\chi \chi = \sum_{\chi \in \widehat{G}} \widehat{f}(\chi)\chi.$$

Therefore, the formula for the *inverse Fourier transform* is

$$f(x) = \sum_{\chi \in \widehat{G}} \widehat{f}(\chi)\chi(x).$$

As a simple corollary we obtain the following. Let $\delta \in \mathbb{C}^G$ be defined as follows:

$$\delta(a) = 0 \text{ if } a \neq 0 \text{ and } \delta(0) = 1.$$

Corollary 11.11. *The following holds:*

- $\widehat{\delta}(\chi) = 1/n$ *for all* $\chi \in \widehat{G}$.
- $\delta = \frac{1}{n} \sum_{\chi \in \widehat{G}} \chi$.

Proof. For the first bullet, note that

$$\widehat{\delta}(\chi) = \frac{1}{n} \sum_x \delta(x)\overline{\chi(x)} = \frac{1}{n}\delta(0)\overline{\chi(0)} = 1/n.$$

To prove the second bullet, note that by the inverse formula we have

$$\delta = \sum_{\chi \in \widehat{G}} \widehat{\delta}(\chi)\chi$$

which by the first bullet equals

$$\frac{1}{n} \sum_{\chi \in \widehat{G}} \chi$$

as desired. □

A useful observation is the following: let C be the character table of G as defined above. Since

$$\widehat{f}(\chi) = \frac{1}{n}\sum_{x} f(x)\overline{\chi(x)},$$

it follows that

$$\widehat{f} = \frac{1}{n} f^t C^*,$$

where $f^t = (f(x))_{x \in G}$ (considered as a row vector). This will be useful in the following theorem (Recall that the inner product over $\widehat{G}$ is defined as a sum rather than an average).

Theorem 11.12 (Plancheral formula)**.** *For any $f, g \in \mathbb{C}^G$ we have*

$$\langle \widehat{f}, \widehat{g} \rangle = \langle f, g \rangle.$$

Proof. Write f, g, $\widehat{f}, \widehat{g}$ as row vectors, and observe that

$$\widehat{f} = \frac{1}{n} f^t C^*, \text{ and } \widehat{g} = \frac{1}{n} g^t C^*.$$

Therefore,

$$\langle \widehat{f}, \widehat{g} \rangle = \widehat{f}^t \overline{\widehat{g}} = \frac{1}{n^2} f^t C^* C \overline{g} = \frac{1}{n} f^t \overline{g} = \langle f, g \rangle$$

as desired. □

Corollary 11.13. $\|\widehat{f}\| = \|f\|$.

Next, let $A \subseteq G$ and we wish to work with the characteristic function $A(x)$ of A. Note that for every $A, B \subseteq G$ we have

$$\langle A, B \rangle = \mathbb{E}_x A(x) B(x) = \frac{1}{n}|A \cap B|,$$

where the first equality holds since B is a real valued function.

Moreover, as in a previous section, observe that

$$\widehat{A}(\chi_0) = \frac{1}{n}|A|.$$

Indeed,

$$\widehat{A}(\chi_0) = \mathbb{E}_x A(x) \overline{\chi_0(x)} = \frac{1}{n}|A|.$$

That is, the first Fourier coefficient gives the 'probability' that a fixed element $x \in G$ belongs to A if A is a randomly (uniformly) chosen subset of size $|A|$. The other coefficients give some measurement about the 'pseudorandomness' structure of A. Let

$$\Phi(A) = \max\{|\widehat{f}(\chi)| : \chi \in \widehat{G}, \chi \neq \chi_0\}.$$

The smaller $\Phi(A)$ is, the more 'random like' A is. Estimating $\Phi(A)$ is usually the main goal when applying Fourier analysis to problems in additive combinatorics.

Let's give a simple lower bound that holds for all $A \subseteq G$.

Proposition 11.14. *For every $A \subseteq G$, if $|A| \leq n/2$, then*

$$\Phi(A) \geq \frac{1}{n}\sqrt{|A|/2}.$$

Proof. Observe that

$$\|\widehat{A}\|^2 = \|A\|^2 = |A|/n.$$

On the other hand,

$$\|\widehat{A}\|^2 = \sum_{\chi} |\widehat{A}(\chi)|^2 \leq |\widehat{A}(\chi_0)|^2 + (n-1)\Phi(A)^2.$$

Combining the above two inequalities we obtain:

$$\Phi(A)^2 \geq \frac{|A|}{2n^2}$$

as desired. □

Some other useful properties of $\Phi(A)$ are summarized as follows:

- $\Phi(A) = \Phi(G \setminus A)$ for every $A \subseteq G$. Indeed, note that $(G \setminus A) = 1 - A$, and it's quite easy to show that for all $\chi \in \widehat{G}$ we have
$$\widehat{f}(\chi) = (\widehat{1-f})(\chi).$$
- If $gcd(k,n) = 1$ then $\Phi(kA) = \Phi(A)$ for every $A \subseteq G$, where $kA = \{ka : a \in A\}$.
- If $\alpha \in Aut(G)$, then $\Phi(A) = \Phi(\alpha A)$.
- $\Phi(A + a) = \Phi(A)$ for al $a \in A$.

11.4 A general application: Equations over finite abelian groups

We shall consider the following general problem (can be seen as a generalization of Roth's theorem): Let $A_1, \ldots, A_k \subseteq G$ and let a be a fixed element of G. We wish to estimate the number of solutions for the equation:

$$x_1 + \ldots + x_k = a, \text{ where } x_i \in A_i \text{ for all } i.$$

Let $|A_i| = m_i$ for all i. Now, assume for a moment that the sets A_i are fixed and $a \in G$ is chosen uniformly at random. This gives an expected number of $\frac{m_1 \ldots m_k}{n}$ solutions. The main goal here is to show that under some reasonable assumptions the expectation is quite close to the actual number of solutions for *every* $a \in G$! (we've already seen it in the brief sketch of Roth's Theorem, but it remains nice even long after you see it for the first time...).

To see what's going on in there, we first observe that replacing A_k by $A_k - a$ and set $x_1 + \ldots + x_k = 0$ doesn't change the number of solutions. Therefore, it is enough to consider the homogeneous equation

$$x_1 + \ldots + x_k = 0, \text{ where } x_i \in A_i \text{ for all } i.$$

Similarly to the sketch of Roth's theorem, we want to write an explicit formula for the number of solutions to this equality, which is denoted by N:

$$N = \sum_{\bar{x} \in \prod A_i} \delta(x_1 + \ldots + x_k) = \frac{1}{n} \sum_{\chi \in \widehat{G}} \sum_{\bar{x}} \chi(x_1 + \ldots + x_k).$$

Since $\chi(x_1 + \ldots + x_k) = \chi(x_1) \ldots \chi(x_k)$, the right hand side can be written as

$$\frac{1}{n} \sum_{\chi \in \widehat{G}} \prod_{i=1}^{k} \left(\sum_{x_i \in A_i} \chi(x_i) \right).$$

Observe that

$$n\widehat{A}_i(\bar{\chi}) = \sum_{x_i \in A_i} \chi(x_i),$$

and therefore

$$N = n^{k-1} \prod_{i=1}^{k} \widehat{A}_i(\chi_0) + R = \frac{m_1 \ldots m_k}{n} + R,$$

where

$$R = n^{k-1} \sum_{\chi \neq \chi_0} \prod_{i=1}^{k} \widehat{A}_i(\chi).$$

The formula we've just obtained is useful only if we can show that $|R|$ is 'small'. To conclude that the equation of interest has a solution at all, we need to prove that $|R| < \frac{m_1 \ldots m_k}{n}$. To this end we need the aid of the following simple and yet powerful tool.

11.5 The Cauchy-Schwarz trick

Let us consider the case $k = 3$ in the above equation we are interested at. We will show that if at least one of the sets A_i is *smooth* (that is, all non-principal Fourier coefficients of A_i are small) and the sets of not too small, then the equation has roughly the 'expected' number of solutions.

Theorem 11.15. *Let $A_1, A_2, A_3 \subseteq G$, $a \in G$, and let N be the number of solutions to*

$$x_1 + x_2 + x_3 = a \text{ where } x_i \in A_i \text{ for all } i.$$

Then,

$$\left| N - \frac{m_1 m_2 m_3}{n} \right| < n\Phi(A_3)\sqrt{|A_1||A_2|}.$$

Proof. As mentioned before, it is enough to solve

$$x_1 + x_2 + x_3 = 0, \text{ where } x_i \in A_i.$$

Observe that $\Phi(A_3) = \Phi(A_3 - a)$, so it doesn't change the conclusion. Indeed,

$$\Phi(A_3) = \max\{|\widehat{A}_3(\chi)| : \chi \neq \chi_0\} = \max\{|\mathbb{E}_x A_3(x)\chi(x)| : \chi \neq \chi_0\},$$

and the result is obtained by a change of variable $x \to x + a$.

Recall that

$$R = n^2 \sum_{\chi \neq \chi_0} \prod_{i=1}^{3} \widehat{A}_i(\chi).$$

Therefore,

$$|R| \leq n^2 \sum_{\chi \neq \chi_0} \prod_{i=1}^{3} |\widehat{A}_i(\chi)| \leq n^2 \Phi(A_3) \sum_{\chi \in \widehat{G}} |\widehat{A}_1(\chi)||\widehat{A}_2(\chi)|.$$

By Cuachy-Schwarz, the right hand side is at most

$$n^2\Phi(A_3)\left(\sum_{\chi}|\widehat{A}_1(\chi)|^2\right)^{1/2}\left(\sum_{\chi}|\widehat{A}_2(\chi)|^2\right)^{1/2} = n\Phi(A_3)\sqrt{|A_1||A_2|}.$$

This completes the proof. □

Corollary 11.16. *If $\Phi(A_3)/|A_3| < \frac{\sqrt{|A_1||A_2|}}{n^2}$, then there is a solution to the above equation.*

Proof. This is just a restatement of $|R| < \frac{|A_1||A_2||A_3|}{n}$. □

So now we're done with the general setting, and it's clear what we need to prove in order to obtain the desired. In the following few sections we will give few examples of complete proofs of this type.

11.6 Roth's Theorem in $\mathbb{Z}_3^n$

In this section we prove Roth's theorem in $\mathbb{Z}_3^n$. This very clean proof was obtained by Meshulam and serves as a very good example for understanding what actually is going on.

First, note that for $\alpha \in \mathbb{Z}_m^n$ and $\omega = \exp(2\pi i/m)$, we can define the character

$$\chi_\alpha(x) = \omega(\alpha^T x).$$

Our main goal is to prove the following theorem:

Theorem 11.17. *There exists an absolute constant $C > 0$ such that for all $A \subseteq \mathbb{Z}_3^n$, if $|A| \geq C3^n/n$, then A contains a 3-AP.*

Note that the bound is much weaker than the one we saw in the section about the Capset problem (which was proved using the polynomial method).

Proof. Let $A \subseteq \mathbb{Z}_3^n$ be a subset of density $\mu = |A|/3^n$, and let $A : \mathbb{Z}_3^n \to \{0,1\}$ be the indicator function for A (so $\widehat{A}(0) = \mu$). Note that a 3-AP in this setting is of the form $x, y, -(x+y)$. The main idea is to estimate the probability that randomly chosen $x, y \in \mathbb{Z}_3^n$ will satisfy $A(x)A(y)A(-(x+y) = 1$. Note that if A is a random subset, then the success probability is μ^3. Our goal is to show that if A is not too small, then even if not random, this probability is not 0. This will give us the desired.

Claim 11.18. *The probability that $x, y, -(x+y) \in A$ is $\sum_\alpha \widehat{A}(\alpha)^3$.*

Proof. Note that the probability that $x, y, -(x+y) \in A$ is exactly $\mathbb{E}_{x,y}[A(x)A(y)A(-(x+y))]$. Since $A(z) = \sum_\alpha \widehat{A}(\alpha)\chi_\alpha(z)$, we obtain that the probability is exactly

$$\mathbb{E}_{x,y}[\sum_{\alpha,\beta,\gamma}\widehat{A}(\alpha)\widehat{A}(\beta)\widehat{A}(\gamma)\chi_\alpha(x)\chi_\beta(y)\chi_\gamma(-(x+y))]$$

which, by changing the summation and using the fact that $\chi(s+t) = \chi(s)\chi(t)$, can be rewritten as

$$\sum_{\alpha,\beta,\gamma}\widehat{A}(\alpha)\widehat{A}(\beta)\widehat{A}(\gamma)\mathbb{E}_{x,y}[\chi_\alpha(x)\chi_\beta(y)\chi_{-\gamma}(x)\chi_{-\gamma}(y)] = \sum_{\alpha,\beta,\gamma}\widehat{A}(\alpha)\widehat{A}(\beta)\widehat{A}(\gamma)\mathbb{E}_x[\chi_{\alpha-\gamma}(x)]\mathbb{E}_y[\chi_{\beta-\gamma}(y)].$$

Now, observe that if $\alpha \neq \gamma$ or $\beta \neq \gamma$, then the corresponding $\mathbb{E}_x$ or $\mathbb{E}_y$ is 0. Therefore, we obtain that the above sum equals

$$\sum_\gamma \widehat{A}(\gamma)^3$$

as desired. □

Corollary 11.19. *Suppose that $|\widehat{A}(\alpha)| < \mu^2/2$ for all $\alpha \neq 0$. Then A contains a 3-AP.*

Proof. Indeed, observe that

$$\Pr_{x,y}[x, y, -(x+y) \in A] = \sum_\alpha \widehat{A}(\alpha)^3 = \mu^3 + \sum_{\alpha\neq 0} \widehat{A}(\alpha)^3 \geq \mu^3 - \sum_{\alpha \neq 0} |\widehat{A}(\alpha)|^3.$$

Using Parseval's we obtain that

$$\sum_{\alpha\neq 0} |\widehat{A}(\alpha)|^3 \leq \max_{\alpha\neq 0} |\widehat{A}(\alpha)| \sum_\alpha |\widehat{A}(\alpha)|^2 = \mathbb{E}_x |A(x)|^2 \leq \mu^3/2.$$

Therefore, the probability that $x, y, -(x+y) \in A$ is at least $\mu^3 - (\mu^3/2) = \mu^3/2 > 0$. This completes the proof. □

Remark 11.20. *Note that we cheated a bit by ignoring the 'trivial' 3-AP's of the form x, x, x. Convince yourself that as there are not too many such sequences, then this cheat can be easily fixed.*

Basically, what we've shown is that if all the Fourier coefficients are small, then A is 'random like'. For convenience, we say that A is *η-uniform* if $|\widehat{A}(\alpha)| \leq \eta$ for all $\alpha \neq 0$. Working with this notation, we proved that if A is $\mu^2/2$-uniform then A contains a 3-AP. What if A is not? in particular, it means that there exists $\beta \neq 0$ for which $|\widehat{A}(\beta)| \geq \mu^2/2 =: \eta$. We now show that it means that A is positively correlated with one of the three hyperplanes $\beta \cdot x = j$, $j \in \mathbb{Z}_3$. This will enable us to iterate using a density increment argument.

Proposition 11.21. *Suppose $f : \mathbb{Z}_p^n \to \mathbb{R}$ is a function satisfying $\mathbb{E}[f] = \mu$ and $|\widehat{f}(\beta)| \geq \eta$ for some $\beta \neq 0$. Then, there exists $c \in \mathbb{Z}_p$ such that*

$$\mathbb{E}_x[f(x) | \beta \cdot x = c] \geq \mu + \eta/2.$$

Proof. Let $g = f - \mu$, so $\mathbb{E}[g] = 0$ and $|\widehat{g}(\beta)| \geq \eta$. Note that

$$|\widehat{g}(\beta)| = |\mathbb{E}_x[g(x)\omega^{-\beta x}]| = |\frac{1}{p}\sum_{j=0}^{p-1} \mathbb{E}_x[g(x)\omega^{-j} | \beta x = j]|$$

which by the triangle inequality is at most

$$\frac{1}{p}\sum_j |\mathbb{E}_x[g(x)|\beta x = j]| =: \frac{1}{p}\sum_j |\delta_j|,$$

where $\delta_j = \mathbb{E}_x[g(x)|\beta x = j]$. Since $|\widehat{g}(\beta)| \geq \eta$ we obtain, by averaging, that for some $j \in \mathbb{Z}_p$ we have $|\delta_j| \geq \eta$. In order to complete the proof we need to get rid of the absolute value. To this end, simply observe that the average of all the δ_i's is clearly $\mathbb{E}[g] = 0$. Therefore, we can deduce that

$$\eta \leq \frac{1}{p}\sum_j (|\delta_j| + \delta_j),$$

so there is a j with $|\delta_j| + \delta_j \geq \eta$. This gives $\delta_j \geq \eta/2$ so we have $\mathbb{E}_x[g(x)|\beta x = j] \geq \eta/2$, which implies $\mathbb{E}_x[f(x)|\beta x = j] \geq \mu + \eta/2$. This completes the proof. □

Based on the above proposition, we know that either A contains a 3-AP or three must be some hyperplane $\beta x = j$ where the density of A restricted to it is at least $\mu + \mu^2/4$. Note that this hyperplane is isomorphic to $\mathbb{Z}_3^{n-1}$ in the sense that there exists some affine transformation which maps it to $\mathbb{Z}_3^{n-1}$. Such a transformation will take A to some $A' \subseteq \mathbb{Z}_3^{n-1}$ of density larger than $\mu+\mu^2/4$, and clearly a 3-AP in A corresponds to a 3-AP in A' and vice versa under such a transformation. Therefore, we can now repeat the argument on $\mathbb{Z}_3^{n-1}$ to obtain A'' and more... at some point, either we find a 3-AP or we obtain a subspace where A (or the image of A under all these transformations along the way) has density larger than 1 on, which is clearly an absurd. This completes the proof. □

11.7 A proof of Roth's Theorem

In this section we present a proof of Roth's theorem that avoids iterations (now we work in $\mathbb{Z}_N$). This proof was obtained by Croot and Sisask not so long ago.

Given an integer N, we let $r_3(N)$ denote the size of any largest subset $S \subseteq [N]$ that contains no 3-term AP. Now Roth's Theorem can be stated as follows:

Theorem 11.22 (Roth's Theorem)**.** *We have that* $r_3(N) = o(N)$.

The theorem will be obtained by showing that $r_3(N)/N$ is asymptotically decreasing. To do so, start with $S \subseteq [N]$ of size $|S| = r_3(N)$ that contains no 3-AP. Then, convolve it with a measure on a carefully chosen 3-AP. The set T where this convolution is positive will be significantly larger than S and yet will have very few 3-AP's. Lastly, using a version of theorem of Varnavides we obtain that $r_3(N)/N$ is much smaller than $r_3(M)/M$ for some $M = (\log N)^{1/16-o(1)}$. This implies the theorem quite easily.

For $f : \mathbb{Z}_N \to [0, 1]$, let $\Lambda(f) := \mathbb{E}_{x,d\in\mathbb{Z}_N} f(x)f(x + d)f(x + 2d)$. In a sightly misleading way, we will refer to this parameter as the 'number of 3AP in f'. When f is the indicator function of a subset $S \subseteq [N]$, then this expression is the density of the number of arithmetic progressions in S. Like in Section 11.1, one can easily show that

$$\Lambda(f) = \sum_{r\in\mathbb{Z}_n} \widehat{f}(r)^2\widehat{f}(-2r).$$

We make use of the notation $\|t\|$ as the distance from t to the nearest integer. Now we are ready to prove Roth's theorem:

Proof. Let $\kappa := \limsup_{N\to\infty} r_3(N)/N$. We will show that $\kappa = 0$ (and then we are done).

We can assume that N is a prime, as there is always a prime of order $(1 + o(1))N$ so we will only pay a constant factor.

Let $S \subseteq [N]$ be a 3-AP free subset with $|S| = r_3(N)$, and let f be its indicator function. Let

$$R := \{r \in \mathbb{Z}_N : |\widehat{f}(r)| \geq (2 \log\log N/ \log N)^{1/2}\}.$$

By Parseval's inequality, this set of large Fourier coefficients cannot be too big. In particular

$$|R| \leq \log N/2 \log\log N.$$

Indeed,

$$1 \geq |S|/N = \langle f, f \rangle = \langle \widehat{f}, \widehat{f} \rangle \geq (2 \log\log N / \log N)|R|,$$

giving the above.

Moreover, we can dilate this set to lie in a short part of $\mathbb{Z}_N$. That is, we can choose an integer dilate x satisfying

$$0 < x < N^{1-1/(|R|+1)} \leq N/\log N,$$

such that for all $r \in R$ we have

$$\|xr/N\| \leq N^{-1/(|R|+1)} \leq 1/\log N.$$

Indeed, enumerate $R = \{r_1, \ldots, r_{|R|}\}$ and define the set of points

$$\{(j, jr_1, \ldots, jr_{|R|}) : j \in [N]\}.$$

This is a set of size N, and therefore, if we split $[N]^{|R|+1}$ into 'boxes' of size $s^{|R|+1}$, then as long as N is larger than the number of boxes, $N^{|R|+1}/s^{|R|+1}$, there must be a box that contains at least two points. Let $s = N^{1-1/(|R|+1)}$ and let $(j, jr_1, \ldots, jr_{|R|})$ and $(k, kr_1, \ldots, kr_{|R|})$ be two such points. Set $x = k - j$ and observe that $x < N^{1-1/(|R|+1)}$ and $\|xr_i/N\| \leq N^{-1/(|R|+1)} \leq 1/\log N$ holds for all i as desired.

Taking such an x, let

$$B := \{0, x, 2x\},$$

and let h be the normalised indicator function for B, given by

$$h(n) := N \cdot B(n)/3.$$

Convolve f with h and obtain

$$g(n) = (f * h)(n) = (f(n) + f(n-x) + f(n-2x))/3,$$

and observe that

$$\widehat{g}(0) = \widehat{f}(0).$$

(that is, f and g are averaged the same.) We now show that the Fourier coefficients of these two functions are quite close to each other, and then we can 'replace' f by g.

Note that

$$\widehat{f}(r) - \widehat{g}(r) = \widehat{f}(r)(1 - \widehat{h}(r)),$$

and therefore, using the claim bellow and some analysis, one can show that for all $r \in \mathbb{Z}_N$ we have

$$|\widehat{f}(r) - \widehat{g}(r)| \leq C(\log\log N/\log N)^{1/2}.$$

Indeed, suppose that $r \notin R$. In this case we have

$$|\widehat{f}(r) - \widehat{g}(r)| = |\widehat{f}(r)(1 - \widehat{h}(r))| \leq (2\log\log N/\log N)^{1/2}.$$

(note that $|\widehat{h}(r)| = |\mathbb{E}_n h(n) e^{-2\pi i r n/N}| \leq 1$.)

Next, if $r \in R$, then

$$|\widehat{h}(r)| = |\mathbb{E}_n h(n) e^{2\pi i r n/N}| = |\frac{1}{N}(h(0) + h(x)e^{2\pi i O(1/\log N)} + h(2x)e^{2\pi i O(1/\log N)})| = 1 + O(1/\log N),$$

giving

$$|\widehat{f}(r) - \widehat{g}(r)| = O(1/\log N).$$

Next we show that

$$|\Lambda(f) - \Lambda(g)| = O((\log\log N/\log N)^{1/2}).$$

Indeed,

$$|\Lambda(f) - \Lambda(g)| = |\sum_r \widehat{f}^2(r)\widehat{f}(-2r) - \sum_r \widehat{g}^2(r)\widehat{g}(-2r)|$$

$$= |\sum_r \widehat{f}^2(r)(\widehat{f}(-2r) - \widehat{g}(-2r)) + \sum_r (\widehat{f}^2(r)\widehat{g}(-2r) - \widehat{g}^2(r)\widehat{g}(-2r))|.$$

Clearly, the right hand side is at most

$$\max_s |\widehat{f}(s) - \widehat{g}(s)| \left(|\sum_r \widehat{f}^2(r)| + |\sum_r (\widehat{f}(r) + \widehat{g}(r))\widehat{g}(-2r)| \right)$$

which by the above estimate, Parseval's and Cauchy-Schwarz can be easily upper bounded as desired.

All in all, observe that as f contains no non-trivial 3AP's, we have

$$\Lambda(f) = O(1/N),$$

which by the above arguments give us

$$\Lambda(g) = O(\log\log N/\log N)^{1/2}.$$

Getting closer to the punch line, let $T := \{n \in \mathbb{Z}_N : g(n) > 0\}$, and note that

- $\Lambda(T) = O(\Lambda(g))$ (which is trivial), and
- $\Lambda(T) = O(\log\log N/\log N)^{1/2}$ (which follows from the above).

Since S is 3AP-free, we have $g(n) \leq 2/3$ for all n. Therefore,

$$T(n) \geq 3g(n)/2$$

for all n, which implies (by the fact that g averages the same as f)

$$|T| \geq 3|S|/2.$$

To summarize our progress: we have managed to find a subset T which is significantly larger than S, but with only few 3AP's. Since S was of size $r_3(N)$, we have that T is of size at least $3r_3(N)/2$, and in order to complete the proof it is enough to show that any set of this size must contain 'many' 3AP's. This was obtained by Varnavides (we will skip its proof at this section)

Lemma 11.23 (Varnavides). *For any $1 \leq M \leq N$, and for any set $A \subseteq [N]$, we have*

$$T_3(A) \geq \left(\frac{|A|/N - (r_3(M)+1)/M}{M^4} \right) N^2,$$

where $T_3(A)$ is the number of non-trivial 3AP's in A.

To complete the proof of the theorem, let

$$M = (\log N / \log\log N)^{1/16},$$

and apply the lemma to T to obtain

$$\Lambda(T) \geq \frac{3|S|/2N - (r_3(M)+1)/M}{M^4}.$$

Using the fact that

$$\Lambda(T) = O(\log\log N / \log N)^{1/2}$$

we conclude that

$$r_3(N)/N = |S|/N \leq 2r_3(M)/3M + O((\log\log N/\log N)^{1/4}),$$

which shows that $r_3(N)/N \leq r_3(M)/M$. □

12 Characteristic functions

In this section we would like to talk about *characteristic functions.* Our goal is to show some other applications of Fourier analysis in what is called *anti-concentration* arguments and other related stuff. This section is based on (or more or less – a copy paste of) the corresponding chapter in the book Probability and Random Processes by Grimmett and Stirzaker.

Let us first recall the definition of the *moment generating function* of a random variable X. This is a function $M_X : \mathbb{R} \to [0,\infty)$ defined by $M_X(t) = \mathbb{E}[e^{tX}]$.

Moment generating functions are proved to be very useful in handling non-negative integral random variables. They also have some nice properties such as: suppose $M_X(t) < \infty$ on some open interval around 0, then

- $\mathbb{E}[X] = M'(0)$, and $\mathbb{E}[X^k] = M^{(k)}(0)$;
- $M_X(t) = \sum_{k=0}^{\infty} \frac{\mathbb{E}[X^k]}{k!} t^k$;
- If X, Y are independent random variables then $M_{X+Y}(t) = M_X(t)M_Y(t)$.

The main disadvantage of moment generating functions is that the integrals that define them are not always finite. Therefore, it makes sense to switch to another class of functions which are potentially at least equally useful and finiteness is guaranteed.

Definition 12.1. *The* characteristic function *of X is the function $\phi_X : \mathbb{R} \to \mathbb{C}$ defined by*

$$\phi_X(t) = \mathbb{E}[e^{itX}].$$

Characteristic functions are related to the Fourier transform since $\phi_X(t) = \int e^{itx} f(x)dx$. Note that this integral is well defined by considering it as

$$\phi_X(t) = \mathbb{E}[\cos(tX)] + i\mathbb{E}[\sin(tX)].$$

Furthermore, ϕ_X is better behaved than M_X.

Theorem 12.2. *The characteristic function ϕ satisfies:*

1. *$\phi(0) = 1$, $|\phi(t)| \le 1$ for all t.*
2. *ϕ is uniformly continuous on $\mathbb{R}$.*
3. *ϕ is non-negative definite. That is,*

$$\sum_{j,k} \phi(t_j - t_k) z_j \bar{z}_k \ge 0$$

for all real numbers $t_1, \dots, t_n$ and complex numbers $z_1, \dots, z_n$.

Proof. We leave 1. and 2. as an exercise and only prove 3. Note that

$$\sum_{j,k} \phi(t_j - t_k) z_j \bar{z}_k = \sum_{j,k} \int (z_j e^{it_j x})(\bar{z}_k e^{-it_k x}) f(x) dx = \mathbb{E}\left[|\sum_j z_j e^{it_j X}|^2 \right] \ge 0.$$

□

Actually, Theorem 12.2 characterizes characteristic functions in the sense that ϕ is characteristic function if and only if it satisfies $1-3$ in the theorem. This result is called Bochner's theorem and we are not going to prove it in these notes.

Now we wish to establish some properties of characteristic functions. First, we show that from a knowledge on ϕ_X we can find the distribution of X. We won't write such a statement in full generality but rather start with the following easier statement whose proof is straightforward from Taylor's theorem for functions of complex variables:

Theorem 12.3. 1. *If $\phi_X^{(k)}(0)$ exists then $\mathbb{E}|X^k| < \infty$ whenever k is even, and $\mathbb{E}|X^{k-1}| < \infty$ whenever k is odd.*

2. *If $\mathbb{E}|X^k| < \infty$ then*

$$\phi_X(t) = \sum_{j=0}^{k} \frac{\mathbb{E}[X^j]}{j!}(it)^j + o(t^k),$$

and so $\phi_X^{(k)}(0) = i^k \mathbb{E}[X^k]$.

Another important property of characteristic functions is that they enable us to handle sums of independent random variables:

Theorem 12.4. *If X and Y are independent then $\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t)$.*

Proof. Indeed, we have

$$\phi_{X+Y}(t) = \mathbb{E}[e^{it(X+Y)}] = \mathbb{E}[e^{itX}]\mathbb{E}[e^{itY}],$$

where the last equality holds by the independence of X and Y. □

Theorem 12.5. *If $a, b \in \mathbb{R}$ and $Y = aX + b$ then $\phi_Y(t) = e^{itb}\phi_X(at)$.*

Proof. Indeed,

$$\phi_{aX+b}(t) = \mathbb{E}[e^{itaX+itb}] = e^{itb}\mathbb{E}[e^{itaX}] = e^{itb}\phi_X(at),$$

as desired. □

In applications that I want to consider in later sections, we might work with random variables which are not independent. For this we need the following definition:

Definition 12.6. *The* joint characteristic function *of X and Y is the function $\phi_{X,Y} : \mathbb{R}^2 \to \mathbb{C}$ defined by*

$$\phi_{X,Y}(s,t) = \phi_{sX+tY}(1) = \mathbb{E}[e^{isX+itY}].$$

The proof of the following theorem is left as an exercise:

Theorem 12.7. *Random variables X, Y are independent if and only if*

$$\phi_{X,Y}(s,t) = \phi_X(s)\phi_Y(t), \text{ for all } s, t.$$

We saw in Theorem 12.3 that to find moments of X one can differentiate $\phi_X(t)$ at $t = 0$. A similar calculation gives the 'joint' moments $\mathbb{E}[X^jY^k]$.

12.1 Examples of characteristic functions

1 **Constant distribution.** Suppose that $X = \mu$ with probability 1. Then

$$\phi_X(t) = e^{it\mu}.$$

2 **Bernoulli distribution.** Suppose that X is Bernoulli with parameter p. Then

$$\phi_X(t) = e^{it0}(1-p) + e^{it}p = 1 - p + pe^{it}.$$

3 **Binomial distribution.** If $X \sim \text{Bin}(n,p)$, then $X = \sum X_i$ where the X_i's are iid Bernoulli with parameter p random variables. Therefore,

$$\phi_X(t) = \prod \phi_{X_i}(t) = (1 - p + pe^{it})^n.$$

4 **Exponential distribution.** If the density function of X is $f(x) = \lambda e^{-\lambda x}$ for $x \geq 0$ (and 0 otherwise) then

$$\phi_X(t) = \int_0^\infty e^{itx}\lambda e^{-\lambda x}dx = \frac{\lambda}{\lambda - it},$$

and the calculations for the last equality are being omitted.

5 **Normal distribution.** Suppose $X \sim N(0,1)$. Then

$$\phi_X(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty e^{itx-x^2/2}dx = e^{-t^2/2},$$

where again, the calculations are being omitted. Note that by the properties of ϕ that we've learnt, we obtain that for $Y \sim N(\mu, \sigma^2)$ we have

$$\phi_Y(t) = \phi_{\sigma X+\mu}(t) = e^{it\mu-\sigma^2t^2/2}.$$

12.2 Inversion and continuity theorems

Here we will see two main reasons for why characteristic functions are useful. The first is that characteristic function uniquely determines the distribution. In fact, there actually exists a formula to recover the distribution from the characteristic function. Let us state a special case:

Theorem 12.8. *If X is continuous with density function f and characteristic function ϕ_X then*

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \phi_X(t) dt$$

at every point x at which f is differentiable.

Proof. Note that this is just the Fourier inversion formula as discussed (for the discrete case) few sections ago. □

The general case is the following:

Theorem 12.9. *Let X have a distribution function F and characteristic function ϕ_X. Define $\bar{F} : \mathbb{R} \to [0,1]$ by*

$$\bar{F}(x) = \frac{1}{2} \left\{ F(x) + \lim_{y \uparrow x} F(y) \right\}.$$

Then

$$\bar{F}(b) - \bar{F}(a) = \lim_{N \to \infty} \int_{-N}^{N} \frac{e^{-iat} - e^{-ibt}}{2\pi i t} \phi_X(t) dt.$$

We won't prove this theorem, but as a corollary (left as an exercise) we obtain

Corollary 12.10. *Random variables X, Y have the same characteristic function if and only if they have the same distribution.*

Exactly similar results hold for jointly distributed random variables. For example, if X and Y have joint density function f and joint characteristic function ϕ then whenever f is differentiable at (x, y), we have

$$f(x, y) = \frac{1}{4\pi^2} \int \int_{\mathbb{R}^2} e^{-isx - ity} \phi(s, t) ds dt.$$

The second thing we want to discuss in this section is the following. Suppose we are now dealing with a sequence of random variables $X_1, X_2, \ldots$. We state a theorem which roughly speaking says that if the distribution functions of the sequence, $F_1, F_2, \ldots$ converge to some limit F, then the characteristic functions approach to the characteristic function of F.

Definition 12.11. *We say that $F_1, F_2, \ldots$ of distribution functions* converges *to the distribution function F and write $F_n \to F$, if $F(x) = \lim_{n \to \infty} F_n(x)$ at each point x where F is continuous.*

Theorem 12.12 (Continuity theorem.). *Suppose $F_1, F_2, \ldots$ is a sequence of distribution functions with corresponding characteristic functions $\phi_1, \phi_2, \ldots$.*

(a) If $F_n \to F$ for some distribution function F with characteristic function ϕ, then $\phi_n(t) \to \phi(t)$ for all t.

(b) Conversely, if $\phi(t) = \lim_{n \to \infty} \phi_n(t)$ exists and is continuous at $t = 0$, then ϕ is the charactersitic function of some distribution function F, and $F_n \to F$.

12.3 Few limit theorems

In this section we prove few limit theorems. The first is the 'law of large numbers' which asserts that if we perform many independent trials, then the average outcome converges to the expectation. Before stating it formally we need the following definition:

Definition 12.13. *If $X, X_1, X_2, \ldots$ is a sequence of random variables with distribution functions $F, F_1, F_2, \ldots$, we say that X_n converges in distribution to X, written $X_n \to^D X$, if $F_n \to F$ as $n \to \infty$.*

Theorem 12.14 (Law of large numbers)**.** *Let $X_1, \ldots, X_n$ be a sequence of iid random variables with finite mean μ. Their partial sums $S_n = \sum_{i=1}^n X_i$ satisfy*

$$\frac{1}{n} S_n \to^D \mu.$$

Proof. The theorem asserts that as $n \to \infty$ we have

$$\Pr[n^{-1} S_n \leq x] \to \begin{cases} 0 & x < \mu \\ 1 & x > \mu \end{cases}$$

The approach is to show that the characteristic function of $n^{-1}S_n$ approaches the characteristic function of the constant random variable μ. Let ϕ be the common characteristic function of the X_i's, and let ϕ_n be the characteristic function of $n^{-1}S_n$. Then, by the properties of characteristic functions that we discussed before, we have

$$\phi_n(t) = \phi_X(t/n)^n.$$

The behavior of $\phi_X(t/n)$ is given by Theorem 12.3

$$\phi_X(t) = 1 + it\mu + o(t).$$

Therefore, we obtain

$$\phi_n(t) = (1 + it\mu + o(t))^n \to e^{it\mu}$$

where the limit is the characteristic function of μ. This completes the proof. □

What we've just shown is that for large enough n we have $S_n \approx n\mu$. Can we say something about $|S_n - n\mu|$? apparently, if the X_i's have finite variance then

- $S_n - n\mu$ is of order $\sqrt{n}$.
- The distribution of $\frac{S_n - n\mu}{\sqrt{n}}$ approaches to the normal distribution.

Theorem 12.15 (Central Limit Theorem (CLT))**.** *Let $X_1, X_2, \ldots$ be a sequence of iid random variables with finite mean μ and finite non-zero variance σ^2, and let $S_n = \sum_{i=1}^n X_i$. Then,*

$$\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \to^D N(0,1).$$

Proof. First, let us write $Y_i = \frac{X_i - \mu}{\sigma}$, and let ϕ_Y be the characteristic function of the Y_i's. Observe that $\mathbb{E}[Y] = 0$ and $\mathbb{E}[Y^2] = 1$. Therefore, by Theorem 12.3 we have

$$\phi_Y(t) = 1 - \frac{1}{2}t^2 + o(t^2).$$

Let $\psi_n(t)$ be the characteristic function of

$$U_n = \frac{S_n - n\mu}{\sqrt{n\sigma^2}} = \frac{1}{\sqrt{n}} \sum Y_i,$$

and observe by Theorems 12.4 and 12.5 that

$$\psi_n(t) = \left(\phi_Y(t/\sqrt{n})\right)^n = \left(1 - \frac{1}{2n}t^2 + o(t^2/n)\right)^n \to e^{-\frac{t^2}{2}}.$$

The right hand side is the characteristic function of $N(0,1)$ so we are done. □

The CLT asserts that the distribution function of S_n (normalized...) converges to the distribution function of $N(0,1)$. A natural thing to ask is whether a corresponding result holds for the density functions and mass functions? We will see that mainly it is, but we need some assumptions about the 'smoothness' of the functions, as it is not necessarily true that $F_n \to F$ implies $F_n' \to F'$. A result of this kind is called 'local central theorem' since it deals with local behavior instead of the cumulative behavior. To simplify the notation we assume that the X_i's have zero mean and variance 1.

Theorem 12.16 (Local CLT)**.** *Let $X_1, X_2, \ldots$ be iid random variables with 0 mean and variance 1. Suppose that their common characteristic function ϕ satisfies*

$$\int_{-\infty}^{\infty} |\phi(t)|^r dt < \infty$$

for some integer $r \geq 1$. Then, the density function g_n of $U_n = \sum X_i/\sqrt{n}$ exists for $n \geq r$ and

$$g_n(x) \to \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \text{ as } n \to \infty \text{ uniformly in } x.$$

Proof. First observe that $|\phi|^r$ being integrable implies that $|\phi|^n$ is integrable for all $n \geq r$, as $|\phi| \leq 1$. Therefore, g_n exists and is given by the inversion formula

$$g_n(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \psi_n(t) dt,$$

where $\psi_n(t) = (\phi(t/\sqrt{n}))^n$ is the characteristic function of U_n. The Fourier inversion theorem is valid for the normal distribution as well, and therefore

$$\left| g_n(x) - e^{-\frac{1}{2}x^2} \right| \leq \frac{1}{2\pi} \left| \int_{-\infty}^{\infty} e^{-itx} \left[\phi(t/\sqrt{n})^n - e^{-\frac{1}{2}t^2} \right] dt \right| \leq I_n,$$

where

$$I_n = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \phi(t/\sqrt{n})^n - e^{-\frac{1}{2}t^2} \right| dt.$$

It thus suffices to show that $I_n \to 0$ as n goes to infinity. From the assumptions on the X_i's and Theorem 12.3 we have

$$\phi(t/\sqrt{n}) = 1 - \frac{1}{2n}t^2 + o(t^2/n), \text{ as } t \to 0.$$

Therefore, there exists some $\delta > 0$ such that for all $|t| \leq \delta$ we have

$$|\phi(t)| \leq e^{-\frac{1}{4}t^2}.$$

Moreover, for any constant $a > 0$ we have $\phi(t/\sqrt{n})^n \to e^{-\frac{1}{2}t^2}$ uniformly for $|t| \leq a$, as n tends to infinity. These two observations enable us to upper bound by $o(1)$ the above integral in the interval $[-\delta\sqrt{n}, \delta\sqrt{n}]$.

To complete the proof, we need to show that

$$\frac{1}{2\pi}\int_{|t|>\delta\sqrt{n}} \left|\phi(t/\sqrt{n})^n - e^{-\frac{1}{2}t^2}\right| dt = o(1).$$

To this end, observe that since g_n exists for $n \geq r$, we can (relatively) easily conclude that $|\phi(t)^r| < 1$ for all $t \neq 0$, and $|\phi(t)|^r \to 0$ as $t \to \pm\infty$. Therefore, $|\phi(t)| < 1$ for $t \neq 0$ and $|\phi(t)| \to 0$ as $t \to \pm\infty$. In particular, we conclude that

$$\eta := \sup\{|\phi(t)| : |t| \geq \delta\}$$

satisfies $\eta < 1$. Now, for $n \geq r$ we have

$$\int_{|t|>\delta\sqrt{n}} |\phi(t/\sqrt{n})^n - e^{-\frac{1}{2}t^2}|dt \leq \eta^{n-r}\int_{|t|>\delta\sqrt{n}} |\phi(t/\sqrt{n})|^r dt + 2\int_{t>\delta\sqrt{n}} e^{-\frac{1}{2}t^2} dt,$$

which clearly approaches 0 as $n \to \infty$. This completes the proof. □

13 Some simple examples using Fourier analysis

In this section we give some simple and less simple examples using Fourier analysis.

13.1 Random sums over a finite field

It is well known (and very easy to prove in many ways) that if $X_1, \ldots, X_n$ (for any $n \geq 1$) are iid where $\Pr[X_1 = 0] = \Pr[X_1 = 1] = 1/2$, then for $S_n = \sum X_i$, working over $\mathbb{Z}_2$, we have

$$\Pr[S_n = 1] = \Pr[S_n = 0] = \frac{1}{2}.$$

In this section we discuss generalizations of the above equality over finite fields $\mathbb{Z}_q$ where $q > 2$.

For convenience, it will be easier for us to consider the case where the variables output the numbers ± 1. Note that it does'nt make sense to consider such a scenario over $\mathbb{Z}_2$, but it is also very easy to convert such a result for $q > 2$ into the 0/1 setting (WHY?). In what follows we use Fourier analysis to settle the general case.

Suppose that we are working over $\mathbb{Z}_q$ for some prime q, and for every $a \in \mathbb{Z}_q$, let $\delta_a(x)$ be the indicator function for the event $x = a$. That is,

$$\delta_a(x) = \begin{cases} 1 & x = a \\ 0 & \text{otherwise} \end{cases}.$$

Recall that $\widehat{\delta_0}(\chi) = \frac{1}{q}$ for all $\chi \in \widehat{\mathbb{Z}_q}$ (this follows immediately from the definition of the Fourier coefficient), and therefore,

$$\delta_0(x) = \frac{1}{q}\sum_{\chi} \chi(x) = \frac{1}{q}\sum_{r \in \mathbb{Z}_q} \omega^{rx},$$

where ω is any qth root of unity (for example, let us set from now on $\omega = e^{-2\pi i/q}$). Suppose that $X_1, \ldots, X_n$ are iid random variables, where $\Pr[X_1 = 1] = \Pr[X_1 = -1] = 1/2$, and we are interested in

$$S_n = \sum_{i=1}^{n} X_i.$$

A natural question to consider is the following:

Question 13.1. *For any fixed s, what is the probability that $S_n = s \mod q$?*

We will answer this question using Fourier analysis. Later on we will extend our argument to vector valued random variables.

Theorem 13.2. *Suppose that $n = \omega(q^2 \log q)$, then for any $s \in \mathbb{Z}_q$ we have*

$$\Pr[S_n = s \mod q] = (1 + o(1))/q.$$

Proof. For simplicity, let us first take care of the case $s = 0$. Note that

$$\Pr[S_n = 0] = \mathbb{E}\left[\delta_0\left(\sum X_i\right)\right] = \mathbb{E}\left[\sum_{r \in \mathbb{Z}_q} \widehat{\delta_0}(r)\omega^{r\sum X_i}\right] = \frac{1}{q} + \frac{1}{q}\mathbb{E}\left[\sum_{r \neq 0} \omega^{r\sum X_i}\right].$$

It is thus suffices to show that

$$\left|\mathbb{E}\left[\sum_{r \neq 0} \omega^{r\sum X_i}\right]\right| = o(1).$$

Observe that by a change of summations and independence we have

$$\mathbb{E}\left[\sum_{r \neq 0} \omega^{r\sum X_i}\right] = \sum_{r \neq 0}\prod_i \mathbb{E}\left[\omega^{rX_i}\right] = \sum_{r \neq 0}\prod_i \left(\frac{e^{2\pi i r/q} + e^{-2\pi i r/q}}{2}\right),$$

and that the right hand side (in absolute value) is at most

$$\sum_{r \neq 0} |\cos(2\pi r/q)|^n = \sum_{r \neq 0} |\cos(\pi r/q)|^n.$$

Note that in the last equality we used the fact that 2 is invertible mod q.

Now, observe that

$$|\cos(\pi x)| \leq e^{-2\|x\|^2},$$

where $\|x\|$ is the distance of x from the nearest integer, and therefore the above is at most

$$\sum_{r \neq 0} e^{-2n\|r/q\|^2} \leq q e^{-2n/q^2}.$$

Therefore, by taking $n \geq mq^2 \log q$ we obtain an upper bound of the form $\left(\frac{1}{q}\right)^{2m-1}$ which is clearly $o(1)$ as m goes to infinity. □

Remark 13.3. *Note that the same result holds if we look at $S_n = \sum a_i X_i$, where all the a_i's are non zero mod q. Moreover, for the case $s \neq 0$, the same proof basically holds by replacing the first line in the proof by*

$$\Pr[S_n = q] = \mathbb{E}[\delta_0(\sum X_i - s)].$$

These two things are left as easy exercises.

Exercise 13.4. *Try to improve the bound on n by a more careful analysis of the last inequality in the above proof. Note that we assumed that as long as $r \neq 0$ we have $\|r/q\|^2 \geq 1/q^2$ which is clearly far from the truth...*

13.2 Random sums of vectors over a finite field

Let us now try to obtain a similar result for vectors. That is, suppose now that $X_1, \ldots, X_n$ are iid random variable, each of which is uniformly chosen in $\{\pm\}^d$. Consider

$$S_n = \sum_{i=1}^{n} X_i$$

and we are interested in the question:

Question 13.5. *Given $v \in \mathbb{Z}_q^d$, what is the probability that $S_n = v$?*

We will show that if n is large enough comparing to d and q, then the answer is of the form $(1 + o(1))\frac{1}{q^d}$. That is, S_n is roughly uniformly distributed.

Theorem 13.6. *Let $q \geq 3$ be a prime, $d \in \mathbb{N}$, and let $n = \omega(q^2 \log(dq))$. Then for any $v \in \mathbb{Z}_q^d$ we have*

$$\Pr[S_n = v] = (1 + o(1))/q^d.$$

Proof. Similarly to the previous proof, observe that for every $x \in \mathbb{Z}_q^d$ we have

$$\delta_0(x) = \frac{1}{q^d} \sum_{r \in \mathbb{Z}_q^d} \omega^{-\sum r_i x_i},$$

where ω is a qth root of unity (from now on assume $\omega = e^{2\pi i/q}$).

Again, for simplicity we assume that $v = 0$ (hopefully you are already convinced that it doesn't matter) and write

$$\Pr[S_n = 0] = \mathbb{E}[\delta_0(\sum X_i)] = \frac{1}{q^d}\mathbb{E}\left[\sum_{r \in \mathbb{Z}_q^d} \prod_{j=1}^{n} \omega^{-\sum_{i=1}^{d} r_i X_{ji}}\right].$$

This equals

$$\frac{1}{q^d} + \frac{1}{q^d}\sum_{r \neq 0} \prod_{j=1}^{n} \prod_{i=1}^{d} \mathbb{E}[\omega^{-r_i X_{ji}}] = \frac{1}{q^d} + \frac{1}{q^d}\sum_{r \neq 0} \prod_{i=1}^{d} \cos(2\pi r_i/q)^n.$$

Therefore, in order to complete the proof, it is enough to show that

$$|\sum_{r \neq 0} \prod_{i=1}^{d} \cos(2\pi r_i/q)^n| \leq \sum_{r \neq 0} \prod_{i=1}^{d} |\cos(2\pi r_i/q)|^n = \sum_{r \neq 0} \prod_{i=1}^{d} |\cos(\pi r_i/q)|^n = o(1).$$

Let us write V_k for the set of all vectors $r \in \mathbb{Z}_q^d$ with support of size exactly k. Clearly,

$$|V_k| = \binom{d}{k}(q-1)^k \leq (dq)^k.$$

We also use (again) the following easy inequality

$$\cos(\pi x) \leq e^{-2\|x\|^2},$$

where $\|x\|$ is the distance of x from the nearest integer.

Now, by rearranging the above sum we obtain

$$\sum_{r \neq 0} \prod_{i=1}^{d} |\cos(\pi r_i/q)|^n \leq \sum_{k=1}^{d} |V_k| e^{-2kn/q^2} \leq \sum_{k=1}^{d} (dq)^k e^{-2kn/q^2}.$$

Therefore, as long as we take $n = \omega(q^2 \log(dq))$, the above sum is $o(1)$. This completes the proof. □

Exercise 13.7. *Try to improve the bound on n by a more careful analysis of the last display (that is, don't use only the support but also the structure of the vectors r).*

13.3 Random sums over the integers

In this section we keep on improving our technique until we could 'see' how to prove stronger theorems. Our aim is to prove the following:

Theorem 13.8. *Let* $X_1, \ldots, X_n$ *be iid random variables with* $\Pr[X_1 = 1] = \Pr[X_1 = -1] = \frac{1}{2}$. *Then, for* $S_n = 1X_1 + 2X_2 + \ldots + nX_n$ *we have*

$$\Pr[S_n = 0] \leq Cn^{-3/2}.$$

Proof. We use the identity (for $k \in \mathbb{Z}$)

$$\delta_0(k) = \int_{-1/2}^{1/2} e^{2\pi i k x} dx,$$

to obtain (skipping some easy steps)

$$\Pr[S_n = 0] = \mathbb{E}[\delta_0(S_n)] = \int_{-1/2}^{1/2} \prod_{k=1}^{n} \cos(2\pi x k) dx.$$

Now, in order to estimate the above integral, observe that for small x, by the Taylor expansion of $\cos x$ we have

$$\cos(x) \approx e^{-x^2/2}.$$

Moreover, it is quite easy to show that the integral

$$\int_{\varepsilon/n < |x| \leq 1/2} \prod_{k=1}^{n} \cos(2\pi x k) dx$$

is exponentially (in n) small, and therefore it is enough to estimate

$$\int_{-\varepsilon/n}^{\varepsilon/n} \prod_{k=1}^{n} \cos(2\pi x k) dx$$

. The advantage is that now we can use the above estimate on $\cos(2\pi x k)$ for all $k \leq n$ and obtain

$$\int_{-\varepsilon/n}^{\varepsilon/n} \prod_{k=1}^{n} \cos(2\pi x k) dx \leq \int_{-\varepsilon/n}^{\varepsilon/n} e^{-Cx^2 \sum_{k=1}^{n} k^2} dx \leq \int_{-\varepsilon/n}^{\varepsilon/n} e^{-C' x^2 n^3} dx.$$

Finally, by a change of variable $t = n^{3/2} x$, using the fact that $\int_{-\infty}^{\infty} e^{-t^2} dt = O(1)$, we obtain the desired bound of the form $O(n^{-3/2})$. This completes the proof. □

14 Higher degree Erdős-Littlewood-Offord type inequalities

Recall that the Erdős-Littlewood-Offord problem deals with linear functions of the form

$$S_n = \sum_{i=1}^{n} a_i X_i,$$

where the X_i are iid Bernoulli random variables. In particular, we are interested at

$$\rho(\bar{a}) = \sup_m \Pr[S_n = m].$$

In this section we discuss generalizations of this problem to non-linear functions.

14.1 A quadratic Erdős-Littlewood-Offord inequality

Define

$$Q := \sum_{1 \leq i,j \leq n} a_{ij} z_i z_j,$$

where the a_{ij}'s are fixed and the z_is are iid Bernoulli random variables. We wish to prove the following theorem (due to Costello, Tau and Vu):

Theorem 14.1. *Let Q be as above, let $U_1 \cup U_2 = [n]$ be any nontrivial partition, and let S be any non-empty subset of U_1. For each $i \in S$, let d_i be the number of indices $j \in U_2$ such that $|a_{ij}| \geq 1$. Suppose that $d_i \geq 1$ for each $i \in S$. Then for any interval I of length* 1 *we have*

$$\Pr[Q \in I] = O\left(|S|^{-1/2} + |S|^{-1} \sum_{i \in S} d_i^{-1/2} \right)^{1/4}.$$

It is very unlikely that the above bound is best possible, and in fact, in later sections we will probably prove some better bounds. The goal is to sketch the main tricks to be used later in similar problems.

Proof. A natural thing to do is to rewrite

$$Q := \sum_{i=1}^{n} Q_i z_i,$$

where $Q_i := \sum_{j=1}^{n} a_{ij} z_j$ for all i. Writing in this form, one would expect to obtain the proof by two applications of the Erdős-Littlewood-Offord inequality. Unfortunately, the Q_i's are not independent, and therefore this plan fails. To overcome this problem, we will use the following *decoupling* lemma.

Lemma 14.2 (Decoupling lemma)**.** *Let X, Y be independent random variables and $\mathcal{E} := \mathcal{E}(X, Y)$ be an event depending on X and Y. Then,*

$$\Pr[\mathcal{E}] \leq \left(\Pr[\mathcal{E}(X, Y) \wedge \mathcal{E}(X', Y) \wedge \mathcal{E}(X, Y') \wedge \mathcal{E}(X', Y')]\right)^{1/4},$$

where X' and Y' are independent copies of X and Y, respectively.

Proof. We will only handle the case where both X and Y take only finite number of values, say, $x_1, \ldots, x_n$ and $y_1, \ldots, y_m$, respectively (the general case is being left as an exercise). Clearly, we have

$$\Pr[\mathcal{E}(X, Y)] = \sum_{i=1}^{n} \Pr[\mathcal{E}(x_i, Y)] \Pr[X = x_i]$$

and

$$\Pr[\mathcal{E}(X, Y) \wedge \mathcal{E}(X, Y')] = \sum_{i=1}^{n} \Pr[\mathcal{E}(x_i, Y)]^2 \Pr[X = x_i].$$

Therefore, by Cauchy-Schwarz we obtain

$$\Pr[\mathcal{E}] \leq \Pr[\mathcal{E}(X, Y) \wedge \mathcal{E}(X, Y')]^{1/2}.$$

Similarly, we have

$$\Pr[\mathcal{E}(X, Y) \wedge \mathcal{E}(X, Y')] = \sum_{j=1}^{m} \sum_{s=1}^{m} \Pr[\mathcal{E}(X, y_j) \wedge \mathcal{E}(X, y_s)] \Pr[Y = y_j] \Pr[Y = y_s],$$

and

$$\Pr[\mathcal{E}(X,Y) \wedge \mathcal{E}(X',Y) \wedge \mathcal{E}(X,Y') \wedge \mathcal{E}(X',Y')] = \sum_{j} \sum_{s} \Pr[\mathcal{E}(X, y_j) \wedge \mathcal{E}(X, y_s)]^2 \Pr[Y = y_j] \Pr[Y = y_s].$$

Therefore, by Cauchy-Schwarz we have

$$\Pr[\mathcal{E}(X, Y) \wedge \mathcal{E}(X, Y')] \leq \Pr[\mathcal{E}(X, Y) \wedge \mathcal{E}(X, Y') \wedge \mathcal{E}(X', Y) \wedge \mathcal{E}(X', Y')]^{1/2}.$$

From here, to complete the proof is a trivial exercise. □

Let $Z \in \{0,1\}^n$ be the random variable $(z_1, \ldots, z_n)$, and consider $Q(Z)$. Fix a non-trivial partition $U_1 \cup U_2 = [n]$ and a non-empty subset $S \subseteq U_1$. Let I be an interval of length 1. We need to show that

$$\Pr[Q(Z) \in I]^4 = O\left(|S|^{-1/2} + |S|^{-1} \sum_{i \in S} d_i^{-1/2}\right).$$

(Recall that d_i is the number of indices $j \in U_2$ for which $|a_{ij}| \geq 1$).

Define $X := (z_i)_{i \in U_1}$ and $Y := (z_i)_{i \in U_2}$, and write $Q(Z) = Q(X, Y)$. Let z_i' be an independent copy of z_i and set $X' = (z_i')_{i \in U_1}$ and $Y' = (z_i)_{i \in U_2}$. Applying the decoupling lemma, it is enough to show

$$\Pr[Q(X, Y), Q(X', Y), Q(X, Y'), Q(X', Y') \in I] = O\left(|S|^{-1/2} + |S|^{-1} \sum_{i \in S} d_i^{-1/2}\right).$$

Observe that $R := Q(X, Y) - Q(X', Y) - Q(X, Y') + Q(X', Y')$ can be written as

$$R = \sum_{i \in U_1} \sum_{j \in U_2} a_{ij}(z_i - z_i')(z_j - z_j') = \sum_{i \in U_1} R_i w_i,$$

where for $i \in [n]$ we have $w_i = z_i - z_i'$, and R_i is the random variable

$$R_i := \sum_{j \in U_2} a_{ij} w_j.$$

Note that now the random variables $(R_i)_{i \in U_1}$ are independent of $(w_i)_{i \in U_1}$.

Consider the events $Q(X, Y) \in I, Q(X', Y) \in I, Q(X, Y') \in I, Q(X', Y') \in I$. If all these hold, then R lies in the interval $J := 2I - 2I$ of length 4. Therefore, it is enough to show that

$$\Pr[R \in J] = O\left(|S|^{-1/2} + |S|^{-1} \sum_{i \in S} d_i^{-1/2}\right).$$

Recall that for each $i \in U_1$, d_i is the number of coefficients $j \in U_2$ for which $|a_{ij}| \geq 1$. Therefore, for each $i \in S \subseteq U_1$ we can apply Erdős-Littlewood-Offord to R_i to obtain

$$\Pr[|R_i| \geq 1 \text{ for all } i \in S] \geq 1 - O(\sum_{i \in S} d_i^{-1/2}).$$

This bound is a bit wasteful, so we will improve it by using higher moments. For each $i \in S$, let I_i be the indicator random variable for the event $|R_i| \geq 1$. Then,

$$\Pr[|R_i| \geq 1] = \mathbb{E}[I_i] = 1 - O(d_i^{-1/2}).$$

By linearity of expectation we obtain

$$\mathbb{E}[\sum_{i \in S} I_i] = |S| - O(\sum_{i \in S} d_i^{-1/2}).$$

Since $d_i \geq 1$, we have at least one $j \in U_2$ for which $|a_{ij}| \geq 1$ which implies that $\mathbb{E}[I_i] \geq 1/2$. Thus, we also have

$$\mathbb{E}[\sum_{i \in S} I_i] \geq |S|/2.$$

Now let us compute the variance of $\sum_{i \in S} I_i$:

$$\operatorname{Var}(\sum_{i \in S} I_i) = \mathbb{E}[(\sum_{i \in S} I_i)^2] - \mathbb{E}[\sum_{i \in S} I_i]^2 \leq |S|^2 - \left(|S| - O\left(\sum_{i \in S} d_i^{-1/2}\right)\right)^2 = O(|S| \sum_{i \in S} d_i^{-1/2}).$$

By Chebyshev's inequality we obtain that

$$\Pr[\sum_{i\in S} I_i \leq |S|/4] = O(\frac{1}{|S|}\sum_{i\in S} d_i^{-1/2}).$$

Therefore, with probability $1 - O(\frac{1}{|S|}\sum_{i\in S} d_i^{-1/2})$ we have $|R_i| \geq 1$ for at least $|S|/4$ values of $i \in S$.

Condition on the R_i's for all $i \in U_1$, and assume that $|R_i| \geq 1$ for at least $|S|/4$ values $i \in S$ (call this event $\mathcal{E}_1$). By the Erdős-Littlewood-Offord inequality we obtain

$$\Pr[R \in J \mid \text{ the } R_i\text{'s are fixed}] = O(|S|^{-1/2}).$$

Therefore,

$$\Pr[R \in J] = \Pr[\neg\mathcal{E}_1] + \Pr[\mathcal{E}_1 \wedge R \in J] \leq O(\frac{1}{|S|}\sum_{i\in S} d_i^{-1/2}) + O(|S|^{-1/2})$$

as desired. □

15 Beating very small probabilities

In this section we illustrate how to beat very small probabilities in certain scenarios. First, we will consider a problem that was recently introduced by Deneanu and Vu regarding the probability of an $n \times n$, ± 1 matrix to be *normal* (will be defined bellow). Later, we will consider the problem of giving a non-trivial upper bound on the probability for a random $n \times n$, ± 1 matrix to be *Hadamard* (will be defined in the relevant subsection).

15.1 Probability for being a normal matrix

Consider an $n \times n$ matrix M with entries from $\{\pm 1\}$. We say that M is *normal* if and only if $MM^T = M^TM$. Clearly, every symmetric matrix is normal, and therefore, the number of $n \times n$, ± 1 normal matrices, denoted by $\mathcal{N}(n)$, is at least $2^{\binom{n+1}{2}}$. It is not too hard to construct normal matrices which are not symmetric (exercise!), and a natural question to ask is whether one can find 'many' non-symmetric, normal matrices? In a recent paper, Deneanu and Vu introduced this problem and conjectured that the answer is 'no':

Conjecture 15.1 (Deneanu-Vu).

$$\mathcal{N}(n) = 2^{(1+o(1))\binom{n+1}{2}}.$$

Translating the above into the language of probability, let $\nu_n := |\mathcal{N}(n)|/2^{n^2}$ be the probability that a random $n \times n$, ± 1 matrix M_n is normal. The above conjecture is equivalent to

Conjecture 15.2 (Deneanu-Vu).

$$\nu_n = 2^{-\frac{1}{2}n^2 + o(n^2)}.$$

Note that the bound we are trying to prove is much smaller than exponential and therefore all standard large deviation inequalities (at least to the best of my knowledge) fail in tackling this problem. Deneanu and Vu have managed to obtain such bounds by using some rank arguments in a clever way and they obtained the first non-trivial upper bound on ν_n.

Theorem 15.3.

$$\nu_n \leq 2^{-(0.302+o(1))n^2}.$$

The proof of Deneanu and Vu is not very long, but is quite sophisticated and technical so we won't give it in full (we will briefly sketch their ideas throughout the section). Instead, we will prove a weaker bound (but yet of the form $2^{-\alpha n^2}$ for some $\alpha > 0$) using a recent (and simple) anticoncentration inequality (which at some point I'll also add to these notes). The advantage is that this proof is based on a quite general principle and can be applied for other counting problems. The disadvantage is that the obtained bound is weaker than the one of Deneanu and Vu. But, at the end of the section we will describe how to combine both approaches to obtain a slight improvement on their bound (so the final bound won't have such a short and elegant bounf...). Before stating the theorem that we wish to prove, we need some notation. Recall that M is normal if and only if $MM^T - M^TM = 0$. For technical reasons in the proof (which is based on an inductive construction) we need a bound on the number of $n \times n$, ± 1 matrices for which $MM^T - M^TM = N$, where N is any arbitrary (but fixed) matrix. This leads us to the following definition:

Definition 15.4. *An $n \times n$ matrix M is said to be N-*normal *if and only if*

$$MM^T - M^TM = N.$$

Note that if we take $N = 0$ then N-normal is just the standard definition of normal.

For an $n \times n$ matrix, we let $\mathcal{N}(N)$ be the set of all $n \times n$, ± 1 matrices which are N-normal. Recall that in the special case where $N = 0$ we have $\mathcal{N}(n) = \mathcal{N}(N)$ is the set of all $n \times n$, ± 1 normal matrices.

Observe that every matrix M is N-normal with respect to the matrix $N := MM^T - M^TM$. The main goal is to show that no family $\mathcal{N}(N)$ is 'too large'. Moreover, note that N is always a symmetric matrix.

Deneanu and Vu actually proved the following:

Theorem 15.5 (Deneanu-Vu)**.** *Let N be any (but fixed) $n \times n$ matrix. Let M be a randomly chosen $n \times n$, ± 1 matrix. Then*

$$\Pr[M \textit{ is } N\textit{-normal}] \leq 2^{-\alpha n^2 + o(n^2)},$$

where $\alpha = 0.302$.

We give a simple proof for Theorem 15.5 with $\alpha =$??. We then sketch the ideas of Deneanu and Vu to obtain the better bound, and discuss how both methods can be combined to obtain slight improvement (but still far from the conjectured $\alpha = 0.5$ unfortunately...). Before proving the theorem we need the following preliminary lemmas.

The basic tool we are going to use is the following simple observation by Odlyzko:

Lemma 15.6. *For any d-dimensional subspace $W \subseteq \mathbb{R}^n$ we have*

$$|W \cap \{\pm 1\}^n| \leq 2^d.$$

Sketch. Note that a d-dimensional space depends on d coordinates. Therefore, it contains at most 2^d ± 1 vectors. □

The following lemma is a trivial exercise in linear algebra NOT SURE YET THAT REALLY NEED IT:

Lemma 15.7. *Suppose that M is a $k \times n$ matrix of rank r, u is any vector in $\mathbb{R}^k$, and consider*

$$Mx = u.$$

Then, there are at most 2^{n-r} solution vectors $x \in \{\pm 1\}^n$.

The next lemma is a simple corollary from Odlyzko's observation (Lemma 15.6).

Lemma 15.8. *Let M be a randomly chosen $m \times m$ matrix with entries taken from $\{\pm 1\}$. Let $0 < \gamma < 1$ be any fixed number. Then,*

$$\Pr[rank(M) \leq \gamma m] \leq 2^{-(1-\gamma)^2 m^2 + O(m)}.$$

Proof. Expose M in m steps, one row at a time. Let $1 \leq r \leq \gamma m$ be any integer, and let $\mathcal{E}_r$ be the event that $rank(M) = r$. By a loss of an $\binom{m}{r}$ factor, we may assume that its first r rows are of rank exactly r and the rest lie in their span. Note that by Lemma 15.6 we have that there are at most 2^r vectors from $\{\pm 1\}^m$ at the span of the first r rows. Therefore, for every $r < k \leq m$, the probability that the kth row lies in this span is at most 2^{r-m}, which gives us

$$\Pr[\mathcal{E}_r] \leq \binom{m}{r} 2^{-(m-r)^2}.$$

All in all, the probability that M has rank at most γm is upper bounded by

$$\sum_{r=1}^{\gamma m} \Pr[\mathcal{E}_r] \leq 2^{-(1-\gamma)^2 m^2 + O(m)},$$

as desired. This completes the proof. □

Next, we wish to show that 'most' matrices satisfy some additional rank-type conditions. This will enable us to use our new anti-concentration inequality. Specifically, let us define:

Definition 15.9. *Let r and ℓ be two integers. An $m \times n$ matrix M is said to admit an (r, ℓ)-*rank partition *if and only if one can partition the columns of M into ℓ disjoint subsets, each of which corresponds to a submatrix of rank at least r.*

In the following lemma we upper bound the probability for a random matrix not to admit an (r, ℓ)-rank-partition.

Lemma 15.10. *Suppose that M is a random $m \times n'$ matrix with all entries in $\{\pm 1\}$, and let $0 < \gamma < 1$ be any fixed constant.*

$$\Pr[M \textit{ does not admit a } (\gamma m, \ell)\textit{-rank partition}] \leq 2^{-(1-\gamma)^2 m n' + (1-\gamma)^2 m^2 \ell + O(n')}.$$

Proof. NOTE SUPER FORMAL – NEED SOME ASSUMTIONS ABOUT n' AS WELL. WAS TO GENEROUS IN WRITING $O(m)$ WHENEVER NEEDED. Before exposing M, let us partition its columns into $t := n'/m$ disjoint subsets of size exactly m each, and let $A_1, \dots, A_t$ denote the corresponding $m \times m$ sub matrices. By Lemma 15.8, for each $1 \le i \le t$ we have

$$\Pr[rank(A_i) \le \gamma m] \le 2^{-(1-\gamma)^2 m^2 + O(m)}.$$

Therefore, the probability to have at least $t - \ell$ indices $1 \le i \le t$ with $rank(A_i) \le \gamma m$ is at most

$$\sum_{k=t-\ell}^{t} \binom{t}{k} \left(2^{-(1-\gamma)^2 m^2 + O(m)}\right)^k \le 2^t 2^{-(1-\gamma)^2 mn' + (1-\gamma)^2 m^2 \ell + O(m)}$$

$$= 2^{-(1-\gamma)^2 mn' + (1-\gamma)^2 m^2 \ell + O(m)}.$$

This completes the proof. □

Finally, the following lemma is our anti-concentration inequality which is the main tool we are going to use.

Lemma 15.11. *Suppose that M is an $m \times n'$ matrix which admits a (γ, ℓ)-rank partition. Let $x \in \{\pm 1\}^{n'}$ be a vector chosen uniformly at random. Then, for all $u \in \mathbb{R}^m$ we have*

$$\Pr[Mx = u] \le \left(2^{-\ell}\binom{\ell}{\ell/2}\right)^r.$$

In particular, if ℓ is large enough, then the right hand side is approximately $\left(\sqrt{\frac{2}{\pi\ell}}\right)^r \le \ell^{-r/2}$.

Now we are ready to prove Theorem 15.5.

NOTE THAT THE FIRST PAGE AND A HALF ARE NOT PART OF THE PROOF BUT MORE A DISCUSSION AND NOTATION. MAYBE WE CAN TAKE IT OUT OF THE PROOF AS A PRELIMINARY DISCUSSION?

Proof of Theorem 15.5. Given any matrix X we let $r_i(X)$ and $c_i(X)$ be its ith row and column, respectively. With this notation, note that for a given matrix M, being N-normal is equivalent to:

$$r_i(M) r_j^T(M) - c(M)_i^T c(M)_j = N_{ij}, \text{ for all } 1 \le i, j \le n. \tag{6}$$

Suppose now that M is N-normal and for all $1 \le k \le n$ let

$$M = \begin{bmatrix} A_k & B_k \\ C_k & D_k \end{bmatrix},$$

where A_k is a $k \times k$ matrix. Observe that with this notation in hands, by distinguishing between few cases, (6) is equivalent to saying that M satisfies all the following:

(*i*) For all $1 \le i, j \le k$ we have

$$r(A_k)_i r(A_k)_j^T + r(B_k)_i r(B_k)_j^T - c(A_k)_i^T c(A_k)_j - c(C_k)_i^T c(C_k)_j = N_{ij}.$$

(*ii*) For all $1 \le i \le k$ and $1 \le j \le n - k$ we have

$$r(A_k)_i r(C_k)_j^T + r(B_k)_i r(D_k)_j^T - c(A_k)_i^T c(B_k)_j - c(C_k)_i^T c(D_k)_j = N_{i,k+j}.$$

(iii) For all $1 \leq i, j \leq n-k$ we have

$$r(C_k)_i r(C_k)_j^T + r(D_k)_i r(D_k)_j^T - c(B_k)_i^T c(B_k)_j - c(D_k)_i^T c(D_k)_j = N_{k+i,k+j}.$$

Now, suppose we want to construct an N-normal matrix in $n-1$ steps. We will do it as follows: For every $1 \leq k \leq n-1$, in Step k we completely reveal all the entries in the kth row and column along with the diagonal element $d_{k+1} := M_{k+1,k+1}$. Let M_k be the structure obtained after $k+1$ steps, then we have

$$M_k = \begin{bmatrix} A_k & B_k & \\ C_k & d_{k+1} & * \\ & * & * \end{bmatrix},$$

where the $*$'s are the parts of D_k which remain unknowns at this step. Observe that A_k, the first column of B_k (together with the diagonal element d_k), and the first row of C_k, form the matrix A_{k+1} as defined above. In particular, the matrix A_{k+1} is already determined after this step. Moreover, both B_{k+1} and C_{k+1} are determined up to the last row and last column, respectively. In step $k+1$ we reveal $r_{k+1}(D_{k+1})$ and $c_{k+1}(D_{k+1})$ (we can forget about d_{k+2} for now). In order to make M_{k+1} 'valid' (that is something that can potentially being extended into an N-normal matrix by filling out D_{k+1}), we need that for all $1 \leq i \leq k$ we have

$$r_{k+1}(A_{k+1})r_i(A_{k+1})^T + r_{k+1}(B_{k+1})r_i^T(B_{k+1}) - c_{k+1}(A_{k+1})^T c_i(A_{k+1}) - c_{k+1}(C_{k+1})^T c_i(C_{k+1}) = N_{k+1,i}$$

Observe that only the second and the fourth summands involve unknowns, and therefore there exists $N'_{k+1,i}$ for which the above condition is equivalent to showing

$$r_{k+1}(B_{k+1})r_i^T(B_{k+1}) - c_{k+1}(C_{k+1})^T c_i(C_{k+1}) = N'_{k+1,i}, \tag{7}$$

where $N'_{t,i}$ is uniquely determined by the previously exposed entries.

Let $N' := (N'_{k+1,i})_{i=1}^k$ (considered as a column vector) be the vector of all 'restriction' coming from the above equality. Moreover, let $T_k = [U\ V]$ be a $k \times 2(n-k-1)$ matrix, where U is the matrix obtained from B_k by deleting its first column, and V is C_k^T minus its first column, and let $x_k := \begin{bmatrix} r_{k+1}^T(B_{k+1}) \\ -c_{k+1}(C_{k+1}) \end{bmatrix}$. The above condition is the same as saying that

$$T_k x_k = N'. \tag{8}$$

Another simple observation to make is the following: Suppose we have already exposed

$$M_k = \begin{bmatrix} A_k & B_k \\ C_k & * \end{bmatrix}.$$

Let $f(M_k)$ be the number of ± 1 matrices D_k for which

$$\begin{bmatrix} A_k & B_k \\ C_k & D_k \end{bmatrix}$$

is N normal. By a similar reasoning as we obtained (7), one can easily show that D_k should be an N'-normal for some $(n-k) \times (n-k)$ matrix N' which is already determined by M_k. Therefore, letting $f(k) = \sup_{M_k} f(M_k)$, for all $1 \leq k \leq n$ we have

$$|\mathcal{N}(n)| \leq f(k)|\mathcal{N}(n-k)|. \tag{9}$$

In order to complete the proof, it is enough to prove that for some $\beta, \delta > 0$ we have

$$f(\beta n) \leq 2^{(2\beta-\beta^2)n^2-\delta n^2}.$$

Indeed, assuming this, using the trivial bound $|\mathcal{N}(n-\beta n)| \leq 2^{(1-\beta)^2 n^2}$, using (9) we obtain

$$|\mathcal{N}(n)| \leq 2^{(2\beta-\beta^2)n^2-\delta n^2+(1-\beta)^2 n^2} = 2^{(1-\delta)n^2},$$

as desired.

Claim 15.12. *There exist $\beta, \delta > 0$ for which $f(\beta n) \leq 2^{(2\beta-\beta^2)n^2-\gamma n^2}$.*

Proof. Our plan is to expose $M_{\beta n}$ in βn steps, where in each step $1 \leq k \leq \beta n$ we expose M_k. Then, we want to show that either (8) has 'not too many' solutions, or M_k has some unlikely structure and therefore there are 'not too many' options for such an M_k.

More formally, for all $1 \leq k \leq \beta m$, we are interested in whether the matrix T_k (which is the $k \times 2(n-k-1)$ matrix define above) admits a $(\gamma k, \ell_k)$-rank-partition (recall Definition 15.9), where $\ell_k = n'/2k$, and $n' = 2(n-k-1)$. If yes, then by Lemma 15.11 we obtain that the number of solutions to $T_k x_k = N'$ is at most

$$2^{2(n-k-1)} \ell_k^{-\gamma k/2} \leq 2^{2(n-k)-\frac{\gamma k}{2}\log\frac{n}{2k}}. \tag{10}$$

Note that in the last inequality we used the fact that $(n-k-1)/k \geq n/2k$ for all $k \leq \beta n$. In particular, we need $\beta \leq 1/2$.

Suppose that for some (say) $k \geq \beta n/2$ the matrix T_k does not admit such a partition. Then, by Lemma 15.10 there are at most

$$2^{kn'-(1-\gamma)^2 kn'+(1-\gamma)^2 k^2 \ell + O(n')} = 2^{kn'-\delta_1 kn'}$$

such matrices, where $\delta_1 > 0$ is some constant depending on β and γ. This immediately implies that the number of options for

$$M_{\beta n} = \begin{bmatrix} A_{\beta n} & B_{\beta n} \\ C_{\beta n} & * \end{bmatrix}$$

with the property that for some $k \geq \beta n/2$ the matrix T_k does not admit a $(\gamma k, \ell_k)$-rank-partition is at most

$$2^{(2\beta-\beta^2)n^2-\delta_2 n^2}$$

for some $\delta_2 > 0$.

Next, suppose that for all $k \geq \beta n/2$ we have that T_k admits a $(\gamma k, \ell_k)$-rank-partition. Then, by (10) we have that the number of matrices $M_{\beta n}$ satisfying this property is at most

$$2^{(2\beta-\beta^2)n^2-\sum_{k=\beta n/2}^{\beta n}\frac{\gamma k}{2}\log\frac{n}{2k}} = 2^{(2\beta-\beta^2)n^2-\delta_3 n^2}$$

for some $\delta_3 > 0$. All in all, the number of possible choices for $M_{\beta n}$ is at most

$$2^{(2\beta-\beta^2)n^2-\delta_2 n^2} + 2^{(2\beta-\beta^2)n^2-\delta_3 n^2} \leq 2^{(2\beta-\beta^2)n^2-\delta n^2},$$

where $\delta = \min\{\delta_2, \delta_3\} - o(1)$, as desired. □

This completes the proof of the theorem. □

Now we wish to discuss the strategy from Deneanu and Vu and explain how to obtain a slight improvement by combining their proof with our technique.

A key idea in their proof is to partition the set $\mathcal{N}(n)$ into equivalent classes of permuted matrices. Before explaining the ideas and the motivation for doing so we start with the following definition:

Definition 15.13. *For any $\sigma \in S_n$ and for any $n \times n$ matrix M, define*

$$M_\sigma = P_\sigma M P_\sigma^T,$$

where P_σ is the matrix representing σ. That is, M_σ is the matrix obtained from M by permuting the row and columns according to σ.

Given the above definition, we can form equivalence classes as follows: for $n \times n$ matrices M and M' define

$$M \leftrightarrow M' \Leftrightarrow \exists \sigma \in S_n \text{ such that } M_\sigma = M'.$$

With this notation in hands, let us update the definition of being N-normal for the equivalence classes.

Definition 15.14. *Let N be a fixed $n \times n$ matrix. We say that an $n \times n$ matrix M is N-*normal-equivalent *if and only if there exists $\sigma \in S_n$ such that $MM^T - M^TM = C_\sigma$.*

The following proposition is trivial.

Proposition 15.15. *Let $\sigma \in S_n$, then M is N-normal-equivalent if and only if M_σ is N-normal-equivalent.*

Note that the above proposition shows that equivalence classes preserve the property of being N-normal-equivalent. Moreover, as every equivalence class is of size at most $n! = 2^{o(n^2)}$, it is enough to count the number of equivalence classes which are N-normal-equivalent for the same matrix N, while considering a carefully selected representative from each such class.

The key idea behind this approach is that given any matrix M, one can find a permutation σ such that by considering $M := M_\sigma$, we have a 'very good control on the ranks' $\text{rank}(T_k)$ for all $1 \leq i \leq n$. In particular, as our goal is to solve something of the form

$$T_k x_k = N',$$

by Lemma 15.7 it seems helpful to have these ranks as large as possible (so the number of solutions is small). Before stating writing what 'very good control on the ranks' means, we need to define the following functions (for all $k \leq t$):

$$R_{k,t}(i) := \begin{cases} i & \text{if } 0 < i \leq k \\ k & \text{if } k < i \leq t \\ k + t - i & \text{if } t < i \leq 2n - k - t \\ 2n - 2i & \text{if } 2n - k - t < i \leq n \end{cases}$$

Having defined $R_{k,t}$ we are ready to state the following key lemma in their proof:

Lemma 15.16 (Permutation Lemma)**.** *Let M be any (fixed) $n \times n$ matrix. Then, there exists $k, t \in \mathbb{N}$ and $\sigma \in S_n$ such that M_σ satisfies:*

$$\textit{rank}(T_i) = R_{k,t}(i) \textit{ for all } 1 \leq i \leq n.$$

From now on, the proof strategy is clear. Expose M (actually M_σ) like discussed above, where at every step i we form the structure M_i. Using the Permutation lemma we know that at every step we have some restrictions on the number of solutions to

$$T_i x_i = N',$$

and clearly they accumulate to an upper bound of the form $2^{n^2-\alpha n^2}$ for some α. It turns our that the bound $\alpha = 0.25$ is quite simple to obtain using this approach, and the main difficulty in their paper is to get something better.

Before diving into more details, we need the following definition:

Definition 15.17. *Let $\mathcal{N}_{k,t}(N)$ be the collection of all N-normal matrices M with ± 1 entries which satisfy the conclusion of the Permutation lemma with k, t.*

Remark 15.18. *As from now on N is fixed, we will simply write $\mathcal{N}_{k,t}$ instead of $\mathcal{N}_{k,t}(N)$.*

In the following lemma they showed how the special form obtained by the Permutation lemma gives tight control on $\Pr[M \text{ is } N\text{-normal-equivalent}]$. We will use D as the diagonal entries of M.

Lemma 15.19 (Recursion Lemma)**.** *For all $i < j$, let $X_{i:j}$ be any specific outcome of the x_ℓ's where $i \leq \ell \leq j$. Then, for any $1 \leq k \leq t \leq n$ and $1 \leq i \leq n$ we have*

$$\sup_{D, X_{1:i-1}} \Pr[M \in \mathcal{M}_{k,t} \mid D, X_{1:i-1}] \leq \begin{cases} 2^{-R_{k,t}(i-1)} \sup_{D,X_{1:i}} \Pr[M \in \mathcal{M}_{k,t} \mid D, X_{1:i}] & \textit{if } 2n - 2i > \textit{rank}_i(M) \\ 2^{-(n-i)^2 + o(n^2)} & \textit{if } 2n - 2i \leq \textit{rank}_i(M). \end{cases}$$

Proof. Without loss of generality we can assume that M itself satisfies the conclusion of the Permutation lemma with k, t (as otherwise define $M := M_\sigma$). Moreover, as $M \in \mathcal{M}_{k,t}$ we have that M is C-normal, and therefore

$$T_{i-1} x_i = c,$$

for some c which is uniquely determined by C, D and $x_1, \dots, x_{i-1}$. Hence, conditioned on $x_1, \dots, x_{i-1}$ and D, the vector x_i belongs to a (shifted) subspace H of dimension $\max\{2n - 2i - \operatorname{rank}_{i-1}(M), 0\}$. Moreover, observe that by the Permutation lemma we have $\operatorname{rank}_{i-1}(M) = R_{k,t}(i-1)$. Using Odlyzko's observation (Lemma 15.6) we obtain that

$$\Pr[M \in \mathcal{M}_{k,t} \mid D, X_{1:i-1}] = \sum_{X_i \in H} \Pr[M \in \mathcal{M}_{k,t} \mid D, X_{1:i}] \Pr[X_i \in H]$$

$$\leq 2^{-(2n-2i) + \max\{2n-2i-\operatorname{rank}_{i-1}(M), 0\}} \sup_{X_i \in \{\pm 1\}^{2n-2i}} \Pr[M \in \mathcal{M}_{k,t} \mid D, X_{1:i}].$$

In order to complete the proof we distinguish between two cases:

Case 1. $2n - 2i > \operatorname{rank}_{i-1}(M)$. In this case, as

$$\max\{2(n-i) - \operatorname{rank}_{i-1}(M), 0\} = 2(n-i) - \operatorname{rank}_{i-1}(M),$$

we immediately get the bound

$$\Pr[M \in \mathcal{M}_{k,t} \mid D, X_{1:i-1}] \leq 2^{-R_{k,t}(i)} \sup_{X_i \in \{\pm 1\}^{2n-2i}} \Pr[M \in \mathcal{M}_{k,t} \mid D, X_{1:i}].$$

Case 2. $2(n-i) \leq \text{rank}_{i-1}(M)$. In this case, as $|R_{k,t}(\ell) - R_{k,t}(\ell-1)| \leq 2$ for all ℓ, we obtain that for all $0 \leq j \leq n-i$ we have

$$2n - 2i - 2j \leq \text{rank}_{i-1+j}(M).$$

Therefore, by the recurrence relation we've obtained, we conclude that

$$\Pr[M \in \mathcal{M}_{k,t} \mid D, X_{1:i-1}] \leq 2^{-2(n-i)-2(n-i-1)} \sup_{X_1,\ldots,X_{i+1}} \Pr[M \in \mathcal{M}_{k,t} \mid D, X_{1:i+1}]$$

$$\leq 2^{-\sum_{j=0}^{n-i} 2j} = 2^{-(n-i)^2 + o(n^2)}.$$

This completes the proof. □

Remark 15.20. *Note that there is a very weak point in the above proof. Namely, whenever we are to expose x_i for some i, we use a bound which is based only on the rank of T_{i-1} and no other structure. Therefore, theoretically, if we can do slightly better, then we will obtain an improvement to Theorem 15.5. Basically, we could rewrite the above lemma as: the number of options for x_i is at most $2^{2(n-i)-R_{k,t}(i-1)}$ if $2n-2i > \text{rank}_i(M)$ or 1 otherwise.*

15.2 Combining all the above

Here we do the actual steps towards proving Theorem 15.5. Fix the matrix C, and our goal is to upper bound $\Pr[M \in \mathcal{M}_{k,t}]$ for all $k \leq t$. Then, as there are at most n^2 options for k and t, by a simple union bound we obtain the desired.

Note that for some specific values of k and t the problem is trivial. For example, it is quite obvious to see (WHY?) that $\mathcal{M}_{k,t}$ is empty whenever $k + t < n - 2, k > 2n/3$ or $t + k/2 > n$ (to see is, recall the definition of $R_{k,t}$). We distinguish between two cases, where in each case we take a slightly different approach. In the first case we provide an upper bound on $\Pr[M \in \mathcal{M}_{t,k}]$ when $2t + k$ is close to $2n$, while in the second case we provide good bounds when $2t + k$ is far from $2n$. Afterwards, we combine the two results to get the desired bound through some optimization.

15.2.1 The first case

Lemma 15.21. *For $1 \leq k \leq 2n/3$ and $\frac{k}{2} < n - t \leq k$ we have*

$$\Pr[M \in \mathcal{M}_{k,t}] \leq \begin{cases} 2^{n^2+k^2+t^2+kt-2kn-2nt+o(n^2)} & \text{if } k \geq n/2 \\ 2^{t^2-3k^2+2kn+kt-2nt+o(n^2)} & \text{if } k \leq n/2 \end{cases}$$

In particular we have

$$\Pr[M \in \mathcal{M}_{k,t}] \leq 2^{-0.25n^2+o(n^2)}.$$

15.2.2 The second case

Let $M \in \mathcal{M}_{k,t}$ and consider T_t. Recall that T_t has t rows and $2(n-t-1)$ columns, rank k and the the additional property that for any $1 \leq i \leq n-t-1$, the deletion of both its ith and $(n-t-1+i)$th columns reduces the rank by at least 1. The following definitions will be convenient for us:

Definition 15.22. *Let M be a fixed $q \times 2m$ matrix. We say that M has the property $\mathcal{P}$ if for any $1 \leq i \leq m$, by deleting both the ith and the $(i+m)$th columns we reduce the rank of M by at least* 1.

Definition 15.23. *Let $A := \{\beta \mid \Pr[M \text{ is } C\text{-normal}] \leq 2^{-(\beta+o(1))n^2}\}$. Define*

$$\alpha = \sup A - 0.0001.$$

Observe that Lemma 15.21 implies that $\alpha \geq 0.2499$. We wish to prove the following:

Lemma 15.24. *Given $1 \leq k \leq t \leq n$ we have that*

$$\Pr[M \in \mathcal{M}_{k,t}] \leq 2^{(1-\alpha)t^2-k^2/2-n^2+nk+o(n^2)}.$$

Intuitively, the above lemma asserts that if we take a $t \times 2(n-t-1)$ ±1 matrix at random, then the probability that it satisfies $\mathcal{P}$ is very small. Note that as was mentioned earlier, the probability is 0 unless $n-k-2 \leq t \leq n-k/2$.

16 Inverse theorems

In this section we consider some inverse theorems in additive number theory. The main theorem we are interested at (but won't prove in full) is the so called Freiman's inverse theorem, that concerns the structure of sets with small *sumsets*. Suppose that A is a subset of an abelian group G, and define

$$A + A := \{a + a' : a, a' \in A\}.$$

Clearly, if $|A| = n$ then $|A + A| \geq n$, and equality can occur for example if A is a subgroup. On the other hand, we trivially have

$$|A + A| \leq \binom{n+1}{2},$$

which is also tight if, for example, $G = \mathbb{Z}$ and $A = \{1, 5, 5^2, \ldots, 5^{n-1}\}$. The following simple exercise is nice to start understanding what's going on here.

Exercise 16.1. *Let $A \subseteq \mathbb{Z}$ be a set of size n. Then $|A + A| \geq 2n - 1$. Moreover, equality holds if and only if A is an arithmetic progression.*

The main question we want to deal with is about the structure of sets with small, but not 'too small' sumsets. For example, what is the structure of a subset $A \subseteq \mathbb{Z}$ with $|A + A| \leq 2000|A|$?

The following notion will play an important role from now on: Let $x_0, \ldots, x_d \in \mathbb{Z}$ and let $M_1, \ldots, M_d$ be positive integers. The set

$$P = \left\{ x_0 + \sum_{j=1}^{d} m_j x_j \mid 0 \leq m_j \leq m_j - 1 \right\}$$

is called a d-dimensional arithmetic progression, or a general arithmetic progression (GAP for short) of dimension d. We say that P is *proper* if $|P| = M_1 \cdots M_d$. It is relatively straight forward to prove that if P is proper then $|P + P| \leq 2^d|P|$.

Freiman's inverse theorem says, more or less, that the converse also holds:

Theorem 16.2 (Freiman's inverse theorem)**.** *Let $A \subseteq \mathbb{Z}$ be a subset of size n. Suppose that $|A+A| \leq C|A|$ for some C. Then A is contained in a proper d-dimensional GAP of size at most Kn, where d and K depend only on C.*

There are many more results of a similar flavor which are based on this theorem. The type of problems we will mostly be interested in this section (after understanding few ingredients in the proof of the above theorem of course...) are an 'inverse Littlewood-Offord'-type statements. That is to say, suppose that $(a_1, \dots, a_n)$ is sum integer valued vector and that $X_1, \dots, X_n$ are iid random variable, each of which is distributed as

$$\Pr[X = 1] = \Pr[X = -1] = 1/2.$$

We've already seen that for every $m \in \mathbb{Z}$ we have

$$\Pr[\sum a_i X_i = m] = O(1/supporta).$$

Moreover, we clearly have

$$\sup_m \Pr[\sum X_i a_i = m] \geq 2^{-n}.$$

The main question is: suppose that

$$\sup_m \Pr[\sum X_i a_i = m] \geq n^{-C}$$

for some C. What is the additive structure of our vector a? note that if this probability is large, then intuitively it means that there are many cancelations in the sums $\sum \pm a_i$, so intuitively one would expect that the set $A = \{a_i \mid i\}$ has a relatively small sumset. We will discuss this problem in a much more detail later.

16.1 Plünnecke's inequalities

One of the main ingredients in the proof of Freiman's theorem is the following theorem:

Theorem 16.3 (Plünnecke-Ruzsa)**.** *Suppose that $|A + A| \leq C|A|$ for some $C > 0$. Then, for any k, ℓ we have*

$$|kA - \ell A| \leq C^{k+\ell}|A|.$$

We won't prove this theorem here but will give some details. We will start with a different looking statement that requires few definitions. A *Plünnecke* graph of level h is a directed graph G on some vertex set $V_0 \cup \dots V_h$ satisfying:

1. All edges in $E(G))$ are edges from some V_i to V_{i+1}.

2. (Forward splitting of paths) Let $0 \leq i \leq h - 2$ and suppose that $u \in V_i$, $v \in V_{i+1}$ and $w_1, \dots, w_k \in V_{i+2}$ are such that uv and all the vw_i's are edges of G. Then, there are distinct $v_1, \dots, v_k \in V_{i+1}$ such that all of the uv_j's and the v_jw_j's are edges in G.

3. (Backward splitting of paths) Let $0 \leq i \leq h-2$ and suppose $u_1, \ldots, u_k \in V_i$, $v \in V_{i+1}$ and $w \in V_{i+2}$, with $u_i v$ and vw being edges for all i. Then there are distinct $v_1, \ldots, v_k \in V_{i+1}$ such that all of the $u_j v_j$ and $v_j w$ are edges.

Now, let $X \subseteq V_0$ and let $N_i(X)$ be the set of all vertices in V_i which can be reached by path starting from some $x \in X$. The ith *magnification ratio* of G, $D_i(G)$, is

$$D_i(G) = \inf_{X \subseteq V_0, X \neq \emptyset} \frac{|N_i(X)|}{|X|}.$$

Proposition 16.4 (Plünnecke)**.** *Let G be a Plünnecke graph of level $h \geq 2$. Then we have the inequalities*

$$D_1 \geq D_2^{1/2} \geq D_3^{1/3} \geq \ldots \geq D_h^{1/h}.$$

The key step in deducing Theorem 16.3 is the following

Proposition 16.5. *Let A, B be subsets of an abelian group with $|A + hB| \leq C|A|$. Then, for any $h' \geq h$, there is a set $A' \subseteq A$ with $|A' + h'B| \leq C^{h'/h}|A'|$.*

Proof. Define a directed graph as follows. Set $V_i = A + iB$, and join $v \in V_i$ to $v' \in V_{i+1}$ if and only if $v' - v \in B$. This is (more or less trivially) a Plünnecke graph. The hth magnification ratio, D_h, is at most C because

$$\inf_{Z \subseteq A} \frac{|N_h(Z)|}{|Z|} \leq \frac{|N_h(A)|}{|A|} \leq \frac{|A + hB|}{|A|} \leq C.$$

In particular, for any $h' \geq h$, from Proposition 16.4 and the above estimate we can write

$$D_{h'} \leq D_h^{h'/h} \leq C^{h'/h}$$

which is equivalent to what we are trying to prove. □

It follows immediately from Proposition 16.5 (taking $h = 1$ and $B = A$) that if $|A + A| \leq C|A|$ then $|kA| \leq C^k|A|$ for any $k \geq 2$. In order to prove Theorem 16.3 we need the following lemma

Lemma 16.6. *Let U, V, W be subsets of an abelian group. Then*

$$|U||V - W| \leq |U + V||U + W|.$$

Proof. For any $d \in V - W$ fix $v(d) \in V$, $w(d) \in W$ with $v(d) - w(d) = d$. Define

$$\Phi : U \times (V - W) \to (U + V) \times (U + W)$$

by

$$\Phi(u, d) = (u + v(d), u + w(d)),$$

and show that this is injective. Indeed, if $\Phi(u, d) = (x, y)$, then by definition we have $x - y = d$. Now, after knowing d, we know $v(d)$ and $w(d)$ and this enables us to recover u as well. □

To complete the proof of Theorem 16.3: Suppose $|A + A| \leq C|A|$ and suppose without loss of generality that $\ell \geq k$. We may apply Proposition 16.5 twice to get $A'' \subseteq A' \subseteq A$ satisfying

$$|A' + kA| \leq C^k|A'|$$

and

$$|A'' + \ell A| \le C^{\ell}|A''|.$$

Then by Lemma 16.6 we have

$$|A''||kA - \ell A| \le |A'' + kA||A'' + \ell A| \le |A' + kA||A'' + \ell A| \le C^{k+\ell}|A'||A''| \le C^{k+\ell}|A||A''|.$$

This gives the desired.

16.2 Inverse Littlewood-Offord inequalities

Here, for convenience, we 'shift' our definition of GAP of dimension d from the previous section as follows:

$$P := \{a + m_1 v_1 + \ldots + m_d v_d \mid -M_j/2 < m_j < M_j/2 \text{ for all } 1 \le j \le d\}.$$

We say P is *symmetric* if $a = 0$. We say that P is *proper* if the map

$$(m_1, \ldots, m_d) \to m_1 v_1 + \ldots + m_d v_d$$

is injective. If P is symmetric and proper, we define the P-*norm* $\|v\|_P$ of a point $v = m_1 v_1 + \ldots + m_d v_d$ in P by

$$\|v\|_P = \left(\sum_{i=1}^{d} \left(\frac{|m_i|}{M_i} \right)^2 \right)^{1/2}.$$

Let $\bar{a} = (a_1, \ldots, a_n)$ be a vector and assume that

$$\Pr[\sum a_i x_i = 0] \approx 2^{-C}.,$$

where the x_i's are iid bernoulli (± 1) r.v. Moreover, assume that for some $\varepsilon > 0$ we have that

$$\Pr[\sum a_i x_i = 0] \ge \varepsilon \Pr[\sum a_i y_i = 0],$$

where the y_i's are iid 'lazy walks' variables.

We wish to prove the following:

Theorem 16.7 (Tao-Vu). *There is a constant C' such that the following holds. There exist integers $1 \le d \le C'$ and $M_1, \ldots, M_d \ge 1$ with*

$$\prod M_i \le C' 2^{n-C}$$

and non-zero elements $v_1, \ldots, v_d \in \mathbb{F}$ such that the following holds.

1. *The corresponding symmetric GAP P is proper and contains all the a_i's.*
2. *The a_i's have small P-norm:*

$$\sum_{j=1}^{n} \|a_j\|_P^2 \le C'.$$

3. The set $\{v_1, \ldots, v_d\} \cup \{a_1, \ldots, a_n\}$ *is contained in the set*

$$\{\frac{p}{q}v_1 : p, q \in \mathbb{Z}, q \neq 0, |p|, |q| \leq n^{o(n)}\}.$$

They work over $\mathbb{F} = \mathbb{Z}_{n^{n/2}}$. We are interested in much smaller primes so we don't really need the third condition in order to overcome union bound. Before turning into the proof we need quite a lot of preparation. First, we will use Freiman's theorem as stated here:

Theorem 16.8 (Freiman's inverse theorem)**.** *For any constant C there are d, δ such that: for any finite set A of integers with $|A + A| \leq C|A|$, there is a proper arithmetic progression P of rank d such that $A \subseteq P$ and $|A|/|P| \geq \delta$.*

We will not try to optimize or calculate the 'correct' dependency in any of the parameters. Moreover, Freiman's theorem can be also phrased in finite fields of a sufficiently large order (this is what we will actually use).

The following lemma (we skip this proof...) shows that given a GAP which is not proper, by a cost of a constant factor, embed it into a proper one.

Lemma 16.9 (GAP lies inside proper GAP)**.** *There is a constant C such that the following holds: Let P be a GAP of dimension d in an abelian group G. Suppose that every non-zero element of G has order at least $d^{Cd^3}|P|$. Then there exists Q which is a proper GAP of dimension at most d for which $P \subseteq Q$ and $|Q| \leq d^{Cd^3}|P|$.*

Another ingredient we are going to use is the following lemma:

Lemma 16.10 (Sumset estimates)**.** *Let A be a symmetric finite subset of an abelian group G such that $|4A| \leq C|A|$ for some $C \geq 1$. Then, for any $k \geq 4$ we have*

$$|kA| \leq \binom{C + k - 3}{k - 2} C|A|.$$

Proof. The proof is based on a covering argument of Ruzsa. Consider the sets $x + A$ as x ranges inside $3A$. Each such set has size exactly $|A|$ and is a subset of $4A$. Let $X \subseteq 3A$ be a largest subset for which $(x+A) \cap (x'+A) = \emptyset$ for all $x \neq x' \in X$. Clearly, $|X| \leq |4A|/|A| \leq C$. In particular, for all $y \in 3A$ there exists $x \in X$ such that $x+A$ intersects $y+A$. This implies that $y \in X+A-A = X+2A$ (we also used the fact that A is symmetric here) and we obtained

$$A + 2A = 3A \subseteq X + 2A.$$

Iterating this argument gives us

$$kA \subseteq 2A + (k-2)X$$

for all $k \geq 2$. Thus

$$|kA| \leq |2A||(k-2)X| \leq C|A||(k-2)X|.$$

Note that we trivially have

$$|\ell X| \leq \binom{|X| + \ell - 1}{\ell},$$

which together with the above completes the argument. □

To prove Theorem 16.7 we want to use Freiman's inverse theorem on the vector $\bar{a}$ (or to a large subset of its coordinates). Our goal now is to show that one can find such a subset A with small doubling.

We start with a Fourier analytic argument: Let $V = \{x \in \mathbb{F}^n : ax = 0\}$. By Fourier expansion we have

$$V(\bar{x}) = \frac{1}{|\mathbb{F}|} \sum_{\xi \in \mathbb{F}} e_p(\sum_i x_i a_i \xi),$$

where $p = |\mathbb{F}|$ and e_p is the primitive character $e_p(x) = e^{2\pi i x/p}$.

Consider X as a vector of ± 1 entries, we clearly have

$$\Pr[X \in V] = \mathbb{E}[1_{X \in V}] = \frac{1}{|\mathbb{F}|} \sum_{\xi \in \mathbb{F}} \prod_{j=1}^{n} \cos(2\pi a_j \xi / p).$$

In particular, after a short manipulation we obtain

$$\Pr[X \in V] \leq \frac{1}{p} \sum_{\xi \in \mathbb{F}} \prod_{j=1}^{n} \left(\frac{1}{2} + \frac{1}{2} \cos(2\pi a_j \xi / p) \right)^{1/2}.$$

Similarly, if we take Y to be a vector of iid r.v where each coordinate i satisfies

$$\Pr[y_i = 0] = 1 - \mu, \text{ and } \Pr[y_i = 1] = \Pr[y_i = -1] = \mu/2,$$

we obtain

$$\Pr[Y \in V] \leq \frac{1}{p} \sum_{\xi \in \mathbb{F}} \prod_{j=1}^{n} \left((1-\mu) + \mu \cos(2\pi a_j \xi / p) \right).$$

Define

$$f(\xi) = \prod_{j=1}^{n} \left(\frac{1}{2} + \frac{1}{2} \cos(2\pi a_j \xi / p) \right)^{1/2} \text{ and } g(\xi) = \prod_{j=1}^{n} \left((1-\mu) + \mu \cos(2\pi a_j \xi / p) \right),$$

and by a simple calculation TO SHOW we see that for all ξ we have $f(\xi) \leq g(\xi)^{1/4\mu}$. This in particular implies that if $\mu < 1/4$ then $f(\xi) \leq g(\xi)$ which give

$$\Pr[X \in V] \leq \Pr[Y \in V].$$

Now we want to understand what to do with the information that

$$\Pr[X \in V] \geq \varepsilon_1 \Pr[Y \in V].$$

Let ε_2 be a sufficiently small constant (depending on ε_1). Define the *spectrum* $\Lambda \subseteq \mathbb{F}$ of $\{a_1, \ldots, a_n\}$ as

$$\Lambda := \{\xi \in \mathbb{F} \mid f(\xi) \geq \varepsilon_2\}.$$

Note that Λ is symmetric. Next, we make the elementary observation that (say)

$$1 - 100\|x\|^2 \leq \cos(2\pi x) \leq 1 - \frac{1}{100}\|x\|^2.$$

This implies that

$$f(\xi) \le \exp\left(-\frac{1}{1000}\sum_{j=1}^{n} \|a_j\xi/p\|^2\right).$$

Therefore, as $f(\xi) \ge \varepsilon_2$ for all $\xi \in \Lambda$, there exists a constant $C(\varepsilon_2)$ such that

$$\left(\sum_{j=1}^{n} \|a_j\xi/p\|^2\right)^{1/2} \le C(\varepsilon_2)$$

for all $\xi \in \Lambda$.

We now show that Λ has small sumsets:

Lemma 16.11. *There is a constant depending on the ε's such that*

$$\Lambda \in [C^{-1}2^{-(n-d_\pm)}|\mathbb{F}|, C2^{-(n-d_\pm)}|\mathbb{F}|].$$

Proof. Note that by assumptions we have

$$\frac{1}{p}\sum_{\xi} f(\xi) \ge \varepsilon_1 \frac{1}{p}\sum g(\xi)$$

. But from the definition of Λ and the fact that we take $\mu < 1/4$ we have

$$\frac{1}{p}\sum_{\xi\notin\Lambda} f(\xi) \le \varepsilon_2^{1-4\mu}\sum_{\xi\notin\Lambda} f^{4\mu}(\xi) \le \varepsilon_2^{1-4\mu}\frac{1}{p}\sum_{\xi\in\mathbb{F}} g(\xi).$$

Now we are going to make an essential use of the fact that $\mu < 1/4$. Assuming this, we can make sure that the contribution of ξ outside Λ is negligible, by choosing ε_2 sufficiently small:

$$\sum_{\xi\in\Lambda} = \Theta(\sum_{\xi\in\mathbb{F}} f(\xi)) = \Theta(\sum_{\xi} g(\xi)) = \Theta(|\mathbb{F}|\Pr[X \in V]) = \Theta(2^{d_\pm - n}|\mathbb{F}|).$$

The bounds on $|\Lambda|$ now follow immediately from the fact that for any $\xi \in \Lambda$ we have $\varepsilon_2 \le f(\xi) \le 1$.

It thus remain to prove that the sumsets of Λ are not too large. Specifically, it is enough to prove that there is a constant C with

$$|4\Lambda| \le C|\Lambda|.$$

Note that for $\xi \in 4\Lambda$ we have by triangle inequality that

$$\left(\sum_{j=1}^{n} \|a_j\xi/p\|^2\right)^{1/2} \le C(\varepsilon_2),$$

for some $C(\varepsilon_2)$. From previous estimates we conclude that

$$f(\xi) \ge c(\varepsilon_2) > 0.$$

Therefore,

$$|4\Lambda| \le c(\varepsilon_2)^{-1}\sum_{\xi} f(\xi) \le c(\varepsilon_2)^{-1}\sum_{\xi\in\mathbb{F}} g(\xi) = \Theta(2^{d_\pm - n}|\mathbb{F}) = \Theta(|\Lambda|).$$

This completes the proof. □

Now we want to go back from Λ to $\bar{a}$ using the inverse Fourier transform. For any $x \in \mathbb{F}$ define $\|x\|_\Lambda$ by

$$\|x\|_\Lambda := \left(\frac{1}{|\Lambda|^2} \sum_{\xi,\xi' \in \Lambda} \|x \cdot (\xi - \xi')/p\|^2 \right)^{1/2}.$$

One can check that this quantity is between 0 and 1 and obeys the triangle inequality. In particular, we obtain

$$\|x\|_\Lambda \leq \left(\frac{1}{|\Lambda|^2} \sum_{\xi,\xi' \in \Lambda} \|x\xi/p\|^2 \right)^{1/2} + \left(\frac{1}{|\Lambda|^2} \sum_{\xi,\xi' \in \Lambda} \|x\xi'/p\|^2 \right)^{1/2} = 2 \left(\frac{1}{|\Lambda|^2} \sum_{\xi,\xi' \in \Lambda} \|x\xi/p\|^2 \right)^{1/2}.$$

Therefore by summing squares over all $\xi \in \Lambda$ we obtain

$$\sum_{j=1}^n \|a_j\|_\Lambda^2 \leq C(\varepsilon_2).$$

Therefore, we expect many of the a_j's to have small Λ-norm. On the other hand we will show that the set of elements with small such norm has constant doubling:

Lemma 16.12. *There is a constant C such that: Let $A \subseteq \mathbb{F}$ denote the 'Bohr set'*

$$A := \{x \in \mathbb{F} \mid \|x\|_\Lambda \leq \frac{1}{100}\}.$$

Then we have

$$C^{-1} 2^{n-d_\pm} \leq |A| \leq |A + A| \leq C 2^{n-d_\pm}.$$

17 Matrix analysis

In this section we discuss some more advance linear algebra. Later, we will show more combinatorial results, based on the tools which we develop here. We are going to assume that the reader is familiar with basic notions of linear algebra though.

17.1 A QR-decomposition of a matrix

The first observation that we make (left as an exercise) is the following: suppose that $X = (x_1, \ldots, x_k)$ is a linear independent k-tuple of vectors in an n-dimensional inner product space (such a space is referred to as *a finite Hilbert space*). Suppose that the x_i's are column vectors, and we say that a k-tuple $Y = (y_1, \ldots, y_k)$ is *biorthogonal* to X if $\langle x_i, y_j \rangle = \delta_{ij}$ for all i, j.

Exercise 17.1. *Given an X as above, there exists a k-tuple Y biorthogonal to X.*

In the above notation, the Gram-Schmidt procedure can be seen as a matrix factorization theorem. Given an n-tuple X of linearly independent vectors, the procedure gives another tuple $Q = (q_1, \ldots, q_n)$ whose entries are orthonormal vectors. For each $k = 1, 2, \ldots, n$, the vectors x_i's and q_i's have the same linear span. In a matrix notation this can be expressed as $X = QR$, where R is an upper triangular matrix and can be chosen in a way that all the diagonal entries are positive. This makes the factors Q and R both uniquely determined. Note that if the x_i's are not independent, then the procedure can be modified as follows: if x_k depends on $x_1, \ldots, x_{k-1}$, then one can choose $q_k = 0$. Otherwise, proceed as in Gram-Schmidt.If the kth column of Q is 0, then set the kth row of R to be 0. This gives $X = QR$, where R is upper triangular and Q has orthogonal columns, some of them are 0. Take the nonzero columns and extend to an orthonormal basis. Then replace the zero columns by the additional elements of the basis. This still gives $X = QR$, and now Q is orthonormal. This is called the *QR decomposition* of X.

Example 17.2. *Suppose that $X = (x_1, \ldots, x_n)$. Then,*

$$|\det(X)| \leq \prod_{j=1}^{n} \|x_j\|.$$

Indeed, write $X = QR$ as above. Then, since Q is unitary, $\det(X) = \det(Q)\det(R) = \det(R)$. *Moreover, as R is upper triangular, we obtain*

$$\det(R) = \prod_{j=1}^{n} R_{jj},$$

and by definition, we have, for all j, that

$$R_{jj} = x_j^T q_j.$$

Therefore, by Cauchy-Schwarz we have

$$|R_{jj}| \leq \|x_j\| \|q_j\| = \|x_j\|.$$

This gives the desired.

17.2 Linear operators and matrices

Let $\mathcal{L}(V, W)$ be the set of all linear operators from V to W. If we fix the bases of V and W, then each operator has a unique matrix associated with it. Note that the matrix representations are super nice if the bases are being chosen to be orthonormal. Indeed, suppose $A \in \mathcal{L}(V, W)$ and $E = (e_1, \ldots, e_n)$, $F = (f_1, \ldots, f_m)$ are orthonormal bases of V and W, respectively. Then, by orthonormality, observe that every vector $v \in W$ can be written as

$$v = \sum_{i=1}^{n} (f_i^* \cdot v) f_i,$$

and therefore, the ijth entry of A relative to these bases is $a_{ij} = f_i^* A e_j$. In particular, we obtain that the matrix of A relative to these bases is $F^* A E$.

18 Singularity over finite fields

In this section we consider a random ±1, $n \times n$ matrix M over a finite field $\mathbb{Z}_p$. The problem we are interested at is

Problem 18.1. *Estimate*

$$\sigma_n(p) := \Pr[M \textit{ is singular over } \mathbb{Z}_p].$$

We prove the following theorem:

Theorem 18.2. *For all $p \leq$ we have*

$$\sigma_n(p) = .$$

Proof. The proof is based on the Fourier method and is one of the rear cases where one can make all the calculations exact. Recall that in $\mathbb{Z}_p$ we have

$$\delta_0(x) = \frac{1}{p} \sum_{k \in \mathbb{Z}_p} \exp\left(\frac{2\pi i k x}{p}\right).$$

We use the notation $e_p(x) = \exp(2\pi i x/p)$ and simply write

$$\delta_0(x) = \frac{1}{p} \sum_{k \in \mathbb{Z}_p} e_p(kx).$$

Our goal is to estimate the probability for M being singular over $\mathbb{Z}_p$. Note that M is singular over $\mathbb{Z}_p$ if and only if there exists $a \in \mathbb{Z}_p^n \setminus \{0\}$ such that $Ma = 0$. Now, observe that for every a we have

$$\Pr[Ma = 0] = \left(\Pr\left[\sum_{i=1}^{n} a_i x_i = 0\right]\right)^n,$$

where the x_i's are iid bernouli ±1 random variables. Since

$$\Pr\left[\sum_i a_i x_i = 0\right] = \mathbb{E}\left[\delta_0\left(\sum_i a_i x_i\right)\right],$$

by the above equalities we obtain

$$\begin{aligned}
\Pr[Ma = 0] &= \left(\mathbb{E}\left[\delta_0(\sum_i a_i x_i)\right]\right)^n \\
&= \left(\mathbb{E}\left[\frac{1}{p}\sum_{k \in \mathbb{Z}_p} e_p\left(k \sum_{j=1}^{n} a_j x_j\right)\right]\right)^n \\
&= \left(\mathbb{E}\left[\frac{1}{p}\sum_{k \in \mathbb{Z}_p} \prod_{j=1}^{n} e_p\left(k a_j x_j\right)\right]\right)^n \\
&= \left(\frac{1}{p}\sum_{k \in \mathbb{Z}_p} \prod_{j=1}^{n} \mathbb{E}\left[e_p\left(k a_j x_j\right)\right]\right)^n \\
&= \left(\frac{1}{p}\sum_{k \in \mathbb{Z}_p} \prod_{j=1}^{n} \cos\left(\frac{2\pi k a_j}{p}\right)\right)^n
\end{aligned}$$

$$= p^{-n}\left(1+\sum_{k\neq 0}\prod_{j=1}^{n}\cos\left(\frac{\pi k a_j}{p}\right)\right)^{n},$$

where the last equality holds by a change of variable.

Next, let A be the set of all vectors in $\mathbb{Z}_p^n$ which have at least εn non zero coordinates. For every $a \in A$ we can write:

$$p(a) := p^{-n}\left(1+\sum_{k\neq 0}\prod_{j=1}^{n}\cos\left(\frac{\pi k a_j}{p}\right)\right)^{n} = p^{-n}\left(1+n\sum_{k\neq 0}\prod_{j=1}^{n}\cos\left(\frac{\pi k a_j}{p}\right)+o(\text{of ugly sum/prod})\right).$$

Our main goal is to give an upper bound to

$$\sum_{a\in A}p(a) = |A|p^{-n} + p^{-n}n\sum_{a\in A}\sum_{k\neq 0}\prod_{j=1}^{n}\cos\left(\frac{\pi k a_j}{p}\right).$$

Write

$$p^{-n}n\sum_{a\in A}\sum_{k\neq 0}\prod_{j=1}^{n}\cos\left(\frac{\pi k a_j}{p}\right) = p^{-n}n\sum_{a\in Z_p^n}\sum_{k\neq 0}\prod_{j=1}^{n}\cos\left(\frac{\pi k a_j}{p}\right) - p^{-n}n\sum_{a\in Z_p^n\setminus A}\sum_{k\neq 0}\prod_{j=1}^{n}\cos\left(\frac{\pi k a_j}{p}\right).$$

We will later show that the second summand is vert small, and therefore it is enough to consider only the first summand

$$p^{-n}n\sum_{a\in Z_p^n}\sum_{k\neq 0}\prod_{j=1}^{n}\cos\left(\frac{\pi k a_j}{p}\right) = p^{-n}n\sum_{k\neq 0}\sum_{a\in \mathbb{Z}_p^n}\prod_{j=1}^{n}\cos\left(\frac{\pi k a_j}{p}\right)$$

Before proceeding with the proof, we need some definitions/notation. Let $\vec{n} = (n_1,\ldots,n_p)$ be a positive integer valued vector such that $\sum_i n_i = n$. A vector $a \in \mathbb{Z}_p^n$ is said to be of *type* $\vec{n}$, if it has exactly n_j entries which are equal j for all $1 \le j \le p$. For every $k \neq 0$ and $\vec{n}$, we let $T_k(\vec{n})$ be the set of all vectors a such that ka is of type $\vec{n}$. Now, rearranging the above expression we obtain

$$\begin{aligned}
p^{-n}n\sum_{k\neq 0}\sum_{a\in \mathbb{Z}_p^n}\prod_{j=1}^{n}\cos\left(\frac{\pi k a_j}{p}\right) &= p^{-n}n\sum_{k\neq 0}\sum_{\vec{n}}|T_k(\vec{n})|\prod_{j=1}^{p}\cos^{n_j}\left(\frac{\pi j}{p}\right)\\
&= p^{-n}n\sum_{k\neq 0}\sum_{\vec{n}}\binom{n}{n_1,\ldots,n_t}\prod_{j=1}^{p}\cos^{n_j}\left(\frac{\pi j}{p}\right)\\
&= p^{-n}n\sum_{k}\left(\sum_{j=1}^{p}\cos\left(\frac{\pi}{p}\cdot j\right)\right)^{n},
\end{aligned}$$

where the last equation follows from the multinomial identity, and this equals 0 (not that the sum of cos is $Re(\sum e^{-\pi k/p})$). □

19 The two families theorem of Bollobás

Theorem 19.1. *Let $(A_i, B_i)_{i\in I}$ be a collection of pairs of sets for which the following hold:*

1. *$A_i \cap B_i = \emptyset$ for all i, and*
2. *$A_i \cap B_j \neq \emptyset$ for all $i \neq j$.*

Then, we have $\sum_i \frac{1}{\binom{a_i+b_i}{b_i}} \leq 1$.

There are many proof for the above theorem and we decided to present the following:

Proof. We will proceed by induction on $n = |\cup_{i\in I}(A_i \cup B_i)|$, and without loss of generality we assume that the ground set is $[n]$. For $n = 1$ the theorem is trivial. For the induction step, we wish to remove an element $x \in [n]$ and induct on $n-1$. For each $x \in [n]$ we let $I_x = \{i \in I \mid x \notin A_i\}$, and $B_i^x = B_i \setminus \{x\}$. Moreover, consider the collection of pairs $(A_i, B_i^x)_{i\in I_x}$, and observe that we cannot have $A_i \cap B_j^x = \emptyset$ for some $i \neq j \in I_x$. Indeed, as $A_i \cap B_j \neq \emptyset$ and $B_j = B_j^x \cup \{x\}$, it follows that $x \in A_i \cap B_j$ which is clearly absurd (by the choice of I_x). Therefore, for each $x \in [n]$, by induction we have

$$\sum_{i\in I_x} \frac{1}{\binom{a_i+b_i^x}{a_i}} \leq 1.$$

Next, observe that by induction we have

$$n = \sum_{x\in[n]} 1 \geq \sum_{x\in[n]} \sum_{i\in I_x} \frac{1}{\binom{a_i+b_i^x}{a_i}},$$

and in order to obtain an equality sign, we need the second summand to be exactly 1 for all x.

Now, let us fix i, and run over all x for which $i \in I_x$ (that is, $x \notin A_i$). There are exactly $n - a_i$ such elements x. It follows that for b_i values of x we have $b_i^x = b_i - 1$ and for $n - a_i - b_i$ values of x we have $b_i^x = b_i$. Therefore, by a change of summation, the right hand side of the above inequality is

$$\sum_{i\in I} \left(\frac{n-a_i-b_i}{\binom{a_i+b_i}{a_i}} + \frac{b_i}{\binom{a_i+b_i-1}{a_i}} \right) = \sum_{i\in I} \left(\frac{(n-a_i-b_i)a_i!b_i!}{(a_i+b_i)!} + \frac{b_i\cdot(b_i-1)!a_i!}{(a_i+b_i-1)!} \right)$$

$$= n \sum_{i\in I} \frac{1}{\binom{a_i+b_i}{a_i}}.$$

This completes the proof of the theorem. □

References