Péter Csikvári

Combinatorics 2

Lecture note

Contents

1. Probabilistic methods

In this chapter we will study the most basic applications of the so-called probabilistic method.

1.1 Diagonal Ramsey numbers

Recall that the Ramsey-number $R(r,b)$ denotes the smallest n such that no matter how we color the edges of the complete graph K_n with red and blue colors it will either contain an induced red K_r or a blue K_b. Note that the definition implies that for $n = R(r,b) - 1$ there is a coloring of K_n without red K_r and blue K_b.

Theorem 1.1.1 (Erdős)**.** *Suppose that the positive integers n, k satisfy the inequality $\binom{n}{k}2^{1-\binom{k}{2}} < 1$. Then $R(k,k) > n$. In particular, $R(k,k) > \lfloor 2^{k/2} \rfloor$ if $k \geq 3$.*

Proof. We need to show that there exists a coloring of the edge set of K_n that does not contain either monochromatic red or blue clique K_k. Let us color each edges with color red or blue with probability $1/2$ independently of each other. Now let us estimate the probability that the coloring is bad, i. e., it contains a monochromatic red or blue K_k. For each $S \subset V(G)$ with $|S| = k$ let A_S be the event the induced subgraph on S is monochromatic. Then

$$\mathbb{P}(\text{coloring is bad}) \leq \sum_{|S|=k} \mathbb{P}(A_S) = \binom{n}{k}\frac{2}{2^{\binom{k}{2}}}.$$

By the condition of the theorem $\binom{n}{k}2^{1-\binom{k}{2}} < 1$, so the probability that the coloring is good is positive.

Next we show that for $k \geq 3$ and $n = \lfloor 2^{k/2} \rfloor$ the condition of the theorem is satisfied. Indeed,

$$\binom{n}{k}2^{1-\binom{k}{2}} < \frac{n^k}{k!}2^{1-\binom{k}{2}} \leq \frac{2^{k^2/2}}{k!}2^{1-\binom{k}{2}} = \frac{2^{(k+2)/2}}{k!} < 1.$$

if $k \geq 3$. □

1.2 First moment method

In the previous section we have seen some very simple ideas how to find a certain structure S by proving that it exists with positive probability just by using union bound. Here we study another very simple technique. This is the so-called first moment method. In many cases the structure S that we need to find is defined through some parameter $f(S)$. For instance, we need to prove that there exists a structure S for which some parameter $f(S)$ satisfies $f(S) \geq \rho$. If we find a probability space in which the expected value of $f(S)$ is bigger or equal to ρ then we can immediately conclude that $f(S) \geq \rho$ with positive probability.

1.2.1 Large bipartite subgraphs

Theorem 1.2.1 ([1])**.** *Let G be a graph with n vertices and $e(G)$ edges. Then G has a bipartite subgraph with at least $e(G)/2$ edges.*

Proof. One can rephrase the statement of the theorem as follows: there exists a cut $(A, V \setminus A)$ of G such that the number of edges $(e(A, V \setminus A))$ contained in the cut is at least $e(G)/2$.

Let us consider the random set A which contains every $v \in V(G)$ with probability $1/2$ independently of each other. (This way we have defined a probability space.) Let us consider the random variable $X = e(A, V \setminus A)$. We have to show that with positive probability $X \geq e(G)/2$. To this end it is enough to show that $\mathbb{E}X = e(G)/2$. This is indeed true. For every edge $f \in E(G)$ let us introduce the indicator random variable X_f which takes value 1 if f is in the cut $(A, V \setminus A)$, and 0 otherwise. Then

$$\mathbb{E}X = \mathbb{E}\left(\sum_{f \in E(G)} X_f\right) = \sum_{f \in E(G)} \mathbb{E}X_f.$$

(Note that the random variables X_f are not necessarily independent, but the linearity of expectation holds true even with non-independent random variables.) For all $f \in E(G)$ we have $\mathbb{E}X_f = 1/2$ since the end points of f are in the same set with probability $1/2$ and they are in different sets with probability $1/2$. Hence

$$\mathbb{E}X = \sum_{f \in E(G)} \mathbb{E}X_f = \sum_{f \in E(G)} \frac{1}{2} = \frac{1}{2}e(G).$$

We are done! □

1.2.2 Independent sets

Theorem 1.2.2 (Caro; Wei)**.** *Let G be a graph with vertex degrees $d_1, \dots, d_n$. Let $\alpha(G)$ be the size of the largest independent set of the graph G. Then*

$$\alpha(G) \geq \sum_{i=1}^{n} \frac{1}{d_i + 1}.$$

Proof. Consider a random permutation of the vertices. Let us encircle all the vertices that precede all their neighbors in the given order. Let $X(\pi)$ be the random variable that counts the number of encircled vertices. For a given vertex $v \in V(G)$ let X_v be the indicator variable that the the vertex v is encircled or not. Then $X = \sum_{v \in V(G)} X_v$, consequently

$$\mathbb{E}X = \sum_{v \in V(G)} \mathbb{E}X_v.$$

Note that for a vertex v we have $\mathbb{E}X_v = \frac{1}{d_v+1}$ since the probability that v precedes its neighbors is the same as saying that v is the first among $d_v + 1$ vertices in a random permutation, and this probability is clearly $\frac{1}{d_v+1}$. Hence

$$\mathbb{E}X = \sum_{v \in V(G)} \mathbb{E}X_v = \sum_{i=1}^{n} \frac{1}{d_i + 1}.$$

With positive probability X is at least as large as this expected value. On the other hand, in an arbitrary order the encircled vertices form an independent set since if two of them were adjacent then the second of the two vertices in the order would not be encircled. Hence

$$\alpha(G) \geq \mathbb{E}X = \sum_{i=1}^{n} \frac{1}{d_i + 1}$$

as required. □

Remark 1.2.3. From the above proof one can easily deduce Turán's theorem.

1.3 Graphs with large chromatic number and girth

Theorem 1.3.1 (Erdős [4])**.** *For arbitrary (k, ℓ) there exists a graph G whose chromatic number is at least k and the length of its shortest cycle is at least ℓ.*

Proof. Let $G(n, p)$ be the random graph with n vertices such that we draw all edges with probability $p = p(n)$ independently of each other. In this proof we will set

$p = n^{-\alpha}$, where $\alpha \geq 0$ is a parameter chosen later. First we estimate the number of cycles shorter than ℓ . Given vertices $v_1 v_2 \dots v_r$ form a cycle if $v_i v_{i+1}$ $(r+1=1)$ are all edges, the probability of this event is p^r. Naturally, we can choose the sequence $v_1 v_2 \dots v_r$ in $n(n-1)\dots(n-r+1)$ ways, we only have to take take into account that we counted the same cycle $2r$ ways (rotated and reflected copies). Let X be the random variable counting the number of cycles of length at most $\ell - 1$. Furthermore, let $X(v_1 \dots v_r)$ $(r \leq \ell - 1)$ be the indicator random variable that the vertices $v_1 \dots v_r$ form a cycle in this order. Then

$$X = \sum_{r, v_1 \dots v_r} X(v_1 \dots v_r).$$

Hence

$$\mathbb{E}X = \sum_{r, v_1 \dots v_r} \mathbb{E}X(v_1 \dots v_r) = \sum_{r=3}^{\ell-1} \frac{n(n-1)\dots(n-r+1)}{2r} p^r \leq \sum_{r=3}^{\ell-1} \frac{(np)^r}{2r}.$$

Set $M = \sum_{r=3}^{\ell-1} \frac{(np)^r}{2r}$. Suppose that with some choice of p we can ensure that M is small then with positive probability the number of cycles of length at most $\ell - 1$ will be at most M and by throwing out one point from each cycle we get a graph on at least $n - M$ vertices that does not contain a cycle of length at most $\ell - 1$. In fact, we need to be a little bit more careful as we need that the number of short cycles is small with large probability. Fortunately, we get it immediately: with probability at least $1/2$ the number of cycles of length at most $\ell - 1$ is at most $2M$. Otherwise the expected value would be bigger than M.

Before we try to chose p appropriately let us see how we can bound the chromatic number $\chi(G)$ of G. Here we use the simple fact that

$$\chi(G) \geq \frac{n}{\alpha(G)}.$$

This is true since all coloring class induces an independent set so its size is at most $\alpha(G)$, so we need at least $\frac{n}{\alpha(G)}$ colors to color G. So to make $\chi(G)$ large, it is enough to ensure that $\alpha(G)$ is small. Let us bound the probability that $\alpha(G) \geq s$. For a set S of size s let A_S be the event that S does not induce any edge. Then

$$\mathbb{P}(\alpha(G) \geq s) \leq \sum_{|S|=s} \mathbb{P}(A_S) = \binom{n}{s} (1-p)^{\binom{s}{2}} \leq n^s (1-p)^{\binom{s}{2}} \leq (ne^{-p(s-1)/2})^s.$$

(In the last step we used the fact that $1 + x \leq e^x$ is satisfied for all x. This is a rather standard bound that is quite good if x is small.)

Now it is clear what we have to keep in mind: let M be small, so we need a small p, but we also need that s is not too large and so we need that $ne^{p(s-1)/2} < 1$. We can easily achieve it as follows: set $p = n^{\theta-1}$ where $\theta = \frac{1}{2(\ell-1)}$ and $s = \lceil \frac{3}{p} \log n \rceil$. Then

$$M = \sum_{r=3}^{\ell-1} \frac{(np)^r}{2r} \leq n^{\theta(\ell-1)} \sum_{r=3}^{\ell-1} \frac{1}{2r} \leq n^{1/2} \log n \leq \frac{n}{4}$$

if n is large enough. On the other hand,

$$\mathbb{P}(\alpha(G) \geq s) \leq (ne^{-p(s-1)/2})^s \leq 1/4$$

if n is large enough. Since $\mathbb{P}(X \geq 2M) \leq 1/2$ and $\mathbb{P}(\alpha(G) \geq s) \leq 1/4$, with positive probability there exists a graph where the number of short cycles is at most $n/2$ and $\alpha(G) \leq s$. Now from all cycles of length at most $\ell - 1$ let us throw out 1 vertex and let G^* be the obtained graph. Then G^* has at least $n/2$ vertices and it does not contain a cycle of length at most $\ell - 1$. Furthermore, $\alpha(G^*) \leq \alpha(G)$ since G^* is an induced subgraph of G. Then

$$\chi(G^*) \geq \frac{|V(G^*)|}{\alpha(G^*)} \geq \frac{n/2}{3n^{1-\theta} \log n} = \frac{n^\theta}{6 \log n}.$$

If n is large enough this is bigger than k. We are done! □

2. Strongly regular graphs

2.1 Introduction

In this chapter we study strongly regular graphs, these are very special simple graphs. Strongly regular graphs are often very symmetric graphs, and linear algebraic tools are particularly amenable to study them.

Definition 2.1.1. A graph G is a strongly regular graph with parameters (n, d, a, b) if it has n vertices, d–regular, two adjacent vertices have exactly a common neighbors, and two non-adjacent vertices have exactly b common neighbors.

For instance, a 4-cycle is a strongly regular graph with parameters $(4, 2, 0, 2)$ while a 5-cycle is a strongly regular graph with parameters $(5, 2, 0, 1)$. Note that a k-cycle is never strongly regular if $k \geq 6$. The Petersen-graph is a strongly regular graph with parameters $(10, 3, 0, 1)$.

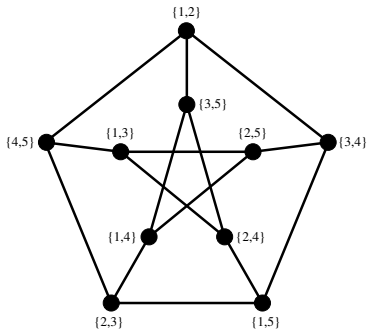


Figure 2.1: Petersen-graph as the Kneser(5,2) graph.

In what follows we try to find necessary conditions for the parameters (n, d, a, b) to enable the existence of a strongly regular graph with parameters (n, d, a, b). The first one is very elementary.

Proposition 2.1.2. *Let G be a strongly regular graph with parameters (n, d, a, b). Then*

$$d(d-1-a) = (n-d-1)b.$$

Proof. Let u be a fixed vertex. Let us count the number of vertex pairs (v_1, v_2) for which the following holds: $(u, v_1) \in E(G)$, $(v_1, v_2) \in E(G)$ and $(u, v_2) \notin E(G)$, and $v_1, v_2 \neq u$.

We can choose v_1 in d different ways, then we can choose v_2 from the d neighbors of v_1, but we cannot choose u, and we cannot choose those a vertices which are connected to u. So the number of these pairs are $d(d-1-a)$.

On the other hand, we can choose v_2 in $n-d-1$ different ways, as we cannot choose u and its neighbors. After choosing v_2, we can choose v_1 in b ways as u and v_1 has b common neighbors.

Hence $d(d-1-a) = (n-d-1)b$. □

2.2 Adjacency matrix and linear algebra

It turns out that strongly regular graphs can be best studied through their adjacency matrix. The adjacency matrix $A(G)$ of a simple graph $G = (V, E)$ is defined as follows: it is a symmetric matrix of size $|V| \times |V|$ labelled by the vertices of the graph G, and

$$A(G)_{u,v} = \begin{cases} 1 & \text{if } (u, v) \in E(G), \\ 0 & \text{if } (u, v) \notin E(G). \end{cases}$$

If the graph G is clear from the context we will simple write A instead of $A(G)$.

Next let us compute the eigenvalues and its multiplicities of a strongly regular graph. It will turn out that a strongly regular graph has only 3 different eigenvalues and the simple fact that the multiplicities of the eigenvalues must be non-negative integers imposes a very strong condition on the parameters (n, d, a, b).

Observe that for a simple graph G, the entries of A^2 can be understood very easily. In the diagonal of A^2 we have the degrees of the vertices, in our case, there will be d everywhere. On the other hand, for $i \neq j$, $(A^2)_{ij}$ counts the number of common neighbors of vertex i and j which is a or b according to i and j are adjacent or not. So in $A^2 + (b-a)A$ we have d's in the diagonal and b's everywhere else. Hence

$$A^2 + (b-a)A - (d-b)I = bJ,$$

where J is the all 1 matrix.

Now let us assume that $A\underline{x} = \lambda\underline{x}$, where $\underline{x} = (x_1, \ldots, x_n)$. Then

$$(A^2 + (b-a)A - (d-b)I)\underline{x} = (\lambda^2 + (b-a)\lambda - (d-b))\underline{x},$$

while

$$bJ\underline{x} = b(\sum_{i=1}^{n} x_i)\underline{1}.$$

Hence by comparing the i'th coordinates we get that

$$(\lambda^2 + (b-a)\lambda - (d-b))x_i = b(\sum_{i=1}^{n} x_i).$$

If $\lambda^2 + (b-a)\lambda - (d-b) \neq 0$, then all x_i must be equal, and we simply get the usual eigenvector belonging to d. Otherwise, $\lambda^2 + (b-a)\lambda - (d-b)$ must be 0, hence

$$\lambda = \lambda_{\pm} = \frac{a - b \pm \sqrt{(a-b)^2 + 4(d-b)}}{2}.$$

It is easy to see that if G is a disconnected strongly regular graph then it must be the disjoint union of some K_{d+1}. Since it is not really interesting, let us assume that G is connected. Then we know that the multiplicity of the eigenvalue d is exactly 1. Let m_+ and m_- be the multiplicities of the other two eigenvalues. Since the number of eigenvalues is n, we know that

$$1 + m_+ + m_- = n.$$

We also know that $TrA = 0$, so

$$0 = TrA = 1 \cdot d + m_+\lambda_+ + m_-\lambda_-.$$

From this we get that

$$m_{\pm} = \frac{1}{2}\left(n - 1 \mp \frac{2d + (n-1)(a-b)}{\sqrt{(a-b)^2 + 4(d-b)}}\right).$$

Let us summarize our results in a theorem.

Theorem 2.2.1. *Let G be a connected strongly regular graph with parameters (n, d, a, b). Then its eigenvalues are d with multiplicity 1, and*

$$\lambda_{\pm} = \frac{a - b \pm \sqrt{(a-b)^2 + 4(d-b)}}{2}$$

with multiplicity

$$m_{\pm} = \frac{1}{2}\left(n - 1 \mp \frac{2d + (n-1)(a-b)}{\sqrt{(a-b)^2 + 4(d-b)}}\right).$$

As an example we can compute the eigenvalues of the Petersen-graph. Recall that this is a strongly regular graph with parameters $(10, 3, 0, 1)$. Then its eigenvalues are $3, 1$ and -2, where the multiplicities are $m_1 = 5$, $m_2 = 4$.

The condition that $m_\pm$ are non-negative integers is a surprisingly strong condition, this is called the *integrality condition*. As an application, let's see which strongly regular graphs have parameters $(n, d, 0, 1)$. We have already seen that the 5-cycle and the Petersen-graph are such graphs with $d = 2$ and $d = 3$. Actually, K_2 is also such graph with $d = 1$, but it's a bit cheating since the fourth parameter hasn't any meaning, not to mention K_1 with $d = 0$... First of all, note that our first proposition implies that $n = d^2 + 1$. Indeed, $d(d - 1 - a) = (n - d - 1)b$ with $a = 0, b = 1$ immediately implies that $n = d^2 + 1$.

Theorem 2.2.2 (Hoffman–Singleton)**.** *Let G be a strongly regular graph with parameters $(d^2 + 1, d, 0, 1)$, where $d \geq 2$. Then $d \in \{2, 3, 7, 57\}$.*

Proof. The eigenvalues of the graph G are d and

$$\lambda_\pm = \frac{-1 \pm \sqrt{4d - 3}}{2}$$

and its multiplicities are

$$m_\pm = \frac{1}{2}\left(d^2 \mp \frac{2d - d^2}{\sqrt{4d - 3}}\right).$$

If $2d - d^2 = 0$, then $d = 0$ or 2. (For $d = 0$, the definition works, but we don't consider it as a strongly regular graph. We simply excluded it by requiring $d \geq 2$.) If $2d - d^2 \neq 0$, then $\sqrt{4d - 3}$ is a rational number. This can only happen if $4d - 3$ is a perfect square. Hence $4d - 3 = s^2$. Then

$$m_\pm = \frac{1}{2}\left(\left(\frac{s^2 + 3}{4}\right)^2 \mp \frac{2\left(\frac{s^2+3}{4}\right) - \left(\frac{s^2+3}{4}\right)^2}{s}\right).$$

Hence

$$m_+ = \frac{s^5 + s^4 + 6s^3 - 2s^2 + 9s - 15}{32s}.$$

Since $32m_+$ is an integer, we get that $s \mid 15$. Hence $s \in \{1, 3, 5, 15\}$. If $s = 1$ then $d = 1$ which we excuded. So $s \in \{3, 5, 15\}$ whence$d \in \{3, 7, 57\}$. Together with $d = 2$ we get that $d \in \{2, 3, 7, 57\}$. □

Remark 2.2.3. One might wonder whether there is such a graph for $d = 7$ and $d = 57$. For $d = 7$ there is such a graph: it is called the Hoffman-Singleton graph. It is the unique strongly regular graph with parameters $(50, 7, 0, 1)$ just as the Petersen-graph and the 5-cycle are the unique strongly regular graphs with parameters $(10, 3, 0, 1)$ and $(5, 2, 0, 1)$. It is not known whether there is a strongly regular graph with parameters $(3250, 57, 0, 1)$.

Remark 2.2.4. The following statement is true in general: the eigenvalues of a strongly regular graph are integers or the parameters of the strongly regular graph satisfies that $(n, d, a, b) = (4k+1, 2k, k-1, k)$ for some k, the latter graphs are called conference graphs, for instance the 5–cycle is a conference graph. This statement can be proved by studying whether $2d + (n-1)(a-b)$ is 0 or not.

Theorem 2.2.5 (Lossers-Schwenk)**.** *One cannot decompose K_{10} into three edge disjoint Petersen-graphs.*

Proof. Suppose for contradiction that we can decompose K_{10} into three edge disjoint Petersen-graphs. Let A_1, A_2 and A_3 be the adjacency matrices of the three Petersen-graphs. Then

$$J - I = A_1 + A_2 + A_3.$$

Note that A_1, A_2 and A_3 has a common eigenvector, namely $\underline{1}$. All other eigenvectors are orthogonal to this vector. In particular, we can consider the eigenspaces of A_1 and A_2 belonging to the eigenvalue 1. Let these eigenspaces be V_1 and V_2. Note that $\dim V_1 = \dim V_2 = 5$ as the multiplicity of the eigenvalue 1 is 5. We know that $V_1, V_2 \subseteq \underline{1}^{\perp}$. Note that $\underline{1}^{\perp}$ is a 9-dimensional vectorspace, so V_1 and V_2 must have a non-trivial intersection: let $\underline{x} \in V_1 \cap V_2$. Then

$$A_3\underline{x} = (J - I)\underline{x} - A_1\underline{x} - A_2\underline{x} = \underline{0} - \underline{x} - \underline{x} - \underline{x} = -3\underline{x}.$$

But this is a contradiction since -3 is not an eigenvalue of the Petersen-graph. □

Remark 2.2.6. It is possible to pack two Petersen-graphs into K_{10}. The above proof shows that the remaining edges form a 3–regular graph H with an eigenvalue -3. This suggests that H should be a bipartite graph. This is indeed true, but one needs to prove first that H is connected. It is quite easy to prove it as the only disconnected 3-regular graph on 10 vertices is $K_4 \cup K_{3,3}$ (why?). It is not hard to show that H cannot be $K_4 \cup K_{3,3}$.

Second proof. Suppose that we can decompose K_{10} into 3 edge-disjoint Petersen-graphs. Let us color the edges of the three Petersen-graphs with blue, red and green colors in order to make it easier to refer to them. Let v be any vertex of K_{10} and let b_1, b_2, b_3 be the neighbors of v in the blue Petersen-graph. Similarly let r_1, r_2, r_3 and g_1, g_2, g_3 be the neighbors of v in the red and green Petersen–graphs.

For a moment, let's put away the green Petersen–graph and let's just concentrate on the bipartite graph induced by the vertices b_1, b_2, b_3 and r_1, r_2, r_3. Note that the edge (v, r_1) is a red edge, so it's not blue! This means that there must be exactly one blue path of length 2 between v and r_1. In other words, r_1 is connected by a blue edge to exactly one of the vertices of b_1, b_2, b_3. Similarly, r_2 and r_3 are connected by a blue edge to exactly one of the vertices of b_1, b_2, b_3. This means that there are exactly 3 blue edges between b_1, b_2, b_3 and r_1, r_2, r_3. By repeating this argument to b_1, b_2, b_3, we find that there are exactly 3 red edges between b_1, b_2, b_3 and r_1, r_2, r_3. This means that there are exactly 3 green edges between b_1, b_2, b_3 and r_1, r_2, r_3.

Now let us consider the green Petersen–graph. If we delete the vertices v, g_1, g_2, g_3 from this graph, the green edges induce a 6-cycle on the remaining vertices. According to the previous paragraph, there is a cut of this 6-cycle which contains exactly 3 edges. But this is impossible: a cut of a 6-cycle always contains even number of edges! Indeed, if we walk around the 6-cycle we need to cross the cut even number of times to get back to the original side of the cut from where we started our walk. $\square$

3. Set systems

Given sets $A_1, \ldots, A_m \subseteq [n]$ one can associate the so-called characteristic vector $\underline{a}_i$ to each set A_i: it is a vector of length n such that the k-coordinate $(\underline{a}_i)_k$ is 1 if $k \in A_i$, and 0 otherwise. Since it is a $0-1$ vector we can consider it in any field $\mathbb{F}$. Note that for the usual scalar product we get that $(\underline{a}_i, \underline{a}_i) = |A_i|$ and $(\underline{a}_i, \underline{a}_j) = |A_i \cap A_j|$. This is the key property that we will use in several proofs.

3.1 The clubs of Oddtown

Theorem 3.1.1. *Let $A_1, \ldots, A_m \subseteq [n]$ such that $|A_i|$ is odd for $i = 1, \ldots, m$ and $|A_i \cap A_j|$ is even for $i \neq j$. Then $m \neq n$.*

Proof. Let $\underline{a}_i \in \mathbb{F}_2^n$ be the characteristic vector of the set A_i. We show that $\underline{a}_1, \ldots, \underline{a}_m$ are linearly independent over $\mathbb{F}_2$. This would immediately give that $m \leq n$. Suppose for contradiction that there are $\lambda_1, \ldots, \lambda_n \in \mathbb{F}_2$ such that not all of them are 0 and

$$\sum_{i=1}^{n} \lambda_i \underline{a}_i = \underline{0}.$$

Take the scalar product of both sides with $\underline{a}_i$. Since $(\underline{a}_i, \underline{a}_i) = 1$ and $(\underline{a}_i, \underline{a}_j) = 0$ we get that

$$0 = \left(\underline{a}_i, \sum_{j=1}^{n} \lambda_j \underline{a}_j \right) = \lambda_j$$

for all $1 \leq i \leq m$ contradicting that not all $\lambda_i = 0$. Hence $\underline{a}_1, \ldots, \underline{a}_m$ are linearly independent over $\mathbb{F}_2$, consequently $m \leq n$.

□

3.2 Same-size intersections

Theorem 3.2.1 (Generalized Fisher inequality)**.** *If the nonempty sets $A_1, \ldots, A_m \subseteq [n]$ are distinct such that all $|A_i \cap A_j|$ have the same size for $i \neq j$, then $m \leq n$.*

Proof. Suppose that $|A_i \cap A_j| = t$ for $i \neq j$. If for some i the set A_i has size t, then $A_i \subseteq A_j$ for all $j \neq i$, and the sets $A_j \setminus A_i$ for $j \neq i$ must be disjoint. This means that $m \leq 1 + n - t \leq n$ since $t \geq 1$ as the sets are non-empty. Thus in what follows we can assume that $|A_i| > t$.

Let $\underline{a}_i \in \mathbb{R}^n$ be the characteristic vector of the set A_i. We show that $\underline{a}_1, \ldots, \underline{a}_m$ are linearly independent over $\mathbb{R}$. This would immediately give that $m \leq n$. Suppose for contradiction that there are $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ such that not all of them are 0 and

$$\sum_{i=1}^{n} \lambda_i \underline{a}_i = \underline{0}.$$

Take the scalar product of both sides with $\underline{a}_i$. Since $(\underline{a}_i, \underline{a}_i) = |A_i|$ and $(\underline{a}_i, \underline{a}_j) = t$ we get that

$$0 = \left(\sum_{j=1}^{m} \lambda_j \underline{a}_j, \sum_{j=1}^{m} \lambda_j \underline{a}_j \right) = \sum_{i=1}^{m} |A_i| \lambda_i^2 + 2t \sum_{1 \leq i < j \leq m} \lambda_i \lambda_j = \sum_{i=1}^{m} (|A_i| - t) \lambda_i^2 + \left(\sum_{i=1}^{m} \lambda_i \right)^2.$$

Since λ_i are not all 0, and $|A_i| > t$ we get that $\sum_{i=1}^{m} (|A_i| - t) \lambda_i^2 + \left(\sum_{i=1}^{m} \lambda_i \right)^2 > 0$, a contradiction that proves that $\underline{a}_1, \ldots, \underline{a}_m$ are linearly independent over $\mathbb{R}$, consequently $m \leq n$.

□

3.3 Forbidden intersections

Theorem 3.3.1. *Let $\mathcal{L} \subseteq [n]$ such that $|\mathcal{L}| = s$. Let $A_1, \ldots, A_m \subseteq [n]$ be a set system satisfying $|A_i| \notin \mathcal{L}$ for $i = 1, \ldots, m$, and $|A_i \cap A_j| \in \mathcal{L}$ for $1 \leq i < j \leq m$. Then*

$$m \leq \binom{n}{s} + \binom{n}{s-1} + \cdots + \binom{n}{1} + \binom{n}{0}.$$

We will actually prove the following stronger statement.

Theorem 3.3.2. *Let p be a prime. Let $\mathcal{L} \subseteq [n]$ such that $|\mathcal{L}| = s$. Let $A_1, \ldots, A_m \subseteq [n]$ be a set system satisfying $|A_i| \notin \mathcal{L} \pmod{p}$ for $i = 1, \ldots, m$, and $|A_i \cap A_j| \in$*

$\mathcal{L} \pmod p$ *for* $1 \le i < j \le m$. *Then*

$$m \le \binom{n}{s} + \binom{n}{s-1} + \cdots + \binom{n}{1} + \binom{n}{0}.$$

Proof. Let $\mathcal{L} = \{\ell_1, \ldots, \ell_s\}$ and let $\underline{a}_i$ denote the characteristic vector of the set A_i. Consider the polynomial $f_i : \{0,1\}^n \to \mathbb{F}_p$ for which

$$f_i(\underline{x}) = \prod_{j=1}^{s}((\underline{a}_i, \underline{x}) - \ell_j).$$

Then $f_i(\underline{a}_i) \neq 0$ in $\mathbb{F}_p$, but for any $k \neq i$ we have $f_i(\underline{a}_k) = 0$. This shows that $f_1, \ldots, f_m$ are linearly independent. Now let us replace x_j^t with x_j for $t \ge 1$ in each monom of each polynomial $f_i(\underline{x})$. Observe that on $\{0,1\}^n$ this does not change the value of any f_i so for the obtained $\widehat{f_i}$ we get the same values. This shows that $\widehat{f_1}, \ldots, \widehat{f_m}$ are also linearly independent over $\mathbb{F}_p$. These are multilinear polynomials of degree at most s, and so they are in a vector space of dimension $\binom{n}{s} + \binom{n}{s-1} + \cdots + \binom{n}{1} + \binom{n}{0}$, that is,

$$m \le \binom{n}{s} + \binom{n}{s-1} + \cdots + \binom{n}{1} + \binom{n}{0}.$$

□

Conjecture 3.3.3 (Erdős)**.** Suppose that $0 < \lambda < 1/2$. Then there exist an $\varepsilon = \varepsilon(\lambda) > 0$ and an $n_0 = n_0(\lambda)$ such that for any $n > n_0$ and integer $t = \lambda n$ if the set system $A_1, \ldots, A_m \subseteq [n]$ satisfies that $|A_i \cap A_j| \neq t$ for $1 \le i < j \le n$, then $m \le (2-\varepsilon)^n$.

Theorem 3.3.4. *Let p be a prime and $n = 4p-1$. Suppose that $A_1, \ldots A_m \subseteq [n]$ such that $|A_i| = 2p-1$ for $1 \le i \le m$, and $|A_i \cap A_j| \neq p-1$ for $1 \le i < j \le m$. Then $m < 1.8^n$.*

Proof. We can apply the previous theorem with $\mathcal{L} = \{0, 1, \ldots, p-2\}$. Then

$$m \le \binom{4p-1}{p-1} + \binom{4p-1}{p-2} + \cdots + \binom{4p-1}{0} < p\binom{4p}{p}.$$

By Stirling's formula we have

$$\binom{4p}{p} \approx \frac{\left(\frac{4p}{e}\right)^{4p}}{\left(\frac{3p}{e}\right)^{3p}\left(\frac{p}{e}\right)^{p}} \cdot \frac{\sqrt{2\pi \cdot 4p}}{\sqrt{2\pi \cdot 3p}\sqrt{2\pi \cdot p}}$$

$$\approx \left(\frac{1}{(1/4)^{1/4}(3/4)^{3/4}}\right)^{4p} \cdot \frac{c}{\sqrt{p}}$$
$$< 1.8^{4p-1}$$

for large enough p. □

Remark 3.3.5. The function

$$h(x) := x \ln\left(\frac{1}{x}\right) + (1-x)\ln\left(\frac{1}{1-x}\right)$$

is called the binary entropy function. From the above computation we can see that

$$\binom{n}{k} \approx \frac{c}{\sqrt{n}} \exp\left(h\left(\frac{k}{n}\right)\right).$$

One can show that for $k \leq n/2$ we have

$$\sum_{j=0}^{k} \binom{n}{j} \leq \exp\left(h\left(\frac{k}{n}\right)\right).$$

3.3.1 Coloring $\mathbb{R}^d$

In this section we study the so-called Hadwiger–Nelson problem. In this problem we aim to find the minimal number of colors for which one can color the points of $\mathbb{R}^d$ such that no two points of unit distance get the same color. One can rephrase this problem as follows. For an integer d let us consider the infinite graph whose vertices are the points of $\mathbb{R}^d$ and two vertices are adjacent if their distance is exactly 1. Let $\chi(d)$ denote the chromatic number of this graph. It is easy to see that $\chi(d) \geq d+1$ because of a regular simplex. In general, it is true that it is sufficient to consider finite subgraphs of $\mathbb{R}^d$ in order to determine the chromatic number of $\mathbb{R}^d$. It is a nice exercise that $\chi(2) \geq 4$, and it was recently shown by Aubrey de Grey that $\chi(2) \geq 5$. On the other hand, there is a construction showing $\chi(2) \leq 7$.

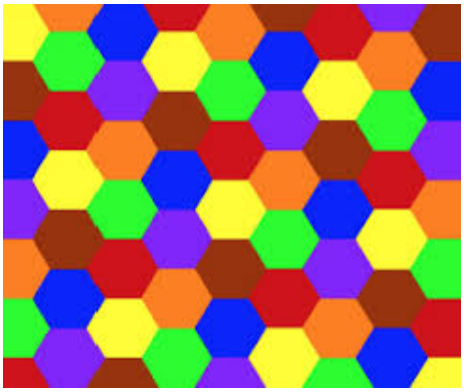


Figure 3.1: Coloring of the plane with 7 colors without unit distance.

Here we prove that $\chi(n)$ grows exponentially.

Theorem 3.3.6. *If $n = 4p - 1$ is large enough and p is a prime, then $\chi(n) > 1.1^n$, and for every large enough n we have $\chi(n) > 1.1^{n/2}$.*

Proof. Instead of distance 1 we will consider distance $\sqrt{2p}$. Consider the points among $\{0,1\}^n$ that contains exactly $2p-1$ pieces of 1's. We can consider them as characteristic vectors of sets. Note that $||v_A - v_B||_2 = \sqrt{2p}$ if and only if $|A \cap B| = p - 1$. So if we color these points, then a color class avoiding distance $\sqrt{2p}$ has at most 1.8^n elements by Theorem 3.3.4. Hence we have at least $\frac{\binom{4p-1}{2p-1}}{1.8^n} > 1.1^n$ color classes. The second statement follows from the fact that for every n there is a prime between n and $2n$. □

3.3.2 Fall of Borsuk's conjecture

Theorem 3.3.7 (Borsuk)**.** *Given the sphere of diameter D in $\mathbb{R}^d$. Then one needs at least $d+1$ closed sets to decompose it into sets of diameter less than D. It suffices to use $d+1$ sets for such a decomposition.*

From this theorem one can deduce the following seemingly stronger result.

Theorem 3.3.8. *Given a bounded, closed, convex set K in $\mathbb{R}^d$ with diameter D such that for each boundary point of K there exists exactly one supporting hyperplane. Then K can be decomposed into $d+1$ closed sets of diameter strictly less than D.*

This theorem lead Borsuk to conjecture the following.

Conjecture 3.3.9 (Borsuk)**.** Every bounded, closed set K of diameter D can be decomposed into $d+1$ closed sets of diameter strictly less than D.

The conjecture is true for $d = 2$, nevertheless the next theorem shows that it is false for large d.

Theorem 3.3.10 (Kahn-Kalai)**.** *Let p be a prime and $d = \binom{4p-1}{2}$, then there exists a bounded, closed set K of diameter D whose decomposition into closed sets of diameter less than D requires at least $1.1^{\sqrt{d}}$ sets.*

Proof. Let us consider the complete graph K_{4p-1} and let us associate a basis vector e_{ij} for each edge of it in $\mathbb{R}^d$ that is 1 at coordinate ij and 0 everywhere else. For each set $S \subseteq [4p-1]$ of size $2p-1$ we associate a point S^* in $\mathbb{R}^d$ for which

$$(S^*)_{ij} = \begin{cases} 1 & \text{if} (i,j) \in E(S, S^c) \\ 0 & \text{otherwis} \end{cases}$$

For $S, T \subseteq [4p-1]$ the distance of S^* and T^* is

$$d(S^*, T^*)^2 = 2k(2p-1-k) + 2(k+1)(2p-1-k) = (4k+2)(2p-1-k)$$

if $|S \cap T| = k$. This has a maximum at $p - \frac{3}{4}$ among real numbers and at $p-1$ among integers. So if we want to avoid putting two points S^* and T^* with diameter D into the same set of the decomposition, then we have to avoid sets S and T with $|S \cap T| = p-1$. It means that by Theorem 3.3.4 we need at least

$$\frac{\binom{4p-1}{2p-1}}{1.8^{4p-1}}$$

sets in the decomposition. This is bigger than $1.1^{4p-1} > 1.1^{\sqrt{d}}$ for large p.

□

4. Packing complete bipartite graphs

4.1 Graham–Pollak theorem

In this chapter we prove the Graham-Pollak theorem and a variant of it.

Theorem 4.1.1 (Graham–Pollak)**.** *If the edge sets of the complete bipartite graphs $H_1, H_2, \ldots, H_m$ decomposes the edge set of the complete graph K_n, then $m \geq n-1$.*

Proof. (Tverberg) Suppose for contradiction that $m < n-1$. Suppose that H_i has vertex set $A_i \cup B_i$. For each vertex j of K_n let us introduce a variable x_j. Let us consider the equations

$$\sum_{j \in A_i} x_i = 0 \quad (i = 1, \ldots, m), \quad \text{and} \quad \sum_{j=1}^{n} x_j = 0.$$

We have n variables and $m+1$ equations. If $m < n-1$, then it has a non-zero solution $\underline{x} \in \mathbb{R}^n$. Observe that

$$0 < \sum_{j=1}^{n} x_j^2 = \left(\sum_{j=1}^{n} x_j\right)^2 - 2 \sum_{1 \leq i < j \leq n} x_i x_j = \left(\sum_{j=1}^{n} x_j\right)^2 - 2 \sum_{i=1}^{m} \left(\sum_{j \in A_i} x_j\right)\left(\sum_{k \in A_i} x_k\right) = 0.$$

This contradiction proves our theorem. □

A generalization of the Graham–Pollak theorem is the following.

Theorem 4.1.2 (Witsenhausen; Graham and Pollak)**.** *Let G be a graph and suppose that its edge set decomposes to m complete bipartite graphs $H_1, \ldots, H_m$. Let $n_+(G)$ and $n_-(G)$ be the number of positive and negative eigenvalues of the adjacency matrix A_G, respectively. Then $m \geq \max(n_+(G), n_-(G))$.*

Remark 4.1.3. The eigenvalues of A_{K_n} are $n-1$ with multiplicity 1, and -1 with multiplicity $n-1$, so this is indeed a generalization of the previous theorem.

Proof. Suppose that H_i has vertex set $A_i \cup B_i$, and their characteristic vectors are u_i and v_i. Consider the matrices $M_i = u_i v_i^T + v_i u_i^T$. Then $A_G = M_1 + \cdots + M_m$. If $w \in \langle u_1, \ldots, u_m \rangle^\perp$, then

$$w^T A_G w = \sum_{i=1}^{m} w^T M_i w = \sum_{i=1}^{k} w^T (u_i v_i^T + v_i u_i^T) w = 0.$$

Clearly, $\dim \langle u_1, \ldots, u_m \rangle^\perp = n - m$. Consider the vector space V_+ spanned by eigenvectors corresponding to the positive eigenvalues. Its dimension is n_+. If $w \in V_+$ a non-zero vector, then $w^T A_G w > 0$. Hence $\langle u_1, \ldots, u_m \rangle^\perp \cap V_+ = \underline{0}$. Hence $n - m + n_+ \leq n$, that is, $n_+ \leq m$. The inequality $n_- \leq m$ follows analogously. □

Let us mention a theorem whose proof is completely analogous.

Theorem 4.1.4. *Let G be a graph, whose largest independent set has size $\alpha(G)$. Let $n'_+(G)$ and $n'_-(G)$ be the number of non-negative and non-positive eigenvalues of the adjacency matrix A_G, respectively. Then $m \leq \max(n'_+(G), n'_-(G))$.*

Remark 4.1.5. Be careful, in Theorem 4.1.2 we consider positive eigenvalues while in Theorem 4.1.4 we consider non-negative eigenvalues, that is why we use $n'_+(G)$ instead of $n_+(G)$. Also the inequality is in the opposite direction.

The problem of packing complete bipartite graphs into the complete graph is originated in another problem about graph addresses. Suppose that we are given a graph G on n vertices, and for each vertex v we want to make an "address", a length k vector with elements $0, 1, *$ satisfying the following property. If u and v has distance d in the graph, then their addresses differ at exactly d places not counting $*$, so we only consider places where one of them is 0 and the other one is 1. A little thinking shows that it is the same as covering the complete graph K_n with k complete bipartite graphs such that each pair u and v is covered by exactly d complete bipartite graphs, their distance in the graph G. One can show that $k = n - 1$ always suffices, and as the above theorem shows you need $k = n-1$ for complete graphs. One can also show that for trees one also needs $n - 1$-long addresses. The underlying fact is that the distance matrix of a tree has $n - 1$ negative eigenvalues and one positive eigenvalue, and so one can modify the proof of Theorem 4.1.2 to that case. Note that in this problem $\sum_{i=1}^{m} M_i$ is equal to the distance matrix, not the adjacency matrix.

5. Enumerative combinatorics

5.1 Generating functions

There are various ways to associate a function to a sequence of numbers. Here we discuss the two most common ways.

Definition 5.1.1. Given a sequence of numbers $a_0, a_1, a_2, \ldots$ the function

$$\sum_{n=0}^{\infty} a_n x^n$$

is called the generating function of the sequence. The function

$$\sum_{n=0}^{\infty} a_n \frac{x^n}{n!}$$

is called the exponential generating functions.

Remark 5.1.2. It is of course a natural question which one to use in a specific problem. If the sequence a_n grows faster than any exponential function, then most likely one needs to use the exponential generating function $\sum_{n=0}^{\infty} a_n \frac{x^n}{n!}$. Note that if

$$A_1(x) = \sum_{n=0}^{\infty} a_n x^n, \quad B_1(x) = \sum_{n=0}^{\infty} b_n x^n, \quad C_1(x) = A_1(x)B_1(x) = \sum_{n=0}^{\infty} c_n x^n,$$

then

$$\sum_{k=0}^{n} a_k b_{n-k} = c_n,$$

while if

$$A_2(x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{n!}, \quad B_2(x) = \sum_{n=0}^{\infty} b_n \frac{x^n}{n!}, \quad C_2(x) = A_2(x)B_2(x) = \sum_{n=0}^{\infty} c_n \frac{x^n}{n!},$$

then

$$c_n = \sum_{k=0}^{n} \binom{n}{k} a_k b_{n-k}.$$

This suggests that if certain expression contains an $\binom{n}{k}$, then one might need to use the exponential generating function.

It is in general true that the functions are designed for a specific kind of multiplication. For instance, in number theory they quite often use the function $F(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$, because the corresponding multiplication is $c_n = \sum_{d \mid n} a_d b_{n/d}$.

5.2 Fibonacci numbers

Let us consider Fibonacci numbers. Recall that $F_0 = 0$, $F_1 = 1$ and $F_n = F_{n-1} + F_{n-2}$ for $n \geq 2$. Consider

$$F(x) = \sum_{n=0}^{\infty} F_n x^n.$$

Proposition 5.2.1. *We have*

$$F(x) = \frac{x}{1 - x - x^2}.$$

Proof. We have

$$\begin{aligned} F(x) &= F_0 + F_1 x + F_2 x^2 + F_3 x^3 + F_4 x^4 + \dots \\ xF(x) &= \quad F_0 x + F_1 x^2 + F_2 x^3 + F_3 x^4 + \dots \\ x^2 F(x) &= \quad\quad F_0 x^2 + F_1 x^3 + F_2 x^4 + \dots \end{aligned}$$

Hence

$$(1 - x - x^2)F(x) = F_0 + (F_1 - F_0)x = x,$$

thus

$$F(x) = \frac{x}{1 - x - x^2}.$$

□

Corollary 5.2.2. *We have*

$$F_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1 + \sqrt{5}}{2} \right)^n - \left(\frac{1 - \sqrt{5}}{2} \right)^n \right).$$

Proof. Let $\varphi_1 = \frac{1+\sqrt{5}}{2}$ and $\varphi_2 = \frac{1-\sqrt{5}}{2}$. Then $1 - x - x^2 = (1 - \varphi_1 x)(1 - \varphi_2 x)$. Hence

$$F(x) = \frac{x}{1 - x - x^2}$$

$$
\begin{aligned}
&= \frac{x}{(1-\varphi_1 x)(1-\varphi_2 x)} \\
&= \frac{1}{\sqrt{5}}\left(\frac{1}{1-\varphi_1 x} - \frac{1}{1-\varphi_2 x}\right) \\
&= \frac{1}{\sqrt{5}}\sum_{n=0}^{\infty}(\varphi_1 x)^n - \frac{1}{\sqrt{5}}\sum_{n=0}^{\infty}(\varphi_2 x)^n \\
&= \sum_{n=0}^{\infty}\frac{1}{\sqrt{5}}\left(\varphi_1^n - \varphi_2^n\right)x^n
\end{aligned}
$$

Since $F(x) = \sum_{n=0}^{\infty} F_n x^n$ we get that $F_n = \frac{1}{\sqrt{5}}\left(\varphi_1^n - \varphi_2^n\right)$.

□

Remark 5.2.3. One can also deduce Corollary 5.2.2 from the fact that for the matrix $M = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ we have $M^n = \begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix}$. This approach also gives an $O(\log n)$ algorithm to compute the Fibonacci numbers: first compute $M, M^2, M^4, M^8, \ldots$, then $M^n = M^{2^{k_0}} M^{2^{k_1}} \ldots M^{2^{k_r}}$ by writing up n in the binary expansion. This approach also provides simple proofs for various identities of the Fibonacci numbers by using easy observations like $M^{n+m} = M^n M^m$.

5.3 Catalan numbers revisited

In this section we revisit Catalan numbers.

Definition 5.3.1. For $n \geq 1$ let C_n be the number of sequences $(s_1, s_2, \ldots, s_{2n})$ satisfying $s_i = \pm 1$, $\sum_{i=1}^{k} s_i \geq 0$ and $\sum_{i=1}^{2n} s_i = 0$. Let $C_0 = 1$.

Remark 5.3.2. In the first semester you learned quite a lot of combinatorial enumeration problems for which the answer is a Catalan number. (For instance, the number of triangulations of a convex $(n+2)$-gon is such a problem.) For a thorough treatment see the book of Richard Stanley: Catalan numbers.

Fun fact: 42 is a Catalan number, so if you want a question for which the answer is 42, just look up Stanley's book.

Proposition 5.3.3. *(a) For $n \geq 1$ we have*

$$C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k}.$$

(b) Let $C(x) = \sum_{n=0}^{\infty} C_n x^n$. *Then*

$$C(x) = \frac{1 - \sqrt{1-4x}}{2x}.$$

Proof. (a) Let $\mathcal{C}_n$ be the set of Catalan sequences with $2n$ elements. Let $s_1 s_2 s_3 \dots s_{2n+1} s_{2n+2}$ a Catalan sequence. Let $2k+2$ be the first time when $s_1 + \dots + s_{2k+2} = 0$. Note that it is possible that $k = 0$ (the sequence starts with 1 and -1) and it is also possible that $k = n$. Then $s_2 s_3 \dots s_{2k+1}$ is again a (possibly empty) Catalan sequence and $s_{2k+3} \dots s_{2n+2}$ is also a Catalan sequence. It is easy to see that it is a bijection between $\mathcal{C}_{n+1}$ and $\bigcup_{k=0}^{n} \mathcal{C}_k \times \mathcal{C}_{n-k}$. Hence

$$C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k}.$$

(b) We have

$$C(x)^2 = \sum_{n=0}^{\infty} \sum_{k=0}^{n} C_k C_{n-k} x^n = \sum_{n=0}^{\infty} C_{n+1} x^n = \frac{C(x) - 1}{x}.$$

Thus $xC(x)^2 - C(x) + 1 = 0$. The solution of this quadratic equation is

$$\frac{1 \pm \sqrt{1-4x}}{2x}.$$

We only need to find out whether we should take positive or negative sign. Since $C(0) = 1$ we need to take the negative sign:

$$C(x) = \frac{1 - \sqrt{1-4x}}{2x}.$$

□

Corollary 5.3.4. *We have*

$$C_n = \frac{\binom{2n}{n}}{n+1}.$$

Proof. Recall that

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n,$$

where

$$\binom{\alpha}{n} = \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!}.$$

Note that

$$\begin{aligned}\binom{\frac{1}{2}}{n} &= \frac{1}{n!}\frac{1}{2}\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\cdots\left(-\frac{3-2n}{2}\right)\\ &= \frac{(-1)^{n-1}}{2^n n!}1\cdot 3\cdot\cdots\cdot(2n-3)\\ &= \frac{(-1)^{n-1}}{(2n-1)2^n n!}\frac{1\cdot 2\cdot 3\cdot\cdots\cdot(2n-1)\cdot 2n}{2\cdot 4\cdot\cdots\cdot 2n}\\ &= \frac{(-1)^{n-1}}{(2n-1)2^n n!}\frac{(2n)!}{2^n n!}\\ &= \frac{(-1)^{n-1}}{4^n(2n-1)}\binom{2n}{n}\end{aligned}$$

Thus

$$C(x) = \frac{1}{2x}\left(1-\sum_{n=0}^{\infty}\frac{(-1)^{n-1}}{4^n(2n-1)}\binom{2n}{n}(-4)^n x^n\right)$$

By comparing the coefficient of x^n we get that

$$C_n = \frac{\binom{2n+2}{n+1}}{2(2n+1)} = \frac{\binom{2n}{n}}{n+1}.$$

□

Remark 5.3.5. Another notable power series using the fact $(1+x)^\alpha = \sum_{n=0}^{\infty}\binom{\alpha}{n}x^n$, is that

$$\frac{1}{\sqrt{1-4x}} = \sum_{n=0}^{\infty}\binom{2n}{n}x^n.$$

Beyond this course 5.3.6. Catalan numbers play an important role in the theory of random matrices. The reason is the following: Catalan numbers are the moments of the so-called Wigner's semicircle distribution. Consider the function

$$f(x) = \frac{1}{2\pi}\sqrt{4-x^2}$$

on the interval $[-2,2]$. Then

$$\int_{-2}^{2} f(x)\,dx = 1,$$

so it is a density function of a probability distribution, this is Wigner's semicircle distribution. Then

$$\int_{-2}^{2} f(x)x^{2n}\,dx = C_n,$$

and

$$\int_{-2}^{2} f(x)x^{2n-1}\,dx = 0.$$

5.4 Snake oil method

Generating functions also provide a powerful tool to evaluate certain sums. The so-called snake oil method is very simple, yet handles various sums. Roughly, the idea is the following. Suppose we have some sum that we would like to evaluate, say $\sum_{k=0}^{n} \binom{n}{k}$. Let us call it A_n. Then we determine $\sum_{n=0}^{\infty} A_n x^n$: generally this requires a change of summation and some simple algebraic manipulation. Once we have the generating function, in our case $\frac{1}{1-2x}$, we start to determine its coefficients: $\frac{1}{1-2x} = \sum_{n=0}^{\infty} 2^n x^n$. From this we conclude that $A_n = 2^n$. Below you can find several examples for this strategy. Can you fill the gaps in the above argument?

Proposition 5.4.1. *We have*

$$\sum_{k=0}^{n} \binom{n+k}{2k} 2^{n-k} = \frac{1}{3}(2 \cdot 4^n + 1).$$

Proof. Let

$$A_n = \sum_{k=0}^{n} \binom{n+k}{2k} 2^{n-k}.$$

Then

$$\begin{aligned}
\sum_{n=0}^{\infty} A_n x^n &= \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} \binom{n+k}{2k} 2^{n-k} \right) x^n \\
&= \sum_{k=0}^{\infty} \frac{1}{2^k} \left(\sum_{n} \binom{n+k}{2k} (2x)^n \right) \\
&= \sum_{k=0}^{\infty} \frac{1}{2^k} \frac{(2x)^k}{(1-2x)^{2k+1}} \\
&= \frac{1}{1-2x} \sum_{k=0}^{\infty} \left(\frac{x}{(1-2x)^2} \right)^k \\
&= \frac{1}{1-2x} \frac{1}{1-\frac{x}{(1-2x)^2}} \\
&= \frac{1-2x}{1-5x+4x^2} = \frac{1-2x}{(1-x)(1-4x)} \\
&= \frac{2}{3} \frac{1}{1-4x} + \frac{1}{3} \frac{1}{1-x} \\
&= \frac{2}{3} \sum_{n} (4x)^n + \frac{1}{3} \sum_{n} x^n.
\end{aligned}$$

Hence

$$\sum_{k=0}^{n} \binom{n+k}{2k} 2^{n-k} = \frac{1}{3}(2 \cdot 4^n + 1).$$

□

Proposition 5.4.2. *We have*

$$\sum_k \binom{m}{k}\binom{n+k}{m} = \sum_k \binom{m}{k}\binom{n}{k} 2^k.$$

Proof. Set

$$A_n = \sum_k \binom{m}{k}\binom{n+k}{m},$$

and

$$B_n = \sum_k \binom{m}{k}\binom{n}{k} 2^k$$

Then

$$\begin{aligned}
\sum_{n=0}^{\infty} A_n x^n &= \sum_{n=0}^{\infty} \left(\sum_k \binom{m}{k}\binom{n+k}{m} \right) x^n \\
&= \sum_{k=0}^{\infty} \binom{m}{k} \left(\sum_n \binom{n+k}{m} x^n \right) \\
&= \sum_{k=0}^{\infty} \binom{m}{k} \frac{x^{m-k}}{(1-x)^{m+1}} \\
&= \frac{x^m}{(1-x)^{m+1}} \sum_k \binom{m}{k} x^{-k} \\
&= \frac{x^m}{(1-x)^{m+1}} \left(1 + \frac{1}{x}\right)^m \\
&= \frac{(1+x)^m}{(1-x)^{m+1}}.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\sum_{n=0}^{\infty} B_n x^n &= \sum_{n=0}^{\infty} \left(\sum_k \binom{m}{k}\binom{n}{k} 2^k \right) x^n \\
&= \sum_{k=0}^{\infty} \binom{m}{k} 2^k \left(\sum_n \binom{n}{k} x^n \right) \\
&= \sum_{k=0}^{\infty} \binom{m}{k} 2^k \frac{x^k}{(1-x)^{k+1}}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{1-x} \sum_{k=0}^{\infty} \binom{m}{k} \left(\frac{2x}{1-x} \right)^k \\
&= \frac{1}{1-x} \left(1 + \frac{2x}{1-x} \right)^m \\
&= \frac{(1+x)^m}{(1-x)^{m+1}}.
\end{aligned}
$$

Hence $A_n = B_n$ for all n. $\square$

6. Number partitions

6.1 Basics

A number partition is a decomposition of a positive number as a sum of non-increasing numbers. For instance, 5 has the following partitions: $5 = 4 + 1 = 3 + 2 = 3 + 1 + 1 = 2 + 2 + 1 = 2 + 1 + 1 + 1 = 1 + 1 + 1 + 1 + 1$. Each partition has a graphical representations called Ferrers diagram. The conjugate of a Ferrers diagramm is a Ferrers diagram obtained from the original one via a reflection to a 45° line. This section is based (more or less) on the treatment of the book A Course in Combinatorics by Van Lint and Wilson (Chapter 15).

Proposition 6.1.1. *Let $p_k(n)$ be the number of partitions of n into at most k parts. Let $p^k(n)$ be the number of partitions of n with at most k parts. Then $p_k(n) = p^k(n)$.*

Proof. There is a natural bijection between the two sets: associate the conjugate partition to a partition. □

Next let us understand the generating function of the sequence $(p_k(n))$.

Theorem 6.1.2. *We have*

$$\sum_{n=0}^{\infty} p_k(n)x^n = \prod_{i=1}^{k} \frac{1}{1-x^i},$$

where $p_k(0) = 1$.

Proof. Note that

$$\prod_{i=1}^{k} \frac{1}{1-x^k} = \prod_{i=1}^{k} (1 + x^i + x^{2\cdot i} + x^{3\cdot i} + \dots).$$

If we expand this product, the coefficient of x^n will come from the products of the form $x^{m_1 \cdot 1} x^{m_2 \cdot 2} \cdots x^{m_k \cdot k}$, where $m_1, \dots, m_k \geq 0$ and $m_1 \cdot 1 + \cdots + m_k \cdot k = n$.

Note that this naturally correspond to the partition in which we have m_1 1's, m_2 2's,...,m_k k's and vice verse each partition naturally corresponds to such a term in the expansion. □

The same idea helps us to understand the generating function of all partitions.

Theorem 6.1.3. *We have*

$$\sum_{n=0}^{\infty} p(n)x^n = \prod_{i=1}^{\infty} \frac{1}{1-x^i},$$

where $p(0) = 1$.

Proof. As before

$$\prod_{i=1}^{\infty} \frac{1}{1-x^i} = \prod_{i=1}^{\infty} (1 + x^i + x^{2\cdot i} + x^{3\cdot i} + \dots).$$

It might be scary to consider an infinite product, but observe that if you want to compute the coefficient of x^n then you always have to choose the term 1 from the terms $1 + x^i + x^{2\cdot i} + x^{3\cdot i} + \dots$ when $i \geq n+1$. Let us introduce the notation $[x^n]f(x)$ for a_n if $f(x) = \sum_n a_n x^n$. Then

$$[x^n] \prod_{i=1}^{\infty} (1 + x^i + x^{2\cdot i} + x^{3\cdot i} + \dots) = [x^n] \prod_{i=1}^{n} (1 + x^i + x^{2\cdot i} + x^{3\cdot i} + \dots) = p_n(n) = p(n)$$

by the previous theorem and the fact that the largest part in a partition of n is at most n. Hence

$$\sum_{n=0}^{\infty} p(n)x^n = \prod_{i=1}^{\infty} \frac{1}{1-x^i}.$$

□

One can think to generating functions $\sum a_n x^n$ in two different ways:

(i) they are algebraic objects which form a ring, you can manipulate them algebraically, but you cannot plug any number (different form 0) into them,

(ii) they are analytic functions with some convergence radius.

The function $\sum n! x^n$ is a good example for the difference between (i) and (ii). Since the convergence radius is 0 for this function, you will hardly be able to do anything with it analytically, but this is a completely eligible algebraic expression, a "prominent" element of a ring.

Theorem 6.1.4. *Let $p_o(n)$ be the number of partitions of n into odd parts. Let $p_u(n)$ be the number of partitions of n into unequal parts. Then*

(a)

$$\sum_{n=0}^{\infty} p_o(n)x^n = \prod_{i=1}^{\infty} \frac{1}{1-x^{2i-1}}.$$

(b)

$$\sum_{n=0}^{\infty} p_u(n)x^n = \prod_{i=1}^{\infty} (1+x^i).$$

(c)

$$p_o(n) = p_u(n).$$

Proof. The proof of part (a) and (b) goes as before. We only concentrate to part (c). Note that

$$1 + x^i = \frac{1-x^{2i}}{1-x^i},$$

hence

$$\prod_{i=1}^{\infty}(1+x^i) = \prod_{i=1}^{\infty} \frac{1-x^{2i}}{1-x^i} = \prod_{i=1}^{\infty} \frac{1}{1-x^{2i-1}}.$$

since the terms $1-x^{2k}$ will cancel from the denominator and the enumerator. Hence $p_o(n) = p_u(n)$. □

Second proof for part (c). We will give a bijection between the set of partitions of n into odd parts and the set of partitions of n into unequal parts. The key ingredient of this bijection will be the observation that any number can be uniquely written in the form $2^k(2t+1)$, where $k, t \geq 0$. So let $(\lambda_1, \ldots, \lambda_m)$ be a partition of n such that $\lambda_1 > \cdots > \lambda_m$. Let $\lambda_i = 2^{k_i}(2t_i+1)$ and replace λ_i by 2^{k_i} pieces of $2t_i+1$. Then clearly we obtained a partition of n to odd parts.

Now we show that we can decode the original partition. Let's count the number of parts $2t_i+1$ in a partition of n into odd parts. Assume that there r_i pieces of $2t_i+1$. Then r_i can be uniquely written in base 2, i.e., there are unique $s_1 > s_2 > \cdots > s_j$ such that $r_i = 2^{s_1} + \cdots + 2^{s_j}$. Now replace the r_i pieces of $2t_i+1$ with elements $2^{s_n}(2t_i+1)$, where $1 \leq n \leq j$.

Hence we gave a bijection between the set of partitions of n to odd parts and the set of partitions of n to unequal parts and so $p_o(n) = p_u(n)$. □

An example for this proof is the following. Consider the partition $8+6+4+3+1$, then $8 = 2^3 \cdot 1$, $6 = 2 \cdot 3$, $4 = 2^2 \cdot 1$, $3 = 3$ and $1 = 1$. Hence the corresponding partition to odd parts will contain $8+4+1 = 13$ pieces of 1's $2+1 = 3$ pieces of 3's. And if you get the partition of 13 1's and 3 pieces of 3's then we know that we have to decompose 13 to 2-powers which can be uniquely done as $8+4+1$, and similarly $3 = 2+1$ so we get back the original partition.

6.2 An upper bound

In this section we give an upper bound for $p(n)$. In this proof we consider the generating function of $p(n)$ as an analytic function.

Theorem 6.2.1. *For $n > 2$ we have*

$$p(n) < \frac{\pi}{\sqrt{6(n-1)}} e^{\pi\sqrt{\frac{2}{3}n}}.$$

Remark 6.2.2. Hardy and Ramanujan proved that $p(n) \asymp \frac{1}{4\sqrt{3}n} e^{\pi\sqrt{\frac{2}{3}n}}$, so $\lim_{n\to\infty} \frac{p(n)}{f(n)} = 1$, where $f(n) = \frac{1}{4\sqrt{3}n} e^{\pi\sqrt{\frac{2}{3}n}}$. As we can see our upper bound agree with this function in the main term (and we won't need to work very hard for this bound).

Proof. (Van Lint) Recall that

$$P(t) = \prod_{k=1}^{\infty} \frac{1}{1-t^k} = \sum_{k=1}^{\infty} p(k) t^k.$$

We will see that $P(t)$ is convergent if $|t| < 1$. Actually, we will choose t to be $0 < t < 1$ later. The idea is the following, we will give an upper bound to $P(t)$ and we will choose a t such that $p(n)t^n$ dominates the terms in $P(t)$.

$$\log P(t) = \log\left(\prod_{k=1}^{\infty} \frac{1}{1-t^k}\right) = \sum_{k=1}^{\infty} \log \frac{1}{1-t^k}.$$

Note that

$$\log \frac{1}{1-t^k} = -\log(1-t^k) = \sum_{j=1}^{\infty} \frac{t^{kj}}{j}.$$

Then

$$\log P(t) = \sum_{k=1}^{\infty}\sum_{j=1}^{\infty} \frac{t^{kj}}{j} = \sum_{j=1}^{\infty} \frac{1}{j} \sum_{k=1}^{\infty} t^{kj} = \sum_{j=1}^{\infty} \frac{1}{j} \frac{t^j}{1-t^j}.$$

Now let $0 < t < 1$, then

$$\frac{1-t^j}{1-t} = 1 + t + t^2 + \cdots + t^{j-1} > jt^{j-1},$$

and so

$$\frac{t^j}{1-t^j} < t^j \frac{1}{(1-t)jt^{j-1}} = \frac{1}{j}\frac{t}{1-t}.$$

Then

$$\log P(t) = \sum_{j=1}^{\infty} \frac{1}{j}\frac{t^j}{1-t^j} < \sum_{j=1}^{\infty} \frac{1}{j}\frac{1}{j}\frac{t}{1-t} = \frac{t}{1-t}\sum_{j=1}^{\infty}\frac{1}{j^2} = \frac{\pi^2}{6}\frac{t}{1-t}.$$

Now we give a lower bound to $P(t)$. Note that $p(n)$ is a monotone increasing sequence (why?), so

$$P(t) = \sum_{k=1}^{\infty} p(k)t^k \geq \sum_{k=n}^{\infty} p(k)t^k \geq p(n)\sum_{k=n}^{\infty} t^k = p(n)\frac{t^n}{1-t}.$$

Hence

$$\log p(n) + \log\frac{t^n}{1-t} < \log P(t) < \frac{\pi^2}{6}\frac{t}{1-t}.$$

In other words,

$$\log p(n) \leq \frac{\pi^2}{6}\frac{t}{1-t} - n\log t + \log(1-t).$$

Now let $u = \frac{1-t}{t}$, then $t = \frac{1}{1+u}$. Then

$$\log p(n) \leq \frac{\pi^2}{6}\frac{1}{u} - n\log\frac{1}{1+u} + \log\frac{u}{1+u} = \frac{\pi^2}{6}\frac{1}{u} + (n-1)\log(1+u) + \log u.$$

Note that $\log(1+u) < u$ as $1 + u < e^u = 1 + u + \frac{u^2}{2} + \ldots$ Hence

$$\log p(n) < \frac{\pi^2}{6}\frac{1}{u} + (n-1)u + \log u.$$

Now let us choose u such a way that $\frac{\pi^2}{6}\frac{1}{u} = (n-1)u$ as it will be the (almost) optimal choice, then

$$u = \frac{\pi}{\sqrt{6(n-1)}}.$$

Then we have

$$\log p(n) < 2(n-1)u + \log u = \pi\sqrt{\frac{2}{3}(n-1)} + \log\frac{\pi}{\sqrt{6(n-1)}}.$$

In other words,

$$p(n) < \frac{\pi}{\sqrt{6(n-1)}} e^{\pi\sqrt{\frac{2}{3}(n-1)}}.$$

□

6.3 Euler's "pentagonal numbers" theorem

Let us consider the partitions of n into unequal parts, and let $p_e(n)$ be the number of partitions of n into even number of unequal parts, and let $p_o(n)$ be the number of partitions of n into even number of unequal parts. The following theorem is due to Euler.

Theorem 6.3.1. *We have*

$$p_e(n) - p_o(n) = \begin{cases} (-1)^k & \text{if } n = \frac{3k^2 \pm k}{2}, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. We define two transformations on partitions with unequal parts. They will be almost bijection between partitions of n into even and odd number of unequal parts.

Let $\lambda_1 > \cdots > \lambda_m$ be a partition of n into unequal parts. The dots in the last row of the Ferrers diagram is called the base. Its size is denoted by b, clearly $b = \lambda_m$. Let s be the largest integer k for which it is true that

$$\lambda_1 + 1 = \lambda_2 + 2 = \cdots = \lambda_k + k.$$

The number s is the size of the slope of the partition: on the Ferrers diagram, the slope can be seen as follows: draw a 45° line in the direction NE-SW through the upper-right dot of the Ferrers diagram, then the dots on this line is the slope. Its size is clearly s.

Now we give the two transformations:

Transformation I: if $b \leq s$ then delete the base from the Ferrers diagram and add $1-1$ dots to the first b rows, this way we created a new slope. This transformation results a new partition into unequal parts except if the original slope and base had a common dot and $b = s$. In this exceptional case: $n = b+(b+1)+\cdots+(2b-1) = \frac{3b^2-b}{2}$, note that the number of parts in this case is b too.

Transformation II: if $b > s$ then delete the slope from the Ferrers diagram and add a new base of size s to the Ferrers diagram. This transformation results a new partition into unequal parts except if the original slope and base had a common dot and $b = s+1$. In this exceptional case: $n = b+(b+1)+\cdots+(2b-3)+(2b-2) = \frac{3(b-1)^2+(b-1)}{2}$, note that the number of parts in this case is $b-1$.

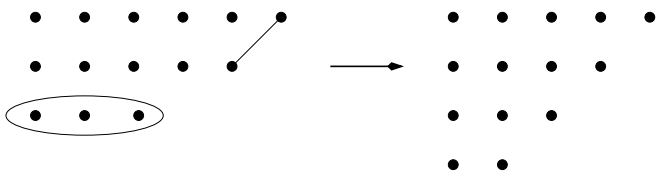


Figure 6.1: Transformation II

Both Transformation I and II change the parity of the number of parts and we can apply exactly one of them to a non-exceptional partition, and for the resulting partition we can only apply the other transformation which gives back the original partition.

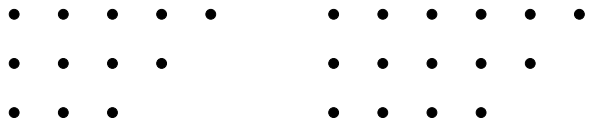


Figure 6.2: Exceptional Ferrers-diagrams

This shows that if $n \neq \frac{3k^2 \pm k}{2}$ then we get a bijection between partitions of n into even and odd number of unequal parts, and if $n \neq \frac{3k^2 \pm k}{2}$ we will have exactly one exceptional partition without pair and it has k parts. Hence we proved the theorem. □

Note that we can easily give the generating function of $p_e(n) - p_o(n)$ as follows:

$$\sum_{n=0}^{\infty}(p_e(n) - p_o(n))x^n = \prod_{i=1}^{\infty}(1 - x^i).$$

Indeed, if we expand the right hand side then a partition of n into k unequal parts will contribute $(-1)^k$ to the coefficient of n. Combining this obseervation with Euler's theorem we get the following corollary.

Corollary 6.3.2. *We have*

$$\prod_{n=1}^{\infty}(1 - x^n) = 1 + \sum_{k=1}^{\infty}(-1)^k \left(x^{(3k^2-k)/2} + x^{(3k^2+k)/2}\right).$$

The corollary of this corollary is a very fast way to compute the sequence $(p(n))$.

Corollary 6.3.3. *For $n \geq 1$ we have*

$$p(n) = \sum_{k=1}^{\infty} (-1)^{k+1} \left(p\left(n - \frac{3k^2 - k}{2} \right) + p\left(n - \frac{3k^2 + k}{2} \right) \right).$$

Proof. Recall that

$$\sum_{n=0}^{\infty} p(n) x^n = \prod_{i=1}^{\infty} \frac{1}{1 - x^i}.$$

Now if we multiply it with

$$\prod_{n=1}^{\infty} (1 - x^n) = 1 + \sum_{k=1}^{\infty} (-1)^k \left(x^{(3k^2-k)/2} + x^{(3k^2+k)/2} \right),$$

and compare the coefficient of n we get that for $n \geq 1$ we have

$$p(n) + \sum_{k=1}^{\infty} (-1)^k \left(p\left(n - \frac{3k^2 - k}{2} \right) + p\left(n - \frac{3k^2 + k}{2} \right) \right) = 0$$

which is equivelent with the statement of the corollary. □

6.4 Partitions fitting into a rectangle and q-binomial numbers

Let $(n)_q = 1 + q + q^2 + \cdots + q^{n-1}$, and $(n)_q! = (n)_q (n-1)_q \dots (1)_q$. Finally, let

$$\binom{n}{k}_q = \frac{(n)_q!}{(k)_q! (n-k)_q!}.$$

Note that

$$(n)_q = \frac{q^n - 1}{q - 1},$$

and so

$$\binom{n}{k}_q = \frac{(q^n - 1) \dots (q^{n-k+1} - 1)}{(q^k - 1) \dots (q - 1)}.$$

It turns out that $\binom{n}{k}_q$ is actually a polynomial in q. This follows from the following recursion formula.

Theorem 6.4.1. *We have*

$$\binom{n}{k}_q = \binom{n-1}{k}_q + q^{n-k} \binom{n-1}{k-1}_q.$$

Proof. This is a straightforward computation. □

Let $p_i(m,n)$ be the number of partitions of i with largest part at most n and at most m parts. So this is the number of partitions of i such that the Ferrers (or Young) diagram of the partition fits into the $m \times n$ rectangle.

Theorem 6.4.2. *We have*

$$\binom{n+m}{m}_q = \sum_{i=0}^{mn} p_i(m,n) q^i.$$

Proof. Observe that

$$p_i(m,n) = p_i(m,n-1) + p_{i-n}(m-1,n)$$

since either the first element of the partition is less than n and then the Young-tableaux fits into a rectangle of $m \times (n-1)$ or the first element is n, and then the Young-tableaux of $i-n$ fits into a rectangle of $(m-1) \times n$.

Note that this recursion is equivalent with

$$P(m,n) := \sum_{i=0}^{mn} p_i(m,n) q^i$$

satisfying the recursion

$$P(m,n) = P(m,n-1) + q^n P(m-1,n).$$

This is the same recursion that is satisfied by $\binom{n+m}{m}_q$. Since we also have

$$\binom{n}{0}_q = 1 = P(n,0) = P(0,n)$$

this shows that $P(m,n) = \binom{n+m}{m}_q$. □

It is easy to see that

$$\binom{n+m}{m}_q = q^{mn} \binom{n+m}{m}_{1/q}$$

from the definition. It means that the coefficients are symmetric: $p_i(m,n) = p_{mn-i}(m,n)$. Actually, this can be seen quite easily by deleting a Young-diagram from an $m \times n$ rectangle and then rotating the obtained diagram by 180°, thus obtaining a Young diagram of $mn - i$.

Theorem 6.4.3. *Let q be a prime power and $\mathbb{F}_q$ be the field with q elements. Then the number $N_q(n,k)$ of k-dimensional subspaces of the n-dimensional vector space $\mathbb{F}_q^n$ is $\binom{n}{k}_q$.*

Proof. Let us compute the number of elements of the following sets in two different ways:

$$S = \{(\underline{v}_1, \ldots, \underline{v}_k) \mid \underline{v}_1, \ldots, \underline{v}_k \in \mathbb{F}_q^n \text{ are linearly independent vectors}\}.$$

We can choose $\underline{v}_1$ in $q^n - 1$ different ways, because we can choose any vector different from $\underline{0}$. Then we can choose $\underline{v}_2$ in $q^n - q = q(q^{n-1} - 1)$ different ways as we can choose everything except $c\underline{v}_1$. Having chosen $\underline{v}_1, \ldots, \underline{v}_{t-1}$ we can choose $\underline{v}_t$ from $\mathbb{F}_q^n - \langle \underline{v}_1, \ldots, \underline{v}_{t-1} \rangle$ so we can choose $\underline{v}_t$ in $q^n - q^{t-1} = q^{t-1}(q^{n-t+1} - 1)$ ways. Hence

$$|S| = q^{k(k-1)/2}(q^n - 1)(q^{n-1} - 1)\ldots(q^{n-k+1} - 1).$$

On the other hand, we can first choose the k-dimensional subspace V induced by $\underline{v}_1, \ldots, \underline{v}_k$ in $N_q(n,k)$ ways and then inside V we can choose $\underline{v}_1, \ldots, \underline{v}_k$ in

$$q^{k(k-1)/2}(q^k - 1)(q^{k-1} - 1)\ldots(q - 1)$$

ways. Hence

$$|S| = N_q(n,k)q^{k(k-1)/2}(q^k - 1)(q^{k-1} - 1)\ldots(q - 1).$$

Hence we get that

$$N_q(n,k) = \binom{n}{k}_q.$$

□

Beyond this course 6.4.4. There is a natural partially ordered set on the Young-tableauxs fitting into the box $n \times m$ by containment. This poset is very reminiscent to the Boolean poset. This is not a coincidence, it can be derived from the the Boolean lattice B_{nm} by factoring out with the group $S_n \wr S_m$, the wreath product of the symmetric groups S_n and S_m. In particular, this is a Sperner poset: the largest antichain is the middle level. We can decompose the poset to chains by proving that there is always a matching between two consecutive levels that covers the smaller level. In particular,

$$p_0(n,m) \leq p_1(n,m) \leq \cdots \leq p_{\lceil nm/2 \rceil}(n,m) \geq \cdots \geq p_{nm}(n,m).$$

A sequence is called unimodal if it is increasing for a while and then it is decreasing. In general, it can be quite hard to prove that a certain combinatorial sequence is unimodal.

6.5 Hook length formula

Let $\lambda = (\lambda_1, \ldots, \lambda_m)$ be a partition of n. Assume that we write the numbers $1, 2, \ldots, n$ into the boxes of the Young-tableaux such that in each row and column the number are monotone increasing from left to right and from top to bottom, and each number appears exactly once. Such a configuration is called standard Young-tableaux.

1	2	6	10	11
3	5	9	12	
4	7			
8				
13				

Figure 6.3: A standard Young-tableaux

The goal of this section is to count the number of standard Young-tableauxs belonging to a given partition λ. The hook length formula of Frame, Robinson and Thrall gives a very elegant formula to determine the number of these standard Young-tableauxs.

For a cell (i, j) of the Young-tableaux let $H_{(i,j)}$ be the set of those cells which are below (i, j) or are left to (i, j) (but not below and left!) including the cell (i, j) itself. Let $h_{ij} = |H_{(i,j)}|$. For instance, the cell $(2, 2)$ containing the number 5 has hook length 4, the cell $(1, 1)$ containing the number 1 has hook length 9.

Theorem 6.5.1 (Frame, Robinson, Thrall)**.** *Let f^λ be the number of standard Young-tableauxs with shape λ. Then*

$$f^\lambda = \frac{n!}{\prod_{i,j} h_{i,j}}.$$

Proof. First of all, it will be a bit more convenient to work with the following formula:

$$g(\lambda_1, \ldots, \lambda_m) = \frac{(\sum_{i=1}^m \lambda_i)! \prod_{i<j}((\lambda_i - i) - (\lambda_j - j))}{\prod_{i=1}^m (\lambda_i + m - i)!}.$$

We will show later that

$$g(\lambda_1, \ldots, \lambda_m) = \frac{n!}{\prod_{i,j} h_{i,j}}.$$

Instead of f^λ, let us write $f(\lambda_1, \ldots, \lambda_m)$. The main idea of the proof is to show that f and g satisfy the same recursion formula with the same boundary conditions which uniquely determine the function f or g, that is, $f = g$. We will show that both f and g satisfy the recursion formula

$$f(\lambda_1, \ldots, \lambda_m) = f(\lambda_1 - 1, \lambda_2, \ldots, \lambda_m) + f(\lambda_1, \lambda_2 - 1, \ldots, \lambda_m) + \cdots + f(\lambda_1, \lambda_2 \ldots, \lambda_m - 1),$$

and

$$g(\lambda_1, \ldots, \lambda_m) = g(\lambda_1 - 1, \lambda_2, \ldots, \lambda_m) + g(\lambda_1, \lambda_2 - 1, \ldots, \lambda_m) + \cdots + g(\lambda_1, \lambda_2 \ldots, \lambda_m - 1).$$

Now we have to stop for a moment, it is clear what $g(\lambda_1, \lambda_2 \ldots, \lambda_k - 1, \ldots, \lambda_m)$ means even if $(\lambda_1, \lambda_2 \ldots, \lambda_k - 1, \ldots, \lambda_m)$ is not a partition, but what does the expression $f(\lambda_1, \lambda_2 \ldots, \lambda_k - 1, \ldots, \lambda_m)$ mean if $(\lambda_1, \lambda_2 \ldots, \lambda_k - 1, \ldots, \lambda_m)$ is not a partition? The trick is that we simply define it as 0 and we will consider it as a boundary condition. Let us call a sequence $(\lambda_1, \lambda_2 \ldots, \lambda_k - 1, \ldots, \lambda_m)$ an almost partition if $(\lambda_1, \lambda_2 \ldots, \lambda_k, \ldots, \lambda_m)$ is a partition, but $(\lambda_1, \lambda_2 \ldots, \lambda_k - 1, \ldots, \lambda_m)$ is not a partition. Clearly, we get an almost partition when $\lambda_k = \lambda_{k+1}$, but $\lambda_k - 1 \not\geq \lambda_{k+1}$. This way we immediately recognise an almost partition: we only have to find the unique element which is 1 less then the next one. Now we are ready to give the boundary conditions:

- $f((n)) = 1$.
- $f(\lambda_1, \ldots, \lambda_{m-1}, 0) = f(\lambda_1, \ldots, \lambda_{m-1})$.
- $f(\lambda'_1, \ldots, \lambda'_m) = 0$ whenever $(\lambda'_1, \ldots, \lambda'_m)$ is an almost partition.

Now we can see that f satisfies the recursion formula with this carefully chosen boundary conditions: we can only write the number n to some end of a row, say k-th row. If deleting this box from the Young-tableaux would result an almost partition then it means that we shouldn't have written n in this box, but this is not a problem since $f(\lambda_1, \lambda_2 \ldots, \lambda_k - 1, \ldots, \lambda_m)$ is 0 anyway by definition. It is also clear that f satisfies the first two boundary conditions. It is also obvious that the recursion formula together with the three boundary conditions completely determine the function f. All we need to show that g also satisfies the recursion formula together with the three boundary conditions.

First, $g((n)) = \frac{n! \cdot 1}{n!} = 1$. Secondly, if $\lambda_m = 0$ then

$$g(\lambda_1, \ldots, \lambda_{m-1}, 0) = \frac{(\sum_{i=1}^m \lambda_i)! \prod_{i<j} ((\lambda_i - i) - (\lambda_j - j))}{\prod_{i=1}^m (\lambda_i + m - i)!} =$$

$$= \frac{(\sum_{i=1}^{m}\lambda_i)!\prod_{i<j\le m-1}((\lambda_i-i)-(\lambda_j-j))\cdot\prod_{i=1}^{m-1}((\lambda_i-i)-(\lambda_m-m))}{\prod_{i=1}^{m-1}(\lambda_i+m-1-i)!\prod_{i=1}^{m-1}(\lambda_i+m-i)\cdot(\lambda_m+m-m)!} =$$

$$= \frac{(\sum_{i=1}^{m-1}\lambda_i)!\prod_{i<j\le m-1}((\lambda_i-i)-(\lambda_j-j))}{\prod_{i=1}^{m-1}(\lambda_i+m-1-i)!} = g(\lambda_1,\dots,\lambda_{m-1})$$

since $(\lambda_i-i)-(\lambda_m-m)=\lambda_i+m-i$ and $\lambda_m!=1$ if $\lambda_m=0$. Next we show that $g(\lambda'_1,\dots,\lambda'_m)=0$ whenever $(\lambda'_1,\dots,\lambda'_m)$ is an almost partition. Since $(\lambda'_1,\dots,\lambda'_m)$ is an almost partition, there exists a k such that $\lambda'_k=\lambda'_{k+1}-1$. Then $(\lambda'_k-k)-(\lambda'_{k+1}-k-1)=0$, but this term appears in the enumerator of the function g. Hence $g(\lambda'_1,\dots,\lambda'_m)=0$. So far we proved that g satisfies the same boundary conditions as f. Next we show that g satisfies the same recursion formula too. Clearly,

$$g(\lambda_1,\dots,\lambda_m)=g(\lambda_1-1,\lambda_2,\dots,\lambda_m)+g(\lambda_1,\lambda_2-1,\dots,\lambda_m)+\cdots+g(\lambda_1,\lambda_2\dots,\lambda_m-1).$$

is equivalent with

$$1=\sum_{k=1}^{m}\frac{g(\lambda_1,\dots,\lambda_k-1,\dots,\lambda_m)}{g(\lambda_1,\dots,\lambda_k,\dots,\lambda_m)}.$$

We have

$$\frac{g(\lambda_1,\dots,\lambda_k-1,\dots,\lambda_m)}{g(\lambda_1,\dots,\lambda_k,\dots,\lambda_m)}=\frac{\frac{(\sum_{i=1}^{m}\lambda'_i)!\prod_{i<j}((\lambda'_i-i)-(\lambda'_j-j))}{\prod_{i=1}^{m}(\lambda'_i+m-i)!}}{\frac{(\sum_{i=1}^{m}\lambda_i)!\prod_{i<j}((\lambda_i-i)-(\lambda_j-j))}{\prod_{i=1}^{m}(\lambda_i+m-i)!}},$$

where $\lambda'_i=\lambda_i$ if $i\neq k$ and $\lambda'_k=\lambda_k-1$. Comparing the two products we get that

$$\frac{g(\lambda_1,\dots,\lambda_k-1,\dots,\lambda_m)}{g(\lambda_1,\dots,\lambda_k,\dots,\lambda_m)}=$$

$$=\frac{\lambda_k+m-k}{n}\prod_{i<k}\frac{(\lambda_i-i)-(\lambda_k-1-k)}{(\lambda_i-i)-(\lambda_k-k)}\prod_{j>k}\frac{(\lambda_k-1-k)-(\lambda_j-j)}{(\lambda_k-k)-(\lambda_j-j)}$$

Now if we introduce the notation $z_j=\lambda_j+m-j$ for all j, then for $i<k$ we have

$$\frac{(\lambda_i-i)-(\lambda_k-1-k)}{(\lambda_i-i)-(\lambda_k-k)}=\frac{1+z_i-z_k}{z_i-z_k}=1+\frac{1}{z_i-z_k},$$

while for $j>k$ we have

$$\frac{(\lambda_k-1-k)-(\lambda_j-j)}{(\lambda_k-k)-(\lambda_j-j)}=\frac{z_k-1-z_j}{z_k-z_j}=\frac{1+z_j-z_k}{z_j-z_k}=1+\frac{1}{z_j-z_k}.$$

Finally, $n=\sum\lambda_i=\sum z_i-\frac{m(m-1)}{2}$. Hence

$$\frac{g(\lambda_1,\dots,\lambda_k-1,\dots,\lambda_m)}{g(\lambda_1,\dots,\lambda_k,\dots,\lambda_m)}=\frac{z_k}{\sum_{i=1}^{m}z_i-\frac{m(m-1)}{2}}\prod_{j\neq k}\left(1+\frac{1}{z_j-z_k}\right).$$

So the claim

$$1 = \sum_{k=1}^{m} \frac{g(\lambda_1, \ldots, \lambda_k - 1, \ldots, \lambda_m)}{g(\lambda_1, \ldots, \lambda_k, \ldots, \lambda_m)}$$

equivalent with

$$\sum_{i=1}^{m} z_i - \frac{m(m-1)}{2} = \sum_{k=1}^{m} z_k \prod_{j \neq k} \left(1 + \frac{1}{z_j - z_k}\right).$$

To prove this identity, let us recall that if $z_1, \ldots, z_m$ are different numbers and $f_1, \ldots, f_m$ are given numbers then there is a polynomial $p(x)$ of degree at most $m-1$ such that $p(z_i) = f_i$ for $i = 1, \ldots, m$. Note that this polynomial is unique: if $p(z_i) = q(z_i) = f_i$ for $i = 1, \ldots, m$ and both $p(x)$ and $q(x)$ have degree at most $m-1$, then the polynomial $p - q$ has degree at most $m-1$ too and it has m zeros hence $p - q \equiv 0$. Lagrange's interpolation gives the polynomial p:

$$p(x) = \sum_{i=1}^{m} f_i \frac{\prod_{j \neq i}(x - z_j)}{\prod_{j \neq i}(z_i - z_j)}.$$

If $m \geq 3$ and $f_i = z_i$, then

$$p(x) = \sum_{i=1}^{m} z_i \frac{\prod_{j \neq i}(x - z_j)}{\prod_{j \neq i}(z_i - z_j)}.$$

On the other hand, the polynomial $q(x) = x$ is clearly satisfies that $q(z_i) = z_i$ and it has degree at most $m-1$. Hence

$$x = \sum_{i=1}^{m} z_i \frac{\prod_{j \neq i}(x - z_j)}{\prod_{j \neq i}(z_i - z_j)}.$$

By comparing the coefficient of x^{m-1} on the two sides we get that

$$0 = \sum_{i=1}^{m} z_i \prod_{j \neq i} \frac{1}{z_i - z_j}.$$

By multiplying by $(-1)^{m-1}$ we get that

$$0 = \sum_{i=1}^{m} z_i \prod_{j \neq i} \frac{1}{z_j - z_i}.$$

Now if we expand the products in

$$\sum_{k=1}^{m} z_k \prod_{j \neq k} \left(1 + \frac{1}{z_j - z_k}\right),$$

we immediately get the terms $\sum_{k=1}^{m} z_k$ by choosing 1 from each term of the products. If we choose exactly one non-1 term from each product we get

$$\sum_{i,j}\left(z_i\frac{1}{z_j-z_i}+z_j\frac{1}{z_i-z_j}\right)=\sum_{i,j}(-1)=-\frac{m(m-1)}{2}.$$

Finally, if choose more than one non-1 term from each product then we simply collect those terms together which contain the same $z_{i_1}, z_{i_2}, \ldots, z_{i_t}$ and apply the identity

$$0=\sum_{j=1}^{t} z_{i_j}\prod_{v\neq j}\frac{1}{z_{i_v}-z_{i_j}}$$

by observing that $t\geq 3$ in this case. Hence

$$\sum_{i=1}^{m} z_i-\frac{m(m-1)}{2}=\sum_{k=1}^{m} z_k\prod_{j\neq k}\left(1+\frac{1}{z_j-z_k}\right)$$

indeed true.

So far we have proved that

$$f(\lambda_1,\ldots,\lambda_m)=\frac{(\sum_{i=1}^{m}\lambda_i)!\prod_{i<j}((\lambda_i-i)-(\lambda_j-j))}{\prod_{i=1}^{m}(\lambda_i+m-i)!}.$$

Now we will show that

$$f(\lambda_1,\ldots,\lambda_m)=\frac{n!}{\prod_{i,j}h_{i,j}}.$$

We only need to observe that

$$\prod_{j}h_{k,j}=\frac{(\lambda_k+m-k)!}{\prod_{k<j}((\lambda_k-k)-(\lambda_j-j))}.$$

This is indeed true: let us start to write the hook lengths from right to left:

$$1\cdot 2\cdot\ldots\cdot(\lambda_k-\lambda_{k+1})\cdot$$

$$(\lambda_k-\lambda_{k+1}+2)\cdot\ldots\cdot(\lambda_k-\lambda_{k+2}+1)\cdot$$

$$(\lambda_k-\lambda_{k+2}+3)\cdot\ldots\cdot(\lambda_k-\lambda_{k+3}+2)\cdot$$

$$\ldots$$

$$(\lambda_k-\lambda_m+m-k+1)\ldots\cdot(\lambda_k+m-k).$$

Note that the missing numbers in this product are $\lambda_k - \lambda_{k+j} + j = (\lambda_k - k) - (\lambda_{k+j} - (k+j))$. Hence indeed we have

$$\prod_j h_{k,j} = \frac{(\lambda_k + m - k)!}{\prod_{k<j}((\lambda_k - k) - (\lambda_j - j))}.$$

This completes the proof of the hook length formula. □

Remark 6.5.2. Recall that the number of sequences $(s_1, s_2, \ldots, s_{2n})$ satisfying the conditions $s_i = \pm 1$, $\sum_{i=1}^{m} s_i \geq 0$ for all $1 \leq m \leq 2n$ and $\sum_{i=1}^{2n} s_i = 0$ is counted by the Catalan-number $C_n = \frac{\binom{2n}{n}}{n+1}$. Now we can give a simple proof for it.

Consider the $2 \times n$ Young-tableax. Note that by the hook length formula we have

$$f^{n,n} = \frac{(2n)!}{(n+1)!n!} = \frac{\binom{2n}{n}}{n+1}.$$

On the other hand, the standard Young-tableauxs with this shape are in one-to-one bijection with the above sequences: let $s_i = 1$ if i is in the first row and $s_i = -1$ if i is in the second row. It is easy to check that this is indeed a bijection (check it!).

Beyond this course 6.5.3. The numbers f^λ play an extremely important role in the representation theory of symmetric groups: they are the dimensions of the irreducible representations. This in particular implies that

$$\sum_{\lambda \vdash n} (f^\lambda)^2 = n!.$$

There is a combinatorial proof of this fact. Indeed, there is a combinatorial bijection between pairs of standard Young-tableauxs of the same shape and the permutations on n elements.

7. Extremal Graph Theory

7.1 Introduction

In this chapter all graphs are simple and finite. If $G = (V, E)$ is graph then $v(G) = |V(G)|$ and $e(G) = |E(G)|$ denotes the number of vertices and edges, respectively. The notation $H \subset G$ means that H is a (not necessarily induced) subgraph of G. Furthermore, C_n will denote the cycle of length n. The goal of this chapter is to prove the following classical theorem of Bondy and Simonovits.

Theorem 7.1.1 (Bondy–Simonovits [3])**.** *For all $k \geq 2$ there exist constants c_k and $n_0(k)$ such that for all graph G_n with $n \geq n_0(k)$ vertices and $e(G_n) \geq c_k n^{1+1/k}$ edges the graph G_n contains a cycle C_{2k}.*

This chapter follows the treatment of the original paper of Bondy and Simonovits [3]. Practically the same proof can be found in the book Extremal graph theory by Béla Bollobás [2]. This book is still a standard reference on this area despite the fact that some parts are superseded by now[1].

This chapter is organized as follows. In the next section we recall the case of C_4. Furthermore, we show a similar, but simpler result that contains the main idea of the Bondy–Simonovits theorem. This way we will be able to separate the main idea from the technical difficulties. In the third part we will prove the Bondy–Simonovits theorem. In fact, we will prove a stronger statement than what we stated.

7.2 Retrospection and the main idea of the proof

First, we study the case of C_4. You might have learned it in your first year, but it is worth a new look. The proof of this theorem is different from the proof of

[1]This is not a big surprise, the book was written in 1978, so for instance it does not contain the regularity lemma at all, a cornerstone of the modern extremal graph theory

Theorem 7.1.1, but its main idea is also very important, and is based on the technique cherry-counting.

Theorem 7.2.1. *Let G_n be a graph on n vertices that does not contain a C_4. Then*

$$e(G_n) \leq \frac{n^{3/2}}{2} + \frac{n}{4}.$$

Proof. Let $d_1, d_2, \ldots, d_n$ be the degree sequence of the graph G_n. Then

$$2e(G_n) = \sum_{k=1}^{n} d_k.$$

Let us count the number of paths of length 2 (i. e., it is the path on 3 vertices, more popularly known as cherry) in the graph G. Let ch be the number of cherries. If the middle vertex of the cherry is the k.th vertex then we can choose the other two vertices in $\binom{d_k}{2}$ ways. Hence

$$\mathrm{ch} = \sum_{k=1}^{n} \binom{d_k}{2}.$$

The crucial observation is that every vertex pair can be the end points of at most one cherry, otherwise two cherries with the same endpoints determine a C_4. So there are at most $\binom{n}{2}$ cherries. Hence

$$\binom{n}{2} \geq \mathrm{ch} = \sum_{k=1}^{n} \binom{d_k}{2}.$$

Next we introduce a useful lower bound for the number of cherries from the exact formula.

$$\mathrm{ch} = \sum_{k=1}^{n} \binom{d_k}{2} = \frac{1}{2}\sum_{k=1}^{n} d_k^2 - \frac{1}{2}\sum_{k=1}^{n} d_k = \frac{1}{2}\sum_{k=1}^{n} d_k^2 - e(G_n) \geq$$

$$\geq \frac{1}{2n}\left(\sum_{k=1}^{n} d_k\right)^2 - e(G_n) = \frac{(2e(G_n))^2}{2n} - e(G_n) = \frac{2e(G_n)^2}{n} - e(G_n).$$

Whence

$$\binom{n}{2} \geq \frac{2e(G_n)^2}{n} - e(G_n).$$

This means that $e(G_n)$ is at most the largest root of the quadratic equation

$$\frac{2}{n}x^2 - x - \frac{n(n-1)}{2} = 0.$$

Hence

$$e(G_n) \leq x_1 = \frac{1+\sqrt{1+4\frac{2}{n}\frac{n(n-1)}{2}}}{2\frac{2}{n}} =$$

$$= \frac{n}{4}\left(1+\sqrt{4n-3}\right) \leq \frac{n}{4}\left(1+\sqrt{4n}\right) = \frac{n^{3/2}}{2} + \frac{n}{4}.$$

This finishes the proof of the theorem. □

The proof of the next theorem is the same as the main idea of the proof of Theorem 7.1.1.

Theorem 7.2.2. *Let G_n be a graph on n vertices without no cycle of length at most $2k$. Then*

$$e(G_n) \leq n^{1+1/k} + n.$$

Proof. Let d be the smallest degree of the graph G, and let x be a vertex of degree d. Let us build a breadth first search tree from the vertex x. (It is not necessarily a spanning tree as G might not be connected.) Let V_j be the set of vertices of distance j from x. Suppose that $j \leq k-1$. Then we claim that if $u, v \in V_j$ then they are not adjacent and they don't have a common neighbor in V_{j+1}. Indeed, if they were adjacent there would be a cycle of length at most $2j+1 < 2k$ in G determined by the spanning tree and the edge (u, v), and if they have a common neighbor in V_{j+1} then there would be a cycle of length at most $2(j+1) \leq 2k$ in G. Hence every vertex in V_j has at least $d-1$ neighbors in V_{j+1}, and for different $u, v \in V_j$ their neighbors are different. This means that $|V_{j+1}| \geq (d-1)|V_j|$ if $j \leq k-1$. This implies that

$$n \geq |V_k| \geq (d-1)|V_{k-1}| \geq (d-1)^2|V_{k-2}| \geq \cdots \geq (d-1)^k.$$

Whence $d \leq n^{1/k} + 1$. Now delete vertex x and let us repeat the argument for the graph on $n-1$ vertices. This graph does not contain a cycle of length at most $2k$ either, so the smallest degree is at most $(n-1)^{1/k} + 1$. We can repeat this process till we get the empty graph. Then we obtain that

$$e(G_n) \leq (n^{1/k}+1) + ((n-1)^{1/k}+1) + \cdots + (1^{1/k}+1) \leq n \cdot n^{1/k} + n.$$

Hence $e(G_n) \leq n^{1+1/k} + n$. □

7.3 General extremal graph theoretic lemmas

In this section we offer a collection of simple observations that comes in handy many times, and we use them in the proof of the Bondy-Simonovits theorem too.

Lemma 7.3.1. *Let G be a graph with $e(G)$ edges. Then there exists a bipartite subgraph H of G such that $d_H(x) \geq \frac{1}{2}d_G(x)$ for all $x \in V(G)$. In particular, $e(H) \geq \frac{1}{2}e(G)$.*

Proof. For a set $A \subseteq V(G)$ let $e(A, V \setminus A)$ denote the number of edges going between the sets A and $V \setminus A$. Let X be a set that maximizes $e(X, V \setminus X)$. Let H be the bipartite graph determined by X and $V \setminus X$. We claim that $d_H(x) \geq \frac{1}{2}d_G(x)$ for all $x \in V(G)$. Indeed, if there were a $v \in V$ violating the inequality then by putting v to the other class the number of edges between the two classes would increase with $(d_G(v) - d_H(v)) - d_H(v) = d_G(v) - 2d_H(v)$ edges. This would contradict the choice of X. This finishes the proof of the lemma. □

Remark 7.3.2. Lemma 7.3.1 is used as follows. Suppose that we wish to prove that if G has no forbidden subgraph K then it has at most $O(f(n))$ edges. If we can prove this statement for bipartite graphs then we can prove it for all graphs using the lemma only losing a factor 2. Since these sorts of lemmas can be useful we mention two more lemmas of this type without proof.

Lemma 7.3.3. *Let G be a graph with $e(G)$ edges. Then there exists a bipartite subgraph $H = (A, B, E')$ of G such that $||A| - |B|| \leq 1$ and $e(H) \geq \frac{1}{2}e(G)$.*

Remark 7.3.4. So we can require that the two classes of the bipartite graph is (almost) the same size. Unfortunately, we cannot require that at the same time the inequality $d_H(x) \geq \frac{1}{2}d_G(x)$ holds true for all vertex x. The lemma can be proved by the first moment method.

Sometimes it can be useful that the degrees are not only bounded below, but also bounded above. The next lemma is an example for such a statement.

Lemma 7.3.5. *Let G be a d–regular graph. Then there exists a bipartite subgraph $H = (A, B, E')$ for which*

$$\left| d_H(x) - \frac{d}{2} \right| \leq 10\sqrt{d \log d}$$

for all $x \in V(G)$.

Remark 7.3.6. The main idea of the proof is: LLL, i. e., Lovász local lemma.

$\star \quad \star \quad \star$

Next we prove another very useful and very simple statement.

Lemma 7.3.7 (Minimum degree-average degree principle)**.** *Let P be a property for which the following is satisfied: if H has the property P then so all graphs G that contains H as an induced subgraph. (Example: let $P(K)$ be the property that H contains K.) Assume that if a graph H on at most n vertices has minimum degree $r = r(n)$ then it has property P. Suppose that the graph G does not satisfy property P, then it has at most $(r-1)n$ edges, and consequently the average degree is at most $2r$.*

Proof. Since G does not satisfy property P, it has a degree at most $r-1$. After deleting this vertex, the new graph again contains a vertex of degree at most $r-1$. By continuing this argument we get that G has at most $(r-1)n$ edges. Since the number of edges is less than rn, the average degree is less than $2r$. □

Remark 7.3.8. Example: if the minimum degree of a graph G is at least 3 then it contains a cycle of even length (why?). Whence if a graph on n vertices does not contain a cycle of even length then it has at most $2n$ edges.

Another way to rephrase Lemma 7.3.7 is that if we wish to prove that a certain graph has at most rn edges, then it is enough to prove that with minimal degree r the graph would have property P.

7.4 Proof of Theorem 7.1.1

In this section we prove the Bondy–Simonovits theorem. As we already mentioned we will prove a slightly stronger statement.

Theorem 7.4.1. *Let G_n be a graph on n vertices. If $e(G_n) > 100kn^{1+1/k}$ then $C_{2\ell} \subseteq G_n$ for all $\ell \in [k, kn^{1/k}]$.*

In fact, it will be more convenient to prove the following statement.

Theorem 7.4.2. *Let $E = e(G_n)$. Then $C_{2\ell} \subseteq G_n$ for all $\ell \geq 2$ satisfying the inequalities $\ell \leq \frac{E}{100n}$ and $\ell n^{1/\ell} \leq \frac{E}{10n}$.*

It is easy to see that Theorem 7.4.2 implies Theorem 7.4.1 which in turn implies Theorem 7.1.1.

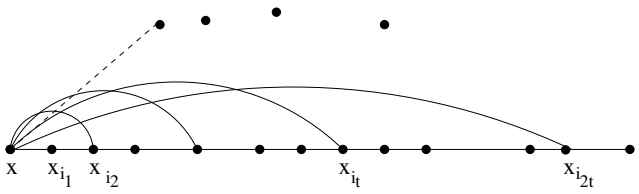
⋆ ⋆ ⋆

First we prove a rather strange lemma that, at least for the first sight, has nothing to do with the theorem. Later this lemma will come in handy. We need the following definition.

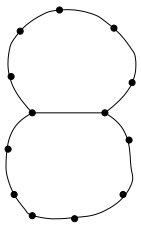
Definition 7.4.3. A (not necessarily proper) coloring of the vertices of a graph G is *t-periodic* if arbitrary path of length t in G has end points of the same color. (A path has length t if it has t edge.)

Lemma 7.4.4. *Let t be a positive integer. Let G be a connected graph with $e(G)$ edges, $v(G)$ vertices such that $e(G) \geq 2tv(G)$. Then any t-periodic coloring of G has at most 2 colors.*

Proof. First we show that if $e(G) \geq 2tv(G)$, then G contains two adjacent vertices that are connected by two paths of length at least t. By Lemma 7.3.7 (minimum degree average degree principle) it is enough to show this claim where the minimum degree is at least $2t$. So assume that the minimum degree is at least $2t$ in the graph G.



Let $P = x_1x_2 \dots x_k$ be a longest path of the graph G. Since P is a longest path all neighbors of x_1 should be on the path otherwise one can extend P with a new vertex. Let the neighbors of x_1 be $x_{i_1} = x_2, x_{i_2}, \dots, x_{i_r}$, where $r \geq 2t$. Let us consider the vertices x_1 and x_{i_t}. They are adjacent and the paths $P_1 = x_1x_2 \dots x_{i_t}$ and $P_2 = x_{i_t}x_{i_t+1} \dots x_{i_{2t}}x_1$ are of length at least t. In what follows we call this special subgraph a Θ-graph.



Since the coloring of G is t-periodic the coloring of this Θ-subgraph is also t-periodic. Let us consider this t-periodic coloring. Let C_1, C_2, C_3 be the three cycles of the Θ-subgraph, and let ℓ_1, ℓ_2, ℓ_3 be their lengths. Furthermore, let t_i be the smallest period with which the cycle C_i is periodic.

It is easy to see that the period should be the same on the three cycles (why?), so $t_1 = t_2 = t_3$. Moreover, $t_i \mid \ell_i$. If C_3 is the longest cycle, then $\ell_1 + \ell_2 - \ell_3 = 2$, and so $t_i = t^* \mid 2$, i. e. $t^* = 1$ or 2. Hence we can use at most 2 colors on the Θ-subgraph.

Since G is connected, for arbitrary vertex v there exists a path connecting it with the Θ-subgraph. Suppose that it is of length $ta + b$, where $b < t$. Then making $t - b$ steps on the Θ-subgraph we can reach a vertex of this subgraph that is connected to v by a path of length divisible by t. Hence their colors must be the same implying that the whole graph is colored with at most 2 colors. □

The next lemma can be considered as the main step of the proof as Theorem 7.4.2 will easily follow from it.

Lemma 7.4.5. *Let G be a bipartite graph on n vertices with minimum degree at least $s = \max(5\ell n^{1/\ell}, 50\ell)$. Then $C_{2\ell} \subseteq G$.*

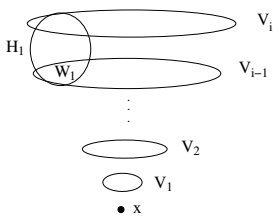
Proof. Let x be an arbitrary vertex of the graph G, and let V_i be the vertices of distance i from x. Since G is bipartite all V_i induces an independent set.

Suppose for contradiction that G does not contain a cycle $C_{2\ell}$. We will show that for all $1 \leq i \leq \ell$ we have

$$\frac{|V_i|}{|V_{i-1}|} \geq \frac{s}{5\ell}. \tag{7.1}$$

This will immediately lead to a contradiction as $s \geq 5\ell n^{1/\ell}$ implies that $|V_i| \geq n^{1/\ell}|V_{i-1}|$, and so $|V_\ell| \geq n$, but the whole graph has n vertices.

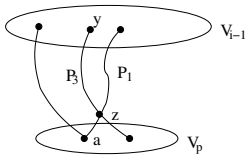
We will prove the statement 7.1 by induction on i. The case $i = 1$ is trivial since $\deg(x) \geq s \geq \frac{s}{5\ell}$. Suppose that we already proved the claim for $i - 1$.



Let $H = G[V_{i-1} \cup V_i]$ be the subgraph induced by the set $V_{i-1} \cup V_i$. Let $H_1, \ldots, H_q$ be the connected components of H. Finally set $W_j = H_j \cap V_{i-1}$. We will call a path $y_1 y_2 \ldots y_r$ monotone if the distance of y_i from x is strictly monotone increasing or decreasing. (In other words, a monotone path has at most one common vertex with each set V_i.)

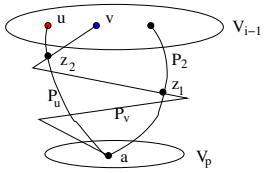
First we show that $e(H_1) < 4\ell v(H_1)$. This is trivial if $|W_1| = 1$, because then for the degree d of this one vertex we have the inequality $d < 4\ell(d + 1)$, and this is trivially satisfied. So let us suppose that W_1 has at least 2 vertices.

Let $a \in V_p$ be a vertex satisfying the following conditions:
(i) there exist two monotone paths P_1 and P_2 from vertex a to W_1 whose only intersection is the vertex a,
(ii) p is minimal with respect to condition (i).



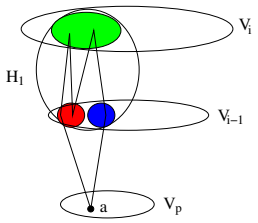
First we show that all vertices of W_1 is connected with vertex a by a monotone path. Let $y \in W_1$ and P_3 be a monotone path connecting x and y. Because of the minimality of p, the path P_3 has to intersect P_1 at some point z, and then we get a monotone path between a and y: aP_1zP_3y.

Next we associate colors red and blue to the vertices of W_1 in such a way that if two vertices, say u and v, have different colors then, there exist monotone paths from a to u and v whose only intersection is the vertex a.



We can do it as follows. If for a vertex u there exists a monotone path from u to a that is disjoint from P_2 (apart from vertex a), then let us color the vertex u to red, otherwise let us color it to blue. We will show that this coloring satisfies the above condition.

Suppose that the color of u is red, and the color of v is blue. Then there exists a monotone path P_u between u and a such that P_2 is disjoint from P_u. Let P_v be any monotone path between v and a. Since v is blue P_v has to intersect P_2 in a vertex different from a. Let the last such intersection point (counted on P_v) be z_1. Suppose that P_u and P_v are not disjoint: they have an intersection, let z_2 be the last intersection again counted on P_v. Now we have to distinguish two cases according to the order of z_1 or z_2 on P_v. If z_2 is closer to v then $vP_vz_2P_ua$ is a path between v and a disjoint from P_2, so the color of v should have been red. If z_1 is closer to v on P_v then the path $vP_vz_1P_2a$ is a monotone path between v and a that is disjoint from P_u. Hence this coloring is indeed good.



Next let us color the vertices of $H_1 \cap V_i$ to green. We claim that this blue-red-green coloring of H_1 is a t-periodic coloring with $t = 2(\ell - i + p + 1)$. Suppose for contradiction that this is not true. Since H_1 is a bipartite graph this can only occur if there is a red $u \in W_1$ and a blue $v \in W_1$ that are connected by a path Q of length $2(\ell - i + p + 1)$. Since u and v have different colors there are disjoint paths P_u and P_v from u and v to the vertex a. Then the paths Q, P_u, P_v form a cycle of length $2(\ell - i + p + 1) + 2(i - 1 - p) = 2\ell$, a contradiction. So this coloring is indeed t-periodic.

Now we can use our strange Lemma 7.4.4:

$$e(H_1) \leq 2tv(H_1) < 4\ell v(H_1).$$

Similarly we get for the connected components $H_2, \ldots, H_q$ that $e(H_i) \leq 4\ell v(H_i)$ $(i = 1, \ldots, q)$. Hence $e(H) < 4\ell v(H)$.

Let $H^* = G[V_{i-2} \cup V_{i-1}]$. Similarly to the previous case we get that $e(H^*) < 4\ell v(H^*)$. By the induction we get that

$$\frac{|V_{i-1}|}{|V_{i-2}|} \geq \frac{s}{5\ell}. \tag{7.2}$$

On the other hand,

$$e(H) + e(H^*) \geq s|V_{i-1}|$$

since the minimum degree is at least s. Whence

$$4\ell(|V_{i-1}| + |V_i| + |V_{i-2}| + |V_{i-1}|) = 4\ell(v(H) + v(H^*)) > e(H) + e(H^*) \geq s|V_{i-1}|.$$

Thus

$$|V_i| \geq \frac{1}{4\ell}((s - 8\ell)|V_{i-1}| - 4\ell|V_{i-2}|).$$

Now let us use inequality 7.2:

$$|V_i| \geq \frac{1}{4\ell}\left((s - 8\ell) - \frac{20\ell^2}{s}\right)|V_{i-1}|.$$

Since $s \geq 50\ell$ we have

$$\frac{|V_i|}{|V_{i-1}|} \geq \frac{1}{4\ell}(s - 9\ell) > \frac{1}{4\ell}\frac{4s}{5} = \frac{s}{5\ell}.$$

This proves the claim of the lemma. □

Now we are ready to prove Theorem 7.4.2. For the sake of convenience we repeat the statement.

Theorem 7.4.2 *Let $E = e(G_n)$. Then $C_{2\ell} \subseteq G_n$ for all $\ell \geq 2$ satisfying the inequalities $\ell \leq \frac{E}{100n}$ and $\ell n^{1/\ell} \leq \frac{E}{10n}$.*

Proof. We prove the statement by induction on n. For $n = 1$ the conditions cannot be satisfied, so the claim is trivially true.

Let us use Lemma 7.3.1: there exists a bipartite subgraph $H_n \subseteq G_n$ such that $e(H_n) \geq e(G_n)/2$ and all degrees in H are at least the half of the corresponding degree in G. If all degree in H_n is at least $\frac{E}{2n}$, then

$$\max(5\ell n^{1/\ell}, 50\ell) \leq \frac{E}{2n}$$

implies that $C_{2\ell} \subseteq H_n$ according to Lemma 7.4.5. Hence $C_{2\ell} \subseteq G_n$.

If there exists a vertex $w \in V(H_n)$ whose degree in H_n is smaller than $\frac{E}{2n}$ then the degree of this vertex in G_n is smaller than $\frac{E}{n}$. Let $G_{n-1} = G_n - w$. We will show that G_{n-1} satisfies the two hypotheses of the induction:

$$\ell \leq \frac{e(G_{n-1})}{100(n-1)} \quad \text{and} \quad \ell(n-1)^{1/\ell} \leq \frac{e(G_{n-1})}{10(n-1)}.$$

Then $C_{2\ell} \subseteq G_{n-1}$ and of course, $C_{2\ell} \subseteq G_n$. Next we show that the hypotheses of the induction are indeed satisfied. Since $\ell \leq \frac{e(G_n)}{100n}$ thus

$$\ell \leq \frac{e(G_n)}{100n} = \frac{e(G_n) - \frac{e(G_n)}{n}}{100(n-1)} \leq \frac{e(G_{n-1})}{100(n-1)}.$$

(Since the degree of w is smaller than $e(G_n)/n$.) On the other hand, because of $\ell n^{1/\ell} \leq \frac{e(G_n)}{10n}$ we have

$$\ell(n-1)^{1/\ell} \leq \ell n^{1/\ell} \leq \frac{e(G_n)}{10n} = \frac{e(G_n) - \frac{e(G_n)}{n}}{10(n-1)} \leq \frac{e(G_{n-1})}{10(n-1)}.$$

This proves that the hypotheses of the induction are satisfied. This proves the theorem. □

7.5 Construction for graphs without cycles of fixed length

In this section we show that for $k = 2, 3, 5$ the Bondy–Simonovits theorem is tight up to constant.

Theorem 7.5.1. *For the values $k = 2, 3, 5$ and for a prime power q there exists a graph G_n on $n = 2q^k$ with at least q^{k+1} edges that does not contain a cycle C_{2k}.*

Proof. (Wenger, Conlon) Let us consider the following bipartite graph $G = (P, L, E)$ that we will call $D_k(q)$. Let $P = \{\underline{x} \in \mathbb{F}_q^k\}$. For $\underline{u} \in \mathbb{F}_q^{k-1}$ and $z \in \mathbb{F}_q$ let us consider the line

$$\ell_{\underline{u},z} = \{(0, \underline{u}) + y(1, z, z^2, \dots, z^{k-1}) \mid y \in \mathbb{F}_q\}.$$

Let $L = \{\ell_{\underline{u},z} \mid \underline{u} \in \mathbb{F}_q^{k-1}, z \in \mathbb{F}_q\}$. Finally, let $E = \{(p, \ell) \mid p \in \ell\}$. Clearly, $|P| = |L| = q^k$ and $|E| = q^{k+1}$.

We call two lines $\ell_{\underline{u}_1,z_1}$ and $\ell_{\underline{u}_2,z_2}$ parallel if $z_1 = z_2$ but $\underline{u}_1 \neq \underline{u}_2$. Observe that parallel lines have no intersection since if we know $p = (0, \underline{u}) + y(1, z, \dots, z^{k-1})$ and z, then we immediately know y from the first coordinate, and then this determine $\underline{u}$.

Let us consider a cycle of length $2t$ in the graph G, let it be $p_1\ell_1p_2\ell_2 \dots p_t\ell_t(p_1)$. First of all, observe that ℓ_i and ℓ_{i+1} cannot be parallel, because parallel lines have no intersection.

On the other hand, we claim that for any ℓ_i there is an ℓ_j parallel with it. The reason is the following: let $\ell_i = \ell_{\underline{u}_i,z_i}$, then since $p_{i+1}, p_i \in \ell_i$ we have

$$\begin{aligned} p_{i+1} - p_i =& ((0, \underline{u}_i) + y_i(1, z_i, z_i^2, \dots, z_i^{k-1})) - ((0, \underline{u}_i) + y_{i+1}(1, z_i, z_i^2, \dots, z_i^{k-1})) \\ =& (y_{i+1} - y_i)(1, z_i, z_i^2, \dots, z_i^{k-1}). \end{aligned}$$

Let $a_i = y_{i+1} - y_i \neq 0$, then with the notation $p_{k+1} = p_1$ we have

$$\underline{0} = \sum_{i=1}^{k} (p_{i+1} - p_i) = \sum_{i=1}^{k} a_i(1, z_i, z_i^2, \dots, z_i^{k-1}).$$

For different $z_1, \dots, z_r$ where $r \leq k$ the vectors $(1, z_i, z_i^2, \dots, z_i^{k-1})$ are linearly independent, so the above sum can only be $\underline{0}$ if for each i there are some $j_1, \dots, j_s$ such that $z_i = z_{j_1} = \dots = z_{j_s}$ and $a_i + a_{j_1} + \dots + a_{j_s} = 0$. This means that for each ℓ_i there is at least one other line that is parallel with it.

In case of $k = 2$ and $k = 3$ we are immediately done since it cannot occur that in $\ell_1\ell_2(\ell_1)$ or $\ell_1\ell_2\ell_3(\ell_1)$ we have no two neighboring indices with parallel lines, but for each line we have another line that is parallel with it. For $k = 5$ this is also a contradiction, because the parallel classes would give a proper coloring of the five cycle $\ell_1\ell_2\ell_3\ell_4\ell_5(\ell_1)$ without a singleton class, but there is no such coloring since you cannot color a C_5 with two colors, but with more than two colors there will be a singleton class.

Hence $D_k(q)$ does not contain a cycle of length $2k$. (It is also true that $D_3(q)$ does not contain a cycle of length 4, and $D_5(q)$ does not contain a cycle of length 4 or 6.) □

8. Schwartz–Zippel Lemma

8.1 Introduction

In this chapter we prove the so-called Schwartz–Zippel lemma and see a typical application of it.

Theorem 8.1.1 (Schwartz–Zippel)**.** *Let $\mathbb{F}$ be an arbitrary field. Let S be a finite subset of $\mathbb{F}$. Suppose that $p(x_1, \ldots, x_m)$ is a polynomial of degree d with coefficients from $\mathbb{F}$. Then the number of $(s_1, \ldots, s_m) \in S^m$ with $p(s_1, \ldots, s_m) = 0$ is at most $d|S|^{m-1}$. In other words, if we choose $s_1, \ldots, s_m \in S$ independently and uniformly at random, then the probability that $p(s_1, \ldots, s_m) = 0$ is at most $\frac{d}{|S|}$.*

Proof. We prove the claim by induction on m. For $m = 1$ the statement claims that a univariate degree d polynomial has at most d zeros, this is well-known. Now suppose that $m > 1$. Let us write $p(x_1, \ldots, x_m)$ in the following form:

$$p(x_1, \ldots, x_m) = \sum_{j=0}^{k} p_j(x_1, \ldots, x_{m-1}) x_m^j,$$

where $k = \deg_{x_m} p$. Note that $\deg p_k(x_1, \ldots, x_{m-1}) = d - k$. Let

$$S_0 = \{(s_1, \ldots, s_{m-1}) \mid s_i \in S, p_k(s_1, \ldots, s_{m-1}) = 0\},$$

and

$$S_1 = \{(s_1, \ldots, s_{m-1}) \mid s_i \in S, p_k(s_1, \ldots, s_{m-1}) \neq 0\},$$

By induction on m we have $|S_0| \leq (d-k)|S|^{m-2}$. If $(s_1, \ldots, s_{m-1}) \in S_1$, then the polynomial

$$p(s_1, \ldots, s_{m-1}, x_m) = \sum_{j=0}^{k} p_j(s_1, \ldots, s_{m-1}) x_m^j,$$

has at most k solutions. Hence the number of solutions of $p(s_1, \ldots, s_m) = 0$ with $s_1, \ldots, s_m \in S$ is at most $|S_0| \cdot |S| + |S_1|k \leq (d-k)|S|^{m-2} \cdot |S| + |S|^{m-1} \cdot k = d|S|^{m-1}$. We are done.

□

8.2 Perfect matchings in bipartite graphs

In this section we show how one can use the Schwartz-Zippel lemma to decide whether a bipartite graph contains a perfect matching. Suppose that $G = (A, B, E)$ is a bipartite graph such that $|A| = |B| = n$. For sake of simplicity we assume that the elements of A and B are labelled by the elements of $\{1, 2, \ldots, n\}$. Let us introduce the matrix R of size $n \times n$ as follows: $R_{ij} = x_{ij}$ if $i \in A$ and $j \in B$ are adjacent, and $R_{ij} = 0$ if they are not adjacent. Here x_{ij} is just a variable. Note that if G does not contain a perfect matching, then $\det(S) = 0$. If it contains a perfect matching, say M, then nothing cancels the term $(-1)^s \prod_{(i,j)\in M} x_{ij}$ in the expansion of $\det(S)$. In this case $\det(S)$ is a multivariate polynomial of degree n. Note that we cannot use Gauss elimination to a matrix containing variables (why?), but we can do the following: we randomly substitute elements of S into x_{ij} and check whether the determinant is non-zero or not. If $\det(R) \neq 0$, then the probability that after the evaluation the result is 0 is at most $\frac{n}{|S|}$. So choose a set S of size $4n$ and do the following algorithm: pick random elements of S and evaluate $\det(R)$. If it is non-zero, then G has a perfect matching. If it is 0, then output that it has no perfect matching. The probability that the algorithm errs, that is, it has a perfect matching, is at most $1/4$. Iterating this process t times the probability that the algorithm errs is at most $1/4^t$.

One detail that might be interesting is that it is worth choosing the set S in a finite field $\mathbb{F}_p$. This way we can save the trouble with counting with fractions or with large numbers. So we choose a prime p bigger than $4n$, and we can even choose S to be the whole $\mathbb{F}_p$.

9. Perfect matchings in planar graphs and grids

9.1 Planar graphs

In this section we show a fast method to determine the number of perfect matchings of planar graphs. We will concentrate to bipartite graphs, but a slight modification of the method is able to deal with non-bipartite graphs too.

Let $G = (X, Y, E)$ be a bipartite graph with parts $|X| = |Y| = n$, and let B be the following $n \times n$ matrix:

$$B_{uv} = \begin{cases} 1 & \text{if } (u,v) \in E(G), \\ 0 & \text{if } (u,v) \notin E(G) \end{cases}$$

for $u \in X, v \in Y$. This matrix is called the bipartite adjacency matrix or incidence matrix. Then the permanent

$$\operatorname{per}(B) = \sum_{\pi \in S_n} \prod_{j=1}^{n} B_{j,\pi(j)}$$

counts the number of perfect matchings of G. The problem with the permanent is that –unlike the determinant– in general there is no fast way to compute it. The idea is that maybe we can put $\pm$ signs into S such that for the obtained signed matrix B^σ we have $\operatorname{per}(B) = |\det(B^\sigma)|$. Honestly, we should regard such a signing a miracle since $n!$ terms should have the same sign in the determinant. We will show that for planar graphs such a miracle happens. We show it through a series of lemma.

Definition 9.1.1. We say that a cycle C of G is evenly placed if $G \setminus C$ contains a perfect matching.

Lemma 9.1.2. *Let G be a bipartite graph. Let $\sigma : E(G) \to \{-1, 1\}$ such that every evenly placed cycle of length $4k$ contains an odd number of negative edges, and every*

evenly placed cycle of length $4k+2$ contains an even number of negative edges. Then $\operatorname{per}(B) = |\det(B^\sigma)|$*, where B and B^σ are the matrices defined above.*

Proof. Every perfect matching corresponds to a permutation. Let π and ρ be two permutations. We need that $\operatorname{sgn}(\pi)\prod_{i=1}^n B_{k,\pi(k)}$ and $\operatorname{sgn}(\rho)\prod_{i=1}^n B_{k,\rho(k)}$ have the same sign. Note that the union of the two perfect matchings can be decomposed to edges and even cycles. These even cycles are trivially evenly placed. A cycle of length $2t$ correspond to a cycle of length t in the permutation $\pi\rho^{-1}$. Since $\operatorname{sgn}(\pi)\operatorname{sgn}(\rho) = \operatorname{sgn}(\pi\rho^{-1})$ we observe that a cycle of length $4k$ in the graph G will give a negative sign for $\prod_{i=1}^n B_{k,\pi(k)}\prod_{i=1}^n B_{k,\rho(k)}$ and also a negative sign in $\operatorname{sgn}(\pi\rho^{-1})$, while a cycle of length $4+2$ in the graph G will give a positive sign for $\prod_{i=1}^n B_{k,\pi(k)}\prod_{i=1}^n B_{k,\rho(k)}$ and also a positive sign for $\operatorname{sgn}(\pi\rho^{-1})$. Hence $\operatorname{sgn}(\pi)\prod_{i=1}^n B_{k,\pi(k)}$ and $\operatorname{sgn}(\rho)\prod_{i=1}^n B_{k,\rho(k)}$ have the same sign.

□

While the above lemma gives a sufficient condition for a good signing, a tiny problem is that a graph generally contains a lot of cycles, and it is not easy to check the condition of the lemma for all (evenly placed) cycles. Fortunately, for planar graphs it is enough to check it for cycles around faces. We call a cycle surrounding an inner face a boundary cycle.

Lemma 9.1.3. *Let G be a bipartite graph. Let $\sigma : E(G) \to \{-1, 1\}$ such that every boundary cycle of length $4k$ contains an odd number of negative edges, and every boundary cycle of length $4k+2$ contains an even number of negative edges. Then every evenly placed cycle of length $4k$ contains an odd number of negative edges, and every evenly placed cycle of length $4k+2$ contains an even number of negative edges.*

Proof. Suppose that C is an evenly placed cycle of length 2ℓ. Let $F_1, F_2, \ldots, F_k$ be the faces inside C. Its boundary cycles are $C_1, \ldots, C_k$. Let $2\ell_j$ be the length of the cycle C_j. Let n_C and n_{C_j} be the number of negative edges along C and C_k respectively. Then we know that n_{C_k} is congruent with $\ell_k - 1$ modulo 2. Then the number of negative edges modulo 2 around C is

$$n_C \equiv \sum_{j=1}^{k} n_{C_j} \equiv \sum_{i=1}^{k} (\ell_i - 1) \pmod 2$$

since we count every edges not in C twice. On the other hand, the number of edges of the graph determined by C and $F_1, F_2, \ldots, F_k$ is

$$\frac{1}{2}(2\ell + 2\ell_1 + \cdots + 2\ell_k) = \ell + \ell_1 + \cdots + \ell_k.$$

If there are r points inside C, then the number of vertices is $2\ell + r$. Then by Euler's formula we have

$$\ell + \ell_1 + \cdots + \ell_k = (k+1) + (r + 2\ell) - 2.$$

Hence

$$\ell - 1 + r = \sum_{i=1}^{k} (\ell_i - 1).$$

Now observe that r is even as the cycle is evenly placed. So

$$n_C \equiv \sum_{i=1}^{k} (\ell_i - 1) \equiv \ell - 1 \pmod 2.$$

This is exactly what we wanted to prove. □

Now we are ready to prove that every bipartite planar graph has a proper edge signing (so-called Kasteleyn signing).

Theorem 9.1.4. *Every planar bipartite graph has a signing σ such that every boundary cycle of length $4k$ contains an odd number of negative edges, and every boundary cycle of length $4k+2$ contains an even number of negative edges.*

Proof. We prove the statement of the theorem by induction on the number of faces. If there is only one face, then the statement is trivial since there is no cycle in the graph. Suppose that we have at least two faces. Let F_{outer} be the outer face and let F be a face that has at least one common edge with it. Let E' be the set of common edges of F_{outer} and F. Delete E' from G, the obtained graph has fewer faces so by induction there is a proper signing of it. Now add back E' and sign them in such a way that F is also properly signed. Note that it cannot ruin the signing of other bounded faces. We are done. □

9.2 Kasteleyn's theorem

Theorem 9.2.1 (Kasteleyn and independently Fisher and Temperley)**.** *Let $Z_{m,n}$ be the number of perfect matchings of the grid of size $m \times n$. Then*

$$Z_{m,n} = \left(\prod_{j=1}^{m} \prod_{k=1}^{n} \left(4\cos^2\left(\frac{\pi j}{m+1}\right) + 4\cos^2\left(\frac{\pi k}{n+1}\right)\right)\right)^{1/4}.$$

Remark 9.2.2. We could easily construct the proper signing described in the previous section: put a negative sign to the edges in every second row. Nevertheless we will use a slightly different strategy: we use complex numbers!

Proof. Note that the grid is a bipartite graph, so we can color the vertices of the grid by black and white such that only vertices of different color are adjacent. Let S be the bipartite adjacency matrix) of the graph: $S_{ij} = 1$ if black and white vertices b_i and w_j are adjacent, and 0 otherwise. Then the number of perfect matchings is exactly $\mathrm{per}(S)$, the permanent of S. We will give a "signing" σ of S such that $\mathrm{per}(S) = |\det(S^\sigma)|$.

Let $S^\sigma_{(x,y),(x,y\pm1)} = i$ and $S^\sigma_{(x,y),(x\pm1,y)} = 1$, and 0 otherwise. We claim that $\mathrm{per}(S) = |\det(S^\sigma)|$. One way to see it is the following: from any perfect matching M_1 we can arrive to any other perfect matching M_2 by a sequence of moves of the following type: choose two edges of the form $e = ((x,y),(x+1,y)), f = ((x,y+1),(x+1,y+1)$ and replace them by $e' = ((x,y),(x,y+1)), f' = ((x+1,y),(x+1,y+1)$, or do the reverse of this operation (why?). In $\det(S^\sigma)$ this operation does the following thing: the sign of th corresponding permutation changes because we did a transposition, but also the weight of the perfect matching changes since we changed two edges of weight 1 to two edges of weight i or vice versa. So every expansion term of $\det(S^\sigma)$ corresponding to a perfect matching will give the same quantity, hence $\mathrm{per}(S) = |\det(S^\sigma)|$.

Next we will compute $\det(S^\sigma)$. It will be more convenient to work with the matrix

$$A = \begin{pmatrix} 0 & S^\sigma \\ (S^\sigma)^T & 0 \end{pmatrix}.$$

Clearly, $\det(A) = \det(S^\sigma)^2$. It turns out that we can give all eigenvectors and eigenvalues explicitly. Note that the vector consisting of the values $f(x,y)$ is an eigenvector of A belonging to the eigenvalue λ if

$$\lambda f(x,y) = f(x+1,y) + f(x-1,y) + if(x,y+1) + if(x,y-1),$$

where $f(r,t) = 0$ if $r \in \{0, m+1\}$ or $t \in \{0, n+1\}$. Let $1 \le j \le m$, $1 \le k \le n$, and $z = e^{\pi i \frac{j}{m+1}}$ and $w = e^{\pi i \frac{k}{n+1}}$. Let us consider the vector $f_{j,k}$ defined as follows:

$$f_{j,k}(x,y) = (z^x - z^{-x})(w^y - w^{-y}) = -4\sin\left(\frac{\pi jx}{m+1}\right)\sin\left(\frac{\pi ky}{n+1}\right).$$

Then with $\lambda_{j,k} = z + \frac{1}{z} + i\left(w + \frac{1}{w}\right)$ we have

$$\lambda_{j,k} f_{j,k}(x,y) = f_{j,k}(x+1,y) + f_{j,k}(x-1,y) + i f_{j,k}(x,y+1) + i f_{j,k}(x,y-1).$$

Indeed,

$$f_{j,k}(x+1,y)+f_{j,k}(x-1,y) = (z^{x+1}-z^{-x-1})(w^y-w^{-y})+(z^{x-1}-z^{-x+1})(w^y-w^{-y}) =$$

$$= (z+z^{-1})(z^x-z^{-x})(w^y-w^{-y}) = (z+z^{-1})f_{j,k}(x,y),$$

and

$$if_{j,k}(x,y+1)+if_{j,k}(x,y-1) = i((z^x-z^{-x})(w^{y+1}-w^{-y-1})+(z^x-z^{-x})(w^{y-1}-w^{-y+1}) =$$

$$= i(w+w^{-1})(z^x-z^{-x})(w^y-w^{-y}) = i(w+w^{-1})f_{j,k}(x,y),$$

It is easy to see that the vectors $f_{j,k}$ are pairwise orthogonal to each other, consequently they are linearly independent. Since A has nm eigenvalues, we have found all of them. Note that

$$\lambda_{j,k} = 2\cos\left(\frac{\pi j}{m+1}\right) + i2\cos\left(\frac{\pi k}{n+1}\right).$$

Hence

$$Z_{m,n} = \left(\prod_{j=1}^{m}\prod_{j=1}^{n}\lambda_{j,k}\right)^{1/2} = \left(\prod_{j=1}^{m}\prod_{j=1}^{n}|\lambda_{j,k}|^2\right)^{1/4} =$$

$$= \left(\prod_{j=1}^{m}\prod_{k=1}^{n}\left(4\cos^2\left(\frac{\pi j}{m+1}\right)+4\cos^2\left(\frac{\pi k}{n+1}\right)\right)\right)^{1/4}.$$

□

Corollary 9.2.3. *We have*

$$\lim_{m,n\to\infty}\frac{1}{mn}\log Z_{m,n} = \frac{4}{\pi^2}\int_0^{\pi/2}\int_0^{\pi/2}\log(4\cos^2(x)+4\cos^2(y))\,dx\,dy.$$

Remark 9.2.4. Surprisingly, there is a nice expression for the above integral:

$$\frac{4}{\pi^2}\int_0^{\pi/2}\int_0^{\pi/2}\log(4\cos^2(x)+4\cos^2(y))\,dx\,dy = \frac{1}{\pi}\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)^2} \approx 0.2915.$$

Beyond this course 9.2.5. Instead of rectangles one can consider other shapes of the grid graphs. One particularly interesting choice is the so-called Aztec diamond (see below).

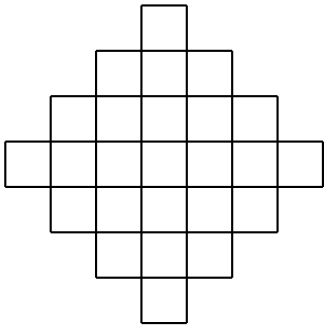


Figure 9.1: Aztec diamond of size 4.

The Aztec diamond of size n denoted by A_n has exactly $2^{n(n+1)/2}$ perfect matchings. The value $\frac{1}{|A_n|} \ln \mathrm{pm}(A_n) = \frac{1}{4} \ln(2) \approx 0.1732$ is smaller than the value we got for rectangles. The reason is that the boundary of the Aztec diamond effects heavily to a random perfect matching. This statement will be more clear if we start to add some colors to the vertices.

Given a perfect matching (denoted with blue edges) we can color the vertices of the Aztec diamond with blue, red, green, yellow as follows. First we color the vertices with black and white corresponding to the bipartite class. Then we color the end vertices of horizontal edges of the perfect matching to blue if the rightmost vertex is black, and to red if the rightmost vertex is white. Similarly, we color the end vertices of vertical edges of the perfect matching to green if the bottom vertex is black, and to yellow if the bottom vertex is white. What we see in the picture below that in a random perfect matching the different corners get different colors. This means that the perfect matching configuration seems to be "frozen" close to the corners.

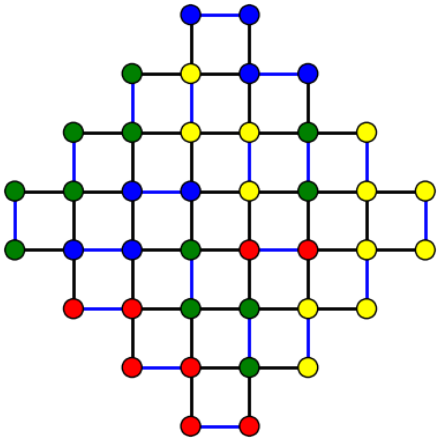


Figure 9.2: Aztec diamond of size 4 with a random perfect matching and the corresponding coloring.

Let use see what happens if we take an Aztec diamond of size 100. Here we only colored the pixels as the drawing of the graph would be too ugly.

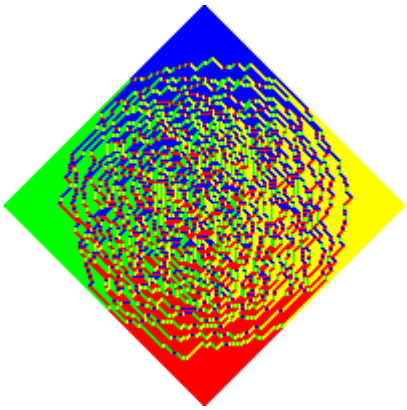


Figure 9.3: Aztec diamond of size 100 with a random perfect matching and the corresponding coloring.

From this picture one can even make a guess what will be the shape of the non-frozen area. Yes, it is a circle! This is the so-called arctic circle theorem.

Bibliography

[1] Noga Alon and Joel H Spencer. *The probabilistic method*. John Wiley & Sons, 2016.

[2] Béla Bollobás. *Extremal graph theory*. Courier Corporation, 2004.

[3] John A Bondy and Miklós Simonovits. "Cycles of even length in graphs". In: *Journal of Combinatorial Theory, Series B* 16.2 (1974), pp. 97–105.

[4] Paul Erdös. "Graph theory and probability". In: *Canadian Journal of Mathematics* 11 (1959), pp. 34–38.