1 Meatspace Stuff

Homework possibly. No exam.

Book: Enumerative and Asymptotic Combinatorics Other recommended books: Stanley's books, Odlyzko's Asymptotic Enumeration Methods (available online), and Flajolet and Sedgewick's Asymptotic Combinatorics (available online).

Onto the Good Stuff

2 Catalan Numbers

Lattice path definition, weakly above x-axis. Formula $C_n = \frac{1}{n+1}\binom{2n}{n}$, proof by reflection method. Consider all paths from $(0,0)$ to $(2n,0)$ of which there are $\binom{2n}{n}$. If such a path hits $(x,-1)$ at any point, reflect the continuation of the path across $y=-1$. Gives a path from $(0,0)$ to $(2n,-2)$, of which there are $\binom{2n}{n-1}$. Therefore $C_n = \binom{2n}{n} - \binom{2n}{n-1}$ and we're done. Asymptotically from Stirling's formula,

$$C_n \sim \frac{c}{n^{3/2}} 4^n. \tag{1}$$

2.1 Plane Trees

Also counted by the Catalan numbers, C_n for an $(n+1)$-vertex tree.

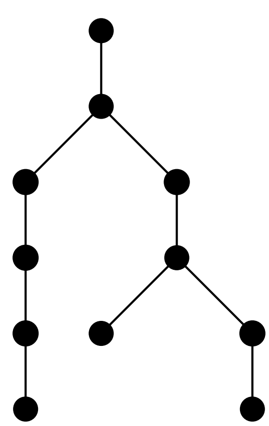


Figure 1: A plane tree.

Let D be the degree of the root in a random plane tree. What's the probability that $D = 1$? Answer is approx. $1/4$, ratio of two successive Catalan numbers. What is the expected value of D? Turns out it actually converges to 3 as $n \to \infty$.

Let's see why. There is the standard bijection from plane trees into Catalan lattice paths using a depth-first search (or preorder traversal). The degree of the root translates into the number of contact points with the x-axis. So at the very least, we doubt this happens a lot, so we expect that $E[D]$ will end up being small or finite.

We're also going to need binary trees. Another bijection exists from binary trees on n vertices to plane trees on $n+1$ vertices (didn't quite catch what he did). Under this bijection, our random variable D becomes Q, the length of the left path of the binary tree. We can use symmetry to help us: if Q' is the length of the right path, clearly Q and Q' have the same expectation. But Q' translates into A in plane trees: A is the length of the rightmost path in the plane tree, minus 1.

In Catalan lattice paths, A translates into the length of the final strict descent of the path, which we write as the variable B. By symmetry, consider the length of the *initial* ascent instead. The first step must be up. The second will be up or down with probability (roughly) $1/2$. The first upward steps all have roughly probability $1/2$, so we see at least that we're going to get a constant for the expectation. You need to be more careful than we are being to get the answer of 3.

$P(D_n = r) = \frac{r}{2^{n+1}}$, as the answer for reference.

3 Probability

What's the probability that a labeled graph on n vertices is connected? It goes to 1 as $n \to \infty$.

Proof. By considering small ways we could break the connectedness of the graph, we see that the number of labeled graphs is bigger than

$$2^{\binom{n}{2}}\left(1 - \frac{n}{2^{n-1}} - \frac{\binom{n}{2}}{2^{2n-4}}\right) = 2^{\binom{n}{2}}(1 - o(1)), \tag{2}$$

since the numerators are polynomial and the denominators are exponential. $\square$

Let L_n be the number of non-isomorphic connected graphs on n vertices. Then we claim

$$L_n \sim \frac{2^{\binom{n}{2}}}{n!}. \tag{3}$$

Turns out that the random graph has with high probability a trivial automorphism group, and this can be used to prove the claim.

1 Partitions

Integer partitions λ of n. Let $p(n)$ be the number of distinct partitions of n.

Theorem 1.1 (Euler) *If $P(t)$ is the generating function*

$$P(t) = 1 + \sum_{n=1}^{\infty} p(n) t^n \tag{1}$$

then we have

$$P(t) = \prod_{i=1}^{\infty} \frac{1}{1 - t^i} \tag{2}$$

Proof: Nothing too scary. Helps to write a partition as $\lambda = 1^{m_1} 2^{m_2} \cdots$ in order to see what's going on. $\square$

Asymptotic perspective: what are the asymptotics of $p(n)$? We will start with easy bounds, and work our way toward harder bounds. Today we will eventually prove

Theorem 1.2 *There is an α and a β such that*

$$e^{\alpha\sqrt{n}} < p(n) < e^{\beta\sqrt{n}}. \tag{3}$$

1.1 Algorithm for $p(n)$

If someone asks me to calculate $p(17522)$, how could I do it fast? A good idea is, take $(\log P(t))'$ to find a recurrence: from theorem 1.1 we have

$$t \cdot \frac{P'(t)}{P(t)} = \sum_{m=1}^{\infty} \frac{m \cdot t^m}{1 - t^m}. \tag{4}$$

Therefore we have $tP'(t) = A(t) \cdot P(t)$, where $A(t)$ is the above sum. Compare coefficients: on the left hand side we have $n \cdot p(n)$, whereas on the right hand side we have

$$\sum_{k=0}^{n} \sigma(k) \cdot p(n-k), \tag{5}$$

where

$$A(t) = \sum_{k} \sigma(k) \cdot t^k = \sum_{m=1}^{\infty} \frac{m \cdot t^m}{1 - t^m}, \tag{6}$$

and it is clear that $\sigma(k)$ is the sum of divisors of k.

This gives a $O(n^2)$ algorithm for $p(n)$. Igor disagrees with this: because the number of digits is going to be $\sqrt{n}$. Additionally, the fact that

$$\sigma(n) \le e^{\gamma} n \log\log n, \tag{7}$$

where γ is the Euler-Mascheroni constant is useful (?). Calculating $\sigma(k)$ is hard, because if $k = pq$ then we have to recover the primes p, q. So Igor views this claim as somewhat bogus.

1.2 Distinct parts

Theorem 1.3 *Let $d(n)$ be the number of partitions of n with distinct parts. Then*

$$\mathcal{D}(t) = 1 + \sum_{k=1}^{\infty} d(n)t^n = \prod_{i=1}^{\infty}(1+t^i). \tag{8}$$

Same proof as theorem 1.1.

Theorem 1.4 (Euler Pentagonal)

$$P(t)^{-1} = \sum_{m=-\infty}^{\infty} (-1)^m t^{\frac{m(3m\pm 1)}{2}} \tag{9}$$

Proof:[Proof (by Franklin)] We have

$$P(t)^{-1} = \prod_{i=1}^{\infty}(1-t^i), \text{ has coefficients } d_e(n) - d_o(n), \tag{10}$$

where $d_e(n)$ is the number of partitions of n into distinct parts with an even number of parts, and $d_o(n)$ the same for odd number of parts. Suffices then to prove that

$$d_e(n) - d_o(n) = (-1)^m \tag{11}$$

whenever $n = \frac{m(3m\pm 1)}{2}$, and 0 otherwise.

This suggests an involutive proof. Given a partition λ, compare the smallest part and the largest decreasing subsequence along the top right diagonal, with size a, b respectively. If $a > b$, pull the diagonal off and place it down as a smaller part. If $a < b$, do the reverse, and fix λ if $a = b$. It's easy to see that this is indeed an involution, and of course we immediately see that it changes the parity of number of parts whenever $a \neq b$. When $a = b$, some analysis gives that n is a pentagonal number by staring at □

[scale=2] (0,0) rectangle (1,1); (1,1) – (2,1) – (1,0); (0.9,-0.1) rectangle (1.1,0.1);

Figure 1: Case when $a = b$.

Corollary 1.1 *We have*

$$p(n) = p(n-1) + p(n-2) - p(n-3) - p(n-5) + \cdots \tag{12}$$

where the subtractands are the pentagonal numbers.

This gives a **much** better algorithm for computing $p(n)$, as the pentagonal numbers are sparse. Igor claims $O(n^{3/2})$, but adds the caveat that writing down a $\sqrt{n}$-digit number really brings it up to $O(n^2)$ or such.

1.3 Back to asymptotics

Proposition 1.1 $p(n) > e^{\alpha\sqrt{n}}$.

Proof: Just look at a really small subset of partitions of n: take a $k \times k$ triangle partition, and assume k even. Let $n = \frac{k(k+1)}{2} + k/2$, and place the $k/2$ extra boxes at different points on the triangle This gives us

[scale=2] (0,0) – (0,1) – (1,1) – (0,0);

Figure 2: Triangle

$$p(n) > \binom{k}{k/2} \simeq \frac{c}{\sqrt{k}} 2^k, \tag{13}$$

with $k \sim \sqrt{2n}$. Thus,

$p(n) > (2^{\sqrt{2})^{\sqrt{n}}} \cdot \frac{c}{\sqrt{k}}.(14)$

$\square$

1 Partitions

On Monday we proved the product formula for the generating function for $p(n)$ and the lower bound for the asymptotics

$$e^{\alpha\sqrt{n}} < p(n) < e^{\beta\sqrt{n}}. \tag{1}$$

Definition 1. $p_k(n)$ is the number of integer partitions of n into k parts.

Lemma 1. $p_k(n) > \frac{n^{k-1}}{(k!)^2}$.

Proof. Let $n = \lambda_1 + \cdots + \lambda_k$. Then $k!p_k(n)$ is clearly greater than the number of k-compositions $n = a_1 + \cdots + a_k$ (strictly more because some numbers can be the same). We know that the number of k-compositions is just

$$\binom{n+k-1}{k-1} = \frac{(n-k-1)(n-k-2)\cdots(n+1)}{(k-1)!}, \tag{2}$$

and this now implies the lemma. □

Consider now when we take $k = \sqrt{n}$. Then we have

$$p_k(n) > \frac{n^{\sqrt{n}-1}}{n(\frac{\sqrt{n}}{e})^{2\sqrt{n}} \cdot (c\sqrt{n})^2}, \tag{3}$$

and some cancellation gives

$$p(n) > p_k(n) > \frac{e^{2\sqrt{n}}}{cn^2}, \tag{4}$$

which Igor just ignores some stuff to get

$$p(n) > (e^2)^{\sqrt{n}(1-O(1))}. \tag{5}$$

Now we get to the upper bound, but first a theorem.

Theorem 1 (Euler)**.**

$$P(t) = 1 + \sum_{k=1}^{\infty} \frac{(t^k)^2}{(1-t)^2(1-t^2)^2\cdots(1-t^k)^2} \tag{6}$$

Proof. Enumerate based on the Durfee square, and use symmetry to get the squares in the denominator. □

Let's examine the denominator more closely: we write

$$\frac{1}{(1-t)^2(1-t^2)^2\cdots(1-t^k)^2} = \sum q_k(n)t^n. \tag{7}$$

Observe that $q_k(n) = q_{k-1}(n) + 2q_{k-1}(n-k) + 3q_{k-1}(n-2k) + \cdots$ by thinking about chopping off the last $(1-t^k)^2$ term.

Claim 1. *We have*

$$q_k(n) \leq \frac{(n+k^2)^{2k-1}}{(2k-1)!(k!)^2}. \tag{8}$$

Proof of claim. By induction. □

Observe that from the above theorem by Euler,

$$p(n) = q_1(n-1) + q_2(n-4) + q_3(n-9) + \cdots, \tag{9}$$

and so we have

$$p(n) \leq \sum \frac{(n+k^2)^{2k-1}}{(2k-1)!(k!)^2}. \tag{10}$$

Use Stirling's formula to obtain the upper bound on $p(n)$.

"Boring proof over," now on to a different proof. This proof was the state of affairs until the 19th century.

Theorem 2.

$$p(n) \sim \frac{1}{4\sqrt{3}n} e^{\pi\sqrt{\frac{2}{3}n}}. \tag{11}$$

We won't prove this exactly, but we'll get close. In fact, Ramanujan came up with a formula $p(n) =$ something $+ o(1)$, but we'll also not go there.

Proof of easier version by Erdős. We prove that

$$p(n) \lesssim e^{\pi\sqrt{\frac{2}{3}n}}, \tag{12}$$

or

$$\log p(n) \leq \pi\sqrt{\frac{2}{3}n}(1 + O(1)). \tag{13}$$

Recall the formulas

$$np(n) = \sum_{r=1}^{n} r \sum_{m=1}^{\lfloor n/r \rfloor} p(n - mr) \tag{14}$$

$$np(n) = \sum_{k} \sigma(k) p(n-k), \tag{15}$$

with $\sigma(k) = \sum_{r|k} r$.

Assume that $p(k) < e^{c\sqrt{k}}$ with $c = \pi\sqrt{\frac{2}{3}}$ (proving by induction). Then by this assumption we have

$$np(n) < \sum_{rm<n} r e^{c\sqrt{n-mr}} < e^{c\sqrt{n}} \sum_{r=1}^{\infty} \sum_{m=1}^{\infty} r e^{(-cm/2\sqrt{n})r}, \tag{16}$$

with the last inequality coming from... actually we're not quite sure, maybe Taylor's theorem? Steven says Taylor's theorem. Now we have

$$\sum_{r=1}^{\infty} rt^r = \frac{t}{(1-t)^2}, \tag{17}$$

and then the (Maple verified!) equation

$$\frac{e^{-x}}{(1-e^{-x})^2} < \frac{1}{x^2}, \tag{18}$$

and we apply these to that big old sum in (16):

$$p(n) < e^{c\sqrt{n}} \sum \frac{e^{-cm/2\sqrt{n}}}{(1-e^{\square})^2} < e^{c\sqrt{n}} \sum_{m=1}^{\infty} \frac{4n}{c^2 m^2} = e^{c\sqrt{n}} \cdot \frac{4n}{c^2} \cdot \frac{\pi^2}{6}, \tag{19}$$

and plugging in our value $c = \pi\sqrt{\frac{2}{3}}$ gives us our desired result. $\square$

1 Partitions

Today, three short stories.

1.1 Arithmetic in partitions

Question: what is $p(n) \mod k$?

Theorem 1. *$p(n)$ is both even and odd infinitely often.*

Proof. Recall the pentagonal formula

$$p(n) = p(n-1) + p(n-2) - p(n-5) - p(n-7) + \cdots . \tag{1}$$

Draw a number line from 1 to n, and consider the pairs of numbers less than n by a pentagonal number, i.e. consider $n - \frac{m(3m\pm 1)}{2}$. Suppose that after some time everything is odd (past the green line). If we let a_m be the + and b_m the −

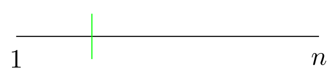


Figure 1: Suppose after some point all $p(n)$ are odd.

of those pentagonal numbers, we have $a_m - b_m = m$, and $b_{m+1} - a_m = 2m+1$. This means we can ensure that an even number of $n - a_m, n - b_m$ lie to the right of the green line, which would make $p(n)$ even, contradicting our hypothesis. Therefore $p(n)$ is even infinitely often.

For the even case, $\square$

Theorem 2. *$p(n)$ is not divisible by 3 infinitely often.*

Conjecture 3. *$p(n) = 1 \mod 3$, infinitely often.*

Igor's comment: it's a miracle that this is still not known.

Theorem 4 (Ramanujan)**.** *We have*

$$p(5n-1) = 0 \mod 5 \tag{2}$$

for all n, and

$$p(7n-1) = 0 \mod 7 \tag{3}$$

for all n.

It would take a whole lecture to prove these guys, so we're not gonna do it now.

Freeman Dyson introduced concept of the rank of a partition. The *rank* of λ is $r(\lambda) = \lambda_1 - \lambda'_1$.

Claim 5. *For $n = 5k-1$, let $p_i(n)$ be the number of partitions λ of n with rank $r(\lambda) = i \mod 5$. Then we have*

$$p_0(n) = p_1(n) = \cdots = p_4(n) \tag{4}$$

Note that it's easy to see that $p_1(n) = p_4(n)$ and $p_2(n) = p_3(n)$ by taking transposes. It took a 40 page paper to prove this, eventually.

1.2 Sampling random partitions

How can we sample them uniformly when calculating $p(n)$ is already tough? Igor reminds us of the formula

$$n \cdot p(n) = \sum \sigma(k)p(n-k) = \sum_{r=1}^{n} r \sum_{m=1}^{\lfloor n/r \rfloor} p(n-mr), \tag{5}$$

and one can somehow use this?

Two algorithms. Part 1: let $p_k(n)$ be the number of partitions of n with $\lambda_1 = k$. Then we have

$$p_k(n) = p_k(n-k) + p_{k-1}(n-k) + \cdots + p_1(n-k) \tag{6}$$

(visually, this corresponds to chopping off the first row of the partition and examining the size of the next part). Now we have

$$p(n) = p_1(n) + p_2(n) + \cdots + p_n(n), \tag{7}$$

So flip a coin that puts us in one of these cases with the proper probability. The top row of your partition is now whatever you've chosen there and it's time to recurse.

This algorithm requires us to precompute the numbers $p_k(n)$, but once we pick our k we are down to picking from the cases in equation (6). It's very easy to implement, but it has some big disadvantages.

Part 2: Fristedt's algorithm. We do the following:

1. Input n, choose q "wisely" (Steven says $q = e^{-c/\sqrt{n}}$).
2. For each $i = 1, \ldots, n$, let m_i be a random variable with distribution $Geo(1-q^i)$ (**important**: you need to verify that these are independent, as Steven points out).
3. Let $N = 1m_1 + 2m_2 + \cdots + nm_n$.
4. If $N = n$, then output $\lambda = 1^{m_1}2^{m_2}\cdots n^{m_n}$. If not, repeat and hope for better luck.

Follow up on the independence of geometric distributions: we have $P(X = k) = p(1-p)^k$ with $k = 0, 1, 2, \ldots$, and these do sum to 1 for what it's worth (?!?).

Theorem 6. *This algorithm generates uniform random partitions of n.*

Proof.

$$
\begin{aligned}
P(\text{output} = \lambda) &= (1-q)q^{1\cdot m_1}(1-q^2)q^{2\cdot m_2}\cdots && (8)\\
&= (1-q)(1-q^2)\cdots(1-q^n)q^{|\lambda|}, && (9)
\end{aligned}
$$

and since $|\lambda| = n$ the probability is the same for every λ, so it is uniformly distributed. □

Fristedt's algorithm says that the number of parts of size 3 is mostly independent of the number of parts of size 15, etc. This is a nice insight and is part of how Steven got his optimal algorithm (proven optimal up to a constant, "and the constant is root 2!").

1 More symmetric group

We recall yesterday's lemmas:

Lemma 1. *Let $X(\sigma)$ be the number of ℓ-cycles in σ. Then*

$$E(X) = \frac{1}{\ell}. \tag{1}$$

Lemma 2. *Suppose $\ell \neq m$ and $\ell + m \leq n$. Then*

$$E[\#\ell\text{-cycles} \cdot \#m\text{-cycles}] = \frac{1}{\ell m}. \tag{2}$$

Lemma 3. *We have*

$$E[(\#\ell\text{-cycles})^2] = \begin{cases} 1/\ell & 2\ell > n, \\ 1/\ell + 1/\ell^2 & 2\ell \leq n. \end{cases} \tag{3}$$

We also had a partially proved main lemma, and today we complete the proof.

Claim 4.

$$E[Z] = V[Z] = \sum_{i=1}^{n} \frac{1}{a_i}. \tag{4}$$

Proof. We have

$$E[Z] = \sum_{i=1}^{k} E[\#a_i\text{-cycles}] = \sum_{i=1}^{k} \frac{1}{a_i} \tag{5}$$

and then

$$\begin{aligned} V[Z] = E[Z^2] - E[Z]^2 =& 2\sum_{i<j} E[\#a_i\text{-cycles} \times \#a_j\text{-cycles}] \\ &+ \sum_{i=1}^{k} E[(\#a_i\text{-cycles})^2] - E[Z]^2 \qquad (6) \\ \leq& 2\sum_{i<j} \frac{1}{a_i a_j} + \sum_{i=1}^{k} \left(\frac{1}{a_i} + \frac{1}{a_i^2}\right) - E[Z]^2 \qquad (7) \\ =& \sum_{i=1}^{k} \frac{1}{a_i}. \qquad (8) \end{aligned}$$

□

This finishes the proof of the main lemma that we started yesterday.

Question: what's the probability than σ has no even cycles? Well, probability that 1 is in an odd cycle is roughly $1/2$, and the same for the next element not in said cycle, etc. Gives us that P is roughly

$$\left(\frac{1}{2}\right)^{\log n} \sim \frac{c}{n^{\alpha}}. \tag{9}$$

We can get more exact than this though.

Lemma 5. *Let p be a positive integer. Then*

$$P(\sigma \text{ has no cycles dividing } p) = \prod_{i=1}^{\lfloor n/p \rfloor} \left(1 - \frac{1}{p \cdot i}\right) \tag{10}$$

Thus for $p = 2$ and n even we have

$$\left(1-\frac{1}{2}\right)\left(1-\frac{1}{4}\right)\cdots\left(1-\frac{1}{n}\right) \sim \exp\left[\sum \log\left(1-\frac{1}{2i}\right)\right] \tag{11}$$

$$\sim \exp\left[\sum \frac{1}{2i}\right] \tag{12}$$

$$\sim \exp\left[\frac{-\log n}{2}\right] \tag{13}$$

$$\sim \frac{1}{\sqrt{n}}. \tag{14}$$

Proof. Via generating functions: let

$$|C_\lambda| = z_\lambda = \frac{n!}{1^{m_1} m_1! \cdot 2^{m_2} m_2! \cdots}, \tag{15}$$

where C_λ is the conjugacy class of σ with cycle type λ. Then

$$1 = \sum_{\lambda \vdash n} \frac{z_\lambda}{n!} = [t^n] \prod_{i=1}^{\infty} \left(1 + \frac{t^i}{1! \cdot i} + \frac{t^{2i}}{2! \cdot i^2} + \cdots\right), \tag{16}$$

and this large product in the latter expression can be seen to be just $\frac{1}{1-t}$.

Let a_n be the number of $\sigma \in S_n$ with no cycle of length ip, $i = 1, \ldots, \lfloor n/p \rfloor$. Then we have

$$\sum \frac{a_n}{n!} t^n = \prod_{\substack{k \neq ip, \\ k \geq 1}} \left(1 + \frac{t^k}{1! \cdot k} + \frac{t^{2k}}{2! \cdot k^2} + \cdots\right). \tag{17}$$

We rewrite as

$$= \prod_{k \neq ip} \exp\left(\frac{t^k}{k}\right) = \frac{\prod_{i=1}^{\infty} \exp\left(\frac{t^i}{i}\right)}{\prod_{i=1}^{\infty} \exp\left(\frac{t^{ip}}{ip}\right)} \tag{18}$$

and then this becomes

$$= \exp\left[\sum_{i=1}^{\infty} \frac{t^i}{i} - \sum_{i=1}^{\infty} \frac{t^{ip}}{ip}\right] = \exp\left[-\log(1-t) + \frac{1}{p}\log(1-t^p)\right] \tag{19}$$

and then

$$= \frac{(1-t^p)^{\frac{1}{p}}}{1-t} = \left(\frac{1-t^p}{1-t}\right)\left(\frac{1}{1-t^p}\right)^{1-\frac{1}{p}}. \tag{20}$$

This first term is easy to expand, and this second term becomes

$$1 + \sum_{m=1}^{\infty} t^{mp}(1-\frac{1}{p})(1-\frac{1}{2p})\cdots \tag{21}$$

completing the proof (See also exercise 5.10 of Stanley). $\square$

Lemma 6.

$$P(\sigma \textit{ has } p\textit{-cycle, } p = \textit{prime, } p < n-2) \to 1 \textit{ as } n \to \infty. \tag{22}$$

Proof. Let A be the set of primes p with $\log^2 n < p < n-2$. Then by our main lemma from last time,

$$P(\sigma \text{ has no } A\text{-cycles}) \leq \left(\sum_{p\in A} \frac{1}{p}\right)^{-1} \sim \left(\log\log x|_{\log^2 n}^{n}\right)^{-1}, \tag{23}$$

the last part being proven on Wikipedia and is due to Euler. Thus this probability is approximately

$$\frac{1}{\log\log n}, \tag{24}$$

which goes to 0 as $n \to \infty$. $\square$

We can get even more info than this. In fact, we have

$$P(\sigma \text{ has exactly one } p\text{-cycle} \mid \sigma \text{ has at least one } p\text{-cycle}) \to 1 \text{ as } n \to \infty. \tag{25}$$

It suffices to prove that this probability is equal to

$$P = \prod_{i=1}^{n-p}\left(1 - \frac{1}{ip}\right). \tag{26}$$

1 More of Wednesday

Recall our main lemma

Lemma 1 (Main Lemma). *$P(\sigma \in S_n$ has a unique cycle of size p, p prime $< n-2) \to 1$ as $n \to \infty$.*

Proof. Let $A = \{p : \log^2 n < p < n-2\}$. Then from last time (see Wednesday for more careful proof),

$$P(\sigma \text{ has at least one } A\text{-cycle}) \to 1. \tag{1}$$

Then we want to show that

$$P(\sigma \text{ has a unique } (pk)\text{-cycle for some } k|\sigma \text{ has some } (pk)\text{-cycle}) \to 1. \tag{2}$$

$$= \prod_{i=1}^{\lfloor (n-p)/p \rfloor} \left(1 - \frac{1}{ip}\right) > \exp\left(\sum_{i=1}^{n} \log\left(1 - \frac{1}{ip}\right)\right) \tag{3}$$

(note the change in bounds), and this is

$$= \exp\left(-\sum \frac{1}{ip} + O(1)\right) = \exp\left(\frac{-\log n + O(1)}{p}\right), \tag{4}$$

and since $\frac{\log n}{p} < \frac{1}{\log n}$,

$$> \exp\left(\frac{-1}{\log n} + o(1)\right) > 1 - \frac{1}{\log n}. \tag{5}$$

□

In particular, σ almost always has some cycle of prime length. We will use this to prove the theorem we mentioned on Monday,

Theorem 2 (Dixon, 1968, Conjectured by Netto, 1896). *$P(\langle \sigma_1, \sigma_2 \rangle = A_n$ or $S_n) \to 1$ as $n \to \infty$.*

We need

Theorem 3 (Jordan ∼1840). *If $G \subset S_n$ is primitive and G has a p-cycle with $p < n-2$ a prime, then $G = A_n$ or $G = S_n$.*

Proof of Dixon's Theorem. It turns out that $P(\langle \sigma_1, \sigma_2 \rangle$ is primitive$) \to 1$, so since with high probability we have a p-cycle (take σ^a for an appropriately chosen a to get rid of everything but the p-cycle), we're done. □

Definition 1. G is *primitive* if

1. G acts transitively on itself.

2. G does not have a "block structure," i.e. nothing like a pair of permutations $\sigma_1 = (12)(34)(56)\cdots$ and $\sigma_2 = (1357\cdots)$. More formally, if $n = k\cdot m$ and we break $[n]$ into k-blocks $B_1, \ldots, B_k$, we have block structure if both σ_i map everything in one block to only things in another specific block.

Claim 4. *$P(\langle\sigma_1, \sigma_2\rangle =$ transitive$) = 1 - 1/n + O(1/n^2)$.*

Proof.

$$1 - P < \sum_{k \le n/2} \binom{n}{k} \cdot P(\langle\sigma_1, \sigma_2\rangle \text{ fixes a specific } k\text{-subset}). \tag{6}$$

The orbit of 1 is some k-subset, so we have

$$< \sum_k \binom{n}{k} \frac{1}{\binom{n}{k}^2} = \sum_k \frac{1}{\binom{n}{k}^2} \sim \frac{1}{n} + O(\frac{1}{n^2}). \tag{7}$$

□

Claim 5. *The probability that $\langle\sigma_1, \sigma_2\rangle$ has a block structure goes to 0 as $n \to \infty$.*

Proof. How many blocks of size k, if $n = km$?

$$\#\{B_1, \ldots, B_m\} = \frac{1}{m!}\binom{n}{k, k, \ldots, k} = \frac{n!}{k!^m m!} \tag{8}$$

$$P(\sigma_1, \sigma_2 \text{ preserve block structure}) < \sum_{k|n} \frac{n!}{k!^m m!} \cdot \left(\frac{n!}{k!^m m!}\right)^{-2} < n \cdot \frac{n!}{2^{n/2}\left(\frac{n}{2}\right)!} \tag{9}$$

□

2 Some history

In 1832 after Galois' death, Jordan received his notes from Galois' family. In 1840 Jordan's use of these notes led to a ("easy") proof of the theorem mentioned today. Next in 1845, Bertrand proved the following fact: if p divides the order of σ, and $n/2 \le p < n - 2$, then same conclusion as Jordan's theorem. Then Bertrand wondered, are there even any primes between $n/2$ and $n - 2$? Bam, Bertrand's postulate, and away we go into number theory.

1 On trees and stuff

1.1 Random Labeled Trees

(Random trees are "a little too broad.") We start with Cayley's formula:

$$\#\text{ labeled trees with } n \text{ vertices} = n^{n-2}. \tag{1}$$

Questions (think of the tree as rooted at 1.):

1. What is the average degree of the root?

   Answer: $\frac{2n-2}{n}$, by considering the total degree of the vertices and dividing by the number of vertices. This goes to 2 as $n \to \infty$.

2. What is the average number of leaves?

   Answer: $\sim \frac{n}{e}$.

3. What is the average height?

   Answer: "I'm not gonna tell you!" Solved by Renyi and Szekeres, who came up with a formula after lots of complex analysis methods.

1.1.1 Prüfer Code

Given a labeled tree, remove the largest leaf and let the first element of the code be its parent (Stanley vol. 2, page 25) This is the *Prf̈er code*, which has length $n-2$ (first element is always the root 1) for a total of n^{n-2} possible codes (i.e, it's easy to show that there's a bijection between labeled trees and Prf̈er codes).

Observation: the number of i's in the code is equal to $\deg(i) - 1$.

We can use this to answer the first question again: we ask for the expected number of 1's in the code, and get $1 + \frac{n-2}{n}$.

Now for the number of leaves:

$$E[\#leaves] = \sum_{i=2}^{n} P(i - leaf) = \left(1 - \frac{1}{n}\right)^{n-2} \cdot (n-1) \sim \frac{n}{e} \tag{2}$$

Sadly, for the height we cannot use Prüfer codes, we'll need a different method.

Claim 1.

$$E[height\ of\ T] = \theta(\sqrt{n}). \tag{3}$$

1.1.2 Loop erased random walks and Wilson's algorithm

The latter uses the former to create random samplings spanning trees. That is, given $G \subset K_n$ as input, return a spanning tree T of G uniformly at random (for arbitrary trees it's easy: just generate the Prüfer code and biject backwards).

Definition 1. A *loop erased random walk* is described as a process in G: first we start a random walk in G, but every time we loop back through a vertex we erase the vertices visited between the two instances, pretending as if it never happened.

There's some delicate stuff here, like if this makes sense on $\mathbb{Z}^2$ (it does, but takes some thought and care). We obtain an algorithm:

Start at vertex 2.
Do a loop erased random walk until we hit 1.
Pick the smallest element not on the path.
Do a loop erased random walk until we connect to the existing path.
Repeat the last two steps until we obtain a tree.

Theorem 2 (Wilson 1995)**.** *This algorithm generates spanning trees uniformly at random in G.*

We're going to rewrite until this becomes clear. We need the following *cycle popping algorithm*:

- For every vertex and each layer(???), choose one outgoing edge.
- If this is not a tree, there is a cycle somewhere (no figure eights or anything, since only one outgoing edge). "Pop" it, whatever that means (rerandomize)

Observation popping cycles commute. Proof by meditation.

Eventually with some probability you get a tree. We claim this algorithm is the same as Wilson's algorithm. Proof by meditation. We also claim this is a uniform algorithm: the cycles don't matter, so it's just whatever the top tree of the stack is.

We now give a lower bound for the height:

Claim 3.
$$E[\textit{distance from } 2 \textit{ to } 1] = \theta(\sqrt{n}). \tag{4}$$

Proof. Since Wilson's algorithm starts at 2, we want to know how long it takes to hit 1. We are writing down numbers and deleting everything everytime we see a repeat. Igor claims this probability is $\sqrt{n}$. $\square$

1 Spanning trees in graphs

We continue last lecture, following up on Wilson's algorithm.

From last time, we had

Corollary 1. *$E[d] = O(\sqrt{n})$, where d is the distance between vertices 1 and 2.*

Our proof was through loop-erased random walks and the birthday paradox.

Theorem 2 (Meir, Moon 1970)**.** *Let z be the max depth of our spanning tree. Then*

$$E[z] = \sqrt{2\pi n}(1 + o(1)). \tag{1}$$

We get this through the following lemma.

Lemma 3.

$$P(z = k) = \frac{k}{n-1} \cdot \frac{n(n-1)\cdots(n-k+1)}{n^k}. \tag{2}$$

To prove *this* lemma, we need another.

Lemma 4. *Suppose T_k is a (not necessarily spanning) tree in K_n with k vertices. Then the number of spanning trees in K_n which contain T_k is equal to $k \cdot n^{n-k-1}$.*

Proof of Lemma 4. Assume without loss of generality that T_k is a tree on the last k vertices and consider a tree that contains T_k. Prüfer's code will at some point start cleaning out T_k, and there will be $k-1$ consecutive positions in the code that correspond to this cleaning. These are set once we pick a root of T_k to attach to the rest of the tree somehow, and we have k choices for the root. The rest of the (length $n-2$ total) Prüfer code gives n^{n-k-1}. □

Proof of Lemma 3. Let T_k be a length k path from 1 to 2. There are $(n-2)\cdots(n-k+1) = (k-2)!\binom{n-2}{k-2}$ such paths. By Lemma 4, there are $k \cdot n^{n-k-1}$ spanning trees containing T_k. Calculate the probability:

$$P = \frac{(k-2)! \cdot \binom{n-2}{k-2} \cdot k \cdot n^{n-k-1}}{n^{n-2}} = \frac{k}{n-1} \cdot \frac{n(n-1)\cdots(n-k+1)}{n^k}. \tag{3}$$

□

Let's see now how Lemma 3 implies the Theorem.

Proof of Theorem. Let $p(n,k) = P(z = k)$ in K_n, and observe that

$$\frac{P(n,k+1)}{p(n,k)} = \frac{k+1}{k} \cdot \frac{n-k}{n}. \tag{4}$$

Let's observe how these numbers change as k increases. For $k = 1$, we have $p(n,k) = \frac{1}{n-1}$. For $k = n$ we have $p(n,k) = \frac{n}{n-1}\frac{n!}{n^n} \sim e^{-cn}$. Thus we expect for $p(n,k)$ to increase for a while but then taper off exponentially as k increases.

The maximum is obtained around $k = \sqrt{n}$. To be more precise, we'll write an estimate. We claim that $p(n,k) \sim e^{-k^2/2n}$. To see this, we have

$$e^{-t/(1-t)} < 1 - t < e^{-t}, \tag{5}$$

for $0 < t < 1$. Then we have

$$\frac{k}{n-1} e^{-k^2/2(n-1)} < p(n,k) < \frac{k}{n-1} e^{-\frac{1}{n}\binom{k}{2}}. \tag{6}$$

Alternatively, observe that

$$p(n,k) = \frac{k}{n-1} \prod_{i=0}^{k-1} \left(1 - \frac{i}{n}\right). \tag{7}$$

Thus we have

$$p(n,k) \sim \frac{k}{n} e^{-\frac{k^2}{2n}} + O\left(\frac{1}{n^{1+\varepsilon}}\right). \tag{8}$$

Our expected value can them be expressed as an integral

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum k \cdot p(n,k) &= \int_{n^{1/2-\alpha}}^{n^{1/2+\alpha}} u^2 e^{-u^2/2}\, \mathrm{d}u + O(?) &(9)\\ &= \frac{1}{\sqrt{n}} \int_0^\infty u^2 e^{-u^2/2}\, \mathrm{d}u + O(??) &(10)\\ &\sim \sqrt{2\pi}. &(11) \end{aligned}$$

□

As a bonus, we can also use this method to calculate the variance of z to be $\sim (2 - \pi/2)n(1 + o(1))$ (proof that $\pi < 4$, Igor points out).

2 Generating functions over spanning trees

Define

$$P_G(x) = \sum_{spanning forests F in G} x^{k(F)-1}, \tag{12}$$

where $k(F)$ is the number of connected components.

Theorem 5. *We have*

1. $P_{\{1\}}(x) = 1$,
2. $P_{G \sqcup H}(x) = x P_G(x) P_H(x)$,
3. $P_{\overline{G}}(x) = (-1)^{n-1} \cdot P_G(-x-n)$,

(organized in increasing order of difficulty, with the only nontrivial one being the third).

Observe also that the number of spanning trees in G $C(G)$ is equal to $P_G(0)/n$. Note that if O_n is the edgeless graph on n vertices then $P_{O_n} = x^{n-1}$. Additionally from the third statement, $P_{K_n} = (x+n)^{n-1}$.

Back in the day, people would just come up to a podium and read a math paper to an audience.

1 Algebraic Methods

1.1 The Matrix-tree Theorem

Goal is to find $c(G) = \#$ spanning trees in G. Last time we defined the generating function

$$P_G(x) = \sum_{\text{spanning rooted forests } F} x^{k(F)-1} \tag{1}$$

and observed that $c(G) = P_G(0)/n$.

Theorem 1. *We have*

$$P_{K_1}(x) = 1 \tag{2}$$
$$P_{G \sqcup H} = xP_G \cdot P_H \tag{3}$$
$$P_{\overline{G}} = (-1)^{n-1} P_G(-x-n) \tag{4}$$

Note that if O_n is the edgeless graph on n vertices then $P_{O_n} = x^{n-1}$. Additionally from the third statement, $P_{K_n} = (x+n)^{n-1}$. We can get the bipartite graph $K_{p,q}$ from this, as it is the complement of $K_p \sqcup K_q$:

$$P_{K_p \sqcup K_q}(z) = z \cdot (z+p)^{p-1}(z+q)^{q-1} \tag{5}$$

and so

$$P_{K_{p,q}} = (-1)^{p+q-1} \cdot (-x-p-q) \cdot (-x-q)^{p-1}(-x-p)^{q-1}. \tag{6}$$

From this it follows at once that

$$c(K_{p,q}) = p^{q-1}q^{p-1};\ c(K_{n/2,n/2}) = \left(\frac{n}{2}\right)^{n-2} = \frac{n^{n-2}}{2^{n-2}}. \tag{7}$$

Interestingly, this implies that

$$\log c(K_n) = \log c(K_{n/2,n/2})(1 + o(1)). \tag{8}$$

Scott sees a good interpretation. Flip a coin on each edge, what's the chance that the tree will survive? For large n, we could pretend it's about independent and get the above formula.

We also have the generating functions

$$\sum_{T \subset K_n} x_1^{d_1-1} \cdots x_n^{d_n-1} = (x_1 + \cdots + x_n)^{n-2} \tag{9}$$

$$\sum_{T \subset K_{p,q}} x_1^{d_1-1} \cdots x_p^{d_p-1} \cdot y_1^{d_1'-1} \cdots = (x_1 + \cdots + x_p)^{q-1}(y_1 + \cdots + y_q)^{p-1} \tag{10}$$

Theorem 2 (Matrix tree theorem)**.**

$$c(G) = |\det(M'_G)|; \;\; x \cdot P_G(x) = \det(M_G + xI) \tag{11}$$

where

$$M'_G = \begin{pmatrix} +d_1 & -\ell_{i,j} & \cdots \\ & & \end{pmatrix} \tag{12}$$

and its augment

$$M_G = \left(\begin{array}{cccc} d_1 + x & \cdots & & x \\ & d_2 + x & \cdots & x \\ \hline x & \cdots & x & nx \end{array}\right) \tag{13}$$

To prove this, we look to the properties of P. The third property in particular would give us

$$\det(M_{\overline{G}} + xI) = \pm \det(M_G - (x+n)I), \tag{14}$$

so let's prove this.

Let's take a look at $M_{\overline{G}}$. By the nature of complementation,

$$M_{\overline{G}} = -J - M_G + nI, \tag{15}$$

where J is the matrix with every entry equal to 1.

Consider $\det(J + A)$. In general, there is no good formula for this. But, suppose that A has 0 as an eigenvalue with eigenvector $w = (1, 1, \ldots, 1)$ (as our M_G does, since sum of all rows is zero). Every vector in the orthogonal complement of w (can consider only because w is the only 0-eigenvector) is $(v_1, \ldots, v_n)$ with $v_1 + \cdots + v_n = 0$, so $Jv = 0$ and if v is an eigenvector,

$$(J + A)v = \lambda v. \tag{16}$$

Additionally, we have

$$(J + A)w = n \cdot w. \tag{17}$$

To finish, note the first determinant is just

$$\lambda_1 \cdots \lambda_{n-1}/n. \tag{18}$$

The latter determinant adds the eigenvalue $\lambda_n = n$, but the $-nI$ term gets rid of it. To wit,

$$\det(-J - M_G + nI + xI) = (-1)^n \det(J + M_G - nI - xI), \tag{19}$$

and we have just argued that $J + M_G - nI$ has the same eigenvalues as M_G, up to signs (not quite clear). More clearly, $J + M_G - (n+x)I$ has eigenvalues $\lambda_1 - (n+x), \lambda_2 - (n+x), \ldots, \lambda_{n-1} - (n+x), -x$, and the product of these is

$$\prod \lambda_i = \det(M_G - (n+x)I), \tag{20}$$

and $n + x$ is just another variable.

1

Igor's favorite type of argument is conceptual rather than technical.

Let $G = H_n$ be the hypercube graph, i.e. $H_n = K_2 \times \cdots \times K_2$.

Theorem 1. *The number of spanning trees $c(H_n)$ is equal to*

$$1/2n \cdot \prod_{i=1}^{n} (2i)^{\binom{n}{i}}. \tag{1}$$

Observe first that $c(H_n) < \binom{n2^{n-1}}{2^n-1}$, since the number of edges is $n2^{n-1}$. Asympototically, this gives an upper bound of $(en)^{2^n} \sim 2^{2^n \log n}$. Igor then writes this equality?

$$c(H_n) = 2^{2^n-n-1} \prod_{k=1}^{n} (k)^{\binom{n}{k}}. \tag{2}$$

And then this guy,

$$\left(\frac{n}{2}\right)^{\binom{n}{n/2}} = \left(\frac{n}{2}\right)^{2^n/\sqrt{n}} \tag{3}$$

Which implies that

Okay, moving on to a proof of the theorem, we have

$$c(H_n) = \frac{1}{N} \prod_{i=1}^{N-1} \lambda_i, \tag{4}$$

where λ_i are the eigevalues of M_{H_n} and M_{H_n} is as last time, n's down the diagonal and $-1, 0$ off diagonal. Consider $-L_n = M_{H_n} - nI$ and $\tilde{L_n} = L_n/n$. Note that $\tilde{L_n}$ is the transition matrix for a uniform random walk on the graph! Of course, $\lambda_0 = (1, 1, \ldots, 1)$ is an eigenvector. Igor argues that since the graph is bipartite and other eigenvectors must be orthogonal to λ_0, every other eigenvector must have each coordinate equal to ± 1. For every subset S of the basis vectors, we can define such an eigenvector by letting the ± 1 color of vertex $v+e_i$ to be the same as v if $e_i \in S$, and the color flips otherwise.

We now have that $\tilde{L_n} w_S = \lambda_S \cdot w_S$, where $\lambda_S = \frac{2|S|-n}{n}$. We need to convert these to the original eigenvalues of M. We have $\binom{n}{k}$ eigenvectors with $\lambda_S = \frac{2k-n}{n}$, and all of these go to an eigenvalue of $2k - n$ when we move to L_n. Since $M = nI - L_n$, we get $\binom{n}{k}$ eigenvectors with eigenvalue $2(n-k)$.

Thus, we get

$$c(H_n) = \frac{1}{2^n} \prod_{k=0}^{n-1} \left[2(n-k)\right]^{\binom{n}{k}}, \tag{5}$$

and the theorem is proved.

Let x_0 be the origin, and let x_t be a random walk on H_n, so $x_{t+1} = x_t + e_i$ for some randomly chosen basis vector e_i. We determine the mixing time. Consider first the simpler walk where half the time we stay and half the time we leave. Then this has transition matrix $A = (\tilde{L_n} + nI)/2$. We are considering

$$A^t \begin{pmatrix} 1 \\ 0 \\ \vdots \end{pmatrix}, \tag{6}$$

and the mixing time is equal to $n \log n + O(n)$ (coupon collector's problem). Because something about $\sum_{i=1}^{n} n/i$. Igor then gave a justification for how to move to our original walk, but I was distracted by reading about Leonard vs. Pepsico, because Igor brought it up after coupon collecting.

1 Polyominoes

"Enumerating them is not a nice number." Restrict to simply connected polyominoes, and let a_n be the number of polyominoes with area n.

So, what is a_n asymptotically? If you're a physicist, $a_n \sim c \cdot n^\gamma \cdot \lambda^n$ (Igor is arguing that it's often of this form, but nobody knows for sure that it always looks like this). Lot of dimer type problems to study and so on.

Theorem 1.

$$\lim (a_n)^{1/n} = \lambda \tag{1}$$

exists and is finite.

Note that $(a_n)^{1/n} \geq 2$, since we can always attach to the topmost rightmost square a new square to the right or above.

Lemma 2. $a_n < 100^n$

Proof of lemma. Travel counterclockwise around our polyomino. Since we choose a new direction every time, we get an absurd overestimate bound of 3^k where k is the perimeter. But $k \leq 2n + 2 < 4n$. □

Lemma 3. $a_n \cdot a_m \leq a_{m+n}$.

Proof. Find a canonical way to attach two polyominoes of size n, m respectively (leftmost lowermost, and then rightmost topmost). □

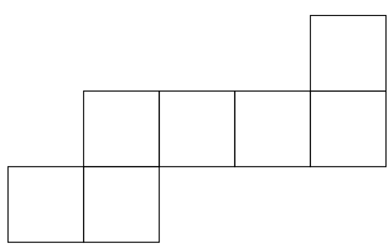
Proof of Theorem. Take the logarithm, and by this last lemma the log-sequence is increasing. □

For specific polyominoes, we can use computer calculation to compute a possible range for λ, "but this is boring."

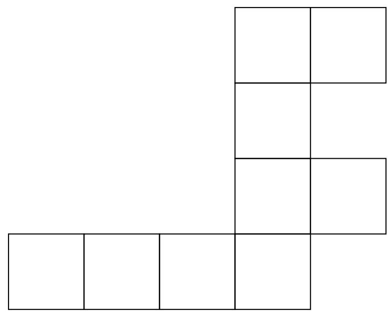
An argument by Klarner-Rivest showed that $\lambda < \frac{27}{4}$. The guts of the argument involve starting at the lower left corner, and then greedily creating a tree by jumping to adjacent squares. The number of trees is some upper bound like $\binom{3n}{n}$ and the result follows.

1.1 Two nice families

First family: directed column-strict polyominoes.



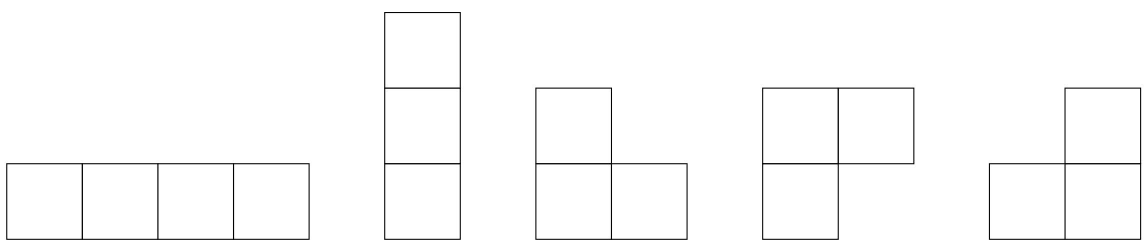
Everything is accessible from the lowerleftmost square by traveling only upwards and rightwards. Column strictness means that every column must be an interval, so the next example breaks this in the last column:



Theorem 4.

$$\#dcsp(n) = F_{2n-1}. \tag{2}$$

For example, when $n = 2$ we have the horizontal and vertical 2-ominoes. For $n = 3$:

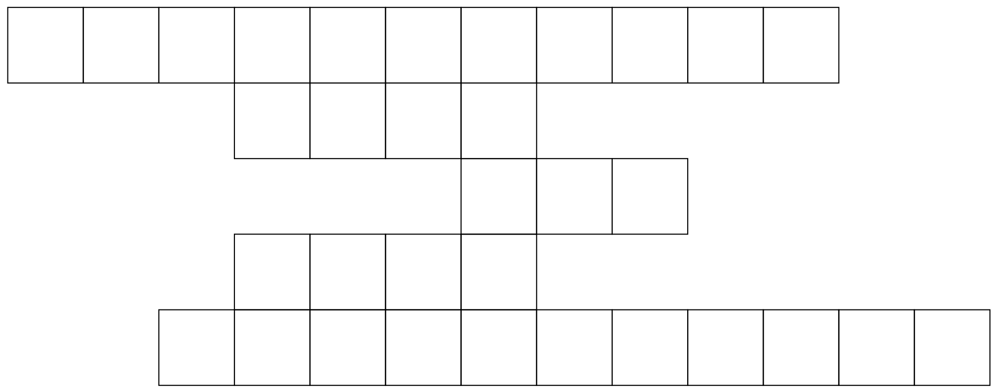


Surprisingly, an inductive proof is not known.

Proof by Deutsch-Prodinger. We prove by exhibiting a bijection between dcsp's and plane trees with height ≤ 3 and $(n+1)$ vertices. "If you wanted to choose your favorite Fibonacci object, this wouldn't be it. But it works." For example, for $n = 2$: For $n = 3$, we get the usual 5 Catalan trees. For $n = 4$, we get all the Catalan trees but we're missing the purely vertical one, so $14 - 1 = 13$. Number the columns of a dcsp 1 to k, and for each column write down the lowest horizontal a_i. Then write down the highest horizontals b_i. Finally, write down their differences $b_{i-1} - a_i$ Now from this we draw a tree: Trees of height ≤ 1 have a generating function $\mathcal{F}_1 = \frac{1}{1-t}$. Let $\mathcal{F}_k$ be the generating function for plane trees of height at most k, indexed by number of edges. Then $\mathcal{F}_2 = 1 + t\mathcal{F}_1 + t^2\mathcal{F}_1^2 + \cdots$, by detaching the root node to create a forest of smaller depth trees. $= \frac{1}{1-t\mathcal{F}_1} = \frac{1}{1-\frac{t}{1-t}} = \frac{1-t}{1-2t}$. By the same argument, $\mathcal{F}_3 = \frac{1}{1-t\mathcal{F}_2} = ugly = \frac{1-2t}{1-3t+t^2}$. This is the generating function for the odd Fibonacci numbers, so we're good. $\square$

In particular, the number of polyominoes grows faster than this

1.2 Horizontally convex polyominoes



Claim that

$$\sum b_n t^n = \frac{t(1-t)^3}{1-5t+7t^2-4t^3} \tag{3}$$

and

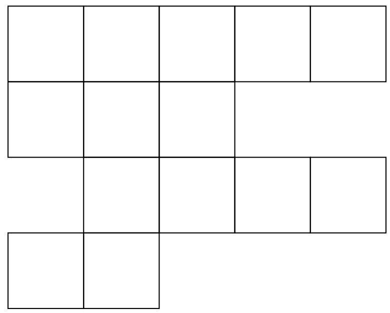
$$b_{n+3} = 5b_{n+2} - 7b_{n+1} + 4b_n \tag{4}$$

for $n \geq 2$ and b_n is the number of these polyominoes of area n. This shows than $a_n \geq b_n \sim 3.2^n$.

This is exercise 4.75 in EC1, exawple 6.5 in Odlyzcko, and V.20 (p. 365) in Flajolet and Sedgewick.

1 Polyominoes

A row-convex polyomino is a stack of intervals with no disjoint intervals on the same level.



Theorem 1 (Klarner 1965)**.** *Denote by a_n the number of row-convex polyominoes of area n. Then*

$$\sum a_n z^n = \frac{z(1-z)^3}{1-5z+7z^2-4z^3} \tag{1}$$

Proof following p.365 of Flajolet and Sedgewick. Let

$$\mathcal{F}_k(z) = \sum [\#\text{ rc poly with top slice} = k \text{ and area } n] z^n, \tag{2}$$

and let

$$\mathcal{F}(z,u) = \sum_{k=1}^{\infty} \mathcal{F}_k(z) u^k. \tag{3}$$

Then we have

$$\mathcal{F}_1 = z + z(\mathcal{F}_1 + 2\mathcal{F}_2 + 3\mathcal{F}_3 + \cdots) \tag{4}$$

$$\mathcal{F}_2 = z^2 + z^2(2\mathcal{F}_1 + 3\mathcal{F}_2 + 4\mathcal{F}_3 + \cdots) \tag{5}$$

$$z^3 + z^3(345etc) \tag{6}$$

To see these equalities, think about going from a top slice of size k to a top slice of size ℓ, then there are $k + \ell - 1$ ways of placing the new row on top of the length k row.

Thus we have an infinite system of equations that we'd like to solve. We could think about z as being a probability and this top row adding as a markov process, but I think he's saying that's not the way to go.

Let's introduce an operator $\mathcal{L} : \mathbb{C}[[u,z]] \to \mathbb{C}[[u,z]]$ defined by

$$\mathcal{L}(u^k) = k(uz) + (k+1)(uz)^2 + \cdots \tag{7}$$

$$= (k-1)\frac{(uz)}{1-uz} + \frac{(uz)}{(1-uz)^2}. \tag{8}$$

Then for $f \in \mathbb{C}[[u, z]]$ we have the general formula

$$\mathcal{L}(f(u)) = \frac{uz}{(1-uz)^2} f(1) + \frac{uz}{1-uz}[f'(1) - f(1)]. \tag{9}$$

We thus obtain the equation

$$\mathcal{F}(z, u) = \frac{uz}{1-uz} + \mathcal{L}(\mathcal{F}(z, u)). \tag{10}$$

Let $\tau(u) = \mathcal{F}(z, u)$, and we wish to obtain an equation for $\tau(1)$. Then plugging in $u = 1$ gives us (just see F&S)

$$\tau(1) = \frac{z}{1-z} + \frac{1}{1-z}\tau'(1) + \frac{(z)^2}{(1-z)^2} \tag{11}$$

and also

$$\tau'(1) = \frac{z}{(1-z)^2} + \frac{z}{(1-z)^2}\tau'(1) + 2\frac{z}{(1-z)^3}\tau(1) \tag{12}$$

Now we have a much easier system of equations in the variables $\tau(1)$ and $\tau'(1)$, from which we get the expression for $\tau(1) = \mathcal{F}(z, 1)$ in equation (1). $\square$

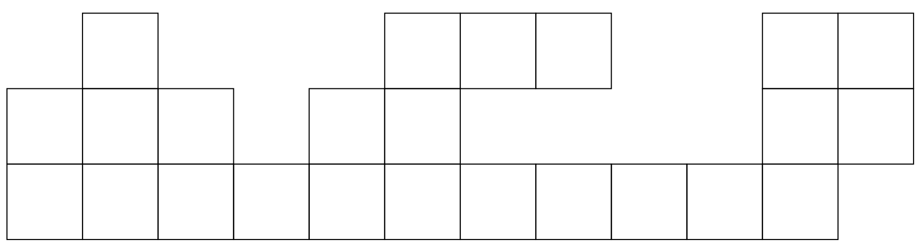
1.1 State of the art

If $p(n)$ is the number of polynomials, we currently know that

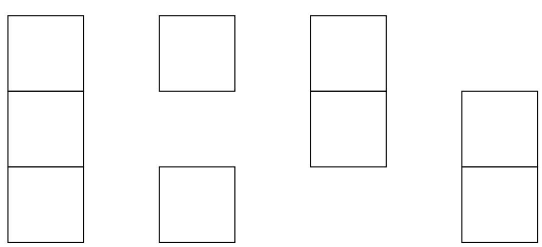
$$p(n) \sim \frac{1}{n}\lambda^n, \tag{13}$$

where it is known that $\lambda > 3.98$ and "$\lambda \approx 4.06$," whatever that means.

For some justification, look at polyominoes that fit into three lines.



Look at vertical cross sections of your polyomino, and note that there's only a finite number of configurations that you can see there, such as the following:



Make a finite graph with acceptable transitions from configurations. You get some awful generating function, but you change the heights and you see that their asymptotic behavior is converging to something like 4.06.

1 Unimodality and Log-Concavity

We'll spend the next few lectures on this. Instead of precise asymptotics on large n, inequalities on all n.

Definition 1. $a_1, \ldots, a_n$ is *symmetric* if $a_i = a_{n+1-i}$ for all i.

Definition 2. $a_1, \ldots, a_n$ is *unimodal* if $a_1 \leq \cdots \leq a_k \geq \cdots \geq a_n$ for some k.

Definition 3. $a_1, \ldots, a_n$ is *log-concave* if $a_i^2 \geq a_{i-1}a_{i+1}$ for $2 \leq i \leq n-1$.

It's easy to see that log-concavity is stronger than unimodality, that is a log-concave sequence is also unimodal. Today we'll focus exclusively on log-concave sequences.

1.1 Analytic Method (real zeros)

Theorem 1. *Suppose we have a polynomial*

$$p(x) = \sum_{j=1}^{n} \binom{n}{j} a_j x^j, \tag{1}$$

and $p(x)$ has only real zeros. Then a_j is log-concave.

Important remark: if we write $b_j = \binom{n}{j} a_j$, then b_j is also a log-concave sequence. This follows from writing out the expression for log-concavity. From this we immediately get that $\binom{n}{j}$ is a log-concave sequence, as $p(x) = (1+x)^n$ has only real zeros.

Corollary 2. *Let $C(n,k)$ be the Stirling numbers of the first kind, the number of permutations in S_n with k cycles (or k left-right minima and so on). Then for a fixed n these form a log-concave sequence.*

Proof of corollary. We claim that

$$p(x) = \sum_{k=1}^{n} C(n,k) t^k = t(t+1)\cdots(t+n-1). \tag{2}$$

To see this, "stare twice" at the formula

$$C(n,k) = C(n-1,k-1) + (n-1)C(n-1,k). \tag{3}$$

On the first stare, we see that the coefficients of the polynomial $t(t+1)\cdots(t+n-1)$ satisfy the same recurrence. On the second stare, we see that this recurrence holds for the Stirling numbers, as n is either in a singleton cycle or we join it into a permutation in S_{n-1} with k cycles, in one of $(n-1)$ ways. $\square$

Our theorem is not however strong enough for Stirling numbers of the second kind, as we do not know the generating functions for them (?).

Proof of theorem. Let $D = \frac{\mathrm{d}}{\mathrm{d}x}$ and let $Q = D^{j-1}p(x)$. We define new polynomials,

$$R(x) = x^{n-j+1}Q\left(\frac{1}{x}\right); \quad A(x) = D^{n-j-1}R(x). \tag{4}$$

Then we have that after some computation,

$$A(x) = \frac{n!}{2}(a_{j-1}x^2 + 2a_jx + a_{j+1}). \tag{5}$$

Now by Rolle's theorem and the like, each of Q, R, A have only real zeros, following from the assumption on P. But A is a quadratic, and so the discriminant must be nonnegative and we have $4(a_j^2 - a_{j-1}a_{j+1}) \geq 0$. □

1.2 Inductive Arguments and Stirling Numbers of the Second Kind

Next, we prove that the Stirling numbers of the second kind are also log-concave. Let $S(n,k)$ be the number of k-partitions of the set $[n]$ (each part must be nonempty). "Think of them as an analogue of binomial coefficients, but without a nice formula for them." However, there is a nice recurrence:

$$S(n,k) = S(n-1,k-1) + kS(n-1,k), \tag{6}$$

which again follows from considering what we do with the number n in our partitioning process.

Theorem 3 (Sagan)**.** *Let $t(n,k)$ be a sequence such that $t(n,k) = c(n,k)t(n-1,k-1) + d(n,k)t(n-1,k)$.*

- *We demand that $c(n,k), d(n,k)$ are log-concave in k for every n.*
- *Additionally, we also want that $c(n,k-1)d(n,k+1)+c(n,k+1)d(n,k-1) \leq 2c(n,k)d(n,k)$.*

With these conditions, $t(n,k)$ is log-concave.

Proof. Need to show that

$$t(n,k)^2 \geq t(n,k-1)t(n,k+1). \tag{7}$$

"We've got no real choice, we just write things down and see what happens." We'll use things like

$$c(n,k-1)c(n,k+1)t(n-1,k-2)t(n-1,k) \leq c(n,k)^2t(n-1,k-1)^2 \tag{8}$$

$$\text{same but with } d \tag{9}$$

$$\begin{aligned} c(n,k-1)d(n,k+1)t(n-1,k-2)t(n-1,k+1) \\ +c(n,k+1)d(n,k-1)t(n-1,k-1)t(n-1,k) \end{aligned} \leq 2c(n,k)d(n,k)t(n-1,k-1)t(n-1,k) \tag{10}$$

Basically, it's an induction on n that's complicated but still straightforward. □

Corollary 4. *$S(n,k)$ forms a log-concave sequence for every n.*

Proof. In the theorem we have $c(n,k) = 1, d(n,k) = k$ and these are clearly log-concave. We just need to check the conditions in the formulation of the theorem. □

1.3 Next Time

Theorem 5. *Denote by $a_G(k)$ the number of rooted spanning forests of G with k components. Then $a_G(k)$ is log-concave.*

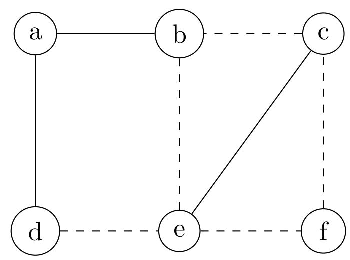


Figure 1: A spanning forest with 3 components (note the singleton tree).

Note that we already studied the generating function for $a_G(k)$ when we looked at the matrix tree theorem. The proof will be linear algebra, "plus epsilon".

Theorem 6 (Huh, 2012)**.** *Same result, but for unrooted spanning forests.*

Proof uses "quite serious" algebraic geometry.

Conjecture 7 (Mason)**.** *Same for all matroids.*

1 Spanning Forests

For a given graph G, let a_i be the number of spanning forests in G with i components.

Theorem 1. *a_i is a log-concave sequence.*

Proof. Consider the matrix tree theorem matrix M_G, and consider the polynomial

$$P(x) = \sum_{i=1}^{n} a_i x^i = \det(M_G + xI). \tag{1}$$

If the matrix is symmetric, this polynomial is the charpoly of a matrix, and by symmetry all the eigenvalues are real. By our theorem from Monday, this implies log-concavity. $\square$

As an immediate consequence, $a_i/\binom{n-1}{i}$ is also log-concave, which is some sort of probablistic interpretation?

1.1 Generalization

Let S be a subset of the edges of E, and let a_i be the number of spanning trees T such that $|T \cap S| = i$.

Theorem 2 (Godsil)**.** *These new a_i are still log-concave.*

Note that this contains our previous theorem, by augmenting our graph G with a special extra vertex (not quite sure I follow what he did). We'll need the following extension of matrix tree theorem:

Theorem 3 (Weighted Matrix Tree Theorem)**.** *(Assume there are no loop edges in G.) Let $w_{i,j}$ be the weight of each edge in G, and define M_G in the natural way:*

$$M_G = \begin{pmatrix} \sum w_{1,j} & -w_{1,2} & \cdots & -w_{1,k} \\ -w_{2,1} & \sum w_{2,j} & \cdots & -w_{2,k} \\ \vdots & & \ddots & \vdots \\ -w_{k,1} & \cdots & -w_{k,k-1} & \sum w_{k,j} \end{pmatrix}. \tag{2}$$

Then

$$\det(M''_G) = \sum_{T \subset G} \prod_{(i,j) \in T} w_{i,j} \tag{3}$$

Where M''_G is the removal of the first row and column (or any k-th row,col).

Proof of Theorem 2. Let

$$w_{i,j} = \begin{cases} q, & (i,j) \in S, \\ 1, & (i,j) \notin S, \\ 0, & (i,j) \notin E. \end{cases} \tag{4}$$

Then we have that $M_G = qM_{SG} + M_{\overline{S}G} = qA + B$.

Claim 4. $\det(M''_G) = \sum b_i q^i$ *has only real zeros.*

Modulo this claim, we are done, again by our theorem from Monday. $\square$

Before proving the above claim, we'll need a lemma.

Lemma 5. *Let A, B, D be $n \times n$ matrices with $A \geq 0$ (positive-definite) and $B = B^T$ symmetric and $D \geq 0$, and finally AB has real eigenvalues. Then $\det(B + qD)$ has real zeros.*

The presence of A is more of an existence statement, you just need to find a matrix that does what you want.

Proof of Lemma. Since $D \geq 0$, $D = C^2$ for some matrix C.

$$\det(B + qD) = \det(B - qC^2)\det = \det(C^{-1}BC^{-1} - qI)\det(C^2), \tag{5}$$

and now CBC is symmetric, so it has real eigenvalues. Augh there's some problems here. $\square$

Proof of claim. Throw our matrices into the lemma. Our matrices in question are positive-semidefinite due to the eigenvalues' relation to degrees. $\square$

Theorem 6. *Let a_i be the number of matchings in G with i edges. Then a_i forms a log-concave sequence. Moreover $\sum a_i x^i$ has real zeros, which is stronger (but it's not important for today).*

Proof by C. Krattenthaler. We prove that for every $k \leq \ell$ we have $a_{k-1}a_{\ell+1} \leq a_k a_\ell$ (Andy points out that this is equivalent to log-concavity, not stronger). We wish to show that there is an injection from

$$\{k - 1 \text{ matchings}\} \times \{\ell + 1 \text{ matchings}\} \tag{6}$$

to

$$\{k \text{ matchings}\} \times \{\ell \text{ matchings}\}. \tag{7}$$

To obtain the injection, color the $k-1$ matchings orange, and the $\ell+1$ matchings green. By the definition of a matching, we see that the matchings create a collection of paths and cycles where the colors alternate along the edges: see Figure 1. Leave the cycles (which are necessarily balanced) and balanced paths alone, and consider the unbalanced paths of matchings (which much exist since $k - 1 < \ell + 1$). Then $b - a = \ell - k + 2$. Switch colors on some of the unbalanced chains, $\binom{a+b}{a} \leq \binom{a+b}{a+1}$ since $a \leq b$. How do we ensure injectivity? I'm not sure I follow his argument. $\square$

1.2 Final Comment

Similar conjectures on Matroids can't use the matrix tree theorem, so you have to do something different. These sat around being open for a while until Stanley proved them.

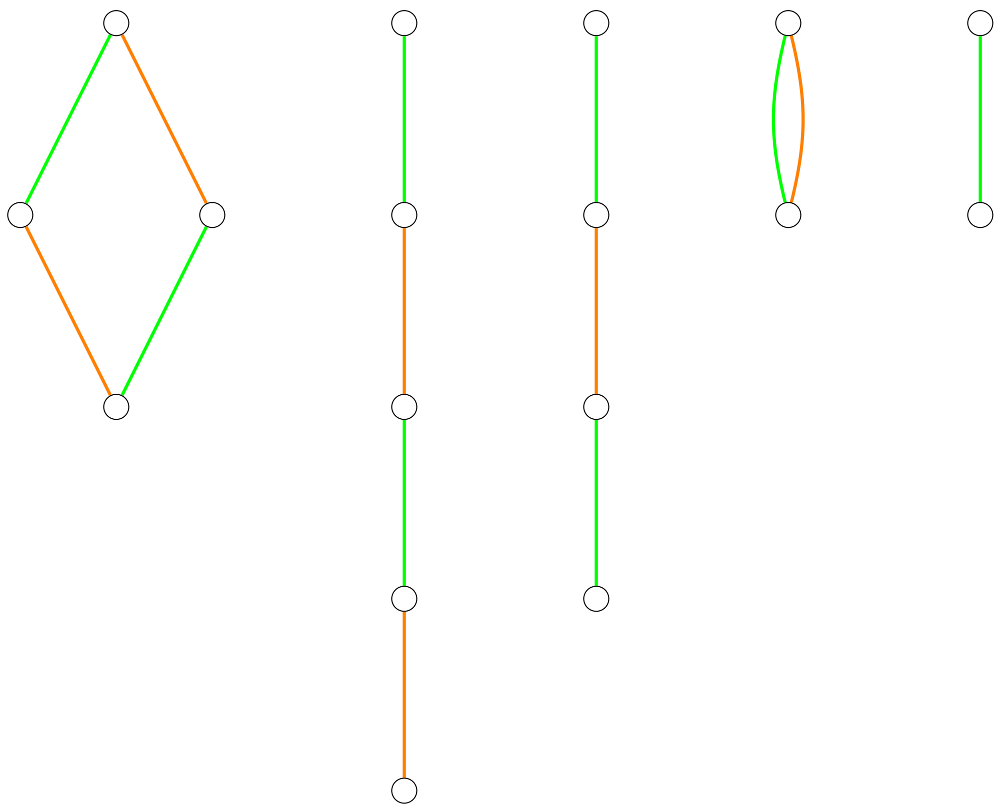


Figure 1: Two matchings induce alternating paths and cycles.

1 Log concavity via geometry

For clarifications on today's lecture, you may wish to track down Stanley's paper *Two Combinatorial Applications of the Alexandrov-Fenchel Inequalities.*

Theorem 1 (Stanley). *Let $\mathcal{P}$ be a poset on $V = \{v_1, \dots, v_{n-1}, w\}$. If a_i is the number of linear extensions f of $\mathcal{P}$; $f(w) = i$, and $f : V \to [n]$ such that $v_1 \prec v_2 \implies f(v_1) < f(v_2)$, and f is a bijection, then a_i forms a log concave sequence.*

An informative example: in Figure 1 we have $a_1 = 2, a_2 = 2, a_3 = 1$ (note

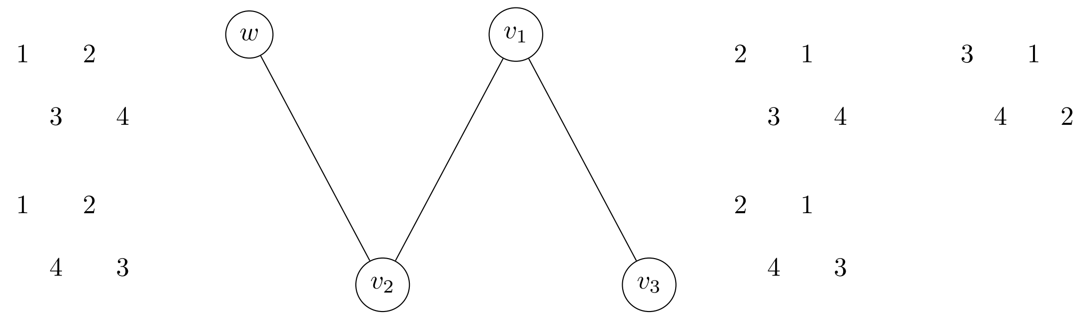


Figure 1: An informative poset

that Igor is ordering downwards instead of the usual upwards).

Corollary 2. *Let $\mathcal{A}_n = \{\sigma \in S_n : \sigma$ is alternating with $\sigma(1) > \sigma(2)$ etc.$\}$, and let $a_i^{(j)} = \#\sigma \in \mathcal{A}_n$ such that $\sigma(j) = i$. Then for a fixed j, $a_i^{(j)}$ is log concave.*

Proof. Apply the above theorem to the poset in Figure 2, with $w = v_j$.

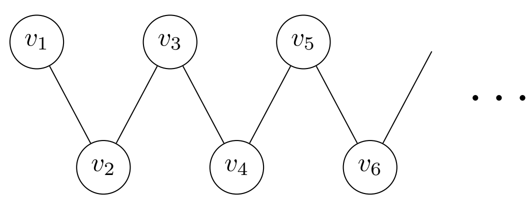


Figure 2: The poset for an alternating permutation structure.

□

Theorem 3 (Alexandrov-Fenchel inequalities). *$\{V_i(K, L) : 0 \le i \le n\}$ is log concave.*

Here, K, L are convex polytopes in $\mathbb{R}^n$, and $V_i(K, L)$ is the i-th *mixed volume* of K, L, defined by the polynomial

$$vol(xK + yL) = \sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i} V_i(K, L), \tag{1}$$

where the above is the Minkowski sum $P + Q = \{p + q : p \in P, q \in Q\}$.

Stanley has found some combinatorial significance to these numbers ($\exists$ a paper with straightforward name on the subject). The proof is a very difficult theorem, even for polytopes. "Not an exercise; I know three proofs none of which are nice and one even uses Hodge theory."

Proof of Theorem 1. We want to show that Alexandrov-Fenchel implies our result. Let $K = K_{\mathcal{P},w} \subset \mathbb{R}^{n-1}$ be defined by if $(z_1, \ldots, z_{n-1}) \in K$,

1. $0 \leq z_i \leq 1$.
2. $z_i \leq z_j$ if $v_i \prec v_j$.
3. $z_i = 0$ if $v_i \prec w$.

Meanwhile, let $L = L_{\mathcal{P},w}$ be the set of points satisfying the first two rules above, but the third rule becomes $z_i = 1$ if $v_i \succ w$.

Easy example: K =square, L =triangle, then what is the sum $K + L$? Then

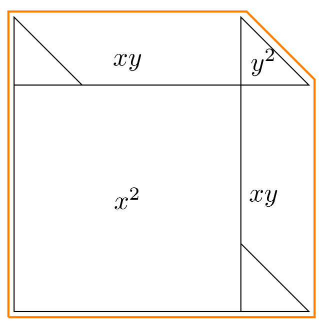


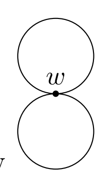
Figure 3: Minkowski sum of triangle and square

$V_0 = area(K), V_2 = area(L), V_1 = area(rectangles)$. Let $Q = \{1, 2\}$, then

$$vol(Q) = \frac{\#linearextensions(\mathcal{P})}{n!}. \tag{2}$$

To see this, split into a bunch of simplices defined by successive inequalities on a permutation of the coordinates, each of which has volume $1/n!$. The second condition on K, L above ensures the permutation must respect poset order, which gives it a direct correspondence to a linear extension of the poset.

These above simplices partition K, L. Since Minkowski sum is additive–$(P_1 \cup P_2) + Q = (P_1 + Q) \cup (P_2 + Q)$, we can split the volume of $xK + yL$ accordingly.



We're supposing for now that w stratifies things above and below Now consider $Q \subset \mathbb{R}^{k+\ell-2}$, with $Q = \Delta_k \times \Delta_\ell$. Determining our simplex looks somehow like a grid walk, which is where the binomial coefficient $\binom{n}{i}$ comes from. To justify this, Igor drew a triangular prism and dissected it into simplices, and related each simplex to a grid walk. The mixed volume turns out to correspond to the number of vertices on top of the shape, versus how many on bottom. "Staircase triangulation" $\square$

Big picture: there is a natural polytope associated to a poset, and its combinatorics give you data about the poset.

1 Gaussian Coefficients

Let $\begin{bmatrix} n \\ k \end{bmatrix}$ be the q-binomial coefficient,

$$\begin{bmatrix} n \\ k \end{bmatrix} = \binom{n}{k}_q = \frac{n!_q}{k!_q(n-k)!_q}. \tag{1}$$

Theorem 1 (Sylvester 1878)**.** *For fixed n, $\begin{bmatrix} n \\ k \end{bmatrix}$ is polynomial with unimodal coefficients.*

For example, consider $n = 4, k = 2$, then

$$\begin{bmatrix} 4 \\ 2 \end{bmatrix} = 1 + q + 2q^2 + q^3 + q^4, \tag{2}$$

and clearly the coefficients are unimodal. Note also that log-concavity is false in general for the q-binomial coefficients.

Sylvester's proof was algebraic, Kathleen O'Hara's proof (which we will discuss today) is a combinatorial/injective proof.

1.1 Greta and Igor's Favorite Quote

In 1852, Cayley conjectured theorem 1. When Sylvester proved it, he wrote "I proved with scarcely any effort something which I thought beyond the realm of human power."

Before we prove this, some basic interpretations of the q-binomial coefficients.

Proposition 1. *We have*

$$\binom{n}{k}_q = \sum_{\lambda \subseteq [k \times n]} q^{|\lambda|}, \tag{3}$$

that is, all partitions λ which fit in the k row by n column box.

This is Proposition 1.7.3 in Stanley (EC1).

Proof. Nothing scary: with basic manipulations we can easily establish the recurrence

$$\binom{n}{k}_q = \binom{n-1}{k-1}_q + q^k \binom{n-1}{k}_q, \tag{4}$$

but one sees immediately that the right hand side also satisfies this recurrence as well, by conditioning on whether $\lambda_k = 0$, and cutting off the leftmost column if it is nonzero. $\square$

Proposition 2. $\binom{n}{k}_q$ *is equal to the number of k-dimensional subspaces of $\mathbb{F}_q^n$.*

This is Proposition 1.7.2 in Stanley.

Proof. Again, a straightforward computation: compute the number of possible bases of a k-subspace, and deal with overcounting by dividing by $|GL(k, q)|$. $\square$

1.2 Proving Theorem 1

We'd love to prove it by induction, but there's a problem: the sum of two unimodal polynomials isn't necessarily unimodal, since their peak points could be extremely far away from each other.

We will still use induction, but we need to be more subtle.

1.3 KOH Identity (Kathleen O'Hara)

We present the following identity:

$$\binom{n+k}{k}_q = \sum_{\nu \vdash k} q^{2n(\nu)} \prod_{i>0} \binom{(n+2)i - 2(\nu_1' + \cdots + \nu_i')}{m_i(\nu)}_q, \tag{5}$$

where $m_i(\nu)$ is the number of i-parts of ν, and

$$n(\nu) = \sum_{i \geq 1} (i-1)\nu_i = \sum_{i>0} \binom{\nu_i'}{2}. \tag{6}$$

Once we prove this, Theorem 1 follows, as it becomes a sum of unimodal polynomials which are all centered on the same point.

1.4 Proof of the KOH Identity

We will slowly work ourselves up to it. First, we present the q-binomial theorem.

Theorem 2 (q-binomial Theorem)**.**

$$\prod_{i=1}^{n-1} (1 + q^i t) = \sum_{k=0}^{n} q^{\binom{k}{2}} \binom{n}{k}_q t^k. \tag{7}$$

This is equation 1.87 in Stanley.

Proof. Trivial induction following from the recurrence previously established in Proposition 1. □

The KOH identity came about when Kathleen O'Hara was trying to prove an old problem:

Conjecture 3. *Lattice partitions $\lambda \subseteq [k \times n]$ have Sperner's property. That is, the inclusion poset can be decomposed into saturated chains.*

Sperner's property implies unimodality, and is stronger than it. Basically, people were hoping there's an injection from $|\lambda| = m$ to $|\lambda| = m + 1$ that just adds a square each time. It is still open today.

2 Next Time

We'll show a different way to prove KOH identity, from MacDonald.

Recall our goal from last time,

Theorem 1. $\binom{n}{k}_q$ *is unimodal.*

Last time we reduced to a main lemma, the KOH inequality:

Lemma 2 (KOH)**.** *We have*

$$\binom{n+k}{k}_q = \sum_{\nu \vdash k} q^{2n(\nu)} \mathcal{F}(n,\nu), \tag{1}$$

where

$$\mathcal{F}(n,\nu) = \prod_{i \geq 1} \binom{m_i + (n+2)i - 2(\nu_1' + \cdots + \nu_i')}{m_i}_q, \tag{2}$$

and

$$n(\nu) = \sum_{i \geq 2} (i-1)\nu_i \tag{3}$$

(be careful to distinguish the two uses of the symbol n here).

Igor re-mentions that this main lemma is enough to prove our theorem: as he and Greta put it, the summands are all unimodal polynomials and you can check that they all have the same mode.

On our quest to prove KOH we'll need some lemmas.

Lemma 3.

$$\binom{n+k}{k}_q = \sum_{r=0}^{k} \frac{q^{nr}}{\varphi_r(q^{-1})\varphi_{k-r}(q)}, \tag{4}$$

where

$$\varphi_k(q) = (1-q)\cdots(1-q^k). \tag{5}$$

Lemma 4.

$$\sum_{\nu \vdash k} q^{2n(\nu)}/b_\nu(q) = \frac{1}{\varphi_k(q)}, \tag{6}$$

where

$$b_\nu(q) = \varphi_{m_1} \cdot \varphi_{m_2} \cdots . \tag{7}$$

Here's how we use Lemmas 3 and 4.

$$\mathcal{F}(n,\nu) = \prod_{i \geq 1} \sum_{r_i=0}^{m_i} \frac{q^{[(n+2)i - 2(\nu_1' + \cdots + \nu_i')]r_i}}{\varphi_{r_i}(q^{-1})\varphi_{m_i - r_i}(q)}, \tag{8}$$

$$q^{2n(\nu)}\mathcal{F}(n,\nu) = \sum_{\alpha \sqcup \beta = \nu} \frac{q^{a(\alpha,\beta)}}{b_\alpha(q^{-1})b_\beta(q)}, \tag{9}$$

where the second equality comes from exchanging sum and product, and if $\alpha = 1^{r_1}2^{r_2}\cdots$ and $\beta = 1^{m_1-r_1}2^{m_2-r_2}\cdots$, then

$$a(\alpha,\beta) = n \cdot |\alpha| - 2n(\alpha) + 2n(\beta) \tag{10}$$

(note we haven't proved this yet).

Now we're looking at

$$\sum_{\nu \vdash k} q^{2n(\nu)} \mathcal{F}(n,\nu) = \sum_{\alpha \sqcup \beta \vdash k} \frac{q^{a(\alpha,\beta)}}{b_\alpha(q^{-1}) b_\beta(q)}, \tag{11}$$

$$= \sum_{r=0}^{k} q^{nr} \left[\sum_{\alpha \vdash r} \frac{q^{-2n(\alpha)}}{b_\alpha(q^{-1})} \right] \left[\sum_{\beta \vdash k-r} \frac{q^{n(\beta)}}{b_\beta(q)} \right], \tag{12}$$

and now we can apply Lemma 4:

$$\sum_{\nu \vdash k} q^{2n(\nu)} \mathcal{F}(n,\nu) = \sum_{r} \frac{q^{nr}}{\varphi_r(q^{-1}) \varphi_{k-r}(q)} \tag{13}$$

$$= \binom{n+k}{k}_q, \tag{14}$$

the last equality being another application of Lemma 3.

"What Zeilberger did is collect KOH's combinatorial proof and show that these identities were the reason why her proof worked." "Lemma 3 is not that bad, but Lemma 4 is."

Proof of Formula for a.

$$a(\alpha,\beta) = 2n(\nu) + \sum [(n+2)i - 2(\nu_1' + \cdots + \nu_i')](\alpha_i' - \alpha_{i+1}') \tag{15}$$

$$= \cdots = n|\alpha| - 2n(\alpha) + 2n(\beta). \tag{16}$$

□

Proof of Lemma 3. Recall from Monday the q-binomial theorem,

$$\prod_{i=1}^{n-1} (1 + q^i t) = \sum_{r=0}^{n} q^{\binom{r}{2}} \binom{n}{r}_q t^r. \tag{17}$$

We make the substitution

$$\binom{n+k}{k}_q = \frac{(1-q^{n+1}) \cdots (1-q^{n+k})}{(1-q)\cdots(1-q^k)} = \frac{\varphi_n(q)}{\varphi_k(q)\varphi_{n-k}(q)}, \tag{18}$$

and apply the binomial theorem to the numerator of this middle expression with $t = -q^{n+1}$. Then we have

$$\binom{n+k}{k}_q = \sum_{r=0}^{k} \frac{q^{\binom{r}{2}} \binom{k}{r}_q (-q^{n+1})^r}{\varphi_k(q)} = \sum_{r=0}^{k} (-1)^r \frac{q^{\binom{r}{2}+(n+1)r}}{\varphi_r(q)\varphi_{k-r}(q)} \tag{19}$$

and with this last term we combine some stuff to turn the denominator term $\varphi_r(q)$ into $\varphi_r(q^{-1})$. □

For the proof of Lemma 4, Igor points to the paper I. Macdonald, An Elementary Proof of a q-Binomial Identity (1989), posted on the course webpage. It can be proven bijectively, or using the q-Chu-VanDerMonde equation.

0.1 The Story of KOH identity

Zeilberger was happy that he understood this identity, so he wrote a paper modulo the proof of KOH. Then he published another paper called "One Line Proof..." which is about 10 pages long and only proves KOH up to 20 or so (?). Zeilberger offered 25 dollars to anyone who could prove KOH in under 2 pages, and MacDonald took up the challenge.

Special guest lecturer Stephen DeSalvo!

1 More on Dickson's Theorem

The theorem:

$$P(\langle\sigma_1, \sigma_2\rangle = S_n or A_n) \to 1. \quad (1)$$

Today though Stephen cares about the lemmas Igor proved along the way.

Let $c_i(n)$ be the number of i-cycles in a random permutation of length n. Let also Z_i be Poisson(λ_i) where $\lambda_i = 1/i$.

Some basic facts about Poisson:

- $P(Y = k) = \frac{\lambda^k e^{-\lambda}}{k!}$
- $P(Y = 0) = e^{-\lambda}$
- $E(Y) = V(Y) = \lambda$.

Theorem 1 (Goncheror 1944, Kolchin 1971)**.**

$$(c_1(n), c_2(n)\dots) \implies (Z_1, Z_2, \dots) \quad (2)$$

(the arrow means "converges in distribution as $n \to \infty$").

Theorem 2 (Shepp Lloyd 1966)**.** *The largest cycle in a random permutation has expected value* (0.62...)n.

Theorem 3 (Arratia, Tarale)**.** *Let d_{TV} be the total variation distance, equal to* $\sup_A |P(X \in A) - P(Y \in A)|$.

$$d_{TV}((c_1(n), \dots, c_b(n)), (Z_1, \dots, Z_b)) \to 0 \iff \frac{b}{n} \to 0. \quad (3)$$

In fact, there is a function F so that $d_{TV}() \leq F(n/b)$.

This is slightly stronger than Theorem 1

1.1 The Lemmas

The above two theorems, and Theorem 3 in particular allow us to prove the lemmas from Igor's earlier lecture with relative ease.

Lemma 4.

$$E(c_j(n)) = \frac{1}{j} \quad (4)$$

Today we're also saying $E(Z_j) = 1/j$, but now we restrict that $j = o(n)$. Stephen explains this restriction: "If we didn't know what we did with Igor, this restriction is necessary to show that $c_j(n) \sim Z_j$."

Lemma 5.

$$E(c_j(n)c_i(n)) \sim E(Z_jZ_i) = E(Z_j) \cdot E(Z_i) = \frac{1}{j}\frac{1}{i}; i, j = o(n) \tag{5}$$

Lemma 6.

$$E(c_j^2(n)) \sim E(Z_j^2) = \lambda_j^2 + \lambda_j = \frac{1}{j^2} + \frac{1}{j}; j = o(n). \tag{6}$$

Igor's version of the following lemma was that $P(\sigma \text{ has no A cycles}) < \frac{1}{\sum \frac{1}{a_i}}$

Lemma 7 (Main Lemma)**.** *Let A be a (functional) k-subset of $[n]$ with each element $o(n)$*

$$P(\sigma \textit{ has no } A \textit{ cycles}) \sim P(Z_{a_i} = 0 \textit{ for each } i) = \prod_{1}^{k} P(Z = 0) = \tag{7}$$

Lemma 8. *Let p be a positive integer. Then*

$$P(\sigma \textit{ has no cycles dividing } p) = \prod_{1}^{n/p\rfloor} \left(1 - \frac{1}{p^i}\right) \tag{8}$$

In our new probabilistic language, letting $p = o(n)$ we have

$$P(\sigma ...) \sim e^{-\sum \frac{1}{p^i}} \tag{9}$$

Lemma 9.

$$P(\sigma \textit{ has only one } p\textit{-cycle} \mid \sigma \textit{ has a } p\textit{-cycle}) \sim P(Z_p = 1 | Z_p \geq 1) = \frac{\frac{1}{p}e^{-\frac{1}{p}}}{1 - e^{-\frac{1}{p}}} \sim 1 - \frac{1}{p}. \tag{10}$$

This is how we'd prove Dixon's Theorem probabilistically.

Why do we need our cycle lengths to be small? Well, large cycle lengths break independence, while small ones will be independent.

Stephen says that if we want to talk about cycles with length $O(n)$, we've got to do something different.

2 Integer Partitions

Let $p(n)$ be the number of partitions of n. Some common mistakes: $np(n) = \sum \sigma(v)p(n-v)$ is an exact statement. But probabilistic statements like $p(n) \sim \frac{e^{\pi\sqrt{2/3}\sqrt{n}}}{4\sqrt{3}n} = HR(n)$ is very vague: "If you try to glean any hard facts from this, you are doomed for failure." In particular this makes the statement $p(n) = HR(n)(1 + o(1))$.

Stephen sees the following statement on the internet: "Is $p(n)$ log-concave?" Answer: yes for $n > 25$. We'd like to use the fact that $HR(n)$ is log-concave. But this is not enough.

Recall that log-concavity is the statement $a_n^2 > a_{n-1}a_{n+1}$, equivalent to

$$-\log a_{n+1} + 2\log a_n - \log a_{n-1} > 0. \tag{11}$$

Online, Mathoverflow saw the following answer accepted:

$$p(n) \sim \frac{e^{c\sqrt{n}}}{4\sqrt{3}n} + O(e^{c\sqrt{n}/2}), \tag{12}$$

then we show that $HR(n)$ is log-concave.

$$p(n) = \frac{e^{c\sqrt{n}}}{4\sqrt{3}n}(1 + e^{-c\sqrt{n}/2}) \tag{13}$$

take the log of both sides:

$$\log p(n) = \log\left(\frac{e^{c\sqrt{n}}}{4\sqrt{3}n}\right) + \log(1 + e^{-c\sqrt{n}/2}), \tag{14}$$

and then we get?

$$logp(n+1) - 2\log p(n) + \log p(n-1) \sim O(n^{-3/2}) + O(e^{-c\sqrt{n}/2}) > 0, \text{ eventually.} \tag{15}$$

The problem lies in the very first statement, equation (12).

The fix?

$$p(n) \sim \frac{e^{c\sqrt{n}}}{4\sqrt{3}n}\left(1 - \frac{1}{c\sqrt{n}}\right) + O(e^{c\sqrt{n}/2}), \tag{16}$$

which adds a $O(n^{-2})$ and that fixes it! Right?

Nope.

Stephen presents us the following fact:

$$p(n) = \frac{d}{n - 1/24} \cdot \sum_{k=1}^{\infty} A_k(n)\left\{\left(1 - \frac{k}{\mu}\right)e^{\mu/k} + \left(1 + \frac{k}{\mu}\right)e^{-\mu/k}\right\}, \tag{17}$$

with $\mu = \sqrt{n - 1/24}$. So in fact we have

$$p(n) = \frac{e^{c\sqrt{n-1/24}}}{4\sqrt{3}(n - 1/24)}\left(1 - \frac{1}{c\sqrt{n - 1/24}}\right) + O(e^{c\sqrt{n}/2}), \tag{18}$$

so unless we include this annoying $1/24$ we can't actually get the exponential bound we claim, and this actually breaks the argument. "Be careful what you read on the internet!"

Special guest lecturer Greta Panova!

1 Longest Increasing Subsequence

Let $\omega \in S_n$ be randomly selected (example 251973468), and let's look at the longest increasing subsequence we can extract from ω (example $1-3-4-6-8$, length 5). Denote by $\ell(\omega)$ the length of the longest increasing subsequence,

$$\ell(\omega) = \max k | i_1 < i_2 < \cdots < i_k and \omega_{i_1} < \cdots < \omega_{i_k}. \tag{1}$$

Similarly, let $\ell'(\omega)$ denote the length of the longest decreasing subsequence. Finially, let $\ell^*(\omega) = \max(\ell, \ell')$ be the length of the longest monotonic subsequence.

So, what can we say about these statistics? S. Ulam thought about this, and Rota mentioned them to Hammersley, who came up with some answers (1968).

1.1 Ulam's Problem

What is the distribution of ℓ, ℓ' for uniform random sampling of permutations $\omega \in S_n$?

In 2000, Baik-Deift-Johansson showed that the distribution is related to the Tracy-Widom distribution, that is

$$P\left(\frac{\ell(w) - 2\sqrt{n}}{n^{1/6}} \leq t\right) \to F(t), \tag{2}$$

where $F(t)$ is the Tracy-Widom distribution, "it has a formula but it's really nasty." In particular, we see that the expectation of $\ell(\omega)$ is $2\sqrt{n}$. We're not going to prove this, instead we're going to stick to the results from Hammersley's paper.

Observe that if $\ell(\omega) = 1$, then $\ell'(\omega) = n$, so they are in some way related.

Theorem 1 (Erdös-Szekeres)**.** *If $n \geq ad + 1$, then $\ell(\omega) > a$ or $\ell'(\omega) > d$.*

Greta describes the following proof algorithm with the example 251973486: we just start at the beginning and try to make an increasing subsequence, and shunt off if we can't. Our process first builds $P_1 = 25$, then $P_2 = 1$, then $P_1 = 259$, and then $P_2 = 17$, next $P_3 = 34$, till we finally end up with

$$P_1 = 259, P_2 = 178, P_3 = 346. \tag{3}$$

If there are fewer than $d+1$ such sequences, then one of them has length greater than a. Conversely if there are $d + 1$ or more, the last element of each forms a decreasing subsequence of our original permutation.

In particular if $n = k^2 + 1$, then either $\ell \geq k + 1$ or $\ell' \geq k + 1$, so $\ell^* \geq k + 1$, and thus

$$\ell^* \geq \sqrt{n}. \tag{4}$$

Theorem 2 (Hammersley)**.** *For $\omega \in S_n$ chosen uniformly at random, we have*

$$\frac{\ell(\omega)}{\sqrt{n}}, \frac{\ell^*(\omega)}{\sqrt{n}} \to^{prob} c, \tag{5}$$

(converges in probability, i.e. $P(|\ell(\omega)/\sqrt{n} - c| \geq \varepsilon) \to 0$ as $n \to \infty$).

The fact that $c = 2$ is attributable to Vershik-Kerov, as well as independently by Logan-Shepp (separated by the iron curtain, as one might guess from the names and time periods).

A formalism: if we have a stochastic process $\omega_{r,s}$ (today, picking random points in the plane), then the following are satisfied (assume $s \geq r$).

1. The process is stationary: $\omega_{r,s} \sim \omega_{r+\delta,s+\delta}$.
2. Expectation only depends on the difference and is finite (which is why it's not just implied by 1): $E[\omega_{r,s}] = g(s-r)$.
3. Subadditive: $\omega_{r,t} \leq \omega_{r,s} + \omega_{s,t}$ for $r < s < t$, or it could be
4. superadditive: same as above with the former inequality flipped.

Together, the second and third facts imply that

$$g(t-r) \leq g(t-s) + g(s-r). \tag{6}$$

This in turn has the consequence

$$\frac{g(x)}{x} \to c = \inf_x \frac{g(x)}{x}. \tag{7}$$

Theorem 3. *If $\omega_{r,s}$ satisfies the above formalism, then*

$$\frac{\omega_{0,t}}{t} \to^{a.s.} c. \tag{8}$$

So, let's build our stochastic process. Take the square and consider the following Poisson process. In the shaded square, look at random points $(x_1, y_1), \ldots, (x_n, y_n)$,

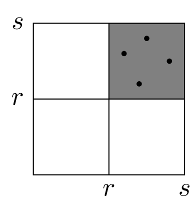


Figure 1: Random points form a permutation.

and read them off in relative order to treat the points like a permutation matrix. Let $\omega_{r,s} = \ell(points)$. We can see from considering the dissected square on r, s, t (Figure 2)that $\omega_{r,t} \geq \omega_{r,s} + \omega_{s,t}$.

Thus we have that $E(\omega_{r,s}) = g(s-r)$, and so

$$\frac{g(s-r)}{s-r} \to c \implies \frac{\omega_{0,t}}{t} \to c. \tag{9}$$

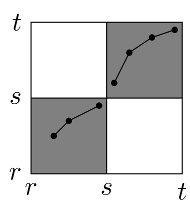


Figure 2: Subadditivity of our process.

1 What Would an Analyst Do?

1.1 Coins in a Fountain

Consider the sequence a_n counting the ways to drop coins into a fountain. This

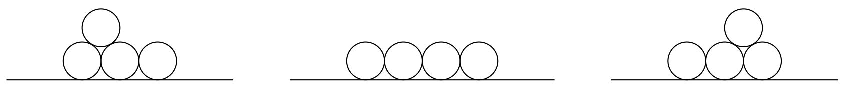


Figure 1: Coins in a fountain: three ways for $n = 4$.

problem is mentioned in 10.7 in Odlyzko, and V.9 in Flajolet-Sedgewick.

Theorem 1. *We have*

$$1 + \sum_{n=1}^{\infty} a_n q^n = \cfrac{1}{1 - \cfrac{q}{1 - \cfrac{q^2}{1 - \cfrac{q^3}{\ddots}}}} = \mathcal{F}(q), \tag{1}$$

the Rogers-Ramanujan continued fraction.

We also consider the extension,

$$\mathcal{F}(z, q) = \cfrac{1}{1 - \cfrac{zq}{1 - \cfrac{zq^2}{1 - \cfrac{zq^3}{\ddots}}}}. \tag{2}$$

Theorem 2. *We have*

$$\mathcal{F}(-1, q) = \prod_{k=1}^{\infty} \frac{(1 - q^{5k-2})(1 - q^{5k-3})}{(1 - q^{5k-1})(1 - q^{5k-4})}. \tag{3}$$

Why are these (in particular the latter theorem) such famous formulas? When Ramanujan first wrote hardy, he wrote with these theorems for the particular value $q = e^{-i\pi/5}$ (?).

Denote by $a_{n,k}$ the number of coin-in-fountain configurations whose bottom row has length k. Then Theorem 1 is equivalent to

Theorem 3.

$$\mathcal{F}(z, q) = \sum a_{n,k} q^n z^k \tag{4}$$

Proof. We will show both sides satisfy the equation

$$\mathcal{F}(z, q) = \frac{1}{1 - zq\mathcal{F}(qz, q)}. \tag{5}$$

For the continued fraction expansion, this is immediate. Consider now the coins in the fountain, and observe the obvious relation to catalan numbers: let $b_{n,k}$

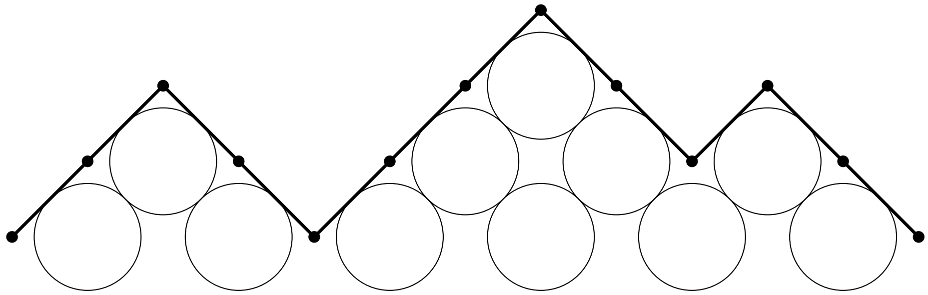


Figure 2: Translating coins to raised Catalan paths (or something like this, anyway).

be the number of Catalan paths from $(0,0)$ to $(2k,0)$ with area n (?). Then it's clear that $a_{n,k} = b_{n,k}$. Then from considering the first time we return to the horizontal (or decomposing on the first "mountain" of coins), we get the equation

$$\mathcal{F}(z,q) = 1 + zq\mathcal{F}(z,q)\mathcal{F}(qz,q), \tag{6}$$

which completes the proof. □

This extra parameter z made all the difference: "there is no easy proof without the z."

Recall that the title of today's lecture is "What Would an Analyst Do?" We treat q as a constant and write

$$\mathcal{F}(z,q) = \frac{A(z)}{B(z)}. \tag{7}$$

We claim that this is a meromorphic function (ratio of two holomorphics), and would like to write a formula for A, B. We use the fraction from Theorem

$$\frac{A(z)}{B(z)} = \frac{1}{1 - zq\frac{A(z)}{B(z)}}, \tag{8}$$

and so we obtain

$$A(z) = B(qz), \tag{9}$$

$$\begin{aligned} B(z) &= B(qz) - qzA(qz) \\ &= B(qz) - qzB(q^2z), \end{aligned} \tag{10}$$

so once we solve for B we'll have A as well.

Let us write B as a Taylor series,

$$B(z) = \sum b_n z^n; \quad b_0 = 1. \tag{11}$$

From equation (10) we obtain that

$$b_n = q^n b_n - q^{2n-1} b_{n-1}, \tag{12}$$

and so we can obtain the closed formula

$$b_n = (-1)^n \frac{q^{n^2}}{(1-q)(1-q^2)\cdots(1-q^n)}. \tag{13}$$

Thus we obtain

$$B(z) = \sum b_n z^n = 1 + \sum_{n=1}^{\infty} \frac{(-1)^n q^{n^2} z^n}{(1-q)\cdots(1-q^n)}, \tag{14}$$

$$A(z) = B(qz) = 1 + \sum_{n=1}^{\infty} \frac{(-1)^n q^{n(n+1)} z^n}{(1-q)\cdots(1-q^n)}. \tag{15}$$

"This is a remarkable result on its own." Igor points out that $\mathcal{F}(-1,1)$ is the golden ratio φ, and if you take the ratio of the partial sums of A, B you get a ratio of two successive Fibonacci numbers. Additionally, $\mathcal{F}(z,1)$ is the generating function for the Catalan numbers.

The following Rogers-Ramanujan formula was proven by Rogers and Schur:

Theorem 4.

$$B(-1) = \prod_{r=1}^{\infty} \frac{1}{(1-q^{5r-1})(1-q^{5r-4})}, \tag{16}$$

$$A(-1) = \prod_{r=1}^{\infty} \frac{1}{(1-q^{5r-2})(1-q^{5r-3})}, \tag{17}$$

so we can understand Rogers-Ramanujan as just taking the ratio of these two expressions.

1 More WWAD

1.1 More Coins in More Fountains

As on Monday, let a_n be the number of ways to place n coins in a fountain. This is equal to

$$\sum_k \#\text{lattice paths } (0,0) \to (2k,0) \text{ with area } n \tag{1}$$

We obtained the following theorems.

Theorem 1.

$$\mathcal{F}(q) = 1 + \sum_{n=1}^{\infty} a_n q^n = \cfrac{1}{1 - \cfrac{q}{1 - \cfrac{q^2}{1 - \cfrac{q^3}{\ddots}}}} \tag{2}$$

Theorem 2 (Rogers Ramanujan identities)**.**

$$\mathcal{F}(q) = \frac{A}{B}, \tag{3}$$

where

$$B(z) = \sum b_n z^n = 1 + \sum_{n=1}^{\infty} \frac{(-1)^n q^{n^2} z^n}{(1-q)\cdots(1-q^n)}, \tag{4}$$

and

$$A(z) = B(qz) = 1 + \sum_{n=1}^{\infty} \frac{(-1)^n q^{n(n+1)} z^n}{(1-q)\cdots(1-q^n)}. \tag{5}$$

1.2 Asymptotics of a_n

First, observe that $a_n > \alpha^n$ for some $\alpha > 1$. This is immediate, since a_n is larger than the number of compositions of n into 1's and 3's, considering only those coin structure of the type in Figure . The generating function for the

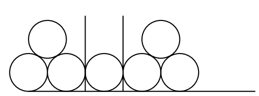


Figure 1: A coin structure representing a composition into parts of size 1 and 3.

number of such compositions b_n is

$$\sum b_n t^n = \frac{1 + c_1 t + c_2 t^2}{1 - t - t^3}, \tag{6}$$

and the denominator has a root at 0.68233, so $b_n \geq 1.46^n$.

Second, we can observe that $a_n \leq \beta^n$, as

$$a_n \leq C_1 + C_2 + \cdots + C_n \leq nC_n < 4^n. \tag{7}$$

Finally, note that $a_{m+n} \geq a_m \cdot a_n$, just by sticking two coin structures of size m, n next to each other. As in Greta's lecture, this actually implies that there is a λ such that

$$\lambda = \lim_{n\to\infty} \frac{\log a_n}{n}. \tag{8}$$

Igor writes that $\lambda = 1.735\ldots$, or $1/\lambda = 0.576148791\ldots$ (Scott asks how many digits we know, Igor says thats actually a complicated question).

Theorem 3. *We have that*

$$a_n = c \cdot \lambda^n + O((1/0.62)^n). \tag{9}$$

This new theorem seems very similar to our theorems from Monday, but it's actually much stronger: even the statement in equation (8) is weak enough that it doesn't forbid things like λ^n/n^ρ, for some positive ρ.

Proof of Theorem. We are going to prove that the function B from Monday "does not have that many zeros." In fact, we shall show that outside a circle of radius 0.62 there is only a zero at $1/\lambda$. Let $A_n(z), B_n(z)$ be the partial sums of

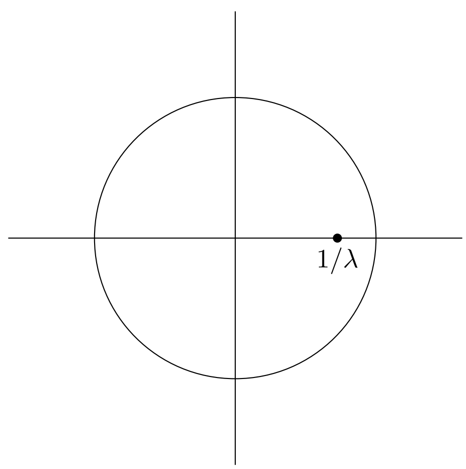


Figure 2: Only one zero inside this small circle of the complex plane.

A, B.

Lemma 4. *There is a unique (simple (?)) zero z_0 of $B(z)$ for $|z| < 0.62$ and $z_0 \approx 0.57$.*

Lemma 5. *We have $A(z_0) > 0$.*

$$B_3(z) = \frac{f(z)}{(1-z)(1-z^2)(1-z^3)}, \tag{10}$$

where $f(z) = 1 - 2z - z^2 + z^3 + 3z^4 + z^5 - 2z^6 - z^7 - z^9$. Define $g(z) \approx f(z)$ by

$$g(z) = \prod_{i=1}^{9} (z - z_i). \tag{11}$$

We can calculate that $z_1 = 0.57577$ (no more digits are needed), and that $z_{2,3} = -0.46997 \pm i0.81792$ (no more digits needed here either) and so on. Each z_i is only an approximation to the roots of f, but it turns out that's all we'll need.

So we have $f(z) = \sum u_k z^k$ and $g(z) = \sum v_k z^k$. Then we claim that

$$\sum_{k=1}^{9} |u_k - v_k| \le 1.7 \cdot 10^{-4}, \tag{12}$$

and so

$$|f(z) - g(z)| \le 1.7 \cdot 10^{-4}, \; |z| < 1. \tag{13}$$

In addition we can calculate that

$$g(z) \ge 8 \cdot 10^{-4}, \; |z| = 0.62. \tag{14}$$

Let $w = |z|$, and consider $0 \le w \le 0.62$. Then

$$\left| \frac{z^{k+1^2 - k^2}}{1 - z^{k-1}} \right| \le \frac{w^{2k+1}}{1 - w^{k+1}} \le \frac{w^9}{1 - w^k}, \; k \ge 5. \tag{15}$$

From this, we obtain that

$$\sum (-1)^k \frac{z^{k^2}}{(1-z^4)\cdots(1-z^k)} \le \frac{w^{19}}{1 - w^4} \tag{16}$$

and

$$\sum_{n=0}^{\infty} \left(\frac{w^4}{1 - w^5} \right)^m \le 7 \cdot 10^{-4} \tag{17}$$

Now we apply Rouchet's theorem: if f, g are holomorphic functions on a region U and $|f(z) - g(z)| < |f(z)| + |g(z)|$ for $z \in \partial U$, then f, g have the same number of roots inside of U. Our above work shows that g is a good enough approximation to f that we can use it to compare with B, and so there is only one zero inside the radius 0.62 disk.

This implies our theorem, for now all other zeros are outside of radius 0.62, hence the specified big-oh term. $\square$

To finish off, here are some calculations that show us :

$$a_{120} = 1.700213368 \cdot 10^{28}, \tag{18}$$

$$a_{120} - c \cdot \lambda^{120} = 1.59 \cdots \cdot 10^9. \tag{19}$$

1.3 Final remarks

Theorems like Theorem 2 are rare. But this type of complex analysis to derive asymptotics is not. Igor mentions that double roots just multiply by n, or other easy things. Furthermore, conformal invariance is useful, as it shows that changing the setting slightly won't change the fundamental behaviour much, because it can't change the nature of the roots.

1 Unimodality of Gauss coefficients via Algebra

Theorem 1 (Sylvester 1878, conjectured by Cayley in 1852)**.** *Let* $\binom{n}{k}_q = \sum a_i q^i$*, then the* a_i *are unimodal.*

Recall that a_i is equal to the number of partitions λ which fit into a $k \times (n-k)$ box.

Theorem 2 (Conjectured by Erdös-Moser in 1963, considered by Lindström in 1969, proved by Stanley in 1978)**.** *Let* $X = \{x_1 < \cdots < x_n\}$ *be a set with* n *elements and each* x_i *real, and let* $S(c) = \{Y \subset X : \sum_{x \in Y} x_i = c\}$*. Then*

$$|S(c)| \leq f(n), \tag{1}$$

where $f(n)$ *is equal to* $S(c)$ *in the case* $Z = \{1, \ldots, n\}$ *and* $c = \lfloor \binom{n+1}{2}/2 \rfloor = \lfloor n(n+1)/4 \rfloor$*.*

In the process of proving it, we will also show that $|S(c)|$ is unimodal for the above set Z.

So for example, if $X = \{1, 3, 2\pi, \pi - 1, \pi + 1\}$ then $|S(2\pi)| = 2$.

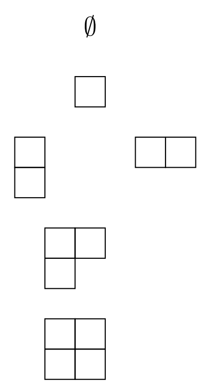
Theorem 3 (Hughes 1977)**.** $\prod(1 + t^i)$ *is unimodal.*

Note that no combinatorial proof of this is known. Stanley used this

1.1 Posets

Consider the poset $\mathcal{L}(m, k)$ be the poset on partitions which fit into a $k \times m$ rectangle, with $\lambda < \mu$ if $\lambda \subset \mu$.

Example :



(2)

Let $M(n)$ be the poset on partitions λ into distinct parts which each part $\leq n$. Example for $n = 3$:

Let $\langle \mathcal{L}(m, k) \rangle$ be the set of $\mathbb{C}$-linear combinations of all λ in the $m \times k$ rectangle. Then

$$\langle \mathcal{L}(m, k) \rangle = \bigoplus_{i=0}^{m \cdot k} \langle \mathcal{L}(m, k) \rangle^i, \tag{3}$$

$$\langle M(m,k)\rangle = \bigoplus_{i=0}^{m\cdot k} \langle M(m,k)\rangle^i, \tag{4}$$

where these are the gradings by $i = |\lambda|$.

We shall define the operators E, F, H on our vector spaces:

- $E : \langle\ \rangle^i \to \langle\ \rangle^{i+1}$
- $F : \langle\ \rangle^i \to \langle\ \rangle^{i-1}$
- $H : \langle\ \rangle^i \to \langle\ \rangle^i$.

The following lemma will be proved next time.

Lemma 4. *We have*

- $[E, H] = 2E$
- $[F, H] = -2F$
- $[E, F] = H$

where $[X, Y] = XY - YX$ *is the standard commutator.*

Today,

Corollary 5. *E has maximal rank.*

Corollary 6. $\dim\langle\ \rangle^i \leq \dim\langle\ \rangle^{i+1}$ *if* $i \leq \dim\langle\ \rangle/2$*, which implies that the dimensions are unimodal.*

"Understanding E is really important." We have that

$$E([\lambda]) = \sum_{\substack{\mu > \lambda, \\ |\mu| = |\lambda|+1}} [\mu], \tag{5}$$

where $[\lambda]$ is just λ as an element of this $\mathbb{C}$-vector space. There are similar formulas for F, H which we can write explicitly. Igor points out we don't really need the formula for H, since $[E, F] = H$.

Igor wants to show that all antichains are smaller than the middle-dimension antichain. I don't think I understand his justification.

Erdös-Moser now comes from here though: associate to each subset Y an element of the vector space. The condition that things add up to c forces us to be considering an antichain, and Igor just argued that the middle antichain is biggest.

1.2 Next Time

Bunch of linear algebra on posets to prove that lemma.

1 Linear Algebra Applications to Unimodality

We'll focus on one of our old problems,

Theorem 1. $\binom{m+n}{m}_q$ *is unimodal as a polynomial in* q.

1.1 Differential Posets

Given a poset and its Young graph $\mathbb{Y}$, let

$$U\lambda = \sum_{\mu=\lambda^+} \mu, \tag{1}$$

$$D\lambda = \sum_{\mu=\lambda^-} \mu. \tag{2}$$

Then if our poset satisfies $UD - DU = I$, we call this a differential poset. Igor claims that partitions in a box is a differential poset. We can add a square on outside corners, or subtract a square on inside corners. Observation: number of outside corners = number of inside corners $+1$.

Proof of Theorem. Recall from last time we introduced the operators E, H, F on

$$\langle \mathcal{L}(m,n) \rangle = \bigoplus_{i=0}^{mn} \langle \mathcal{L}(m,n) \rangle^i \tag{3}$$

where $\mathcal{L}(m,n)$ is the poset on partitions $\lambda \subseteq [m \times n]$.

Note that

$$\binom{m+n}{n}_q = \sum_{i=0}^{mn} \dim \langle \mathcal{L}(m,n) \rangle^i q^i. \tag{4}$$

Denote a partition by $\overline{a} = (a_1, \ldots, a_n)$ where $0 \leq a_1 \leq \cdots \leq a_n \leq m$, writing $\lambda = \overline{a}$ from the bottom up. Then

$$E(\overline{a}) = \sum_{\overline{a}'=(a_1,\ldots,a_i+1,\ldots,a_n)} \overline{a}', \tag{5}$$

$$H(\overline{a}) = (2 \cdot |\overline{a}| - mn) \cdot \overline{a}, \tag{6}$$

and then the difficult one,

$$F(\overline{a}) = \sum_{\overline{a}'=(a_1,\ldots,a_i-1,\ldots,a_n)} (m + n - (a_i + i))(a_i + i) \cdot \overline{a}' \tag{7}$$

(note that in the above sums, we only sum over the things that are still partitions after adding or removing the designated square).

Lemma 2. *We have the following commutator relations.*

- $[E, H] = 2E$

- $[F, H] = -2F$
- $[E, F] = H$.

Proof of Lemma. If we let $k = |\overline{a}|$ we have

$$H(\overline{a}) = (2k - mn)\overline{a}, \tag{8}$$

$$H(E(\overline{a})) = (2k + 2 - mn)(E(\overline{a})), \tag{9}$$

$$E(H(\overline{a})) = (2k - mn)(E(\overline{a})). \tag{10}$$

Therefore we have that when we apply $[H, E] = HE - EH$ to $\overline{a}$ we get

$$[H, E](\overline{a}) = (2k + 2 - mn)(E(\overline{a})) - (2k - mn)(E(\overline{a})) = 2E(\overline{a}), \tag{11}$$

which proves the first desired equation. The proof for $[F, H] = -2F$ is similar. The hard one is showing that $[E, F]$, and Igor meditates aloud about it. Note that our reasoning in Section 1.1 applies (modulo the constants) to E, F as E steps up by one and F steps down by one. Thus we obtain the modification

$$[E, F](\overline{a}) = \Bigg(\sum_{i: a_{i-1} < a_i} ((m+n) - (a_i + i) + 1)((a_i + i) - 1) - \sum_{i: a_i < a_{i+1}} ((m+n) - (a_i + i))(a_i + i) \Bigg) \cdot \overline{a}. \tag{12}$$

Opening this up we get that everything from the subtraction disappears, or something? Anyway it turns out to be $H(\overline{a})$. □

It turns out that we can look instead at E, F, H as matrices in $SL(2, \mathbb{Z})$:

$$E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad F = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{13}$$

Consider as well

$$V = \mathbb{C}\langle x^i y^{k-i}, i = 1, \ldots, k \rangle. \tag{14}$$

In this vector space, we have

$$Ef = x \frac{\partial}{\partial y} f, \tag{15}$$

$$Ff = y \frac{\partial}{\partial x} f, \tag{16}$$

$$Hf = \left(x \frac{\partial}{\partial x} - y \frac{\partial}{\partial y} \right) f. \tag{17}$$

Since all of our formulas for E, F, H are symmetric around $mn/2$, when we decompose into chains we obtain a decomposition of *symmetric* chains. These chains correspond to irreducible representations? Then this implies unimodality (?). □

"Think back to 1828. You're Sylvester and you don't know what $SL(2, \mathbb{Z})$ is, but you still manage to come up with this."

1 One of the Oldest Combinatorial Theorems

"Old combinatorial theorems tend to be trivial, but today an old yet nontrivial one."

1.1 Cayley Compositions (1857)

"Cayley called them x-partitions." Sequences of integers $(a_1, \ldots, a_n)$ such that $1 \leq a_1 \leq 2$, and $1 \leq a_i \leq 2a_{i-1}$ (note if we set $a_0 = 1$ our first thing is a special case of the second).

Theorem 1 (Cayley)**.** *The number of Cayley compositions $c(n)$ of length n is equal to the number of partitions of $2^n - 1$ into parts $\{1, 1', 2, 4, 8, \ldots, 2^{n-1}\}$ (two different colors of 1).*

"Cayley wrote an average of 20 papers per year. Seriously hard-working guy, even when you account for old papers being shorter." We claim that

$$c(n) = \left[t^{2^n-1}\right] \frac{1}{1-t} \prod_{i=1}^{\infty} \frac{1}{1-t^{2^i}}. \tag{1}$$

Igor thinks that the $1, 1'$ business above is annoying, so let's get rid of it:

Definition 1. Let $p(n)$ denote for today the number of partitions of n into $\{1, 2, 4, 8, \ldots\}$.

Then the above equation is saying that

$$c(n) = p(0) + p(1) + p(2) + p(4) + \cdots + p(2^n - 1), \tag{2}$$

and

$$\sum_{n=0}^{\infty} p(n) t^n = \prod_{i=1}^{\infty} \frac{1}{1-t^{2^i}}. \tag{3}$$

Observe that since every number has a unique decomposition into distinct powers of 2, we have

$$\prod_{i=1}^{\infty} \left(1 + t^{2^i}\right) = \frac{1}{1-t}. \tag{4}$$

"Try to figure out the product when the exponent is t^{F_n}, the Fibonacci numbers."

Bijective Proof (Konvalinka, Pak). For the case $n = 2$, we have the compositions

$$(1,1), (1,2), \tag{5}$$
$$(2,1), (2,2), (2,3), (2,4). \tag{6}$$

Consider also the *Cayley partitions* as above of $0 \leq i \leq 3 = 2^n - 1$ for $n = 2$. These are

$$21, 2, 1^3, 1^2, 1, \emptyset \tag{7}$$

Note that there are 6 of each.

We are going to encode the partitions with two numbers. In the above, we correspond a partition to (i, j) with i the number of 2-parts and j the number of 1-parts. Then our partitions become

$$(1,1),(1,0)(0,3)(0,2)etc \tag{8}$$

Define a map

$$\varphi : (a_1, \ldots, a_n) \to (2 - a_1, 2a_1 - a_2, \ldots, 2a_{n-1} - a_n). \tag{9}$$

Claim 2. *The map φ is a bijection from Cayley compositions to Cayley partitions.*

Proof of Claim. Think of the Cayley compositions as points sitting inside some convex polytope (note their defining inequalities are convex).

Note also that for the correspondence in equation (8) we have a sequence $(b_1, \ldots, b_n)$ with $b_i \geq 0$ and

$$2^n - 1 \geq \sum_{i=1}^{n} b_i 2^{n-1}. \tag{10}$$

Combinatorially, our first polytope is a cube. Our second polytope is a simplex. And yet we claim that there is an affine linear map taking the cube's points to the simplex's. "This works, because the convex hull of the integer points (b_i) is not actually a simplex, despite the simplex-like condition on them." More specifically, let B_n be the convex hull of the integer points (b_i), and A_n the same for the points (a_i). Then B_n has the equation

$$0 \leq y_1 \leq 1 \tag{11}$$
$$0 \leq 2y_1 + y_2 \leq 3 \tag{12}$$
$$0 \leq 4y_1 + 2y_2 + y_3 \leq 7 \tag{13}$$
$$\ldots \tag{14}$$
$$0 \leq 2^{n-1}y_1 + \cdots + 2y_{n-1} + y_n \leq 2^n - 1 \tag{15}$$

The claim is now killed by observing that $\varphi : A_n \to B_n$ is a bijection as an affine linear map. $\square$

$\square$

"How do you come up with this proof? Usually there is one or zero affine maps between two polytopes, so it's worth it to look."

1.2 Growth Rate of $c(n)$

What's the growth of $c(n)$? It's got to be at least exponential as $c(n+1) \geq 2c(n)$.

Theorem 3 (Mahler, 1940)**.** *As $n \to \infty$ we have*

$$\log p(n) \sim \frac{\log^2 n}{2 \log 2}. \tag{16}$$

Corollary 4. *The same asymptotic is true for* $\log c(n)$

This last corollary follows immediately from our earlier equation (2).

Proof of Theorem. Note that for n even,

$$p(n) - p(n-1) = p\left(\frac{n}{2}\right), \tag{17}$$

As this left hand side is the number of Cayley partitions with no 1's. Hit this with finite calculus: consider $p(x)$ with

$$p'(x) = p\left(\frac{x}{2}\right). \tag{18}$$

Let

$$p(x) = e^{c \log^2 x}. \tag{19}$$

Then

$$p'(x) = e^{c \log^2 x} \cdot \frac{2c \log x}{x}, \tag{20}$$

and

$$p\left(\frac{x}{2}\right) = e^{c(\log x - \log 2)^2} = e^{c \log^2 x - 2c \log 2 \log x + c(\log 2)^2}. \tag{21}$$

Now the middle term of this last exponent becomes

$$x^{-2c \log 2} = \frac{2c \log x}{x}. \tag{22}$$

Pick

$$c = \frac{1}{2 \log 2}, \tag{23}$$

and we're done. □

Theorem 5 (de Bruijn, 1948)**.**

$$\begin{aligned}\log p(n) = &\frac{1}{2\log 2}\left(\log\left(\frac{n}{\log n}\right)\right)^2 \\ &+ \left(\frac{1}{2} + \frac{1}{\log 2} + \frac{\log\log 2}{\log 2}\right)\log n \\ &- \left(1 + \frac{\log\log 2}{\log 2}\right)\log\log n \\ &+ O(1),\end{aligned}$$

where

$$
\begin{aligned}
O(1) \leftarrow & -\frac{1}{2}\log 2\pi \\
& + \sum_{-\infty}^{\infty} \alpha_k \exp\left\{2\pi i k \frac{\log n - \log\log n + \log\log 2}{\log 2}\right\} \\
& + O\left(\frac{(\log\log n)^2}{\log n}\right),
\end{aligned}
$$

where

$$\alpha_k = \Gamma\left(\frac{2\pi i k}{\log 2}\right)\zeta\left(1 + \frac{2\pi i k}{2\log 2}\right) / \log 2.$$

Then after a while,

Theorem 6 (Knuth,1966)**.**

$$\log p(n) \sim \frac{(\log n)^2}{2\log 2}. \tag{24}$$

A trip back in history, an old nontrivial combinatorial problem.

1 Probléme des ménages

We have $2n$ couples, who go to dinner. They sit at a circular table with $2n$ seats. We want to seat everyone so that no couple sits together. We also want the table to be alternating man-woman-man-etc. Let a_n be the number of such seatings. Let's say that

$$a_n = 2 \cdot n! \cdot b_n, \tag{1}$$

where b_n is the number of $\sigma \in S_n$ such that $\sigma(i) \neq i, i+1 \bmod n$.

This problem is universally attributed to Lucas in 1891. Tate did something similar in knot theory around the same time as well.

An easy exercise: calculate the number of Hamiltonian cycles in $K_{2n} \setminus \{\text{perfect matching}\}$, in terms of b_n.

Theorem 1.1 (Touchard 1934)**.** *We have*

$$b_n = \sum_{k=0}^{n} (-1)^k \frac{2n}{2n-k} \binom{2n-k}{k} (n-k)!, \tag{2}$$

and thus asymptotically,

$$b_n \sim \frac{n!}{e^2}. \tag{3}$$

If we didn't have the $\sigma(i) \neq i+1$ condition we would be looking at derangements, which have asymptotic $\frac{n!}{e}$. This can be viewed as an extension. "In reality, there are better formulas."

Theorem 1.2. $b_n = nb_{n-1} + \frac{1}{n-2}[nb_{n-2} - 4(-1)^n]$.

"Even better:"

Theorem 1.3. $b_n = nb_{n-1} + 2b_{n-2} - (n-4)b_{n-3} - b_{n-4}$.

We won't prove these, but Igor says a computer can do it pretty easily.

Proof of Theorem 1.1. Label the seats $1, 1', 2, 2', \ldots, n, n'$. Let property P_i be that i is in position i'. Let property Q_i be that i is in position $(i+1)'$. Then we apply an inclusion-exclusion:

$$n! - \alpha_1(n-1)! + \alpha_2(n-2)! - \cdots, \tag{4}$$

and now we should figure out these α_k, where α_k is the number of ways to choose k non-adjacent spaces. But this is a standard exercise: for example in Stanley Lemma 2.3.4 he derives

$$a_k = \frac{2n}{2n-k} \binom{2n-k}{k}. \tag{5}$$

"Stanley gives two proofs, one the usual boring one and the other a gorgeous one." Igor likes Stanley's first proof, the linearization method/double counting. $\square$

2 Latin Squares

People were looking at Latin squares before Euler (1782). Why are they called Latin squares? Euler wrote the paper in French, which is very unusual (Igor points out that a student probably wrote it). Since he was looking at double Latin squares, he used a Latin letter and a Greek letter, hence the name. Put down k letters in a square so that each row and column has exactly one of each. Existence is easy: just use ROT-1. Let $L(n)$ be the number of Latin squares of size n.

Theorem 2.1 (van Lint, Wilson). $L(n)^{1/n^2} \sim \frac{n}{e^2}$.

In 1890 Cayley wrote a paper on Latin squares, but this above theorem is recent.

Let $L(n,k)$ be the number of $n \times k$ Latin rectangles, where $n \geq k$ and each column is a permutation, while each row is made of distinct numbers. Then $L(n,2) \sim \frac{(n!)^2}{e}$.

Theorem 2.2 (Riordan). $L(n,3) \sim \frac{(n!)^3}{e^3}$.

Conjecture 2.3. *Let $k = O(n^{6/7})$. Then*

$$L(n,k) = \frac{(n!)^k}{e^{\binom{k}{2}}}. \tag{6}$$

Proof. We determine some bounds on

$$\frac{L(n,k+1)}{L(n,k)} \tag{7}$$

If we start with k columns and we want to add a new one, what are the constraints on our new column? Associate to a $L(n,k)$ Latin square P a square matrix M_p of dimension (?). We do so in a way that each row and column sum to $n-k$. Then our ratio, the number of ways to extend it, is equal to the permanent of M_p.

We quote a theorem: if A is bistochastic, then $Per(A) \geq \frac{n!}{n^n}$. A corollary: if all rows and columns of A sum to $n-k$, then $Per(A) \geq (n-k)^n \frac{n!}{n^n}$.

Thus we get

$$L(n) \geq n! \prod_{k=1}^{n-1} (n-k)^n \frac{n!}{n^n} = \frac{(n!)^{2n}}{n^{n^2}}. \tag{8}$$

This immediately implies the lower bound for our theorem, just by taking the log of both sides and applying Stirling.

For the upper bound: we quote another theorem. If A is a matrix and the rows sum to r_1 through r_n, then

$$Per(A) \leq \prod_{i=1}^{n} (r_i!)^{1/r_i}. \tag{9}$$

Corollary, let all $r_i = n - k$. By similar reasoning we get

$$L(n) \leq \prod_{m=1}^{n} (m!)^{n/m}, \tag{10}$$

and again we take the log of both sides, apply Stirling, and we get to our result. $\square$

1 What we know and what we don't

We've talked about lots of asymptotic results for partitions:

- $\log p(n) = \theta(\sqrt{n})$,
- $p(n) = \frac{c}{n^\alpha} e^{\beta\sqrt{n}} **$

Igor claims that although $p(n)$ has some fantastic formulas, it's not terrible useful. A natural (better) question to ask is to determine

$$P(\lambda_i \geq k), i = \alpha\sqrt{n}, k = \beta\sqrt{n}. \tag{1}$$

It's a limit shape question.

Igor claims that there is a curve γ, given by $1 = e^{-c\alpha} + e^{-c\beta}$, where $c = \frac{\pi}{\sqrt{6}}$, so that the above probability goes to 1 inside γ and goes to 0 outside. In other words, this curve is the shape of a random partition.

Ordinary partitions have $\lambda_i \geq \lambda_{i+1}$. Convex partitions also have $\lambda_i - \lambda_{i+1} \geq \lambda_{i+1} - \lambda_{i+2}$. Our limit shape for convex partitions is significantly different; "the scaling factor is altered"- Stephen Scaling factor becomse $n^{1/3}$ vertically and $n^{2/3}$ horizontally.

We have no nice description, however, for partitions into powers of 2, for example.

1.1 Catalan numbers

We know height, number of peaks, etc. (If we scale horizontally by n and vertically by $\sqrt{n}$, we get a semicircle) But it turns out that we don't know a lot of similar things for coins in a fountain: "the difference is that we fix length for Catalan, and fix area for coins." "I don't think it's impossible, or even difficult, but I don't think anyone's done it."

1.2 Unimodality

We know it for $\binom{n}{k}_q$ in a number of ways: by MacDonald's generating functions using KOH, by linear algebra, and there's a third proof by Pak and Panova using representations of S_n.

Some things we don't know:

- write
$$\binom{n}{k}_q = \sum a_i q^i, \tag{2}$$
then what are the asymptotics of $a_{i+1} - a_i$? "Even small cases are very interesting, it's difficult."

- Is there an analytic proof for unimodality?

Conjecture 1.1 (q-Cat 1)**.** *Let the q-Catalan numbers be*

$$\frac{1}{(n+1)_q}\binom{2n}{n}_q. \tag{3}$$

Then for $n \geq 20$ this is unimodal.

Despite the similarity, it's still not known, because "our methods are not very robust." We could try to use the fact that these numbers can be written as a difference

$$\binom{2n}{n}_q - q^n\binom{2n}{n-1}_q, \tag{4}$$

but there're still complications.

Conjecture 1.2 (q-Cat 2)**.** *Let*

$$c_n(q) = \sum_{P \in Dyck(n)} q^{area(P)}. \tag{5}$$

Then this q-analogue is also unimodal.

1.3 Loop-erased random walks

What people *really* study are LERW on a grid. Can we study the length, territory occupied, Hausdorff dimension of the boundary, etc.? In recent years, one of the big breakthroughs was understanding these guys (Fields medal level work)

But much less is known for Self-avoiding random walks. We conjecture that the behavior is $n^\alpha \cdot \lambda^n$, but nobody has proven it. "People know what α should be, but still nothing."

1.4 Alternating-sign matrices

Number of matrices with entries $0, \pm 1$. Row has to start with 1, end with 1, and instances of ± 1 alternate.

Theorem 1.3 (Zeilberger)**.**

$$a_n = \#ASM's = \prod \frac{}{} \tag{6}$$

We can look at the shape of random ASM's:

I've drawn a circle, but it's actually not, it's some first-order curve. Physicists claim that the points of intersection p, q are not infinitely differentiable.

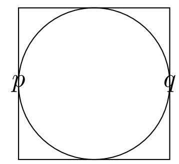
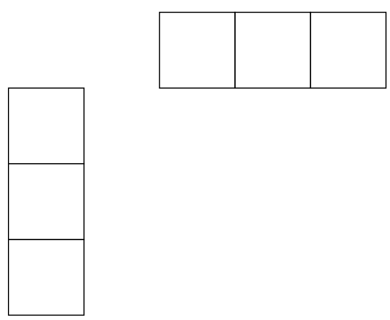


Figure 1: Shape

1.5 Tilings

One can look at a shape like, say, the Aztec diamond. For every point inside, ask what the probability of a horizontal or vertical domino is. One can look for positions where we see controllable behavior: the topmost row is unlikely to have vertical dominoes because it forces everything down one side. You get the "arctic circle" result, where everything outside the circle is frozen for a random tiling of the Aztec diamond. It turns out that for most any growing region, you see an analogue of the arctic circle, which is kind of cool.

What we don't know: let's take trominoes instead:



Can we do the same thing for trominoes? Nobody knows.