Combinatorics

Lecture 1

The Pigeonhole Principle

Let $k, n \in \mathbb{Z}^+$. If more than kn objects are placed into n boxes, then at least one box contains more than k objects. Related result: If exactly kn objects are placed into n boxes and no box has more than/less than k objects in it then every box has exactly k objects in it. Rephrasing, when $k = 1$: if A and B are finite sets of equal size and $f : A \to B$ is a function, then f is one-to-one iff it is onto.

Example 1.1. Suppose $A \subset \{1, 2, \ldots, 2n\}$ with $|A| = n + 1$. Then there are two distinct numbers in A such that one divides the other.

Proof. Write each number $a \in A$ as $a = 2^k m$, where m is odd and $k \in \mathbb{N}$. Since $m \in \{1, 3, 5, \ldots, 2n-1\}$, there are at most n possible values for the odd part m. By the pigeonhole principle there are two numbers in A with the same odd part. The smaller of these divides the larger. □

Example 1.2. Let $a_1, a_2, \ldots, a_n$ be n integers, not necessarily distinct. Then there is a set of consecutive numbers $a_{k+1}, a_{k+2}, \ldots, a_\ell$ (for some $k < \ell$) whose sum is a multiple of n, i.e. $\sum \equiv 0 \bmod n$.

Proof. Let $N = \{0, 1, \ldots, n\}$ and $R = \{0, 1, \ldots, n-1\}$. Define the function $f : N \to R$ with $f(m)$ the remainder of $\sum_{i=1}^{m} a_i \bmod n$ (so $f(0) = 0$). Since $|N| > |R|$, by the pigeonhole principle, there are two values k, ℓ such that $f(k) = f(\ell)$. That is,

$$\sum_{i=1}^{\ell} a_i \equiv \sum_{i=1}^{k} a_i \mod n,$$

so n divides the difference. □

Example 1.3. Chinese Remainder Theorem.
Let m, n be coprime positive integers, and let a, b be integers such that $a \in \{0, 1, \ldots, m-1\}$ and $b \in \{0, 1, \ldots, n-1\}$. Then there exists a positive integer x such that $x \equiv a \bmod m$ and $x \equiv b \bmod n$.

Proof. Define the n integers

$$d_j = a + jm$$

for $j = 0, 1, \ldots, n-1$. Each is congruent to a mod m. First suppose that two of these, say d_i and d_j, have the same remainder when divided by n ($0 \leq i < j \leq n-1$). Then there exists integers $r \in \{0, 1, \ldots, n-1\}$, and q_i, q_j such that $d_i = q_i n + r$, and $d_j = q_j n + r$. Subtracting gives $d_j - d_i = (j - i)m = (q_j - q_i)n$, which is divisible by n. But m and n are coprime, so $n \mid j - i$. This is impossible, as $1 \leq j - i \leq n - 1$. By the pigeonhole principle, all possible remainders occur in $\{d_j \bmod n : j = 0, \ldots, n-1\}$ exactly once; in particular, there is a unique index $p \in \{0, \ldots, n-1\}$ such that $d_p \equiv b \bmod n$. □

Lecture 2

More pigeonhole stuff.

Theorem 2.1. *(The Erdős-Szekeres Theorem.)*
Let m, n be positive integers. In any sequence $a_1, a_2, \ldots, a_{mn+1}$ of $mn+1$ distinct real numbers, there exists an increasing subsequence $a_{i_1} < a_{i_2} < \cdots < a_{i_{m+1}}$ of length $m+1$ (here $i_1 < i_2 < \cdots < i_{m+1}$) or there is a decreasing subsequence $a_{j_1} > a_{j_2} > \cdots > a_{j_{n+1}}$ of length $n+1$ (here $j_1 < j_2 < \cdots < j_{n+1}$).

Proof. For $i = 1, \ldots, mn+1$, let t_i be the length of the longest increasing subsequence starting at a_i (not necessarily contiguous). If $t_i \geq m+1$ for some i, then we are done, so suppose that $t_i \in \{1, \ldots, m\}$ for each $i = 1, \ldots, mn+1$. This defines a function $f : \{1, 2, \ldots, mn+1\} \to \{1, 2, \ldots, m\}$, where $f(i) = t_i$.

By the pigeonhole principle, there exists some $s \in \{1, 2, \ldots, m\}$ such that $f(i) = s$ for at least $n+1$ values of i, say, $f(j_1) = f(j_2) = \cdots = f(j_{n+1}) = s$ where $j_1 < \cdots < j_{n+1}$. Consider a_{j_i} and $a_{j_{i+1}}$, consecutive elements of $a_{j_1}, a_{j_2}, \ldots, a_{j_{n+1}}$. If $a_{j_i} < a_{j_{i+1}}$ then we have an increasing subsequence of length s starting from $a_{j_{i+1}}$ and hence an increasing subsequence of length $s+1$ starting from a_{j_i}, contradicting the fact that $f(j_i) = s$. Therefore

$$a_{j_1} > a_{j_2} > \cdots > a_{j_{n+1}},$$

giving a decreasing subsequence of length $n+1$. □

Proposition 2.1. *(Generalised pigeonhole principle.)*
If more than $a_1+a_2+\cdots+a_t-t$ objects are placed into t boxes, then there is at least one $i \in \{1, 2, \ldots, t\}$ with at least a_i objects in it.

Proof. Otherwise, at most $a_i - 1$ objects in the i^{th} box means there are at most $a_1 + \cdots + a_t - t$ objects in total. □

Double Counting

(Count stuff in two ways.)

A more glorified way to say this:

Suppose that R, C are finite sets and $S \subseteq R \times C$. If $(p, q) \in S$ then we say that P and Q are incident. Let r_p be the number of elements of C incident with $p \in R$. Let c_q be the number of elements of R incident with $q \in C$. Then

$$\sum_{p \in R} r_p = |S| = \sum_{q \in C} c_q.$$

In graph theory we use double counting to prove the "handshaking lemma":

$$\sum_{v \in V} \deg_G(v) = |\{(v, e) : v \in V, e \in E, v \in e\}| = \sum_{e \in E} 2 = 2|E|.$$

Example 2.1. Define the $n \times n$ matrix $A_n = (a_{ij})$ by

$$a_{ij} = \begin{cases} 1 & \text{if } i \mid j, \\ 0 & \text{otherwise,} \end{cases}$$

for positive integers i, j. For example,

$$A_8 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}.$$

This is an upper-triangular 0-1 matrix with 1s on the diagonal. The number of 1s in column j is the number of divisors of j. Denote this number by $t(j)$. This function is quite erratic: $t(j) = 2$ for j prime, and $t(2^k) = k + 1$ for any $k \geq 1$, for example. We want to investigate the average of the first n of these values. Let

$$\bar{t}(n) = \frac{1}{n}\sum_{j=1}^{n} t(j).$$

Note that $n\bar{t}(n)$ is the number of 1s in A_n, and $\sum t(j)$ corresponds to counting these 1s by column. The 1s in row i occur in columns $i, 2i, 3i, \ldots, \left\lfloor \frac{n}{i} \right\rfloor$, so there are $\left\lfloor \frac{n}{i} \right\rfloor$ 1s in row i of A_n. So, by double counting,

$$\bar{t}(n) = \frac{1}{n}\sum_{j=1}^{n} t(j) = \frac{1}{n}\sum_{i=1}^{n}\left\lfloor \frac{n}{i} \right\rfloor \leq \frac{1}{n}\sum_{i=1}^{n}\frac{n}{i} = \sum_{i=1}^{n}\frac{1}{i},$$

The n^{th} harmonic number H_n. Since the error in each summand in passing from $\left\lfloor \frac{n}{i} \right\rfloor$ to $\frac{n}{i}$ is less than 1, it follows that for $n \geq 2$,

$$H_n - 1 < \bar{t}(n) < H_n.$$

But

$$\ln(n) + \frac{1}{n} < H_n < \ln(n) + 1,$$

so

$$\ln(n) - 1 < H_n - 1 < \bar{t}(n) < H_n < \ln(n) + 1.$$

Lecture 3

Extension of Double Counting

If you only have upper bounds on the set S of interest when counting one way, and only lower bounds when counting the other way, then double counting will give you an inequality. We'll now see some applications in *extremal set theory.*

Let $N = \{1, 2, \ldots, n\}$ and let $\mathscr{F} \subseteq 2^N$, i.e. $\mathscr{F}$ is a set of subsets of N. Call $\mathscr{F}$ an antichain if no element of $\mathscr{F}$ is contained in any other. What is the size of the largest antichain? Let $\mathscr{F}_k$ be the set of all k-subsets of N (subsets of size k). Then $\mathscr{F}_k$ is an antichain and $|\mathscr{F}_k| = \binom{n}{k}$.

Fact:

$$\max_k \binom{n}{k} = \binom{n}{\lfloor n/2 \rfloor}$$

(To prove this, find an appropriate set S such that double counting S gives

$$k\binom{n}{k} = (n - k + 1)\binom{n}{k-1}.$$

This implies

$$\binom{n}{k} = \frac{n-k+1}{k}\binom{n}{k-1}$$

for $k \geq 1$, which implies unimodality of binomial coefficients.)

Turns out we can't do better:

Theorem 3.1. *(Sperner's Theorem.)*
The size of a largest antichain of an n-set is $\binom{n}{\lfloor n/2 \rfloor}$.

Proof. Let $\mathscr{F}$ be an antichain. We consider chains of subsets

$$\varnothing = C_0 \subset C_1 \subset C_2 \subset \cdots \subset C_n = N,$$

where $|C_i| = i$ for $i = 0, 1, \ldots, n$. We apply double counting to the set

$$S = \{(A, \mathscr{C}) : \mathscr{C} \text{ is a chain}, C_k = A \text{ for some } k = 0, \ldots, n \text{ and } A \in \mathscr{F}\}.$$

A given chain $\mathscr{C}$ contains at most one element of $\mathscr{F}$, since $\mathscr{F}$ is an antichain. There are exactly $n!$ chains (fix an ordering on N and add elements to the chain in this order). So

$$|S| = \sum_{\text{chains } \mathscr{C}} \left(\# \text{ of elements } A \in \mathscr{F} \text{ which show up in } \mathscr{C}\right) \leq \# \text{ chains} = n!.$$

Next, suppose that there are m_k elements of $\mathscr{F}$ of size k, for $k = 0, 1, \ldots, n$. Let $A \in \mathscr{F}$ have size k. How many chains contain A as an element? ($A = C_k$ for some k?) To build a chain containing A, shove in each element of A first in any order, then pile in the rest of them, giving exactly $k!(n-k)!$ chains. Hence

$$S = \sum \underbrace{m_k}_{\text{number of elements of } \mathscr{F} \text{ of size } k} k!(n-k)!,$$

where the sum is taken from $k = 0, \ldots, n$. By double counting, we conclude that

$$\sum_{k=0}^{n} m_k k!(n-k)! \leq n!.$$

Rearranging gives

$$\sum_{k=0}^{n} \frac{m_k}{\binom{n}{k}} = \sum_{k=0}^{n} m_k \frac{k!(n-k)!}{n!} \leq 1.$$

Replacing each binomial coefficient by the largest one gives

$$\frac{1}{\binom{n}{\lfloor n/2 \rfloor}} \sum_{k=0}^{n} m_k = \frac{|\mathscr{F}|}{\binom{n}{\lfloor n/2 \rfloor}} \leq 1,$$

that is,

$$|\mathscr{F}| \leq \binom{n}{\lfloor n/2 \rfloor},$$

as claimed. □

Now call $\mathscr{F}$ an intersecting family if any two elements of $\mathscr{F}$ have non-empty intersection.

Exercise: Find an intersecting family of size 2^{n-1} — an extremal family (see Prob. Set 1).

Change the question slightly: an intersecting family is called an intersecting k-family if every element of $\mathscr{F}$ has size k. Here we assume $n \geq 2k$, or else it's quite boring.

Example 3.1. Let $\mathscr{F}$ consist of all k-subsets of N which contain 1. Then $\mathscr{F}$ is an intersecting k-family with size $\binom{n-1}{k-1}$. Turns out we can't do better (again):

Theorem 3.2. *(Erdős, Ko, Rado Theorem.)*
The largest size of an intersecting k-family in an n-set is $\begin{pmatrix} n-1 \\ k-1 \end{pmatrix}$ *when* $n \geq 2k$.

Proof. First we prove a seemingly unrelated lemma.

Lemma 3.1. *Let C be a circle divided by n points into n edges. An arc of length k consists of $k+1$ consecutive points and the edges between them. Let $n \geq 2k$ and suppose that we have t distinct arcs $A_1, A_2, \ldots, A_t$ of length k such that any two arcs have an edge in common. Then $t \leq k$.*

Lecture 4

Proof of Lemma 3.1. First we observe that any point on C is the endpoint of at most one arc: if A_i and A_j share an endpoint v then they must proceed in opposite directions, since they are distinct. But $n \geq 2k$, so these arcs cannot have an edge in common, which is a contradiction.

Fix A_1. Since any A_i $(i = 2, \ldots, t)$ has an edge in common with A_1, and all arcs have length k, one of the endpoints of A_i is an "inner point" of A_1. These endpoints must be distinct, but since A_1 has $k-1$ inner points, there can be at most $k-1$ points other than A_1, so $t \leq k$. □

Proof of Theorem 3.2. Let C be a circle with n points and n edges, as above. Let π be a cyclic permutation $\pi = (a_1\ a_2\ \ldots\ a_n)$ of $\{1, 2, \ldots, n\}$. There are $(n-1)!$ such permutations. Given π, we label the *edges* of C with the numbers $a_1, a_2, \ldots, a_n$ clockwise. (Since we choose a cyclic permutation, the starting edge doesn't matter.) We say that $A \in \mathscr{F}$ appears in π if the elements of A appear on C as a block of k consecutive numbers, when C is labelled according to π.

Let $S = \{(A, \pi) : A \in \mathscr{F},\ \pi \text{ is a cyclic permutation of } \{1, 2, \ldots, n\} \text{ and } A \text{ appears in } \pi\}$. We apply double counting. First, we sum over π. Given a cyclic permutation π_0, it follows from Lemma 3.1 that at most k elements of $\mathscr{F}$ can appear in π_0 (since each corresponds to an arc of length k on C, and these all have an edge in common as $\mathscr{F}$ is an intersecting k-family). Hence $|S| \leq (n-1)! \times k$.

Next, we sum over $A \in \mathscr{F}$. For a given $A_0 \in \mathscr{F}$, there are $\underbrace{k!}_{\substack{\text{place elements of } A_0 \text{ with an ordering,} \\ \text{then everything else}}}(n-k)!$ cyclic permutations π for which A_0 appears in π. Hence, by double counting,

$$|\mathscr{F}| k!(n-k)! = |S| \leq k(n-1)!.$$

Therefore,

$$|\mathscr{F}| \leq \frac{k(n-1)!}{k!(n-k)!} = \begin{pmatrix} n-1 \\ k-1 \end{pmatrix},$$

as claimed. □

Lecture 5

Ramsey Theory

How many people do you need to have in a room before you are guaranteed to be able to find either three people who all know each other or three people who are all strangers to each other? We rephrase this in terms of edge colourings of complete graphs.

The complete graph K_n on n vertices has all $\binom{n}{2}$ edges present. An edge colouring is an assignment of colours to the edges. We ask for the smallest positive integer n such that in any red-blue colouring of the edges of K_n, there is a monochromatic triangle (either red or blue). We denote this smallest value of n as $R(3,3)$. (Exercise: $R(3,3) > 5$. Find a red-blue colouring of the edges of K_5 with no monochromatic triangle.)

Lemma 5.1. *$R(3,3) = 6$.*

Proof. It suffices to prove that in any red-blue edge colouring of K_6, there is a monochromatic triangle. Fix an arbitrary red-blue colouring of K_6. Let v be any vertex of K_6. There are 5 edges incident with v, so at least 3 of these edges have the same colouring, by the pigeonhole principle. WLOG, suppose that va, vb and vc are all red. If at least one of the edges ab, ac and bc is red, then adding v gives a red triangle. Otherwise, they are all blue, and abc is a blue triangle. □

Ramsey Numbers

For integers $s, t \geq 2$, let $R(s,t)$ be the smallest positive integer n such that in any red-blue colouring of the edges of K_n, there is iether a red copy of K_s or a blue copy of K_t. Let $R(s,t) = \infty$ if no such n exists. Some true facts: $R(s,t) = R(t,s)$ and $R(s,2) = s$, for all integers $s, t \geq 2$.

Lemma 5.2. *For all integers $s, t \geq 2$, $R(s,t)$ is finite. If $s > 2$ and $t > 2$, then*

$$R(s,t) \leq R(s-1,t) + R(s,t-1) \tag{1}$$

and hence

$$R(s,t) \leq \binom{s+t-2}{s-1}. \tag{2}$$

Proof. We have $R(s,2) = R(2,s) = s$ for all $s \geq 2$. So $R(s,t)$ is finite if $s = 2$ or $t = 2$. Assume by induction that $R(s-1,t)$ and $R(s,t-1)$ are both finite. Let $N = R(s-1,t) + R(s,t-1)$ and fix a red-blue colouring of the edges of K_n. Let x be a vertex of K_n. There are

$$n - 1 = R(s-1,t) + R(s,t-1) - 1$$

edges incident with x, and hence there are either

$$\begin{gathered}\text{either} \geq R(s-1,t) \text{ of red edges incident with } x,\\ \text{or} \geq R(s,t-1) \text{ of blue edges incident with } x.\end{gathered}$$

Without loss of generality, assume the former. Let $n_1 = R(s-1,t)$ and consider a set of n_1 vertices $a_1, \ldots, a_{n_1}$ with xa_i red for $i = 1, \ldots, n_1$. There may be a blue copy of K_t within the edges joining $a_1, \ldots, a_{n_1}$. If so, then we're done. Otherwise, since $n_1 = R(s-1,t)$, there must exist a red copy of K_{s-1} within these edges. Adding x gives a red copy of K_s. Finally, we can check by induction on $(s+t)$ that (1) $\implies$ (2). □

Proposition 5.1. *Lower bound. If $t \geq 2$, then*

$$R(s,t) > (s-1)(t-1).$$

Proof. We arrange $n = (s-1)(t-1)$ vertices into $s-1$ rows and $t-1$ columns. Join two vertices with a blue edge if they belong to the same row, and join them with a red edge otherwise. There is no red K_s because any set of S vertices includes two from a common row, which will have a blue edge between them. Similarly, at most $t-1$ vertices can be chosen from the same row. So, any set of t vertices includes two from distinct rows, and the edge between these is red. Hence there is no blue K_t. (Alternatively, observe largest blue complete subgraph is K_{t-1}). Therefore,

$$R(s,t) > (s-1)(t-1).$$

□

Lecture 6

Some Extensions/Generalisations of Ramsey Theory

1. *"Multicolour Ramsey Theory."*

   Let $R(p_1, \ldots, p_t)$ be the least value of n such that in any colouring of the edges of K_n with t colours, there is a copy of K_{p_i} coloured with colour i, for at least one i.

   (Exercise: When $t \geq 3$ we have

   $$R(p_1, \ldots, p_t) \leq R(p_1, R(p_2, \ldots, p_t)).$$

   It follows from this that $R(p_1, \ldots, p_t)$ is finite for all $t \geq 2$ and all $p_1, \ldots, p_t \geq 2$.

2. We could instead colour k-subsets of $\{1, \ldots, n\}$. (Classical Ramsey Theory corresponds to $k = 2$.)

   Define $R_k(p_1, \ldots, p_t)$ to be the least positive integer n such that in any colouring of the set of $\binom{n}{k}$ k-subsets of $\{1, \ldots, n\}$, there is a set $T \subseteq \{1, \ldots, n\}$ of size p_i such that every k-subset of T is coloured i, for at least one $i \in \{1, \ldots, t\}$. (Won't really look into this.)

   Example 6.1. What is $R_1(p_1, \ldots, p_t)$?
   It is the smallest $n \in \mathbb{Z}^+$ such that in any t-colouring of $\{1, \ldots, n\}$, there is a subset $S \subseteq \{1, \ldots, n\}$ of size p_i coloured i, for at least one $i \in \{1, \ldots, t\}$. By the generalised pigeonhole principle,

   $$R_1(p_1, \ldots, p_t) = p_1 + \cdots + p_t - t + 1.$$

3. "Ramsey-type problem."

   Theorem 6.1. *(Schur's Theorem, in a modern formulation.)*
   Given an integer $t \geq 2$, there exists a positive integer n such that in any colouring of $\{1, \ldots, n\}$ with t colours, there exists $x, y, z \in \{1, \ldots, n\}$, all of the same colour, such that $x + y = z$. (Note, x, y, z do not have to be distinct.)

Proof. Choose n such that $n+1 \geq R(\underbrace{3,3,\ldots,3}_{t \text{ colours}})$, called $r(3;t)$. Let $\chi : \{1,\ldots,n\} \to \{1,\ldots,t\}$ be a given t-colouring of $\{1,\ldots,n\}$. This induces a t-colouring χ^* of the edges of K_{n+1} where the vertices of K_{n+1} are labelled $\{0,1,\ldots,n\}$, defined by

$$\chi^*(i,j) = \chi(|i-j|),$$

for all distinct $i,j \in \{0,1,\ldots,n\}$. By choice of n, there is a monochromatic triangle on i,j,k with $0 \leq i < j < k \leq n$. That is, $\chi^*(i,j) = \chi^*(i,k) = \chi^*(j,k)$. Let $x = j-i, y = k-j$ and $z = k-i$. Note $x,y,z \in \{1,\ldots,n\}$, and $x+y = (j-i)+(k-j) = k-i = z$. Also, $\chi(x) = \chi(y) = \chi(z)$ by definition of χ^*. □

The smallest value of n satisfying the above property is $S(t)$, the Schur number.

4. Schur was originally motivated by Fermat's Last Theorem. What he originally proved was the following:

Theorem 6.2. *For all integers $m \geq 2$, if p is prime and sufficiently large then the equation*

$$x^m + y^m = z^m$$

has a non-zero solution in the integers modulo p.

Proof. Choose p to be sufficiently large such that $p-1 \geq S(m)$ (that is, in any m-colouring of $\{1,\ldots,p-1\}$, there exists $a,b,c \in \{1,\ldots,p-1\}$ with the same colour such that $a+b=c$). Let $\mathbb{Z}_p^* = \{1,\ldots,p-1\}$ with multiplication performed modulo p. Let $H = \{x^m : x \in \mathbb{Z}_p^*\}$. It is a subgroup of $\mathbb{Z}_p^*$ of index $t = \gcd(m,p-1)$. Define a colouring $\chi : \mathbb{Z}_p^* \to \{1,\ldots,t\}$ such that $\chi(a) = \chi(b)$ if and only if $a^{-1}b \in H$. (For example, fix an ordering of hte cosets and let $\chi(a) = i$ if a belongs to the i^{th} coset.)

Then χ is a t-colouring of $\mathbb{Z}_p^*$, and $t \leq m$. By choice of p, there exists $a,b,c \in \{1,\ldots,p-1\}$ such that $\chi(a) = \chi(b) = \chi(c)$ and $a+b=c$. In $\mathbb{Z}_p*$, we can multiply this eqation to conclude that $1 + a^{-1}b = a^{-1}c$. We have $\chi(1) = \chi(a^{-1}b) = \chi(a^{-1}c)$. So $1, a^{-1}b, a^{-1}c \in H$, and hence they are non-zero m^{th} powers in $\mathbb{Z}_p$, showing that the equation $x^m + y^m = z^m$ has a non-zero solution, as required. □

Lecture 7

Inclusion-Exclusion

What we should have seen before:

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

Theorem 7.1. *(Inclusion-Exclusion)*
Let $A_1,\ldots,A_m$ be finite sets, $m \in \mathbb{Z}^+$. Then

$$|A_1 \cup \cdots \cup A_m| = \sum_{i=1}^{n} |A_i| - \sum_{1 \leq i_1 \leq i_2 \leq m} |A_{i_1} \cap A_{i_2}| + \sum_{1 \leq i_1 < i_2 < i_3 \leq m} |A_{i_1} \cap A_{i_2} \cap A_{i_3}| - \ldots$$
$$+ (-1)^{m+1} |A_1 \cap \cdots \cap A_m|.$$

Proof. If $x \notin A_1 \cup \cdots \cup A_m$ then x does not belong to any of the sets $A_{i_1} \cap \cdots \cap A_{i_\ell}$ which appear on the RHS. Hence x contributes 0 to the RHS. Next, we show that if $x \in A_1 \cup \cdots \cup A_m$ then x contributes exactly 1 to the RHS.

Suppose that x belongs to precisely r of the sets $A_1, \ldots, A_m$. Then x will be counted in:

- $\binom{r}{1}$ of the terms $|A_i|$,
- $\binom{r}{2}$ of the terms $|A_{i_1} \cap A_{i_2}|$,
- $\vdots$
- $\binom{r}{s}$ of the terms $|A_{i_1} \cap \cdots \cap A_{i_s}|$, where $1 \leq i_1 < \cdots < i_s \leq m$.

Note that if $s > r$ then the contribution is zero. Hence the total contribution of x to the RHS is

$$\begin{aligned}\binom{r}{1} - \binom{r}{2} + \cdots + (-1)^{r+1}\binom{r}{r} &= \sum_{i=1}^{r}(-1)^{i+1}\binom{r}{i} \\ &= -\sum_{i=1}^{r}(-1)^{i}\binom{r}{i} \\ &= 1 - \sum_{i=0}^{r}(-1)^{i}\binom{r}{i}(1^{r-i}) \\ &= 1 - (1 + (-1))^{r} \\ &= 1.\end{aligned}$$

□

We often want to know the number of elements in some finite set S which satisfy at least one of the given properties $P_1, \ldots, P_m$. Here we let

$$A_i = \{x \in S : x \text{ satisfies } P_i\}$$

for $i = 1, \ldots, m$ and apply Theorem 7.1.

If we instead want to calculate the number of elements of S which satisfy none of the properties $P_1, \ldots, P_m$, then we use the fact that (writing $\overline{B} = S - B$):

$$\begin{aligned}|\overline{A_1} \cap \overline{A_2} \cap \cdots \cap \overline{A_m}| &= |\overline{A_1 \cup \cdots \cup A_m}| \\ &= |S - (A_1 \cup \cdots \cup A_m)| \\ &= |S| - \underbrace{|A_1 \cup \cdots \cup A_m|}_{\text{calculate using Theorem 7.1}}.\end{aligned}$$

If we let $I \subseteq \{1, \ldots, m\}$ and define

$$A_I = \bigcap_{i \in I} A_i,$$

then we can write

$$|\overline{A_1} \cap \cdots \cap \overline{A_m}| = \sum_{I \subseteq \{1,\ldots,m\}} (-1)^{|I|}\, |A_I|.$$

Example 7.1. Find the number of integers in the set $S = \{1, 2, \ldots, 1000\}$ which are divisible by none of $5, 6, 8$.

Let $A_1 = \{x \in S : 5 \mid x\}$, $A_2 = \{x \in S : 6 \mid x\}$ and $A_3 = \{x \in S : 8 \mid x\}$. We want $|\overline{A_1} \cap \overline{A_2} \cap \overline{A_3}|$. Now,

$$\begin{aligned} |A_1| &= \left\lfloor \frac{1000}{5} \right\rfloor = 200, \\ |A_2| &= \left\lfloor \frac{1000}{6} \right\rfloor = 166, \text{ and} \\ |A_3| &= \left\lfloor \frac{1000}{8} \right\rfloor = 125. \end{aligned}$$

Next, pairwise intersections: note $x \in A_1 \cap A_2$ if and only if x is divisible by $\operatorname{lcm}(5, 6) = 30$, for example. So

$$\begin{aligned} |A_1 \cap A_2| &= \left\lfloor \frac{1000}{30} \right\rfloor = 33, \\ |A_2 \cap A_3| &= \left\lfloor \frac{1000}{24} \right\rfloor = 41, \text{ and} \\ |A_1 \cap A_3| &= \left\lfloor \frac{1000}{40} \right\rfloor = 25. \end{aligned}$$

Finally, since $\operatorname{lcm}(5, 6, 8) = 120$, we have

$$|A_1 \cap A_2 \cap A_3| = \left\lfloor \frac{1000}{120} \right\rfloor = 8.$$

By inclusion-exclusion, the number of integers between 1 and 1000 which are divisible by none of $5, 6, 8$ is

$$\begin{aligned} &|S| - (|A_1| + |A_2| + |A_3|) + (|A_1 \cap A_2| + |A_2 \cap A_3| + |A_1 \cap A_3|) - |A_1 \cap A_2 \cap A_3| \\ &= 1000 - 200 - 166 - 125 + 33 + 25 + 41 - 8 \\ &= 600. \end{aligned}$$

Special case: in combinatorics, it often happens that the size of $|A_{i_1} \cap \cdots \cap A_{i_k}|$ depends only on k (not on $i_1, \ldots, i_k$ or $A_{i_1}, \ldots, A_{i_k}$). Then we can define constants

$$\begin{aligned} \alpha_0 &= |S|, \\ \alpha_1 &= |A_1| = |A_2| = \cdots = |A_m|, \\ \alpha_2 &= |A_{i_1} \cap A_{i_2}| \text{ for all } i_1 < i_2, \\ &\vdots \\ \alpha_m &= |A_1 \cap \cdots \cap A_m|. \end{aligned}$$

The inclusion-exclusion formula simplifies to

$$|\overline{A_1} \cap \cdots \cap \overline{A_m}| = \sum_{k=0}^{m} (-1)^k \alpha_k \binom{m}{k}.$$

Example 7.2. How many integers in $S = \{0, 1, \ldots, 99999\}$ have among their digits each of $2, 5$ and 8?

We consider every element of S to be a 5 digit number, by prepending leading zeros if necessary. Let

$$\begin{aligned} A_1 &= \{x \in S : 2 \text{ is not a digit of } x\}, \\ A_2 &= \{x \in S : 5 \text{ is not a digit of } x\}, \text{ and} \\ A_3 &= \{x \in S : 8 \text{ is not a digit of } x\}. \end{aligned}$$

We want $|\overline{A_1} \cap \overline{A_2} \cap \overline{A_3}|$. We calculate, with notation from above:

$$\begin{aligned} \alpha_0 &= 10^5, \\ \alpha_1 &= 9^5, \\ \alpha_2 &= 8^5, \text{ and} \\ \alpha_3 &= 7^5. \end{aligned}$$

So the answer is

$$\alpha_0 - 3\alpha_1 + 3\alpha_2 - \alpha_3 = 10^5 - 3 \cdot 9^5 + 3 \cdot 8^5 - 7^5.$$

Lecture 8

Derangements

Definition 8.1. A *derangement* is a permutation $\pi \in S_n$ such that $\pi(i) \neq i$ for all $i = 1, \ldots, n$.

Let D_n denote the number of derangements of the set $\{1, \ldots, n\}$. Then $D_1 = 0$, $D_2 = 1$ and $D_3 = 2$ (the 3-cycles). (To find a derangement, we need a partition of n into a sum of positive integers, all *greater* than 1.)

Theorem 8.1. *For $n \geq 1$,*

$$D_n = n!\left(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!}\right).$$

Proof. For $j = 1, \ldots, n$, let

$$A_j = \{\sigma \in S_n : j \text{ is a fixed point of } \sigma\}.$$

We calculate

$$D_n = |\overline{A_1} \cap \cdots \cap \overline{A_m}|$$

using inclusion-exclusion. Consider A_1. If $\sigma \in A_1$, then

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ 1 & \sigma(2) & \sigma(3) & \ldots & \sigma(n) \end{pmatrix},$$

where $\{\sigma(2), \ldots, \sigma(n)\} = \{2, \ldots, n\}$. Hence $|A_1| = (n-1)!$. More generally, $|A_j| = (n-1)!$ for all $j = 1, \ldots, n$. Now, suppose $\sigma \in A_1 \cap A_2$. Then

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ 1 & 2 & \sigma(3) & \ldots & \sigma(n) \end{pmatrix}.$$

Then $\{\sigma(3), \ldots, \sigma(n)\} = \{3, \ldots, n\}$, so $|A_1 \cap A_2| = (n-2)!$. Again, $|A_{i_1} \cap A_{i_2}| = (n-2)!$ for all $i_1 < i_2$. More generally,

$$|A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_s}| = (n-s)!$$

for all $i_1 < i_2 < \cdots < i_s$ — specify s fixed points and permute the rest of the symbols. Hence, by the "special case" of inclusion-exclusion, we have

$$\begin{aligned} D_n &= \sum_{k=0}^{n} (-1)^k (n-k)! \binom{n}{k} \\ &= \sum_{k=0}^{n} (-1)^k \frac{n!}{k!} \\ &= n! \sum_{k=0}^{n} \frac{(-1)^k}{k!}. \end{aligned}$$

□

Recall that

$$e^{-1} = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} = \frac{D_n}{n!} + \sum_{k=n+1}^{\infty} \frac{(-1)^k}{k!}.$$

True fact about alternating series: truncating an infinite alternating series with strictly decreasing terms (in absolute value) gives an error term of at most the absolute value of the first omitted term.

In our case, this gives

$$-\frac{1}{(n+1)!} \leq e^{-1} - \frac{D_n}{n!} \leq \frac{1}{(n+1)!}.$$

Hence

$$-\frac{1}{n+1} \leq \frac{n!}{e} - \frac{D_n}{\leq} \frac{1}{n+1}.$$

Since D_n is an integer, we conclude that D_n equals the closest integer to $\frac{n!}{e}$.

Note $\frac{D_n}{n!}$ is the proportion of elements of S_n which are derangements. This proportion tends to e^{-1}. We can also say that $\frac{D_n}{n!}$ is the probability that a permutation, chosen uniformly at random, is a derangement.

Lecture 9

Möbius Inversion

Inclusion-exclusion (also called the "sieve method") is a special case of Möbius inversion on a finite partially ordered set (also called a poset).

Quick Revision on Posets

Definition 9.1. A *relation* R on a set S is a set of ordered pairs $R = \{(a, b) : a, b \in S\}$.

Write $a\ R\ b$ if $(a, b) \in R$, a slight abuse of notation. Important properties of a relation R:

- R is *reflexive* if $(a, a) \in R$ for all $a \in S$.
- R is *symmetric* if $(a, b) \in R \implies (b, a) \in R$ for all $a, b \in S$.
- R is *antisymmetric* if $(a, b) \in R$ and $(b, a) \in R \implies a = b$ for all $a, b \in S$.
- R is *transitive* if $(a, b) \in R$ and $(b, c) \in R \implies (a, c) \in R$ for all $a, b, c \in S$.

Definition 9.2. If R is reflexive, symmetric and transitive, then R is an equivalence relation. If R is reflexive, antisymmetric and transitive, then R is a partial order.

Example 9.1. On a fixed set of integers, say, $\{2, 4, 6, 9, 12, 18\}$, divisibility is a partial order.

A poset can be represented using a Hasse diagram. I suck too much to draw these, go to

`http://en.wikipedia.org/wiki/Hasse_diagram`

if you're keen.

Definition 9.3. An element $x \in S$ is *minimal* in $(S, \leq)$ if there is no $y \in S$ such that $y \leq x$. It is *minimum* if $x \leq y$ for all $y \in S$.

First glimpse of Möbius Inversion

Fix $n \in \mathbb{Z}^+$ and let $N = \{1, 2, \ldots, n\}$. Then $(2^N, \subseteq)$ is a partially ordered set.

Let $F = 2^N \to \mathbb{R}$ be a real valued function on 2^N and define $G : 2^N \to \mathbb{R}$ by

$$G(B) = \sum_{A \subseteq B} F(A) \tag{3}$$

for all $B \in 2^N$. Möbius inversion will allow us to invert (3) to write in terms of G:

$$F(B) = \sum_{A \subseteq B} (-1)^{|B|-|A|} G(A) \tag{4}$$

for all $B \in 2^N$.

Note, (3) and (4) differ only by signs of each term. We prove this later, but first we look at a consequence. Let $A_1, \ldots, A_n$ be subsets of a finite set S, and for $K \in 2^N$ let $F(K)$ be the number of elements of S that belong to only those A_i with $i \notin K$. Then

$$G(K) = \sum_{L \subseteq K} F(L)$$

counts the number of elements of S that belong to all A_i with $i \notin K$, and possibly some others. We call this set

$$A_{\overline{K}} = \bigcap_{i \in \overline{K}} A_i.$$

So

$$G(K) = \left|\bigcap_{i\notin K} A_i\right|.$$

Using (4) we conclude that

$$F(K) = \sum_{L\subseteq K} (-1)^{|K|-|L|} G(L).$$

Taking $K = N = \{1, \ldots, n\}$ in this formula gives

$$F(N) = \sum_{L\subseteq N} (-1)^{n-|L|} G(L),$$

and $F(N) = |\overline{A_1} \cap \cdots \cap \overline{A_n}|$, which is the number of elements of S that only belong to A_i where $i \notin N$; that is, to none of them. Therefore,

$$\begin{aligned} |\overline{A_1} \cap \cdots \cap \overline{A_n}| &= \sum_{L\subseteq N} (-1)^{n-|L|} |A_I| \qquad \left(\text{here } A_I = \bigcap_{i\notin L} A_i\right) \\ &= \sum_{J\subseteq N} (-1)^{|J|} |A_J|, \qquad (\text{letting } J = I), \end{aligned}$$

which is the familiar inclusion-exclusion formula.

We wish to generalise further to arbitrary partial orders.

True fact: Every partial order $(X, \leq)$ on a finite set X has a linear extension $\leq^*$; that is, a linear ordering

$$x_1 \leq^* x_2 \leq^* \cdots \leq^* x_n,$$

where $n = |X|$, such that if $x \leq y$ in the original partial order, then $x \leq^* y$ ($\leq^*$ respects $\leq$).

Sketch proof. Choose a minimal element of X to be x_1. Let $X_1 = X - \{x_1\}$. Repeatedly choose x_j to be a minimal element of X_{j-1} and let $X_j := X_{j-1} - \{x_j\}$. This gives a linear extension of $(X, \leq)$. □

Example 9.2. A linear extension of the partial order in Example 9.1 is $9, 2, 6, 4, 18, 12$ (check).

True fact about the true fact: The Hasse diagram of a linear extension is a line.

Given a finite poset $(X, \leq)$, let

$$\mathscr{F}(X) = \{f : X \times X \to \mathbb{R} : f(x, y) = 0 \text{ if } x \not\leq y\}.$$

Exercise: If $n = |X|$ and $f \in \mathscr{F}(X)$, then f can be represented by an $n \times n$ upper-triangular real matrix A_f.
Hint: index the rows and columns of A_f by X, ordered according to any linear extension of $(X, \leq)$.

Definition 9.4. The *convolution product* of $f, g \in \mathscr{F}(X)$, denoted by $h = f * g$, is defined by

$$h(x, y) = \begin{cases} \displaystyle\sum_{\substack{z\in X,\\ x\leq z\leq y}} f(x, z) g(z, y) & \text{if } x \leq y, \\ 0 & \text{otherwise.} \end{cases}$$

Then $h \in \mathscr{F}(X)$.

Exercise: $A_h = A_f A_g$ when all these matrices are defined with respect to the same linear extension of $(X, \leq)$.

Lecture 10

Three special functions. Given a finite poset $(X, \leq)$, fix a linear extension.

1. The Kronecker delta:
$$\delta(x,y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{otherwise.} \end{cases}$$
Note that $\delta \in \mathscr{F}(X)$, and the matrix $A_\delta = I$.

2. The zeta function, ζ, is defined by
$$\zeta(x,y) = \begin{cases} 1 & \text{if } x \leq y \\ 0 & \text{otherwise.} \end{cases}$$
Note, $\zeta \in \mathscr{F}(X)$.

   True fact: The set of all invertible upper-triangular real matrices (of size $n \times n$, where $n = |X|$) forms a group.

   Hence we can find an inverse for $f \in \mathscr{F}(X)$ if and only if A_f is invertible, that is, $f(x,x) \neq 0$ for all $x \in X$.

   Assume that $f \in \mathscr{F}(X)$ such that $f(x,x) \neq 0$ for all $x \in X$. Inductively define $g \in \mathscr{F}(X)$ by:
$$g(y,y) = \frac{1}{f(y,y)} \qquad \text{for all } y \in X,$$
and then, if $y \neq x$,
$$g(x,y) = -\sum_{\substack{z \in X, \\ x \leq z < y}} g(x,z) \frac{f(z,y)}{f(y,y)} \qquad \text{for all } x < y,$$
and $g(x,y) = 0$ if $x \not\leq y$. (Note in the above sum, $z < y$ means $z \leq y$ and $z \neq y$.) Then $g \in \mathscr{F}(X)$. This implies that when $x < y$,
$$\sum_{\substack{z \in X, \\ x \leq z \leq y}} g(x,z) f(z,y) = 0,$$
and if $x = y$, then
$$\sum_{\substack{z \in X, \\ x \leq z \leq y}} g(x,z) f(z,y) = g(x,x) f(x,x) = 1,$$
by definition. Hence $g * f = \delta$, so g is the unique inverse of f with respect to the convolution product, and $A_g = A_f^{-1}$, so we can find g from A_f using matrix inversion.

3. The Möbius function $\mu \in \mathscr{F}(X)$ is the inverse of the zeta function ζ with respect to the convolution product. (Note, ζ is invertible because $\zeta(x,x) = 1$ for all $x \in X$, as $(X, \leq)$ is reflexive.) The matrix A_μ of μ is A_ζ^{-1}. Since $\mu * \zeta = \delta$, we have
$$\delta(x,y) = \sum_{\substack{z \in X, \\ x \leq z \leq y}} \mu(x,z) \zeta(z,y) = \sum_{\substack{z \in X, \\ x \leq z \leq y}} \mu(x,z)$$

by definition of ζ. Hence

$$\mu(x,x) = 1 \tag{5}$$

for all $x \in X$, and

$$\mu(x,y) = -\sum_{\substack{z \in X,\\ x \leq z < y}} \mu(x,z) \tag{6}$$

for all $x < y$. To calculate μ, use (5) and (6) inductively, or work from the matrices (if we have a concrete poset).

Example 10.1. We compute the Möbius function of the poset $(2^N, \subseteq)$ where $N = \{1, 2, \ldots, n\}$. Let $A, B \in 2^N$ with $A \subseteq B$. We prove by induction on $|B| - |A|$ that

$$\mu(A,B) = (-1)^{|B|-|A|}. \tag{7}$$

Proof. We know $\mu(A,A) = 1$, so (7) holds when $B = A$. Assume now that $B \neq A$, and let $P = |B - A| = |B| - |A|$. From (6) and the induction hypothesis,

$$\begin{aligned}
\mu(A,B) &= -\sum_{\substack{S \in 2^N,\\ A \subseteq S \subset B}} \mu(A,S) \\
&= -\sum_{\substack{S \in 2^N,\\ A \subseteq S \subset B}} (-1)^{|S|-|A|} \quad \text{by induction hypothesis} \\
&= -\sum_{k=0}^{P-1} (-1)^k \binom{P}{k} \quad \text{as } S = A \cup (S - A) \text{ with } k = |S - A| \\
&= (-1)^P \binom{P}{P},
\end{aligned}$$

since

$$\sum_{j=0}^{P} (-1)^j \binom{P}{j} = (1 + (-1))^P = 0.$$

That is, $\mu(A,B) = (-1)^P \binom{P}{P} = (-1)^P = (-1)^{|B|-|A|}$, as required. □

Theorem 10.1. *(Möbius Inversion.)*
Let $(X, \leq)$ be a finite poset with a smallest element, and let μ be the Möbius function of $(X, \leq)$. Given $F : X \to \mathbb{R}$, define $G : X \to \mathbb{R}$ by

$$G(x) = \sum_{\substack{z \in X,\\ z \leq x}} F(z)$$

for all $x \in X$. Then

$$F(x) = \sum_{\substack{y \in X,\\ y \leq x}} G(y)\mu(y,x)$$

for all $x \in X$.

Proof. Fix a linear extension of $(X, \leq)$ and let A_μ be the matrix of μ with respect to this linear extension. Also, let v_F, v_G be $1 \times n$ row vectors corresponding to F and G respectively (with respect to the same linear ordering of X). Here $n = |X|$. The definition of G can be rewritten as

$$v_G = v_F A_\zeta,$$

where ζ is the zeta function for $(X, \leq)$. Hence $v_F = v_G A_\zeta^{-1} = v_G A_\mu$, completing the proof.

Exercise: Check that when specialised to the poset $(2^N, \subseteq)$, we get the formula

$$F(B) = \sum_{A \subseteq B} (-1)^{|B|-|A|}\, G(A)$$

for all $B \in 2^N$, stated earlier. □

Revision Stuff

	order matters	order does not matter
with replacement	n^k, number of sequences in $\{1, \ldots, n\}^k$.	(*)
without replacement	$n(n-1)\cdots(n-k+1) = (n)_k$, counts sequences in $\{1, \ldots, n\}^k$ with distinct entries.	$\binom{n}{k}$, number of subsets of $\{1, \ldots, n\}$ of size k.

For (*): number of multisets of k elements using n distinct elements. Let x_i be the multiplicity of i in the multiset of size k; we count solutions to $x_1 + \cdots + x_n = k$, $x_i \geq 0$. Picture k stars, $n-1$ bars:

$$* \;*\; * \,|\,|\, * \cdots * \,|\, * \cdots *.$$

Choose positions of bars (or stars) to get $\binom{n+k-1}{n-1} = \binom{n+k-1}{k}$ choices.

Lecture 11

Example 11.1. Consider an $n \times n$ chessboard, with some positions marked as forbidden. Calculate the number of ways to place n non-attacking rooks on the board, avoiding forbidden positions.

Store the allowable positions in an $n \times n$ binary matrix, $A = (a_{ij})$, defined by

$$a_{ij} = \begin{cases} 1 & \text{if position } (i,j) \text{ is allowed,} \\ 0 & \text{if position } (i,j) \text{ is forbidden.} \end{cases}$$

Choosing n non-attacking rooks in allowed positions corresponds to choosing n "1"s in A, no two in the same row or column.
For example, with

$$A = \begin{pmatrix} 0 & \boldsymbol{1} & 0 & 1 \\ \mathit{1} & 1 & \mathbf{1} & 0 \\ 1 & 0 & \mathit{1} & \mathbf{1} \\ \mathbf{1} & 1 & 0 & \mathit{1} \end{pmatrix},$$

the italic "1"s correspond to one choice, the bold to another. These correspond to permutations $(\mathbf{1\ 2\ 3\ 4})$ and $(\mathit{1\ 2})(\mathit{3})(\mathit{4})$.

Hence we can count all allowed configurations by calculating

$$\sum_{\sigma \in S_n} a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} = \begin{cases} 1 & \text{if } a_{i\sigma(i)} = 1 \text{ for all } i, \\ 0 & \text{otherwise.} \end{cases}$$

We call this the *permanent* of the matrix A (call this $\operatorname{per}(A)$). (Here, $\sigma \in S_n$ is defined by $\sigma(i) = j$ if and only if entry (i, j) is chosen as the unique position of the rook in row i.]]

Note, the permanent of a binary $n \times n$ matrix also counts the number of perfect matchings in the bipartite graph with adjacency matrix A. Also, $\det(A)$ has the same formula, but with $\operatorname{sgn}(\sigma) \in \{-1, 1\}$ factor on each term. $O(n^3)$ to compute, #P-complete.

Consider the poset $(2^N), \subseteq)$ where $N = \{1, 2, \ldots, n\}$. Each set $S \in 2^N$ with size $|S| = k$ selects a set of K columns of A. Denote the $n \times k$ submatrix of A formed by these columns as $A[S]$. Let $\mathscr{F}_n(S)$ be the set of all functions $f : N \to S$ and let $G_n(S)$ denote the set of surjective (onto) functions in $\mathscr{F}_n(S)$.

Then

$$\mathscr{F}_n(S) = \bigsqcup_{T \subseteq S} G_n(T) \tag{8}$$

Define $F : 2^N \to \mathbb{R}$ by

$$F(S) = \sum_{f \in G_n(S)} \prod_{i=1}^{n} a_{if(i)}$$

for all $S \in 2^N$. Here $F(\varnothing) = 0$. Since permtations are onto surjective functions $N \to N$, we see that our desired answer, $\operatorname{per}(A) = F(A)$.

Define

$$G(S) = \sum_{T \subseteq S} F(T)$$

for all $S \in 2^N$. Then

$$\begin{aligned} G(S) &= \sum_{T \subseteq S} \sum_{f \in G_n(T)} \prod_{i=1}^{n} a_{if(i)} \\ &= \sum_{f \in \mathscr{F}_n(S)} \prod_{i=1}^{n} a_{if(i)} \quad \text{by (8)} \\ &= \prod_{i=1}^{n} \sum_{j \in S} a_{ij}. \end{aligned}$$

By Möbius inversion, we have

$$\begin{aligned} F(N) &= \sum_{S \subseteq N} (-1)^{n-|S|} G(S) \\ &= \sum_{S \subseteq N} (-1)^{n-|S|} \prod_{i=1}^{n} \left(\sum_{j \in S} a_{ij} \right). \end{aligned}$$

Note: There are 2^n terms, so this is not computationally efficient. But it does only require simple calculations.

Direct Product Construction for Finite Posets

Let $(X, \leq_1)$ and $(Y, \leq_2)$ be finite posets, and define $\leq$ on $X \times Y = \{(x, y) : x \in X, y \in Y\}$ by

$$(x, y) \leq (x', y') \iff x \leq_1 x' \text{ and } y \leq_2 y'.$$

We can generalise to the direct product of any finite number of finite posets.

Theorem 11.1. *Let $(X, \leq_1)$ and $(Y, \leq_2)$ be two finite posets with Möbius functions μ_1 and μ_2 respectively. Then*

$$\mu((x, y), (x', y')) = \mu_1(x, x')\mu_2(y, y')$$

for all $(x, y), (x', y') \in X \times Y$.

Proof. If $(x, y) \not\leq (x', y')$ then either $x \not\leq x'$ or $y \not\leq y'$, which implies that $\mu_1(x, x') = 0$ or $\mu_2(y, y') = 0$, respectively. In either case we have $\mu((x, y), (x', y')) = 0$, as required.

If $(x, y) = (x', y')$ then $x = x'$ and $y = y'$. So $\mu_1(x, x') = 1$ and $\mu_2(y, y') = 1$. Therefore $\mu((x, y), (x', y')) = 1 \times 1 = 1$ whenever $(x, y) = (x', y')$.

Now assume that $(x, y) \neq (x', y')$, and proceed by induction.

$$\begin{aligned}
\mu((x, y), (x', y')) &= -\sum_{\substack{(u,v) \in X \times Y, \\ (x,y) \leq (u,v) < (x',y')}} \mu((x, y), (u, v)) && \text{by inductive definition of Möbius function} \\
&= -\sum_{\substack{(u,v) \in X \times Y, \\ (x,y) \leq (u,v) < (x',y')}} \mu_1(x, u)\mu_2(y, v) && \text{by inductive hypothesis} \\
&= -\left(\sum_{\substack{u \in X, \\ x_1 \leq_1 u \leq_1 x'}} \mu_1(u, x') \right) \left(\sum_{\substack{v \in Y, \\ y \leq_2 v \leq_2 y'}} \mu_2(v, y') \right) \\
&= -0 \times 0 + \mu_1(x, x')\mu_2(y, y') \\
&= \mu_1(x, x')\mu_2(y, y')
\end{aligned}$$

as required. □

More generally, the Möbius function of a direct product of a finite number of posets is the product of the corresponding Möbius functions.

Lecture 12

Example 12.1. Let $N = \{1, 2, \ldots, n\}$, where $n \in \mathbb{Z}^+$. Consider the poset $D_n = (N, |)$ given by divisibility. We want to compute the Möbius function $\mu(1, n)$ for this poset.

True fact: If $a, b \in \mathbb{N}$ and $a|b$ then $\mu(a, b) = \mu(1, \frac{b}{a})$.

Now, n has a unique prime power factorisation $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$ where $p_1 < p_2 < \cdots < p_k$ are distinct primes and $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{Z}^+$.

Since

$$\mu(1, n) = -\sum_{\substack{m \in N, \\ m \mid n, \\ m \neq n}} \mu(1, m)$$

for all $n \geq 2$. We need only consider $(N^*, |)$ where $N^* = \{k \in N : k \mid n\}$. Let $r, s \in N^*$. Then $r = p_1^{\beta_1} p_2^{\beta_2} \cdots p_k^{\beta_k}$ and $s = p_1^{\gamma_1} p_2^{\gamma_2} \cdots p_k^{\gamma_k}$, where $\beta_i, \gamma_i \in \{0, 1, \ldots, \alpha_i\}$ for $i = 1, \ldots, k$ (prime factorisations of r and s). Therefore $r \mid s$ if and only if $\beta_i \leq \gamma_i$ for $i = 1, \ldots, k$.

Hence $(N^*, |)$ is precisely the direct product of k linear orders of sizes $\alpha_1 + 1, \ldots, \alpha_k + 1$ respectively. Hence, by the direct product theorem,

$$\mu(1, n) = \prod_{i=1}^{k} \underbrace{\mu(1, p_i^{\alpha_i})}_{\substack{\text{divisibility poset of } ((p_i^{\alpha_i})^*, |) \\ \text{isomorphic to linear order on } \{0, \ldots, \alpha_i + 1\}}}$$

In Problem Sheet 3 Q6(a), we are asked to find the Möbius function of a linear order. This gives:

True fact:

$$\mu(1, p_i^{\alpha_i}) = \begin{cases} 1 & \text{if } \alpha_i = 0, \\ -1 & \text{if } \alpha_i = 1, \\ 0 & \text{if } \alpha_i \geq 2. \end{cases}$$

Therefore

$$\mu(1, n) = \begin{cases} 1 & \text{if } n = 1, \\ (-1)^k & \text{if } n \text{ is the product of } k \text{ distinct primes} \\ 0 & \text{otherwise.} \end{cases}$$

Now we can prove the classical Möbius inversion formula.

Theorem 12.1. *Let $F : \mathbb{Z}^+ \to \mathbb{R}$ be a given function, and define*

$$G(n) = \sum_{\substack{d \in \mathbb{Z}^+ \\ d \mid n}} F(d)$$

for all $n \in \mathbb{Z}^+$. Then for $n \in \mathbb{Z}^+$, we have

$$F(n) = \sum_{\substack{d \in \mathbb{Z}^+ \\ d \mid n}} \mu\left(\frac{n}{d}\right) G(k),$$

where we write $\mu(\frac{n}{k})$ instead of $\mu(1, \frac{n}{k})$.

Proof. For any $n \in \mathbb{Z}^+$, the definition of $G(n)$ depends only on values of $F(d)$ with $d \in N = \{1, \ldots, n\}$. So we restrict our attention to $D_n = (N, |)$. Applying Möbius inversion gives

$$F(n) = \sum_{\substack{d \in \mathbb{N} \\ d \mid n}} \mu(d, n) G(d) = \sum_{\substack{d \in \mathbb{N} \\ d \mid n}} \mu\left(1, \frac{n}{d}\right) G(d),$$

using Problem Sheet 3 Q7. □

We can use this theorem to prove that the Euler ϕ function satisfies

$$\phi(n) = n \prod_{\substack{p \text{ prime}, \\ p \mid n}} \left(1 - \frac{1}{p}\right).$$

Or, prove by straight inclusion-exclusion (Problem Sheet 3 Q4).

Note, Problem Sheet 3 Q8(c): we prove a statement which is equivalent to this *useful* true fact:

$$\phi(n) = \sum_{\substack{d \in \{1,\dots,n\} \\ d \mid n}} d \cdot \mu\left(\frac{n}{d}\right).$$

Example 12.2. We count the number of circular arrangements of "beads" on a "necklace". There are n beads of k distinct types, denoted $a_1, \dots, a_k$. Each type of bead can be used any number of times. The period of an arrangement is the number of clockwise shifts by one position required to leave the circular word unchanged. Denote an arrangement in cycle notation, going clockwise. For example, $(a_1a_2a_1a_2)$ has period 2, and $(a_1a_2a_3a_3)$ has period 4. The period d of an arrangement of N beads satisfies $d \mid n$, as an arrangement of period d consists of a particular pattern of length d repeated $\frac{n}{d}$ times.

Map each arrangement to a linear string, e.g. $a_1a_2a_3a_1a_2a_3$. An n-arrangement with period d corresponds to precisely d dsitinct linear strings. Choosing all n starting points produces each one $\frac{n}{d}$ times. Let $h_k(n)$ be the number of circular arrangements of n beads, with k types available. For $m \in \mathbb{Z}^+$, let $f_k(m)$ denote the number of linear strings of m symbols from $\{a_1, \dots, a_k\}$ which do not consist of a smaller substring repeated more than once. For example, for $f_k(6)$,

$$\begin{array}{l} a_1 \; a_2 \; a_1 \; a_2 \; a_3 \; a_4 \;\; \checkmark \\ a_1 \; a_2 \; a_3 \; a_1 \; a_2 \; a_3 \;\; \times \end{array}$$

(So you can have a repeated substring, but you can't have the *whole* thing as a repeated substring.) Then

$$h_k(n) = \sum_{\substack{d \in \{1,\dots,n\}, \\ d \mid n}} \frac{1}{d} f_k(d),$$

since each arrangement has a period d such that $d \mid n$, and corresponds to precisely d linear strings of length n: each fo these is fully determined by the sequence of the first d symbols, which then repeats.

So we now aim to calculate $f_k(d)$. Let

$$g_k(m) = \sum_{\substack{a \in \{1,\dots,m\}, \\ a \mid m}} f_k(a).$$

Then $g_k(m)$ counts all strings of length m with symbols $\{a_1, \dots, a_k\}$. (Find the "period" of the linear string: if it is a then this string is counted by $f_k(a)$, and $a \mid d$.) Hence $g_k(m) = k^m$. By classical

Möbius inversion,

$$\begin{aligned} f_k(m) &= \sum_{\substack{a\in\{1,\dots,m\},\\ a\mid m}} \mu\left(\frac{m}{a}\right) g_k(a) \\ &= \sum_{\substack{a\in\{1,\dots,m\},\\ a\mid m}} \mu\left(\frac{m}{a}\right) k^a. \end{aligned}$$

(It's $\mu(\frac{m}{a})$ since $\mu(a,m)=\mu(1,\frac{m}{a})$, which we write as $\mu(\frac{m}{a})$. This gives

$$\begin{aligned} h_k(n) &= \sum_{\substack{d\in\{1,\dots,m\},\\ d\mid n}} \frac{1}{d} f_k(d) \\ &= \sum_{\substack{d\in\{1,\dots,m\},\\ d\mid n}} \frac{1}{d} \sum_{\substack{a\in\{1,\dots,d\},\\ a\mid d}} \mu\left(\frac{d}{a}\right) k^a \\ &= \sum_{\substack{a\in\{1,\dots,n\},\\ a\mid n}} \left(\sum_{\substack{m\in\{1,\dots,n\}\\ m\mid \frac{n}{a}}} \frac{\mu(m)}{am} \right) k^a. \end{aligned}$$

In the inner sum, $m \mid \frac{n}{a}$, so write $\frac{n}{a} = mr$. Then

$$\begin{aligned} h_k(n) &= \sum_{\substack{a\in\{1,\dots,n\},\\ a\mid n}} k^a \left(\sum_{\substack{r\in\{1,\dots,\frac{n}{a}\},\\ r\mid \frac{n}{a}}} \frac{r}{n}\, \mu\left(\frac{n/a}{r}\right) \right) \\ &= \sum_{\substack{a\in\{1,\dots,n\},\\ a\mid n}} k^a \frac{\phi(\frac{n}{a})}{n} \\ &= \frac{1}{n} \sum_{\substack{a\in\{1,\dots,n\},\\ a\mid n}} \phi\left(\frac{n}{a}\right) k^a. \end{aligned}$$

Here, recall, ϕ is the Euler ϕ function

$$\phi(n) = \sum_{\substack{d\in\{1,\dots,n\},\\ d\mid n}} d\cdot\mu\left(\frac{n}{d}\right).$$

(Equivalent to Problem Sheet 3 Q8(c).)

Lecture 13

Generating Functions

Say we are interested in a sequence (g_n) for some range of n. We can store the sequence in a generating function

$$G(x) = \sum_n g_n x^n,$$

an "ordinary generating function". Write $[x^n]G(x)$ for the coefficient of x^n in $G(x)$. Note

$$[x^n](x^a G(x)) = [x^{n-a}]G(x),$$

and

$$[cx^n]G(x) = \frac{1}{c}[x^n]G(x)$$

for any non-zero constant c. So (g_n) can be recovered from G.

$$\begin{tikzcd} \underset{\text{discrete}}{g_n} \arrow[r, leftrightarrow, "\text{generating}", "\text{functions}"'] & \underset{\text{continuous}}{G} \end{tikzcd}$$

Example 13.1. Fibonacci numbers.
Write $F_0 = 0, F_1 = 1, F_{n+1} = F_n + f_{n-1}$ for $n \geq 1$. Let $F(x) = \sum_n F_n x^n$. Claim:

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

Proof. Start with the recurrence $F_{n+1} = F_n + F_{n-1}$. Multiply by x^n and sum over $n \geq 1$. (Make sure recurrence is true when we sum.) This gives

$$\sum_{n=1}^{\infty} F_{n+1}x^n = \sum_{n=1}^{\infty} F_n x^n + \sum_{n=1}^{\infty} F_{n-1}x^n.$$

Rewrite this as

$$\underbrace{\frac{F(x) - x}{x}}_{\text{as } F_0 = 0 \text{ and } F_1 = 1} = \underbrace{F(x)}_{\text{as } F_0 = 0} + xF(x).$$

Alternative: guess $F_n = A \cdot \lambda^n$ and find

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

Then initial conditions give solutions to A and B and lead to

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

□

So, why use generating functions? They can

- Help us find an expression for coefficients
- Sometimes find a new recurrence for the coefficients, prove new identities
- Easily find averages
- Sometimes find asymptotics, i.e. limiting behaviour of a_n as $n \to \infty$ (we'll need convergence for this, not covered in this course.)

Lecture 14

Example 14.1. Suppose that $a_{n+1} = 2a_n + n$ for $n \geq 0$, $a_0 = 1$. Find an expression for a_n.

Let $A(x) = \sum_n a_n x^n$. Then

$$\sum_{n\geq 0} a_{n+1}x^n = 2\sum_{n\geq 0} a_n x^n + \sum_{n\geq 0} nx^n.$$

So

$$\begin{aligned}
\underbrace{\frac{A(x)-1}{x}}_{\text{as } a_0=1} &= 2A(x) + \sum_{n\geq 0} x\frac{d}{dx}(x^n) \\
&= 2A(x) + x\frac{d}{dx}\sum_{n\geq 0} x^n \\
&= 2A(x) + x\frac{d}{dx}\frac{1}{1-x} \\
&= 2A(x) + \frac{x}{(1-x)^2}.
\end{aligned}$$

Therefore

$$A(x) - 1 = 2xA(x) + \frac{x^2}{(1-x)^2},$$

so

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

Hence

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

To find the coefficients, use partial fractions. Solve

$$A(x) = \frac{\alpha}{(1-x)^2} + \frac{\beta}{1-x} + \frac{\gamma}{1-2x}$$

to find

$$\begin{aligned}
A(x) &= -\frac{1}{(1-x)^2} + \frac{2}{1-2x} \\
&= 2\sum_{n\geq 0}(2x)^n - \left(\sum_{n\geq 0} x^n\right)^2 \\
&= 2\sum_{n\geq 0} 2^{n+1}x^n - \sum_{n\geq 0}(n+1)x^n.
\end{aligned}$$

Therefore

$$a_n = [x^n]A(x) = 2^{n+1} - n - 1,$$

for $n \geq 0$.

We want to formalise this: treat generating functions as formal power series over $\mathbb{R}$, that is, an expression of the form

$$a_0 + a_1x + a_2x^2 + \dots,$$

where $a_i \in \mathbb{R}$. Given $(a_n)_{n=0}^{\infty}$, define

$$A(x) = a_0 + a_1x + a_2x^2 + \dots$$

to be the ordinary generating function of (a_n). Write $(a_n) \overset{ogf}{\longleftrightarrow} A$. Operations and stuffs:

- If $(a_n) \overset{ogf}{\longleftrightarrow} A$ and $(b_n) \overset{ogf}{\longleftrightarrow} B$, then $A + B \overset{ogf}{\longleftrightarrow} (a_n + b_n)$.
- If $\lambda \in \mathbb{R}$, then $\lambda A \overset{ogf}{\longleftrightarrow} (\lambda a_n)$
  (So the set of all formal power series over $\mathbb{R}$ forms a ring, denoted $\mathbb{R}[[x]]$.)
- Multiplication: the convolution or Cauchy product rule:
  $$AB \overset{ogf}{\longleftrightarrow} \left(\sum_{r=0}^{n} a_r b_{n-r}\right)_{n\geq 0}.$$
- Differentiation: Let $D : f \mapsto \frac{d}{dx}f$ be the differential operator. We *define* its action on a formal power series by
  $$DA(x) = \frac{d}{dx}(A(x)) = \sum_{n=0}^{\infty} na_nx^{n-1}.$$
  So $DA \overset{ogf}{\longleftrightarrow} ((n+1)a_{n+1})_{n=0}^{\infty}$, and hence $(xD)A \overset{ogf}{\longleftrightarrow} (na_n)_{n=0}^{\infty}$.
- From now on, if omitted, assume the summation is taken over $n = 0, \dots, \infty$.

Lemma 14.1. *Suppose that $A \overset{ogf}{\longleftrightarrow} (a_n)$. Then for any $m \in \mathbb{N}$,*

1. $\dfrac{A(x) - (a_0 + a_1x + \dots + a_{m-1}x^{m-1})}{x^m} \overset{ogf}{\longleftrightarrow} (a_{n+m})_{n=0}^{\infty}$.
2. *If P is is a polynomial, then $P(xD)A \overset{ogf}{\longleftrightarrow} (P(n)a_n)$.*

Proof. Problem Sheet 4. □

One more true fact:

$$\frac{A}{1-x} \overset{ogf}{\longleftrightarrow} \left(\sum_{j=0}^{n} a_j\right).$$

So $\frac{A}{1-x}$ generates partial sums of (a_n). Why?

$$\frac{A}{1-x} = A \cdot \frac{1}{1-x} \quad \text{and} \quad \frac{1}{1-x} \overset{ogf}{\longleftrightarrow} (1)_{n=0}^{\infty}.$$

So by the convolution rule for products,

$$A \cdot \frac{1}{1-x} \overset{ogf}{\longleftrightarrow} \left(\sum_{r=0}^{n} a_r \cdot 1\right).$$

Example 14.2. Calculate

$$\sum_{n=0}^{\infty} \frac{n^2+4n+5}{n!}.$$

Let $P(n) = n^2 + 4n + 5$ and $a_n = \frac{1}{n!}$. Then $A(x) \overset{ogf}{\longleftrightarrow} (a_n)$ is just $A(x) = \sum_n \frac{x^n}{n!} = e^x$. Therefore, using part (2) of the above lemma,

$$\begin{aligned}(P(n)a_n)_{n=0}^{\infty} \overset{ogf}{\longleftrightarrow} P(xD) \cdot A \\ &= ((xD)^2 + 4xD + 5)e^x \\ &= (x + x^2)e^x + 4xe^x + 5e^x \\ &= (x^2 + 5x + 5)e^x.\end{aligned}$$

So

$$\sum_{n=0}^{\infty} \frac{n^2+4n+5}{n!} x^n = (x^2 + 5x + 5)e^x.$$

Therefore

$$\sum_{n=0}^{\infty} \frac{n^2+4n+5}{n!} = 11e.$$

Example 14.3. We have non-commuting variables $x_0, \ldots, x_n$, under some non-associative multiplication. The expression

$$x_0 \bullet x_1 \bullet \cdots \bullet x_n$$

is ambiguous. How many ways can we insert parentheses to completely specify the order of multiplication? Let this number be C_n. Then $C_0 = C_1 = 1$, $C_2 = 2$ (either $x_0 \bullet (x_1 \bullet x_2)$ or $(x_0 \bullet x_1) \bullet x_2$). Check that $C_3 = 5$.

Find a recurrence: There is a unique multiplication symbol outside all brackets. This is the final multiplication to be performed. If this final multiplication symbol lies between x_k and x_{k+1}, then

$$\underbrace{(x_0 \bullet \cdots \bullet x_k)}_{C_k \text{ ways}} \bullet \underbrace{(x_{k+1} \bullet \cdots \bullet x_n)}_{C_{n-k-1} \text{ ways}}.$$

(There are C_k ways to parenthesise $x_0 \bullet \cdots \bullet x_k$ and C_{n-k-1} ways to parenthesise $x_{k+1} \bullet \cdots \bullet x_n$.) Hence

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

for $n \geq 1$. Let $C \overset{ogf}{\longleftrightarrow} (C_n)_{n \geq 0}$. Then

$$\begin{aligned}C(x) = \sum_{n=0}^{\infty} C_n x^n &= 1 + \sum_{n=1}^{\infty} C_k C_{n-1-k} x^n \\ &= 1 + x \sum_{m=0}^{\infty} \sum_{k=0}^{m} C_k C_{m-k} x^m \qquad [m = n-1] \\ &= 1 + x(C(x)^2,\end{aligned}$$

using the convolution rule. Therefore $C(x)^2 x - C(x) + 1 = 0$, and hence

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

Since $C_0 = 1$, we must take the " + " option in the second expression, which corresponds to

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

Lecture 15

We continue on from last time.

Definition 15.1. The generalised binomial coefficient is given by

$$\binom{r}{k} = \frac{(r)_k}{k!} = \frac{r(r-1)\cdots(r-k+1)}{k!},$$

for $r \in \mathbb{R}$ and $k \in \mathbb{N}$.

Proposition 15.1. *(Newton's Generalised Binomial Theorem.)*

$$(x+y)^r = \sum_{k=0}^{\infty} \binom{r}{k} x^{r-k} y^k,$$

for $r, x, y \in \mathbb{R}$.

Using this, in Problem Set 4 Q6, we will prove that

$$C(x) = \sum_{n=0}^{\infty} \frac{1}{n+1} \binom{2n}{n} x^n.$$

Hence

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

called the n^{th} Catalan number.

Special case of Multiplication

If $(a_n)_{n=0}^{\infty} \overset{ogf}{\longleftrightarrow} A$, then

$$A^k \overset{ogf}{\longleftrightarrow} \left(\sum_{\substack{r_1+\cdots+r_k=n,\\ r_i \geq 0}} a_{r_1} a_{r_2} \cdots a_{r_k} \right)_{n=0}^{\infty}. \tag{9}$$

The number of terms in this sum is the number of ordered partitions of n into k non-negative parts. Let this number be $f(n,k)$. We derive $f(n,k)$ from the ordinary generating function

$$(1)_{n=0}^{\infty} \overset{ogf}{\longleftrightarrow} \frac{1}{1-x}.$$

Then $(f(n,k))_{n=0}^{\infty} \overset{ogf}{\longleftrightarrow} \frac{1}{(1-x)^k}$, by (9). Hence

$$f(n,k) = [x^n](1-x)^{-k} = \binom{-k}{n}(-1)^n$$

by the generalised binomial theorem. So

$$\begin{aligned}f(n,k) &= \frac{(-k)(-k-1)(-k-2)\cdots(-k-n+1)}{n!}\cdot(-1)^n\\ &= \frac{k(k+1)\cdots(k+n-1)}{n!}\\ &= \binom{k+n-1}{n},\end{aligned}$$

which, as expected, corresponds to $k-1$ 'bars' and n 'stars'.

This gives the useful identity

$$\frac{1}{(1-x)^k} = \sum_n \binom{n+k-1}{n} x^n. \tag{10}$$

Example 15.1. Calculate

$$\sum_{i=0}^{n} i^2.$$

From part (2) of Lemma 14.1,

$$(xD)^2\frac{1}{1-x} \overset{ogf}{\longleftrightarrow} (n^2)_{n=0}^{\infty}.$$

We want partial sums of this. Therefore

$$\frac{1}{1-x}(xD)^2\frac{1}{1-x} \overset{ogf}{\longleftrightarrow} \left(\sum_{i=0}^{n} i^2\right)_0^{\infty}.$$

(See true fact after Lemma 14.1). Now we can check that

$$\frac{1}{1-x}(xD)^2\frac{1}{1-x} = \frac{x(1+x)}{(1-x)^4}.$$

We want the coefficient of x^n. Now

$$[x^n]\frac{1}{(1-x)^4} = \binom{n+3}{3},$$

by (10). Therefore

$$\begin{aligned}[x^n]\frac{x(1+x)}{(1-x)^4} &= [x^n]\frac{x}{(1-x)^4} + [x^n]\frac{x^2}{(1-x)^4}\\ &= [x^{n-1}]\frac{1}{(1-x)^4} + [x^{n-2}]\frac{1}{(1-x)^4}\\ &= \binom{n+1}{3} + \binom{n+2}{3}\\ &= \frac{n(n+1)(2n+1)}{6}.\end{aligned}$$

Example 15.2. Harmonic Numbers.
Let $H_n = \sum_{i=1}^{n} \frac{1}{i}$ for $n \geq 1$. Define

$$H(x) = \sum_{n=1}^{\infty} H_n x^n = \frac{1}{1-x}\cdot\left\{\text{the ogf for } \left(\frac{1}{n}\right)_{n=1}^{\infty}\right\}.$$

(Again by true fact after Lemma 14.1.) Now

$$\sum_{n=1}^{\infty} \frac{x^n}{n} = -\ln(1-x),$$

so

$$H(x) = -\frac{1}{1-x}\ln(1-x) = \frac{1}{1-x}\ln\left(\frac{1}{1-x}\right).$$

Another useful trick: if $A \overset{ogf}{\longleftrightarrow} (a_n)_{n=0}^{\infty}$, then

$$\frac{A(x)+A(-x)}{2} = a_0 + a_2x^2 + \ldots,$$

and

$$\frac{A(x)-A(-x)}{2} = a_1x + a_3x^3 + \ldots.$$

“Snake Oil”

This is a method for manipulating combinatorial sums, and/or for proving combinatorial identities.

Conventions, where $x \in \mathbb{R}$ and $m \in \mathbb{Z}$:

- $\binom{x}{m} = 0$ if $m < 0$ or x is a non-negative integer with $x < m$.
- If summation range is not given, assume it is $-\infty, .., \infty$.

For example,

$$\sum_k \binom{n}{k} = 2^n,$$

for $n \in \mathbb{Z}^+$.

Example 15.3. (Showcasing snake oil.)
Evaluate

$$\sum_{k=0}^{\infty} \binom{k}{n-k}$$

for $n = 0, 1, 2, \ldots$.

Identify the free variable in the sum; it is n. So multiply the sum by x^n adn sum over n. Define

$$f(n) = \sum_{k=0}^{\infty} \binom{k}{n-k}$$

and

$$F(x) = \sum_n \left(x^n \sum_{k=0}^{\infty} \binom{k}{n-k} \right) \overset{ogf}{\longleftrightarrow} (f(n))_n.$$

Plan: find F and extract coefficient of x^n. Now

$$\begin{aligned}
F(x) &= \sum_n \left(x^n \sum_{k=0}^{\infty} \binom{k}{n-k} \right) \\
&= \sum_{k=0}^{\infty} \sum_n \binom{k}{n-k} x^n \qquad \textit{(interchange order of summation).}
\end{aligned}$$

We want to make the inner sum look like the binomial theorem.

$$\begin{aligned}
F(x) &= \sum_{k=0}^{\infty} x^k \sum_n \binom{k}{n-k} x^{n-k} \\
&= \sum_{k=0}^{\infty} x^k \sum_r \binom{k}{r} x^r \qquad [r = n-k] \\
&= \sum_{k=0}^{\infty} x^k (1+x)^k \\
&= \sum_{k=0}^{\infty} \big(x(1+x)\big)^k \\
&= \frac{1}{1-x-x^2}.
\end{aligned}$$

But we recall that

$$\sum_{n=1}^{\infty} F_n x^n = \frac{x}{1-x-x^2},$$

where F_n is the Fibonacci sequence. So

$$\frac{1}{1-x-x^2} = \sum_{n=1}^{\infty} F_n x^{n-1} = \sum_{k=0}^{\infty} F_{k+1} x^k.$$

Therefore

$$\sum_{k=0}^{\infty} \binom{k}{n-k} = F_{n+1},$$

for $n = 0, 1, 2, \ldots$.

Lecture 16

Recall, for fixed $k \in \mathbb{N}$,

$$\frac{1}{(1-x)^{k+1}} = \sum_n \binom{n+k}{k} x^n.$$

Let $r = n+k$ to obtain

$$\frac{1}{(1-x)^{k+1}} = \sum_r \binom{r}{k} x^{r-k}.$$

So (another useful identity):

$$\frac{x^k}{(1-x)^{k+1}} = \sum_r \binom{r}{k} x^r. \tag{11}$$

Example 16.1. (More “snake oil”.)
Prove that, for $m, n \in \mathbb{N}$,

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

Free variables are m, n. We choose to work with n. Cos y not. So, multiply the identity by x^n and sum. Let

$$a_n = \sum_k \binom{m}{k}\binom{n+k}{m}$$

and

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

Define $A \overset{ogf}{\longleftrightarrow} (a_n)_{n=0}^{\infty}$ and $B \overset{ogf}{\longleftrightarrow} (b_n)_{n=0}^{\infty}$. Then

$$\begin{aligned}
A(x) &= \sum_{n=0}^{\infty} x^n \sum_k \binom{m}{k}\binom{n+k}{m} \\
&= \sum_k \binom{m}{k} \sum_{n=0}^{\infty} \binom{n+k}{m} x^n \\
&= \sum_k \binom{m}{k} x^{-k} \sum_{n=0}^{\infty} \binom{n+k}{m} x^{n+k} \\
&= \sum_k \binom{m}{k} x^{-k} \frac{x^m}{(1-x)^{m+1}} \qquad \text{by (11)} \\
&= \frac{x^m}{(1-x)^{m+1}} \sum_k \binom{m}{k} x^{-k} \\
&= \frac{x^m}{(1-x)^{m+1}} \left(1 + \frac{1}{x}\right)^m \\
&= \frac{(1+x)^m}{(1-x)^{m+1}}.
\end{aligned}$$

Next,

$$\begin{aligned}
B(x) &= \sum_{n=0}^{\infty} x^n \sum_k \binom{m}{k}\binom{n}{k} 2^k \\
&= \sum_k \binom{m}{k} 2^k \sum_{n=0}^{\infty} \binom{n}{k} x^n \\
&= \sum_k \binom{m}{k} 2^k \frac{x^k}{(1-x)^{k+1}} \qquad \text{by (11)} \\
&= \frac{1}{1-x} \sum_k \binom{m}{k} \left(\frac{2x}{1-x}\right)^k \\
&= \frac{1}{1-x} \left(1 + \frac{2x}{1-x}\right)^m \\
&= \frac{(1+x)^m}{(1-x)^{m+1}} \\
&= A(x).
\end{aligned}$$

Therefore $a_n = [x^n]A(x) = [x^n]B(x) = b_n$ for all $n \in \mathbb{N}$.

Exponential Generating Functions

For a sequence $(a_n)_{n=0}^{\infty}$, we sometimes work with this generating function:

$$A(x) = \sum_{n=0}^{\infty} \frac{a_n}{n!} x^n.$$

This is the *exponential generating function* of $(a_n)_{n=0}^{\infty}$. We write

$$A \overset{egf}{\longleftrightarrow} (a_n).$$

Note: $A \overset{egf}{\longleftrightarrow} (a_n) \iff A \overset{ogf}{\longleftrightarrow} \left(\frac{a_n}{n!}\right)$.

True facts about egf

Proposition 16.1. *If* $A \overset{egf}{\longleftrightarrow} (a_n)$ *and* $B \overset{egf}{\longleftrightarrow} (b_n)$, *then*

$$AB \overset{egf}{\longleftrightarrow} \sum_k \binom{n}{k} a_k b_{n-k}.$$

Proof. Treat A and B as ordinary generating functions of $\left(\frac{a_n}{n!}\right)$ and $\left(\frac{b_n}{n!}\right)$ respectively, and use convolution. This gives

$$\begin{aligned}
A(x)B(x) &= \sum_{k=0}^{n} \frac{a_k}{k!} \frac{b_{n-k}}{(n-k)!} x^n \\
&= \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} \cdot a_k b_{n-k} \frac{x^n}{n!} \\
&= \sum_k \binom{n}{k} a_k b_{n-k} \frac{x^n}{n!} \overset{egf}{\longleftrightarrow} \left(\sum_k \binom{n}{k} a_k b_{n-k}\right)_n .
\end{aligned}$$

□

Lemma 16.1. *Let* $A \overset{egf}{\longleftrightarrow} (a_n)_{n=0}^\infty$. *Then*

$$D^r A \overset{egf}{\longleftrightarrow} (a_{n+r})_{n=0}^\infty.$$

So if P *is any polynomial then*

$$P(D)A \overset{egf}{\longleftrightarrow} (P(n)a_n)_{n=0}^\infty.$$

Proof. Problem Set 5. □

Exponential or Ordinary?

Think about the "meaning" of multiplication.

- If "objects" of "weight" n are obtained by gluing together an object of weight k and one of weight $n-k$ for some k, then use ordinary generating functions, e.g.

$$(x_0 \bullet \cdots \bullet x_k) \bullet (x_{k+1} \bullet \cdots \bullet x_n).$$

  Importantly, there's no relabelling.

- If objects of weight n are formed by gluing together an object of weight k and an object of weight $n-k$, for some k, *and then relabelling*, then we use exponential generating functions.

Example 16.2. Let $b(n)$ be the number of partitions of the set $\{1, \ldots, n\}$ into non-empty unordered parts. These $b(n)$ are called the Bell numbers. By convention, $b(0) = 1$. Also, $b(1) = 1$ and $b(2) = 2$, $b^3 = 5$, $b^4 = 15$ and $b(5) = 52$.

True fact: The Bell numbers satisfy

$$b(n+1) = \sum_k \binom{n}{k} b(k).$$

Let $B \overset{egf}{\longleftrightarrow} (b(n))_{n=0}^\infty$. By inspection, the exponential generating function of the LHS (of the true fact) is $B'(x)$. The exponential generating function of the RHS is the product of $B(x)$ and the exponential generating function for $(1)_{n=0}^\infty$, which is e^x. So the egf for the RHS is $e^x B(x)$. This gives

$$B'(x) = e^x B(x),$$

with solution

$$B(x) = c \exp\left(e^x\right),$$

for some constant c. Since $B(0) = 1$, we find that $c = e^{-1}$ and hence

$$B(x) = \exp\left(e^x - 1\right).$$

Lecture 17

We have just shown that $(b(n)) \overset{egf}{\longleftrightarrow} \exp(e^x - 1)$. Note: we can derive recurrences from exponential generating functions using the $x\frac{d}{dx}$ log trick. For example, start with

$$\sum_{n=0}^\infty \frac{b(n)}{n!} x^n = \exp(e^x - 1).$$

Take the log of both sides and differentiate, then multiply by x:

$$\begin{aligned} x\frac{d}{dx}\log\left(\sum_{n=0}^{\infty}\frac{b(n)}{n!}x^n\right) &= x\frac{d}{dx}(e^x-1) \\ &= xe^x. \end{aligned}$$

So

$$\begin{aligned} xe^x &= x \ \cdot \ \frac{\displaystyle\sum_{n=1}^{\infty}\frac{b(n)}{(n-1)!}x^{n-1}}{\displaystyle\sum_{n=0}^{\infty}\frac{b(n)}{n!}x^n} \\ &= \frac{\displaystyle\sum_{n=1}^{\infty}\frac{b(n)}{(n-1)!}x^{n}}{\displaystyle\sum_{n=0}^{\infty}\frac{b(n)}{n!}x^n}. \end{aligned}$$

Rearranging gives

$$\begin{aligned} \sum_{n=1}^{\infty}\frac{b(n)}{(n-1)!}x^n &= xe^x\cdot\sum_{n=0}^{\infty}\frac{b(n)}{n!}x^n \\ &= e^x\sum_{n=0}^{\infty}\frac{b(n)}{n!}x^{n+1} \\ &= e^x\sum_{n=1}^{\infty}\frac{b(n-1)}{(n-1)!}x^n \\ &= \sum_{n=1}^{\infty}\left(\sum_k\binom{n-1}{k}b(k)\right)\frac{x^n}{(n-1)!}, \end{aligned}$$

by convolution (note e^x is the egf of $(1)_{n=0}^{\infty}$). Taking the coefficient of x^n gives the recurrence

$$b(n)=\sum_k\binom{n-1}{k}b(k),$$

for $k\geq 1$ — this is where we started.

Consequence of product rule for exponential generating functions:

If $A\overset{egf}{\longleftrightarrow}(a_n)_{n=0}^{\infty}$, then

$$A^k\overset{egf}{\longleftrightarrow}\left(\sum_{\substack{r_1+\cdots+r_k=n,\\ r_i\in\mathbb{N}}}\frac{n!}{r_1!r_2!\cdots r_k!}a_{r_1}a_{r_2}\cdots a_{r_k}\right)_{n=0}^{\infty}.$$

Cards, Decks and Hands

General problem: Count structures which are built up from smaller "atoms" and which are labelled with $1, 2, \dots, n$. For example, permutations can be built up by multiplying disjoint cycles, e.g. (1 5 4)(2 3 7 6). The "atoms" are permutations with a single cycle: a unique unlabelled cycle on k points, or $(k-1)!$ distinct labellings of that cycle with labels $\{1, 2, \dots, k\}$ (picture the cycles as on a directed graph).

Definition 17.1. A *card* C is a "picture" which is labelled with a set $S(C)$ of distinct integers, called *labels.* A card is *standard* if its label set is $S(C) = \{1, 2, \dots, k\}$ for some $k \in \mathbb{N}$. Write $|S(C)| = k$, the *weight* of the card. A *hand* is a set of cards whose labels form a partition of $\{1, 2, \dots, n\}$ for some $n \in \mathbb{Z}^+$. Here n is the *weight* of the hand.

Example 17.1. (I'll write the labelled cycles as a permutation kind of notation... hope it makes sense.)

Cards

$$\begin{array}{cc} (1\ 5\ 4) & (2\ 3\ 7\ 6) \\ (1\ 3\ 2) & (1\ 2\ 4\ 3). \end{array}$$

The two cards in the top row form a hand of weight 7. The two cards in the bottom row are standard cards that preserve the relative order of the labelling.

The last bit seems important:

Definition 17.2. A *relabelling* of a card C is a card C' obtained from C by replacing the labels of C with some other set of labels of the same size, *preserving the relative ordering of the labels.*

In particular, each card can be relabelled in a unique way to give a standard card.

Definition 17.3. A *deck* is a set of standard cards, all distinct, such that the number d_n of cards in the deck with weight n is finite, for all $n \in \mathbb{N}$.

Definition 17.4. The *deck enumerator* $D(x)$ is the exponential generating function for $(d_n)_{n=0}^{\infty}$, that is,

$$D(x) = \sum_{n=0}^{\infty} \frac{d_n}{n!} x^n.$$

We assume that $d_0 \in \{0, 1\}$. If $d_0 = 1$, then there is a unique card with no labels in the deck.

Example 17.2. Consider the deck with cards (1), (1 2), (1 2 3), (1 3 2) consisting of all permutations with a single cycle on $\{1, \dots, k\}$ for some $k \in \mathbb{Z}^+$. Then $d_n = (n-1)!$ for $n \geq 1$, and we take $d_0 = 0$. So

$$\begin{aligned} D(x) &= \sum_{n=1}^{\infty} \frac{d_n}{n!} x^n \\ &= \sum_{n=1}^{\infty} \frac{(n-1)!}{n!} x^n \\ &= \sum_{n=1}^{\infty} \frac{x^n}{n} \\ &= -\log(1-x). \end{aligned}$$

Lemma 17.1. *(Labelled counting lemma.)*
Suppose that we have two decks with empty intersection with deck enumerators $D_1(x)$ and $D_2(x)$. Let a_n be the number of hands of weight n consisting of exactly two cards, one a relabelling of a card from the first deck, the other a relabelling of a card from the second deck. Then

$$(a_n)_{n=0}^{\infty} \overset{egf}{\longleftrightarrow} D_1 D_2.$$

Proof. Let $D_1 \overset{egf}{\longleftrightarrow} (d_k)_{k=0}^{\infty}$, and $D_2 \overset{egf}{\longleftrightarrow} (\widetilde{d}_k)_{k=0}^{\infty}$. A hand is a relabelling of:

$$\text{a card from the first deck with weight } k$$

$$\text{and}$$

$$\text{a card from the second deck with weight } n-k,$$

for some $k \in \{0, \dots, n\}$. Once the set of k labels for the first card is specified, the relabelling is uniquely determined (as relative ordering is respected). There are d_k choices for the first card, $\widetilde{d}_{n-k}$ choices for the second card and $\begin{pmatrix} n \\ k \end{pmatrix}$ choices for the label set for the first card.

Hence

$$a_n = \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} d_k \widetilde{d}_{n-k},$$

which implies that

$$(a_n) \overset{egf}{\longleftrightarrow} D_1 D_2,$$

as required. □

Lecture 18

Corollary. *The exponential generating function for the number of hands of weight n from k disjoint decks with deck enumerators $D_1, \dots, D_k$ respectively is $D_1 D_2 \cdots D_k$.*

Proof. (Sketch.)
By induction on k. True for $k = 2$, since it's just the Labelled Counting Lemma. If $k \geq 3$, treat a hand

$$\text{deck } 1 \qquad \text{deck } 2 \qquad \dots \qquad \text{deck } k$$

as

$$\text{deck } 1 \qquad \underbrace{\text{deck } 2 \qquad \dots \qquad \text{deck } k}_{\text{deck } 0}.$$

Then apply the Labelled Counting Lemma. □

Example 18.1. Stirling number of the second kind.

For $k, n \in \mathbb{N}$, let

$$\begin{Bmatrix} n \\ k \end{Bmatrix} = \text{number of partitions of the set } \{1, 2, \dots, n\} \text{ into } k \text{ nonempty parts.}$$

Theorem 18.1. *The exponential generating function for* $\left(\left\{ n \atop k \right\}\right)_{n=0}^{\infty}$ *is*

$$\frac{1}{k!}(e^x - 1)^k,$$

for fixed $k \in \mathbb{N}$.

Proof. For $i = 1, 2, \ldots, k$, define deck i to be the deck with the following cards:

$$[1]^i \qquad [1\ 2]^i \qquad [1\ 2\ 3]^i \qquad \ldots$$

(The superscript is just a "picture".)

Let D_i be the deck enumerator of deck i. Then $D_i \overset{egf}{\longleftrightarrow} (1)_{n=1}^{\infty}$, so $D_i(x) = e^x - 1$ for $i = 1, \ldots, k$. The number of hands of exactly k cards, one from each deck, equals the number of partitions of the set $\{1, 2, \ldots, n\}$ into k non-empty parts which have been ordered. Hence

$$D_1 D_2 \cdots D_k \overset{egf}{\longleftrightarrow} \left(k! \left\{ n \atop k \right\}\right)_n .$$

Therefore

$$\left(\left\{ n \atop k \right\}\right)_{n=0}^{\infty} \overset{egf}{\longleftrightarrow} \frac{1}{k!} D_1 \cdots D_k = \frac{1}{k!}(e^x - 1)^k,$$

as claimed. □

Note: the Bell numbers satisfy

$$b(n) = \sum_{k=0}^{\infty} \left\{ n \atop k \right\},$$

and

$$(b(n))_n \overset{egf}{\longleftrightarrow} \exp(e^x - 1) = \sum_{k=0}^{\infty} \frac{(e^x - 1)^k}{k!}.$$

Definition 18.1. Let h_n denote the number of hands (from a given deck) of weight n that can be formed from the given deck, and let $h_{n,k}$ be the number of hands of weight n consisting of exactly k cards from the deck. These hands are formed by choosing cards from the deck with replacement (repeats allowed), and then relabelling.

The hand enumerator is

$$H(x) = \sum_{n=0}^{\infty} \frac{h_n}{n!} x^n \overset{egf}{\longleftrightarrow} (h_n)_{n=0}^{\infty}$$

with $h_0 = 1$. The *two variable hand enumerator* is

$$H(x, y) = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \frac{h_{n,k}}{n} ! x^n y^k,$$

with $h_{0,0} = 1$. Note: $H(x, y)$ is an egf with respect to x, and is an ogf with respect to y. Also: $H(x) = H(x, 1)$.

Theorem 18.2. *(Exponential Formula.)*
Take a deck with deck enumerator $D(x)$ *such that* $[x^0]D9x) = 0$ *(all cards have at least one label). Then*

$$H(x) = e^{D(x)} \qquad \text{and} \qquad H(x, y) = e^{yD(x)}.$$

Proof. Since $H(x) = H(x, 1)$, it suffices to prove the second statement. Take k copies of the decks and order them. There are $k!$ distinct orderings. All cards are now distinct. Take a card from each deck to make a hand. Then, by the corollary to Lemma 17.1,

$$D(x)^k \overset{egf}{\longleftrightarrow} (\text{\# of hands of weight } n \text{ with one card each from the ordered decks}).$$

So

$$\frac{1}{k!} D(x)^k \overset{egf}{\longleftrightarrow} (h_{n,k})_{n=0}^{\infty},$$

for $k \geq 1$. Hence

$$\begin{aligned} H(x, y) &= 1 + \sum_{k=1}^{\infty} \left(\sum_{n=0}^{\infty} \frac{h_{n,k} x^n}{n!} \right) y^k \\ &= 1 + \sum_{k=1}^{\infty} \frac{D(x)^k}{k!} y^k \\ &= e^{yD(x)}, \end{aligned}$$

as required. □

From exponential generating functions we can derive recurrences.

Lemma 18.1. *If* $H(x) = e^{D(x)}$ *where* $H \overset{egf}{\longleftrightarrow} (h_n)_{n=0}^{\infty}$ *and* $D \overset{egf}{\longleftrightarrow} (d_n)_{n=0}^{\infty}$ *where* $d_0 = 0$*, then* $h_0 = 1$ *and*

$$nh_n = \sum_k \binom{n}{k} k d_k h_{n-k}$$

for $n \geq 1$*.*

Proof. deferred. □

Example 18.2. Permutations.
We already saw the deck enumerator

$$D(x) = -\log(1 - x) = \log\left(\frac{1}{1 - x}\right)$$

for the deck of cycles with standard labels. Now

$$\begin{aligned} h_n &= \text{\# of hands of weight } n \\ &= n!, \end{aligned}$$

since we get all the permutations. So

$$H(x) = \sum_{n=0}^{\infty} \frac{n!}{n!} x^n = \frac{1}{1 - x} = e^{D(x)}.$$

More interestingly, the two variable hand enumerator is

$$\begin{aligned} H(x, y) &= e^{yD(x)} \\ &= e^{-y \log(1-x)} \qquad \text{(by Exp. Theorem)} \\ &= \frac{1}{(1 - x)^y}. \end{aligned}$$

Here

$$\begin{aligned}h_{n,k} &= \#\text{ of permutations of } \{1,\ldots,n\} \text{ with exactly } k \text{ disjoint cycles}\\ &= \begin{bmatrix} n \\ k \end{bmatrix},\end{aligned}$$

the Stirling number of the first kind.

So

$$\begin{aligned}H(x,y) &= \sum_{n=0}^{\infty}\sum_{k=0}^{\infty}\frac{h_{n,k}}{n!}x^n y^k\\ &= \sum_{n=0}^{\infty}\sum_{k=0}^{\infty}\begin{bmatrix} n \\ k \end{bmatrix}\frac{x^n y^k}{n!}\\ &= e^{yD(x)}\\ &= \frac{1}{(1-x)^y}.\end{aligned}$$

Note, $\begin{bmatrix} n \\ 0 \end{bmatrix} = 0$.

So

$$\begin{aligned}\sum_{k=1}^{n}\begin{bmatrix} n \\ k \end{bmatrix} y^k &= n![x^n](1-x)^{-y}\\ &= n!\binom{y+n-1}{n},\end{aligned}$$

by the generalised binomial theorem. So

$$\begin{aligned}\begin{bmatrix} n \\ k \end{bmatrix} &= [y^k]n!\binom{y+n-1}{n}\\ &= [y^k](y+n-1)(y+n-2)\cdots(y+1)y,\end{aligned}$$

for $k \in \mathbb{N}$.