MATH 249 NOTES

Ian Goulden

April 4, 2008

1 Lecture of January 9

The first six weeks of the course will be concerned with *Enumerative Combinatorics*, also referred to as *Enumeration*, *Combinatorial Analysis* or, simply, *Counting*. This subject concerns the basic question of determining the number of elements in a finite set of mathematical objects.

Let $[\mathbf{n}] = \{1, \ldots, n\}$, for each $n \geq 1$, and $[\mathbf{0}]$ denote the empty set. A *permutation* of $[\mathbf{n}]$ is an ordered list of the elements of $[\mathbf{n}]$, each element appearing once in the list. For example, there are 6 permutations of $[\mathbf{3}]$, namely 123, 132, 213, 231, 312, 321. When $n = 0$, we say that there is a single permutation, which happens to be an empty list. We begin by answering a basic counting question: how many permutations are there of $[\mathbf{n}]$? The answer, given below, can be compactly expressed using *factorial* notation. For each nonnegative integer n, define $n!$, by $0! = 1$, and

$$n! = \prod_{i=1}^{n} i, \quad \text{for } n = 1, 2, \ldots.$$

We say "n factorial" for $n!$.

Example 1.1 *The number of permutations of* $[\mathbf{n}]$ *is* $n!$*, for* $n = 0, 1, \ldots$.

Proof. For $n = 0$, the result is true by the conventions above. For $n \geq 0$, each permutation of $[\mathbf{n}]$ is of the form $a_1 a_2 \ldots a_n$, where $a_1, a_2, \ldots, a_n$ are distinct elements of $[\mathbf{n}]$. There are therefore n choices for a_1. No matter what choice is made for a_1, there are then $n-1$ choices for a_2, since a_2 cannot equal a_1. In fact, iteratively choosing a_{j+1}, since $a_1, \ldots, a_j$ are all different, then a_{j+1} must be chosen from the remaining $n-j$ elements, for $j = 0, 1, \ldots, n-1$, and so the number of choices for the permutation $a_1 a_2 \ldots a_n$ is $\prod_{j=0}^{n-1}(n-j) = n!$, giving the result.

A k-subset of $[\mathbf{n}]$ is a subset of $[\mathbf{n}]$ of size k, for $k = 0, 1, \ldots, n$ (when $k = 0$, the empty set is such a subset, for any $n \geq 0$). For example, there are 6 2-subsets of $[\mathbf{4}]$, namely $\{1, 2\}$, $\{1, 3\}$, $\{1, 4\}$, $\{2, 3\}$, $\{2, 4\}$, $\{3, 4\}$. We now consider another basic counting question: how many k-subsets are there of $[\mathbf{n}]$? The answer, given below, can be compactly expressed

using *binomial coefficient* notation. For nonnegative integers n and $k = 0, 1, \ldots, n$, define $\binom{n}{k}$ by

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

We say "n choose k" for $\binom{n}{k}$. Note that we immediately have the symmetry result

$$\binom{n}{k} = \binom{n}{n-k}.$$

Example 1.2 *The number of k-subsets of $[\mathbf{n}]$ is $\binom{n}{k}$, for $n = 0, 1, \ldots,$ and $k = 0, 1, \ldots, n$.*

PROOF. Let x be the number of k-subsets of $[\mathbf{n}]$. We determine x indirectly, by counting the permutations of $[\mathbf{n}]$. For any permutation $a_1 \ldots a_n$ and any fixed $k = 0, \ldots, n$, the elements in $a_1, \ldots, a_k$ form a k-subset of $[\mathbf{n}]$, call it α. Then the elements in $a_{k+1}, \ldots, a_n$ form the complement, $\overline{\alpha}$, of α with respect to $[\mathbf{n}]$. For example, with $n = 7$ and $k = 3$, for the permutation 5247316 we have $\alpha = \{2, 4, 5\}$ and $\overline{\alpha} = \{1, 3, 6, 7\}$. For each fixed α, there are $k!$ choices of $a_1 \ldots a_k$, since it is a permutation of α, and $(n-k)!$ choices of $a_{k+1} \ldots a_n$, since it is a permutation of $\overline{\alpha}$. Since there are x choices of α, we conclude that the number of permutations of $[\mathbf{n}]$ is equal to

$$k!\,(n-k)!\,x$$

But this is also equal to $n!$, and we have proved that $x = \binom{n}{k}$.

The binomial coefficients provide the answer to many counting questions. One method of proof is to find a 1:1 correspondence between the objects being counted and an appropriate set of subsets. We give two examples of this.

Example 1.3 *Prove that the number of k-subsets of $[\mathbf{n}]$ with no consecutive pairs of elements is $\binom{n-k+1}{k}$.*

PROOF. Let $\mathcal{A}$ be the set of k-subsets of $[\mathbf{n}]$ with no consecutive pairs of elements, and let $\mathcal{B}$ be the set of k-subsets of $[\mathbf{n - k + 1}]$. For example, with $n = 7$ and $k = 3$, we have

$$\mathcal{A} = \{135, 136, 137, 146, 147, 157, 246, 247, 257, 357\},$$

$$\mathcal{B} = \{123, 124, 125, 134, 135, 145, 234, 235, 245, 345\},$$

where we have written each of the subsets as an increasing list of its elements, with no brace brackets. We know that $|\mathcal{B}| = \binom{n-k+1}{k}$, so we establish the result by giving a 1:1 correspondence between $\mathcal{A}$ and $\mathcal{B}$, since then we know that $|\mathcal{A}| = |\mathcal{B}|$. We claim that $f : \mathcal{B} \to \mathcal{A}$ defined by

$$f(\alpha) = 0\,1 \ldots k-1 \;\; + \;\; \alpha,$$

is a 1:1 correspondence, where the addition above means that 0 is added to the first element of α, 1 is added to the second element, and so on, until $k-1$ is added to the last (and biggest) element of α. We leave the proof that f is $1:1$ as an exercise.

In the second example, we consider *lattice paths*, which are paths on the integer lattice in two dimensions, with steps either North by one unit ("N") or East by one unit ("E").

Example 1.4 *Prove that the number of lattice paths from* $(0,0)$ *to* (m,n) *is equal to* $\binom{m+n}{n}$, *or* $\binom{m+n}{m}$.

PROOF. Each lattice path from $(0,0)$ to (m,n) contains exactly $m+n$ steps, with n up and m right. Therefore we can represent them uniquely as an ordered list $s_1 \ldots s_{m+n}$, in which $s_i = N$ for n choices of i, and $s_i = E$ for the remaining m choices of i. Let α denote the set of all i for which $s_i = N$. Then α is an n-subset of $[\mathbf{n} + \mathbf{m}]$, and this is a 1:1 correspondence. The result follows, since there are $\binom{m+n}{n}$ choices of α.

For example, there are $\binom{5}{2}$ paths from $(0,0)$ to $(3,2)$, given by $NNEEE, NENEE$, $NEENE, NEEEN, ENNEE, ENENE$, $ENEEN, EENNE, EENEN, EEENN$.

When $m = n = 0$, we say that there is exactly one path, with no steps, which agrees in this case with the value of the binomial coefficient in Example 1.4.

From Example 1.4 with $m = n$, we know that there are $\binom{2n}{n}$ lattice paths from $(0,0)$ to (n,n). We are now going to consider the number c_n, $n = 0, 1, \ldots$ of these paths that never go below the line $y = x$. These paths are called *Catalan paths*. For example, we have $c_3 = 5$, since the Catalan paths with $n = 3$ are given by $NNNEEE, NNENEE$, $NNEENE, NENNEE, NENENE$.

We are going to prove that

$$c_n = \frac{1}{n+1}\binom{2n}{n}, \quad n = 0, 1, \ldots.$$

These numbers are called *Catalan numbers*.

2 Lecture of January 11

Let $\mathcal{P}_n$, $n \geq 0$, be the set of lattice paths from $(0,0)$ to (n,n), represented as a string of N's and E's (which are the steps of the path) Let $\mathcal{P}'_n$, $n \geq 0$, be the set of lattice paths from $(0,-1)$ to (n,n), starting with N. Then each path in $\mathcal{P}'_n$ has $n+1$ N's and n E's, and $|\mathcal{P}_n| = |\mathcal{P}'_n|$.

The *length* of a path $\pi = \pi_1 \ldots \pi_m$ is $|\pi| = m$, equal to the number of steps in π. We let $\Delta_0(\pi) = -1$, and $\Delta_i(\pi)$ equal the number of N's in the first i steps of π minus the number of E's in the first i steps of π, minus 1, for $i = 1, \ldots, m$. For $\pi \in \mathcal{P}'_n$, we have $\Delta_{2n+1}(\pi) = 0$. The following result is easy to verify.

Proposition 2.1 *Suppose* $\pi = \alpha\beta \in \mathcal{P}'_n$, *where* $|\alpha| = k$, $|\beta| = j$, *where* $k, j \geq 0$ *(and of course* $|\pi| = k + j = 2n + 1$*). Let* $\omega = \beta\alpha$ *(called a* cyclic rearrangement *of* π*), where* $\Delta_k(\pi) = m$. *Then* $\Delta_i(\omega) = \Delta_{i+k}(\pi) - m - 1$, *for* $i = 0, \ldots, j$, *and* $\Delta_i(\omega) = \Delta_{i-j}(\pi) - m$, *for* $i = j+1, \ldots, 2n+1$.

Now let $\mathcal{C}_n$ be the set of paths in $\mathcal{P}_n$ which never go below the line $y = x$. Consider $\pi \in \mathcal{P}'_n$, and let M be the minimum value of $\Delta_i(\pi)$, for $i = 0, \ldots, 2n+1$. Let L be the maximum value of i such that $\Delta_i(\pi) = M$. Then clearly $M \leq -1$, and $0 \leq L \leq 2n$, since $\Delta_0(\pi) = -1$, and $\Delta_{2n+1}(\pi) = 0$. Let $\pi = \alpha\beta$, where $|\alpha| = L$, and $\omega = \beta\alpha$. Then Proposition 2.1 proves immediately that $\Delta_0(\omega) = -1$, and $\Delta_i(\omega) \geq 0$, for $i = 1, \ldots, 2n+1$.

But this means that $\omega = N\psi$, where $\psi \in \mathcal{C}_n$. For $\pi \in \mathcal{P}'_n$, this proves that ω is the *unique* path among the $n+1$ cyclic rearrangements of π starting with N, with $\omega = N\psi$, $\psi \in \mathcal{C}_n$. But there are $n+1$ distinct cyclic rearrangements of any path in $\mathcal{P}'_n$, so we conclude that

$$|\mathcal{C}_n| = \frac{1}{n+1}\binom{2n}{n}, \quad n = 0, 1, \ldots.$$

3 Lecture of January 14

The proof that we'll consider carefully here involves a recurrence for the sequence $\{c_n\}_{n\geq 0}$, and the *generating series*

$$C(x) = \sum_{n\geq 0} c_n x^n$$

for this sequence, where $c_n = |\mathcal{C}_n|$, $n \geq 0$.

Example 3.1 *The sequence $\{c_n\}_{n\geq 0}$ satisfies the recurrence*

$$c_n = \sum_{k=0}^{n-1} c_k\, c_{n-k-1}, \quad n = 1, 2, \ldots, \tag{1}$$

with initial condition $c_0 = 1$.

PROOF. For $n \geq 1$, let π be a Catalan path from $(0,0)$ to (n,n). Then the first step in π must be up. Now, π must end on the line $y = x$, and consider the first time after the initial up-step that π returns to the line $y = x$, which must be with a right-step. Then we can write $\pi = N\pi_1 E\pi_2$, where π_1 and π_2 are Catalan paths with a total of $2n-2$ steps, taken together. Thus if π_1 has $2k$ steps, then there are c_k choices for π_1, and c_{n-k-1} choices for π_2. The result follows by summing over $k = 0, \ldots, n-1$.

Note that the recurrence in Example 3.1, together with the initial condition, uniquely generates the sequence $\{c_n\}_{n\geq 0}$. For example, we have $c_0 = 1$ from the initial condition, then successively compute from the recurrence:

$$\begin{aligned} c_1 &= c_0^2 = 1, \\ c_2 &= c_0\, c_1 + c_1\, c_0 = 2, \\ c_3 &= c_0\, c_2 + c_1^2 + c_2\, c_0 = 5. \end{aligned}$$

We now solve the recurrence (1). The first step is to show that the generating series $C(x)$ satisfies a simple equation.

Example 3.2 *The generating series $C(x)$ satisfies the quadratic equation*

$$x\, C(x)^2 - C(x) + 1 = 0.$$

PROOF. Multiply (1) by x^n and sum for $n \geq 1$, to obtain

$$\sum_{n\geq 1} c_n x^n = \sum_{n\geq 1} \sum_{k=0}^{n-1} c_k c_{n-k-1} x^n.$$

On the LHS of this equation we have $C(x) - c_0 = C(x) - 1$, and on the RHS we change variables of summation from k, n to k, j, where $j = n - k - 1$. Then $n = k + j + 1$, and the summation range becomes $j \geq 0, k \geq 0$, so we have the equation

$$C(x) - 1 = \sum_{j\geq 0} \sum_{k\geq 0} c_k c_j x^{k+j+1} = x\, C(x)^2,$$

giving the result.

Solving the quadratic equation for $C(x)$, we obtain

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

In order to deal with the square root in this expression, we use the *Binomial Theorem*, which says that for all real a and $|x| < 1$, we have

$$(1+x)^a = 1 + \sum_{k\geq 1} \frac{a(a-1)\ldots(a-k+1)}{k!} x^k. \tag{3}$$

This is a Maclaurin series, that you will have encountered in MATH 148. We often write

$$\frac{a(a-1)\ldots(a-k+1)}{k!} = \binom{a}{k},$$

and call this the *binomial coefficient*, even when a is not a positive integer.

Applying the Binomial Theorem, we have

$$(1-4x)^{\frac{1}{2}} = 1 + \sum_{k\geq 1} \binom{\frac{1}{2}}{k} (-4)^k x^k,$$

where

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

and so, from (2) we obtain

$$C(x) = \frac{1}{2x} \pm \left(\frac{1}{2x} - \frac{1}{x} \sum_{k \geq 1} \frac{1}{k} \binom{2k-2}{k-1} x^k \right).$$

Now, in "$\pm$", we cannot select the "+" since this would mean that $C(x)$ has the term x^{-1} with negative exponent (and also that all other coefficients would be negative). Therefore, we must select "$-$", and thus have

$$C(x) = \sum_{k \geq 1} \frac{1}{k} \binom{2k-2}{k-1} x^{k-1} = \sum_{n \geq 0} \frac{1}{n+1} \binom{2n}{n} x^n,$$

from which we conclude that $c_n = \frac{1}{n+1}\binom{2n}{n}$, $n \geq 0$.

4 Lecture of January 16

Now we shall consider a different approach to generating series. Let $\mathcal{C}$ be the set of all Catalan paths from $(0,0)$ to any point on the line $y = x$. For any path π in $\mathcal{C}$, let $wt(\pi)$ denote the number of up-steps in π. Then the generating series for $\mathcal{C}$ with respect to the weight function wt, in variable x, is given by

$$\Phi_{\mathcal{C}}(x) = \sum_{\pi \in \mathcal{C}} x^{wt(\pi)}.$$

In general, a *weight function* on a set is any function whose values are restricted to nonnegative integers. The generating series above is defined for any weight function on any set, as long as the sets $\{\pi \in \mathcal{C} : wt(\pi) = n\}$ are finite for all nonnegative integer choices of n.

Let ψ denote the construction that we carried out in the proof of Example 3.1 – that is suppose that $\psi(\pi) = (\pi_1, \pi_2)$, where $\pi \in \mathcal{C} \setminus \{\varepsilon\}$. Then it is easy to easy to see that

$$\psi : \mathcal{C} \setminus \{\varepsilon\} \to \mathcal{C} \times \mathcal{C} : \pi \mapsto (\pi_1, \pi_2)$$

is a bijection, and moreover that $wt(\pi) = wt(\pi_1) + wt(\pi_2) + 1$. Here we use the notation $\mathcal{C} \times \mathcal{C}$, for ordered pairs of elements of $\mathcal{C}$. In general, for sets $\mathcal{A}$ and $\mathcal{B}$, we define the *Cartesian product* to be the set of ordered pairs

$$\mathcal{A} \times \mathcal{B} = \{(a, b) : a \in \mathcal{A}, b \in \mathcal{B}\}.$$

From the bijection ψ, we deduce that

$$\sum_{\pi \in \mathcal{C} \setminus \{\varepsilon\}} x^{wt(\pi)} = \sum_{(\pi_1, \pi_2) \in \mathcal{C} \times \mathcal{C}} x^{wt(\pi_1) + wt(\pi_2) + 1}$$

Continuing, we obtain

$$x^1 \sum_{\pi_1 \in \mathcal{C}} x^{wt(\pi_1)} \sum_{\pi_2 \in \mathcal{C}} x^{wt(\pi_2)} = x \Phi_{\mathcal{C}}(x)^2.$$

But we also have

$$\begin{aligned}\sum_{\pi\in\mathcal{C}\setminus\{\varepsilon\}} x^{wt(\pi)} &= \sum_{\pi\in\mathcal{C}} x^{wt(\pi)} - 1\\ &= \Phi_{\mathcal{C}}(x) - 1,\end{aligned}$$

and we conclude that

$$\Phi_{\mathcal{C}}(x) - 1 = x\Phi_{\mathcal{C}}(x)^2.$$

Note that this is the same quadratic equation as for $C(x)$ a lecture or two ago. This is no accident, since actually $\Phi_{\mathcal{C}}(x) = C(x)$. To see this, we have

$$\Phi_{\mathcal{C}}(x) = \sum_{n\geq 0} x^n \sum_{\substack{\pi\in\mathcal{C},\\ wt(\pi)=n}} 1 = \sum_{n\geq 0} |\{\pi\in\mathcal{C} : wt(\pi) = n\}| x^n = C(x),$$

as required.

Now we give a general instance of this way of thinking. For any set $\mathcal{A}$, with weight function $\omega : \mathcal{A} \to \{0, 1, \ldots\}$, we define

$$\Phi_{\mathcal{A}}(x) = \sum_{a\in\mathcal{A}} x^{\omega(a)}$$

to be the generating series for $\mathcal{A}$ with respect to weight function ω, in variable x. We denote the coefficient of x^n in $\Phi_{\mathcal{A}}(x)$ by

$$[x^n]\Phi_{\mathcal{A}}(x), \quad n = 0, 1, \ldots$$

We treat "$[x^n]$" as an operator, acting on the left. In terms of this coefficient notation, we have the following fundamental result for enumerative significance of a generating series.

Proposition 4.1 *If $\mathcal{A}$ is a set, with weight function $\omega : \mathcal{A} \to \{0, 1, \ldots\}$, then the number of $a \in \mathcal{A}$ with $\omega(a) = n$ is equal to*

$$[x^n]\Phi_{\mathcal{A}}(x), \quad n = 0, 1, \ldots$$

PROOF. We have

$$\Phi_{\mathcal{A}}(x) = \sum_{a\in\mathcal{A}} x^{\omega(a)},$$

and the result follows immediately.

The most important rule for generating series is the *Product Rule*.

Theorem 4.2 *If $\mathcal{A}, \mathcal{B}, \mathcal{A}\times\mathcal{B}$ have weight functions $\omega_1, \omega_2, \omega$, respectively, and the condition*

$$\omega((a, b)) = \omega_1(a) + \omega_2(b) + c$$

holds for all $(a, b) \in \mathcal{A} \times \mathcal{B}$, then

$$\Phi_{\mathcal{A}\times\mathcal{B}}(x) = x^c\Phi_{\mathcal{A}}(x)\Phi_{\mathcal{B}}(x).$$

PROOF. We have

$$\begin{aligned}\Phi_{\mathcal{A}\times\mathcal{B}}(x) &= \sum_{(a,b)\in\mathcal{A}\times\mathcal{B}} x^{\omega((a,b))} \\ &= \sum_{a\in\mathcal{A}}\sum_{b\in\mathcal{B}} x^{\omega_1(a)+\omega_2(b)+c} \\ &= x^c \sum_{a\in\mathcal{A}} x^{\omega_1(a)} \sum_{b\in\mathcal{B}} x^{\omega_2(b)},\end{aligned}$$

and the result follows.

For sets $\mathcal{A}_1, \ldots, \mathcal{A}_k$, and fixed positive integer k, the product rule extends easily to sets of k-tuples

$$\mathcal{A}_1 \times \ldots \times \mathcal{A}_k = \{(a_1, \ldots, a_k) : a_1 \in \mathcal{A}_1, \ldots, a_k \in \mathcal{A}_k\}.$$

Theorem 4.3 *(Product Rule for k-tuples) If $\mathcal{A}_1, \ldots, \mathcal{A}_k, \mathcal{A}_1 \times \ldots \times \mathcal{A}_k$ have weight functions $\omega_1, \ldots, \omega_k, \omega$, respectively, and the condition*

$$\omega((a_1, \ldots, a_k)) = \omega_1(a_1) + \ldots + \omega_k(a_k) + c$$

holds for all $(a_1, \ldots, a_k) \in \mathcal{A}_1 \times \ldots \times \mathcal{A}_k$, then

$$\Phi_{\mathcal{A}_1\times\ldots\times\mathcal{A}_k}(x) = x^c \Phi_{\mathcal{A}_1}(x) \ldots \Phi_{\mathcal{A}_k}(x).$$

Example 4.4 *Find the number of solutions to $t_1 + \ldots + t_k = n$, where $t_1, \ldots, t_k$ are non-negative integers.*

SOLUTION. Let $\mathcal{S} = \mathcal{A}_1 \times \ldots \times \mathcal{A}_k$, where $\mathcal{A}_i = \{0, 1, 2, \ldots\}$, for $i = 1, \ldots, k$. For $(a_1, \ldots, a_k) \in \mathcal{S}$, let $\omega((a_1, \ldots, a_k)) = a_1 + \ldots + a_k$. Then the answer to this problem is precisely the number of elements in $\mathcal{S}$ with weight function value equal to n, which is equal to

$$[x^n]\Phi_{\mathcal{S}}(x).$$

Now note that $\omega((a_1, \ldots, a_k)) = \tau(a_1) + \ldots + \tau(a_k)$, where τ is the identity function, so the product rule for k-tuples gives

$$\Phi_{\mathcal{S}}(x) = \Phi_{\mathcal{A}_1}(x) \ldots \Phi_{\mathcal{A}_k}(x) = \Phi_{\{0,1,2,\ldots\}}(x)^k,$$

where

$$\Phi_{\{0,1,2,\ldots\}}(x) = \sum_{i\in\{0,1,2,\ldots\}} x^i = 1 + x^1 + x^2 + \ldots,$$

since the weight function is the identity. Thus the answer is given by

$$[x^n](1 + x^1 + x^2 + \ldots)^k = [x^n]\left((1-x)^{-1}\right)^k = [x^n](1-x)^{-k},$$

by summing the geometric series.

5 Lecture of January 18

But

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

from the negative binomial theorem, given below, and the Example is finished.

We now consider the *negative binomial theorem*, the binomial theorem in the case that the exponent a is a negative integer. If m is a positive integer, then we have

$$(1-x)^{-m} = 1 + \sum_{i \geq 1} \binom{-m}{i} (-1)^i x^i,$$

where

$$\begin{aligned} \binom{-m}{i}(-1)^i &= \frac{-m(-m-1)\ldots(-m-i+1)}{i!}(-1)^i \\ &= \frac{m(m+1)\ldots(m+i-1)}{i!} \\ &= \binom{m+i-1}{i} = \binom{m+i-1}{m-1}, \end{aligned}$$

so we have, for any positive integer m,

$$(1-x)^{-m} = \sum_{i \geq 0} \binom{m+i-1}{i} x^i = \sum_{i \geq 0} \binom{m+i-1}{m-1} x^i. \tag{4}$$

As a generalization of Example 4.4, we have the following result. The proof is omitted, since it is identical to the proof given in Example 4.4.

Lemma 5.1 *The number of solutions to $t_1 + \ldots + t_k = n$, where $t_i \in \mathcal{A}_i$, $i = 1, \ldots, k$, for given subsets of the nonnegative integers $\mathcal{A}_i$, is given by*

$$[x^n] \prod_{i=1}^{k} \Phi_{\mathcal{A}_i}(x),$$

where $\Phi_{\mathcal{A}_i}(x) = \sum_{j \in \mathcal{A}_i} x^j$.

For example, from this result we deduce that the number of solutions to $t_1 + \ldots + t_k = n$, for positive integers t_i, $i = 1, \ldots, k$, is equal to

$$\begin{aligned} [x^n]\left(x + x^2 + \ldots\right)^k &= [x^n]\left(x(1-x)^{-1}\right)^k = [x^n]x^k(1-x)^{-k} \\ &= [x^n] \sum_{i \geq 0} \binom{k+i-1}{k-1} x^{k+i} = \binom{n-1}{k-1}, \end{aligned}$$

where we have used the geometric series for the first equality, and the value $i = n - k$ for the last equality.

Of course, this problem can be solved without generating series, using various more elementary methods, such as set bijections. For example, we can note that if $t_1+\ldots+t_k=n$, where $t_1,\ldots,t_k$ are positive integers, then $\{t_1, t_1+t_2,\ldots,t_1+\ldots+t_{k-1}\}$ is a $(k-1)$-subset of $\{1,\ldots,n-1\}$. Moreover, if $\{\alpha_1,\ldots,\alpha_{k-1}\}$, with $1\le\alpha_1<\ldots<\alpha_{k-1}\le n-1$, is a $(k-1)$-subset of $\{1,\ldots,n-1\}$, then $(\alpha_1,\alpha_2-\alpha_1,\ldots,\alpha_{k-1}-\alpha_{k-2},n-\alpha_{k-1})$ is a solution to the equation $t_1+\ldots+t_k=n$. It is straightforward then to check that we have a bijection, which proves that the number of solutions in this case is $\binom{n-1}{k-1}$.

However, when we modify such problems even in a simple way, they can become very complicated to deal with by elementary means, yet the generating series methodology handles them straightforwardly, using perhaps more binomial expansions. Consider the following example.

Example 5.2 *Find the number of solutions to $t_1+\ldots+t_k=n$, where $t_1,\ldots,t_k$ are positive integers not equal to 3.*

SOLUTION. From Lemma 5.1, with $\mathcal{A}_i=\{1,2,4,5,\ldots\}$, for $i=1,\ldots,k$, the answer is given by

$$\begin{aligned}[x^n]\left(\frac{x}{1-x}-x^3\right)^k &= [x^n]\sum_{i=0}^{k}\binom{k}{i}\left(\frac{x}{1-x}\right)^{k-i}(-x^3)^i\\ &= [x^n]\sum_{i=0}^{k}\binom{k}{i}(-1)^i x^{k+2i}(1-x)^{-(k-i)}\end{aligned}$$

Continuing with our solution, from the negative binomial theorem we obtain

$$[x^n]\sum_{i=0}^{k}\sum_{j\ge0}\binom{k}{i}(-1)^i\binom{k-i+j-1}{j}x^{k+2i+j}=\sum_i\sum_j\binom{k}{i}(-1)^i\binom{k-i+j-1}{j},$$

where the double sum on the righthandside is over all $j\ge0$, $0\le i\le k$, subject to the restriction that $k+2i+j=n$. Thus we can replace j by $n-k-2i$, and write this as a single sum

$$\sum_i\binom{k}{i}(-1)^i\binom{n-3i-1}{k-i-1},$$

where this sum ranges from 0 to $\min\{k,\lfloor\frac{n-k}{2}\rfloor\}$. We have used the *floor function* $\lfloor x\rfloor$, for the real number x, whose value is the greatest integer not greater than x. The inequality $i\le\lfloor\frac{n-k}{2}\rfloor$ arises because $j\ge0$, which implies that $n-k-2i\ge0$. The lower index $k-i-1$ in the last binomial coefficient arises from the identity $\binom{a+b}{a}=\binom{a+b}{b}$, for nonnegative integers a,b.

6 Lecture of January 21

Let's write the summation above explicitly, to obtain

$$\binom{n-1}{k-1}-\binom{k}{1}\binom{n-4}{k-2}+\binom{k}{2}\binom{n-7}{k-3}-\ldots.$$

An elementary way of deriving this is to note that the first term, $\binom{n-1}{k-1}$, is the total number of solutions in positive integers. Then, for the second term, we have $\binom{n-4}{k-2}$ as the number of solutions for which $t_j = 3$ for any fixed j (since, with $j = k$, we have $t_1 + \ldots + t_{k-1} + 3 = n$, so $t_1 + \ldots + t_{k-1} = n - 3$). But there are $\binom{k}{1} = k$ choices for j. For the third term, we have $\binom{n-7}{k-3}$ as the number of solutions for which $t_j = t_m = 3$ for any fixed $j < m$ (since, with $j = k-1, m = k$, we have $t_1 + \ldots + t_{k-2} + 3 + 3 = n$, so $t_1 + \ldots + t_{k-2} = n - 6$). But there are $\binom{k}{2}$ choices for j, m. All terms arise in this way, giving the answer required.

Example 6.1 *Find the number of solutions to $t_1 + \ldots + t_k = n$, where $t_1, \ldots, t_k$ are positive integers less than or equal to 6 (this arises as the number of ways of getting n as the sum of k rolls of a die).*

SOLUTION. From Lemma 5.1, with $\mathcal{A}_i = \{1, 2, \ldots, 6\}$, for $i = 1, \ldots, k$, the answer is given by

$$\begin{aligned}
[x^n]\left(x + x^2 + \ldots + x^6\right)^k &= [x^n]\left(\frac{x - x^7}{1 - x}\right)^k = [x^n]x^k(1 - x^6)^k(1 - x)^{-k} \\
&= [x^n]\sum_{i=0}^{k}\sum_{j\geq 0}\binom{k}{i}(-1)^i\binom{k + j - 1}{j}x^{k+6i+j} \\
&= \sum_{i}\sum_{j}\binom{k}{i}(-1)^i\binom{k + j - 1}{j},
\end{aligned}$$

where the double sum on the righthandside is over all $j \geq 0$, $0 \leq i \leq k$, subject to the restriction that $k + 6i + j = n$. For the first equality above, we have evaluated a finite geometric sum. The general formula is

$$a + ax + \ldots + ax^{n-1} = \frac{a - ax^n}{1 - x},$$

where it is often convenient to notice that ax^n is equal to $ax^{n-1} \cdot x$, which can be regarded as the "next " term in the geometric sum if it were to extend to infinity.

Another general formula that has been used in the last two examples is the Binomial theorem for nonnegative integer exponent, which gives

$$(A + B)^k = \sum_{i=0}^{k}\binom{k}{i}A^iB^{k-i} = \sum_{i=0}^{k}\binom{k}{i}A^{k-i}B^i.$$

We now consider *compositions* of a integer. For positive integers n, k, a composition of n with k parts is a k-tuple $(c_1, \ldots, c_k)$ of positive integers such that $c_1 + \ldots + c_k = n$. We call $c_1, \ldots, c_k$ the *parts* of the composition. In addition, by convention, we also say that there is single (empty) composition of 0, with 0 parts, and denote this composition by ε.

Now note that a composition of n with k parts is precisely a solution to the equation $t_1 + \ldots + t_k = n$ in the positive integers, and that this is bijective, so the number of compositions of n with k parts is given by $\binom{n-1}{k-1}$, from the first example following Lemma 5.1

above. Then, summing over all choices of k, for $n \geq 1$, we obtain that the number of compositions of n is given by

$$\sum_{k=1}^{n} \binom{n-1}{k-1} = \sum_{i=0}^{n-1} \binom{n-1}{i} 1^i = (1+1)^{n-1} = 2^{n-1},$$

from the Binomial Theorem (or, combinatorially, by counting the subsets of an $(n-1)$-set).

We now take a direct generating series approach, by considering the set of all compositions (i.e., any k, any n) in which the parts are restricted to a subset $\mathcal{A}$ of the positive integers $\mathcal{N}$. Then this set of compositions is given by

$$\mathcal{S} = \{\varepsilon\} \cup \mathcal{A} \cup \mathcal{A}^2 \cup \ldots.$$

Note that the sets on the RHS above are pairwise disjoint (i.e., every element of $\mathcal{S}$ is contained in exactly one of the sets on the RHS). To emphasize that all sets in the union are disjoint, we write

$$\mathcal{S} = \{\varepsilon\} \,\dot\cup\, \mathcal{A} \,\dot\cup\, \mathcal{A}^2 \,\dot\cup\, \ldots,$$

and refer to "$\dot\cup$" as "disjoint union". Now define a weight function for $\mathcal{S}$, by

$$wt((c_1, \ldots, c_k)) = c_1 + \ldots + c_k,$$

for any $k \geq 0$ (when $k = 0$, the element of $\mathcal{S}$ is ε, consistent with the convention that the empty sum above is 0). Then the compositions of n with parts in $\mathcal{A}$ are precisely the elements of $\mathcal{S}$ of weight n, so our basic enumerative result for generating series implies immediately that the number of compositions of n is

$$[x^n]\Phi_{\mathcal{S}}(x).$$

To deal with the disjoint union in $\mathcal{S}$, we use the following result, the *Sum Rule*.

Theorem 6.2 *For any weight function defined on $\mathcal{A} \,\dot\cup\, \mathcal{B}$, we have*

$$\Phi_{\mathcal{A} \,\dot\cup\, \mathcal{B}}(x) = \Phi_{\mathcal{A}}(x) + \Phi_{\mathcal{B}}(x),$$

(where, on the RHS above, the weight function is simply the restriction of the weight function to the subsets $\mathcal{A}$, $\mathcal{B}$, respectively.)

Proof.

$$LHS = \sum_{c \in \mathcal{A} \,\dot\cup\, \mathcal{B}} x^{wt(c)} = \sum_{a \in \mathcal{A}} x^{wt(a)} + \sum_{b \in \mathcal{B}} x^{wt(b)} = RHS.$$

Applying the Sum Rule to $\mathcal{S}$, we obtain

$$\Phi_{\mathcal{S}}(x) = \sum_{k \geq 0} \Phi_{\mathcal{A}^k}(x). \tag{5}$$

Note that, in order to extend the Sum Rule to the infinite disjoint union in $\mathcal{S}$, we need to ensure that the elements of any fixed weight in $\mathcal{S}$ will only appear in a finite number of sets in the disjoint union. But this is immediate, since elements of weight n can only appear in the sets $\mathcal{A}^k$ with $k \leq n$. Now note that, for $(c_1, \ldots, c_k) \in \mathcal{S}$, we have

$$wt((c_1, \ldots, c_k)) = \tau(c_1) + \ldots + \tau(c_k),$$

where τ is the identity function, so we can apply the Product Rule to (5), to obtain, using the geometric series,

$$\Phi_{\mathcal{S}}(x) = \sum_{k \geq 0} \left(\Phi_{\mathcal{A}}(x)\right)^k = \frac{1}{1 - \Phi_{\mathcal{A}}(x)}. \tag{6}$$

7 Lecture of January 23

As a first example of (6), consider the total number of compositions of n. Here we have $\mathcal{A} = \mathcal{N}$, and $\Phi_{\mathcal{N}}(x) = \sum_{n \geq 1} x^n = \frac{x}{1-x}$, so the total number of compositions of n is given by

$$[x^n] \frac{1}{1 - \frac{x}{1-x}} = [x^n] \frac{1-x}{1-2x} = [x^n] \left(1 + \frac{x}{1-2x}\right) = [x^n] \left(1 + \sum_{i \geq 0} 2^i x^{i+1}\right),$$

and this gives 2^{n-1} for $n \geq 1$ (since we choose $i + 1 = n$), and 1 for $n = 0$ (which checks our first solution above).

For a second example of (6), for $n \geq 0$ let a_n be the number of compositions of n in which no part is equal to 3. In this case, we have $\mathcal{A} = \{1, 2, 4, 5, \ldots\} = \mathcal{N} \backslash \{3\}$, so $\Phi_{\mathcal{A}}(x) = \frac{x}{1-x} - x^3$, and we obtain

$$\sum_{n \geq 0} a_n x^n = \frac{1}{1 - \left(\frac{x}{1-x} - x^3\right)} = \frac{1-x}{1 - 2x + x^3 - x^4}.$$

This could be expanded in powers of x to get an explicit formula for a_n. For example, the first step in obtaining such an expansion is to use a geometric series, giving

$$\sum_{n \geq 0} a_n x^n = (1-x) \sum_{k \geq 0} (2x - x^3 + x^4)^k.$$

However, instead of giving an explicit formula for a_n, we are going to obtain a recurrence equation. Multiplying on both sides by the denominator $1 - 2x + x^3 - x^4$, we obtain

$$\left(1 - 2x + x^3 - x^4\right) \sum_{n \geq 0} a_n x^n = 1 - x + 0\, x^2 + \ldots,$$

an equality of two power series in x. But this means that the coefficient of each power of x on the LHS must equal the coefficient of the corresponding power of x on the RHS, and this gives us a system of equations for $\{a_n\}_{n \geq 0}$, as follows:

$$[x^0] : a_0 = 1,$$

$$[x^1] : a_1 - 2a_0 = -1, \text{ which gives } a_1 = 1,$$

$$[x^2] : a_2 - 2a_1 = 0, \text{ which gives } a_2 = 2,$$
$$[x^3] : a_3 - 2a_2 + a_0 = 0, \text{ which gives } a_3 = 3,$$
$$[x^m] : a_m - 2a_{m-1} + a_{m-3} - a_{m-4} = 0, \text{ for } m \geq 4.$$

This gives us the *recurrence equation*

$$a_m = 2a_{m-1} - a_{m-3} + a_{m-4}, \qquad m \geq 4, \tag{7}$$

with *initial conditions* $a_0 = a_1 = 1, a_2 = 2, a_3 = 3$. This is called a *linear* recurrence with constant coefficients, because a_m is determined by a linear function of $a_{m-1}, a_{m-2}, \ldots$, and the coefficients in this linear function are constants (i.e., not functions of m).

Note that the initial conditions could also be obtained by elementary counting: we have $a_i =$ the number of compositions of i in which 3 is never a part, so $a_0 = a_1 = 1$, $a_2 = 2$ (equal to the total number of compositions of i in each case, since no part can be equal to 3 in a composition of i when $i < 3$). Also $a_3 = 3$, since the compositions in this case are $(1,1,1)$, $(1,2)$, $(2,1)$.

We are now going to give an alternative proof of recurrence equation (7), by giving a direct bijection between appropriate sets of compositions.

We begin with something easier, by considering the recurrence

$$c_m = 2c_{m-1}, \qquad m \geq 1, \tag{8}$$

where c_m is equal to the number of compositions of m. Define $\mathcal{C}_m$ to be the set of compositions of m, for $m \geq 0$. Also, define $\mathcal{C}_{m,i}$ to be the set of compositions of m in which the last part is equal to i, for $m, i \geq 1$, and let $\mathcal{C}'_m = \dot{\bigcup}_{i \geq 2} \mathcal{C}_{m,i}$, $m \geq 1$. Then we clearly have

$$\mathcal{C}_m = \mathcal{C}_{m,1} \,\dot{\cup}\, \mathcal{C}'_m, \qquad m \geq 1. \tag{9}$$

Now we give two combinatorial bijections. The first is

$$\mathcal{C}_{m,1} \cong \mathcal{C}_{m-1}, \quad m \geq 1,$$

which simply says that if the last part (which equals 1) is removed from a composition in $\mathcal{C}_{m,1}$, then we obtain a composition in $\mathcal{C}_{m-1}$, and that this is bijective. The second bijection is

$$\mathcal{C}'_m \cong \mathcal{C}_{m-1}, \quad m \geq 1,$$

which says that if we subtract one from the last part (which equals 2 or more) of a composition in $\mathcal{C}'_m$, then we obtain a composition in $\mathcal{C}_{m-1}$, and that this is bijective. But $|\mathcal{C}_m| = c_m$, $m \geq 0$, so the first bijection above gives $|\mathcal{C}_{m,1}| = c_{m-1}$, $m \geq 1$, and the second bijection above gives $|\mathcal{C}'_m| = c_{m-1}$, $m \geq 1$, and now (8) follows immediately from (9).

For (7), define $\mathcal{A}_m$, for $m \geq 0$, to be the set of compositions of m in which no part is equal to 3, and let $\mathcal{A}_{m,i}$, for $m, i \geq 1$, be the set of compositions in $\mathcal{A}_m$ in which the last part is equal to i. Also, let $\mathcal{A}'_m = \mathcal{A}_{m,2} \,\dot{\cup}\, \mathcal{A}_{m,4} \,\dot{\cup}\, \mathcal{A}_{m,5} \,\dot{\cup}\, \ldots$, $m \geq 1$. Then we immediately have

$$\mathcal{A}_m = \mathcal{A}_{m,1} \,\dot{\cup}\, \mathcal{A}'_m, \qquad m \geq 1, \tag{10}$$

Again we give two combinatorial bijections. The first is identical as for all compositions; it is

$$\mathcal{A}_{m,i} \cong \mathcal{A}_{m-i}, \quad m \geq i,$$

which simply says that if the last part (equal to i) is removed from a composition in $\mathcal{A}_{m,i}$, then we obtain a composition in $\mathcal{A}_{m-i}$, and that this is bijective. The second bijection is more complicated than for all compositions; it is

$$\mathcal{A}'_m \setminus \mathcal{A}_{m,4} \cong \mathcal{A}_{m-1} \setminus \mathcal{A}_{m-1,2}, \qquad m \geq 2,$$

which says that if we subtract one from the last part of a composition on the LHS, then we obtain a composition on the RHS, and that this is bijective. But $|\mathcal{A}_m| = a_m$, $m \geq 0$, so (10) gives

$$a_m = |\mathcal{A}_{m,1}| + |\mathcal{A}'_m|, \qquad m \geq 1,$$

and (7) follows immediately from the two bijections.

Before moving on to our next topic, we record the general result that gives a linear recurrence equation with constant coefficients for the sequence of coefficients in any rational function (*ratio* of polynomials).

Lemma 7.1 *Suppose that*

$$\sum_{i \geq 0} c_i x^i = \frac{P(x)}{1 + \sum_{j=1}^{k} q_j x^j},$$

where $P(x)$ is a polynomial of degree less than k. Then

$$c_m + \sum_{j=1}^{k} q_j c_{m-j} = 0, \qquad m \geq k,$$

with initial conditions determined by the coefficients of $P(x)$.

8 Lecture of January 25

As is often done in mathematics, we will now introduce some notation that allows compact expression for sets of combinatorial objects. We represent compositions in string notation, as strings (ordered lists) of the parts (simply by removing parentheses and commas). We then use the notation

$$\mathcal{D}^* = \{\varepsilon\} \cup \mathcal{D} \cup \mathcal{D}^2 \cup \mathcal{D}^3 \cup \ldots,$$

for any set of strings $\mathcal{D}$. Thus, the set of all compositions corresponds to the case $\mathcal{D} = \mathcal{N}$, and the set of compositions with no parts equal to 3 corresponds to the case $\mathcal{B} = (\mathcal{N} \setminus \{3\})^*$. Here we use the notation

$$\mathcal{A}\mathcal{B} = \{ab : a \in \mathcal{A}, b \in \mathcal{B}\},$$

for sets of strings $\mathcal{A}$, $\mathcal{B}$. For strings a, b, we call ab the *concatenation* product; this product has an identity element, the empty string ε, and it is associative. However, it is not commutative, and in general and the only string with a multiplicative inverse is ε.

We now consider *partitions* of an integer. For positive integers n, k, a partition of n with k parts is a string $a_1 \ldots a_k$ of positive integers, weakly ordered with $a_1 \leq \ldots \leq a_k$, and such that $a_1 + \ldots + a_k = n$. We call $a_1, \ldots, a_k$ the *parts* of the partition, and we also have a single empty partition of 0, with 0 parts, denoted by ε. Let p_n be the number of partitions of n. Then we have $p_0 = 1$, $p_1 = 1$ (partition 1), $p_2 = 2$ (partitions 11 and 2), $p_3 = 3$ (partitions 3, 12 and 111). Let $\mathcal{P}$ be the set of all partitions (i.e., any n, any k). Now define a weight function for $\mathcal{P}$, by

$$wt(a_1 \ldots a_k) = a_1 + \ldots + a_k,$$

for any $k \geq 0$. The the partitions of n are precisely the elements of $\mathcal{P}$ of weight n, so our basic enumerative result for generating series implies immediately that the number of partitions of n is

$$p_n = [x^n]\Phi_{\mathcal{P}}(x).$$

But, using string notation, we have

$$\mathcal{P} = \{1\}^*\{2\}^*\{3\}^* \ldots,$$

and using the product and sum rules, we obtain

$$\Phi_{\mathcal{P}}(x) = \prod_{m \geq 1} \Phi_{\{m\}^*}(x) = \prod_{m \geq 1} \frac{1}{1 - x^m}.$$

In general, it is not easy to obtain a compact formula for the coefficients in this infinite product, and indeed, there is no nice formula for p_n, the number of partitions of n. However, we can still obtain difficult facts about various sets of partitions from their generating series.

Consider the set $\mathcal{D}$, consisting of partitions with *distinct* parts, and the set $\mathcal{O}$, consisting of partitions with *odd* parts. Let d_n be the number of partitions of n in $\mathcal{D}$, and o_n be the number of partitions of n in $\mathcal{O}$, $n \geq 0$. Then, using string notation, we have

$$\mathcal{D} = \{\varepsilon, 1\}\{\varepsilon, 2\}\{\varepsilon, 3\} \ldots, \qquad \mathcal{O} = \{1\}^*\{3\}^*\{5\}^* \ldots,$$

so we have, from the product and sum rules,

$$\sum_{n \geq 0} d_n x^n = \Phi_{\mathcal{D}}(x) = \prod_{m \geq 1} (1 + x^m),$$

and

$$\sum_{n \geq 0} o_n x^n = \Phi_{\mathcal{O}}(x) = \prod_{j \geq 1} \frac{1}{1 - x^{2j-1}}.$$

But, we then obtain

$$\Phi_{\mathcal{D}}(x) = \prod_{m \geq 1} \frac{1 - x^{2m}}{1 - x^m} = \frac{\prod_{i \geq 1}(1 - x^{2i})}{\prod_{m \geq 1}(1 - x^m)} = \Phi_{\mathcal{O}}(x),$$

and we conclude that $d_n = o_n$, for each $n \geq 0$. Since our derivation of this result relies on cancellations in infinite products forms of generating series, we shall now describe a completely elementary bijection that gives an alternative proof of this result (to help give

some confidence in the generating series methods). First, to illustrate the result itself, consider the case $n = 6$. Then, in $\mathcal{D}$, the partitions of 6 are given by 6, 15, 24, 123, and, in $\mathcal{O}$, the partitions of 6 are given by 15, 33, 1113, 111111, so here we have $d_6 = 4 = o_6$. As another example, consider $n = 11$, where, in $\mathcal{D}$, the partitions of 11 are given by

$$11,\ 1\ 10,\ 2\ 9,\ 3\ 8,\ 4\ 7,\ 5\ 6,\ 1\ 2\ 8,\ 1\ 3\ 7,\ 1\ 4\ 6,\ 2\ 3\ 6,\ 2\ 4\ 5,\ 1\ 2\ 3\ 5,$$

and, in $\mathcal{O}$, the partitions of 11 are given by

$$11,\ 1\ 1\ 9,\ 1\ 1\ 1\ 1\ 7,\ 1\ 3\ 7,\ 1\ 5\ 5,\ 3\ 3\ 5,\ 1\ 1\ 1\ 3\ 5,\ 1\ 1\ 1\ 1\ 1\ 1\ 5,$$

$$1\ 1\ 3\ 3\ 3,\ 1\ 1\ 1\ 1\ 1\ 3\ 3,\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 3,\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1,$$

so here we have $d_{11} = 12 = o_{11}$. Now, for the bijection: we describe a mapping $\psi : \mathcal{D} \to \mathcal{O}$ that is *weight-preserving*. For a partition $\delta \in \mathcal{D}$, consider each part d in δ. Write $d = 2^a\, b$, where a is a nonnegative integer, and b is an odd positive integer. Then create 2^a parts in $\psi(\delta)$, each equal to b, and repeat this for all parts d in δ. This mapping is weight-preserving, since if δ is a partition of n, then $\psi(\delta)$ is also a partition of n. For example, we have

$$\psi(2\ 3\ 8\ 12\ 14\ 21) = 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 3\ 3\ 3\ 3\ 3\ 7\ 7\ 21,$$

where both are partitions of 60.

Now, ψ is actually a bijection, since it is easy to describe its inverse: consider an arbitrary partition $\theta \in \mathcal{O}$. For each odd part b that appears in θ, let $N_b \geq 1$ be the number of times that b appears as a part. Now write N_b as a sum of distinct nonnegative powers of 2 (this is unique, since binary representations of positive integers are unique), say as

$$N_b = \sum_{i=1}^{m_b} 2^{a_{b,i}}$$

Then the parts that appear in $\psi^{-1}(\theta)$ are $2^{a_{b,i}}b$, for $i = 1, \ldots, m_b$. For example,

$$\psi^{-1}(3\ 3\ 3\ 3\ 3\ 3\ 7\ 7\ 7) =\ 6\ 7\ 12\ 14,$$

since $N_3 = 6 = 2^1 + 2^2$, and $N_7 = 3 = 2^0 + 2^1$.

The above bijective proof relies on the uniqueness of binary representations, a well known fact that we now prove using generating series and partitions. Let $\mathcal{B}$ be the set of partitions with distinct parts, that are all nonnegative powers of 2, and let b_n be the number of partitions of n in $\mathcal{B}$. Now,

$$\mathcal{B} = \{\varepsilon, 1\}\{\varepsilon, 2\}\{\varepsilon, 4\}\{\varepsilon, 8\}\ldots,$$

and b_n is exactly the number of binary representations of n, so we have

$$\begin{aligned}
\sum_{n\geq 0} b_n x^n &= \Phi_{\mathcal{B}}(x) = \prod_{m\geq 0}(1 + x^{2^m}) \\
&= \prod_{m\geq 0} \frac{1 - x^{2^{m+1}}}{1 - x^{2^m}} \\
&= \frac{1}{1-x} = \sum_{n\geq 0} x^n,
\end{aligned}$$

so we conclude that $b_n = 1$ for each $n \geq 0$, and we have completed a generating series proof of the uniqueness of binary representations.

9 Lecture of January 28

We now consider the set $\mathcal{A}$ of k-subsets $\{\alpha_1, \ldots, \alpha_k\}$ of $\{1, \ldots, n\}$, with the convention that $\alpha_1 < \ldots < \alpha_k$. Now define $d_1 = \alpha_1$, $d_2 = \alpha_2 - \alpha_1$, ..., $d_k = \alpha_k - \alpha_{k-1}$, $d_{k+1} = n - \alpha_k$. Note that $d_1 \geq 1$, ..., $d_k \geq 1$, $d_{k+1} \geq 0$, and that $d_1 + \ldots + d_{k+1} = n$. In fact, let $\mathcal{B}$ be the set of $(k+1)$-tuples $(d_1, \ldots, d_{k+1})$, such that $d_1 \geq 1$, ..., $d_k \geq 1$, $d_{k+1} \geq 0$, and $d_1 + \ldots + d_{k+1} = n$. Then the above mapping from $\mathcal{A}$ to $\mathcal{B}$ is a bijection, since we can invert it by $\alpha_i = d_1 + \ldots + d_i$, $i = 1, \ldots, k$. We call this the *difference-partial sum* bijection. Now, clearly we have $|\mathcal{A}| = \binom{n}{k}$, and as an exercise, we also determine $|\mathcal{B}|$, to check that these are equal. First note that $\mathcal{B}$ consists of the elements of $\mathcal{N}^k\mathcal{N}_0$ of weight n, where $\mathcal{N}_0$ consists of the nonnegative integers, and the weight of a $(k+1)$-tuple $(d_1, \ldots, d_{k+1})$ is equal to $d_1 + \ldots + d_{k+1}$. Then, as in our earlier work on the number of solutions to equations, we use the product rule to obtain

$$\begin{aligned}
|\mathcal{B}| &= [x^n]\Phi_{\mathcal{N}^k\mathcal{N}_0}(x) \\
&= [x^n](x^1 + x^2 + \ldots)^k(1 + x^1 + x^2 + \ldots) \\
&= [x^n]\left(\frac{x}{1-x}\right)^k \frac{1}{1-x} \\
&= [x^{n-k}](1-x)^{-k-1} \\
&= \binom{n-k+k+1-1}{n-k} = \binom{n}{n-k} = \binom{n}{k},
\end{aligned}$$

as we expect. How can this be useful ? Consider a modification of this problem.

Example 9.1 *Determine the number of k-subsets $\{\alpha_1, \ldots, \alpha_k\}$ of $\{1, \ldots, n\}$, such that $\alpha_i \equiv i(\mod 3)$, $i = 1, \ldots, k$ (where we have $\alpha_1 < \ldots < \alpha_k$).*

SOLUTION. Let $\mathcal{N}_{1,3}$ denote the set $\{1, 4, 7, \ldots\}$ of positive integers congruent to $1(\mod 3)$. Then, applying the difference-partial sum bijection, the required number is equal to the number of $(k+1)$-tuples in $\mathcal{N}_{1,3}^k\mathcal{N}_0$ of weight n, and by the product rule this equals

$$[x^n](x^1+x^4+x^7+\ldots)^k(1+x^1+x^2+\ldots) = [x^n]\left(\frac{x}{1-x^3}\right)^k \frac{1}{1-x} = [x^{n-k}](1-x^3)^{-k}(1-x)^{-1}.$$

If we write $(1-x)^{-1}$ as $(1+x+x^2)(1-x^3)^{-1}$, then the answer becomes

$$[x^{n-k}](1+x+x^2)(1-x^3)^{-k-1} = \binom{k+1+\lfloor\frac{n-k}{3}\rfloor - 1}{\lfloor\frac{n-k}{3}\rfloor} = \binom{k+\lfloor\frac{n-k}{3}\rfloor}{k}.$$

Now consider a further modification, in which the subset does not have fixed size.

Example 9.2 *Determine the number of subsets $\{\alpha_1, \ldots, \alpha_k\}$ of $\{1, \ldots, n\}$, such that $\alpha_i \equiv i(\mod 3)$, $i = 1, \ldots, k$ (where we have $\alpha_1 < \ldots < \alpha_k$), where k is not fixed, and can be any nonegative integer.*

SOLUTION. Here, applying the difference-partial sum bijection, the required number is equal to the number of elements of $\mathcal{N}_{1,3}^*\mathcal{N}_0$ of weight n, and by the sum and product rules this equals

$$[x^n]\Phi_{\mathcal{N}_{1,3}^*\mathcal{N}_0}(x) = [x^n]\frac{\Phi_{\mathcal{N}_0}(x)}{1-\Phi_{\mathcal{N}_{1,3}}(x)} = [x^n]\frac{\frac{1}{1-x}}{1-\frac{x}{1-x^3}} = [x^n]\frac{1+x+x^2}{1-x-x^3},$$

for $n \geq 0$.

10 Lecture of January 30

We are now going to consider strings more formally. Initially, the alphabet will be $\{0,1\}$. A $\{0,1\}$-string is a finite, ordered list of 0's and 1's, for example $a = 10011$ and $b = 0001$. We have the *concatenation* product, for example $ab = 100110001$ and $ba = 000110011$, and the *empty* string ε is the identity element for this multiplication, so $\varepsilon a = a = a\varepsilon$ for all strings a. (This multiplication is closed and associative, with identity, but it is not commutative, and in general there is no inverse. Such an algebraic system is called a *monoid*.) The *length* of a string is the number of symbols in it, so length(0010)=4, and length(ε)=0, for example. If $\mathcal{A}$ and $\mathcal{B}$ are sets of strings, then we define

$$\mathcal{AB} = \{ab : a \in \mathcal{A}, b \in \mathcal{B}\}, \quad \mathcal{A}^* = \{\varepsilon\} \cup \mathcal{A} \cup \mathcal{AA} \cup \mathcal{AAA} \cup \ldots,$$

and we usually write powers with exponent notation, e.g., $\mathcal{AAA} = \mathcal{A}^3$. Note the use of set notation above. Consider $\mathcal{A} = \{0, 00\}$, $\mathcal{B} = \{1, 11\}$, $\mathcal{C} = \{\varepsilon, 0\}$. Then we have $\mathcal{AB} = \{01, 011, 001, 0011\}$, but we have $\mathcal{AC} = \{0, 00, 000\}$, since in $\mathcal{AC}$, the same string 00 is formed in two ways. We say that the elements of $\mathcal{AB}$ are *uniquely created*, and that the elements of $\mathcal{AC}$ are not uniquely created.

For an arbitrary set of $\{0,1\}$-strings $\mathcal{A}$, consider the generating series

$$\Phi_{\mathcal{A}}(x) = \sum_{a \in \mathcal{A}} x^{\text{length}(a)}.$$

Then we have the following results.

Theorem 10.1 *(a) If the elements of $\mathcal{AB}$ are uniquely created, then $\Phi_{\mathcal{AB}}(x) = \Phi_{\mathcal{A}}(x)\Phi_{\mathcal{B}}(x)$.*
(b) If the elements of $\mathcal{A}^*$ are uniquely created, then $\Phi_{\mathcal{A}^*}(x) = (1-\Phi_{\mathcal{A}}(x))^{-1}$.

The proofs are omitted, since they are simply the translation of the sum and product rules for the weight function "length", and using the fact that length(ab)= length(a)+ length(b) for all strings a and b. Note in (b) that since elements of $\mathcal{A}^*$ are uniquely created, then the unions in $\mathcal{A}^*$ are all disjoint, so we can use the sum rule.

As a first example, we determine the number of $\{0,1\}$-strings of length n, for each $n \geq 0$. Clearly this number is 2^n, since there are 2 choices in each of the n positions. Using the generating series method, we have that this number is the number of elements in $\{0,1\}^*$ of weight n, and that the elements of $\{0,1\}^*$ are uniquely created, so it is equal to

$$[x^n]\Phi_{\{0,1\}^*}(x) = [x^n]\left(1-\Phi_{\{0,1\}}(x)\right)^{-1} = [x^n](1-2x)^{-1} = 2^n, \quad n \geq 0,$$

from Theorem 10.1, agreeing with our answer above.

As a second example, let $\mathcal{S}$ be the set of $\{0,1\}$-strings in which there is no substring "111", and find the number of elements of $\mathcal{S}$ of length n, $n \geq 0$.

Now, by considering the 0's, we have

$$\mathcal{S} = \{0, 10, 110\}^\star \{\varepsilon, 1, 11\},$$

and the elements of $\mathcal{S}$ are uniquely created in this decomposition, so we conclude that the number of strings in $\mathcal{S}$ of length n is equal to

$$[x^n] \frac{\Phi_{\{\varepsilon,1,11\}}(x)}{1 - \Phi_{\{0,10,110\}}(x)} = [x^n] \frac{1 + x + x^2}{1 - x - x^2 - x^3}, \qquad n \geq 0,$$

from Theorem 10.1.

The decomposition of $\mathcal{S}$ above is a special case of the 0-*decomposition* for the set of all $\{0,1\}$-strings, given by

$$\{0,1\}^* = (\{1\}^*\{0\})^* \{1\}^* = \{1\}^* (\{0\}\{1\}^*)^*,$$

where the strings in $\{0,1\}^*$ are uniquely created in this decomposition. To prove this, simply note that every string in $\{0,1\}^*$ has k 0's for some unique nonnegative integer k, and that these separate $k+1$ possibly empty strings consisting entirely of 1's, ordered from left to right.

Of course, by interchanging 0's and 1's, we obtain the 1-*decomposition* for the set of all $\{0,1\}$-strings, given by

$$\{0,1\}^* = (\{0\}^*\{1\})^* \{0\}^* = \{0\}^* (\{1\}\{0\}^*)^*,$$

where the strings in $\{0,1\}^*$ are uniquely created in this decomposition.

As a third example, let $\mathcal{T}$ be the set of $\{0,1\}$-strings in which there is no substring "000" or "111", and find the number of elements of $\mathcal{T}$ of length n, $n \geq 0$.

In this case, we have

$$\mathcal{T} = \{\varepsilon, 1, 11\} (\{0, 00\}\{1, 11\})^* \{\varepsilon, 0, 00\},$$

and the elements of $\mathcal{T}$ are uniquely created in this decomposition, so we conclude that the number of strings in $\mathcal{T}$ of length n is equal to

$$[x^n] \frac{(1 + x + x^2)^2}{1 - (x + x^2)} = [x^n] \frac{1 + x + x^2}{1 - x - x^2}, \qquad n \geq 0,$$

from Theorem 10.1.

Some useful terminology for dealing with strings is *block*. A block in a $\{0,1\}$-string is a maximal nonempty substring consisting entirely of 0's or of 1's. For example, the blocks of 0011101100 are 00, 111, 0, 11, 00.

11 Lecture of February 4

The second decomposition for $\mathcal{S}$ above is a special case of the *block decomposition*, which for all strings in $\{0,1\}^*$ gives

$$\{0,1\}^* = \{1\}^* (\{0\}\{0\}^*\{1\}\{1\}^*)^* \{0\}^* = \{1\}^* ((\{0\}^* \setminus \{\varepsilon\}) (\{1\}^* \setminus \{\varepsilon\}))^* \{0\}^*.$$

Again, the elements of $\{0,1\}^*$ are uniquely created, and the proof of this is straightforward. Of course, we can interchange the 0's and 1's in such decompositions, as required.

Now we introduce additional complexity, by using more than one weight function, and generating series in more than one variable. Suppose that we have m weight functions $\omega_1, \ldots, \omega_m$, defined on a set $\mathcal{A}$ (we'll assume that m is finite for now, but m need not be finite, as we shall see later). Then we define the generating series in m variables $x_1, \ldots, x_m$, for $\mathcal{A}$ with respect to these weight functions by

$$\Phi_{\mathcal{A}}(x_1, \ldots, x_m) = \sum_{a \in \mathcal{A}} x_1^{\omega_1(a)} \ldots x_m^{\omega_m(a)}.$$

The Sum Rule for generating series in m variables says that for any weight functions defined on $\mathcal{A} \dot{\cup} \mathcal{B}$, we have

$$\Phi_{\mathcal{A} \dot{\cup} \mathcal{B}}(x_1, \ldots, x_m) = \Phi_{\mathcal{A}}(x_1, \ldots, x_m) + \Phi_{\mathcal{B}}(x_1, \ldots, x_m),$$

and the proof is exactly the same as for the case $m = 1$, since it simply uses the fact that

$$\sum_{b \in \mathcal{A} \dot{\cup} \mathcal{B}} = \sum_{b \in \mathcal{A}} + \sum_{b \in \mathcal{B}},$$

independently of the summand. For the product rule for pairwise Cartesian Products, we suppose that there are m weight functions defined for each of $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{A} \times \mathcal{B}$; for $\mathcal{A}$ let these weight functions be ω_{1i}, $i = 1, \ldots, m$, for $\mathcal{B}$ they are ω_{2i}, $i = 1, \ldots, m$, and for $\mathcal{A} \times \mathcal{B}$ they are ω_{3i}, $i = 1, \ldots, m$. If the condition

$$\omega_{3i}((a,b)) = \omega_{1i}(a) + \omega_{2i}(b)$$

holds for all $i = 1, \ldots, m$ and all $(a,b) \in \mathcal{A} \times \mathcal{B}$, then

$$\Phi_{\mathcal{A} \times \mathcal{B}}(x_1, \ldots, x_m) = \Phi_{\mathcal{A}}(x_1, \ldots, x_m) \Phi_{\mathcal{B}}(x_1, \ldots, x_m).$$

The proof of this follows the proof for the case $m = 1$, as given below:

$$\begin{aligned}
\Phi_{\mathcal{A} \times \mathcal{B}}(x_1, \ldots, x_m) &= \sum_{(a,b) \in \mathcal{A} \times \mathcal{B}} \prod_{i=1}^{m} x_i^{\omega_{3i}((a,b))} \\
&= \sum_{a \in \mathcal{A}} \sum_{b \in \mathcal{B}} \prod_{i=1}^{m} x_i^{\omega_{1i}(a) + \omega_{2i}(b)} \\
&= \sum_{a \in \mathcal{A}} \left(\prod_{i=1}^{m} x_i^{\omega_{1i}(a)} \right) \sum_{b \in \mathcal{B}} \left(\prod_{i=1}^{m} x_i^{\omega_{2i}(b)} \right) \\
&= \Phi_{\mathcal{A}}(x_1, \ldots, x_m) \Phi_{\mathcal{B}}(x_1, \ldots, x_m),
\end{aligned}$$

as required. The extension of the product rule to k-tuples for an arbitrary $k \geq 2$ is straightforward.

Now, as an example, let $c_{n,k}$ be the number of $\{0,1\}$-strings of length n, with k occurrences of "00" as a block, and let

$$C(x,y) = \sum_{n\geq 0} \sum_{k\geq 0} c_{n,k} y^k x^n.$$

Define the weight function $\omega(a)$ to equal the number of occurrences of "00" as a block in the string a. Then it is straightforward that $C(x,y) = \Phi_{\{0,1\}^*}(x,y)$, where

$$\Phi_{\{0,1\}^*}(x,y) = \sum_{a\in\{0,1\}^*} x^{\text{length}(a)} y^{\omega(a)}.$$

To determine this generating series, we consider the 1-decomposition

$$\{0,1\}^* = (\{0\}^*\{1\})^* \{0\}^*.$$

Now it is straightforward to determine this generating series from the 1-decomposition, using the sum and product rules for generating series in more than one variable. We have

$$\{0,1\}^* = (\{\varepsilon, 0, 00, 000, \ldots\}\{1\})^* \{\varepsilon, 0, 00, 000, \ldots\},$$

and thus obtain

$$\begin{aligned} C(x,y) &= \frac{1 + x + yx^2 + x^3 + \ldots}{1 - x(1 + x + yx^2 + x^3 + \ldots)} \\ &= \frac{\frac{1}{1-x} + (y-1)x^2}{1 - x\left(\frac{1}{1-x} + (y-1)x^2\right)} \\ &= \frac{1 + (y-1)x^2(1-x)}{1 - 2x - (y-1)x^3(1-x)}. \end{aligned}$$

In general, we do not obtain nice closed formulas for coefficients with many parameters like $c_{n,k}$, in a many variable generating series like $C(x,y)$. However, there are other ways in which additional parameters might enter a counting question, like averaging. For example, define μ_n, $n \geq 0$, to be the average number number of occurrences of "00" as a block among all $\{0,1\}$-strings of length n. For this "average", we consider all $\{0,1\}$-strings to be equiprobable. Then we immediately have

$$\mu_n = \frac{N_n}{D_n},$$

where

$$N_n = \sum_{k\geq 0} k\, c_{n,k}, \quad D_n = \sum_{k\geq 0} c_{n,k}.$$

For example, when $n = 5$, we have $c_{5,i} = 0$, for $i \geq 3$; $c_{5,2} = 1$, for the string 00100; $c_{5,1} = 10$, for the strings $00111, 00101, 00110, 10010, 10011$ and their reverse; and finally $c_{5,0} = 32 - 10 - 1 = 21$, since there are $2^5 = 32$ strings of length 5. Then we have

$$\mu_5 = \frac{0 \cdot 21 + 1 \cdot 10 + 2 \cdot 1}{32} = \tfrac{12}{32} = \tfrac{3}{8}$$

in this case.

12 Lecture of February 8

How do we determine μ_n for arbitrary n ? We can evaluate N_n and D_n from the generating series $C(x,y)$, by observing that

$$C(x,1) = \sum_{n\geq 0}\sum_{k\geq 0} c_{n,k}x^n, \qquad \left(\frac{\partial}{\partial y}C(x,y)\right)\Big|_{y=1} = \sum_{n\geq 0}\sum_{k\geq 0} k\, c_{n,k}x^n,$$

and hence we obtain

$$D_n = [x^n]C(x,1), \qquad N_n = [x^n]\left(\frac{\partial}{\partial y}C(x,y)\right)\Big|_{y=1}.$$

Applying this to the example given, we have

$$C(x,1) = \frac{1}{1-2x} = \sum_{i\geq 0} 2^i x^i,$$

giving

$$D_n = [x^n]C(x,1) = 2^n, \quad n \geq 0.$$

For N_n, we use the quotient rule for differentiating in y, which says

$$\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2},$$

to obtain

$$\begin{aligned}\left(\frac{\partial}{\partial y}C(x,y)\right)\Big|_{y=1} &= \frac{x^2(1-x)(1-2x)+x^3(1-x)}{(1-2x)^2}\\ &= \frac{x^2(1-x)^2}{(1-2x)^2} = \frac{x^2}{1-2x} + \frac{x^4}{(1-2x)^2}\\ &= \sum_{i\geq 0} 2^i x^{i+2} + \sum_{j\geq 0}(j+1)2^j x^{j+4},\end{aligned}$$

where for the last summation over j, we have used the negative binomial theorem

$$(1-y)^{-2} = \sum_{j\geq 0}\binom{2+j-1}{j}y^j = \sum_{j\geq 0}\binom{j+1}{j}y^j = \sum_{j\geq 0}(j+1)y^j.$$

Thus we have $N_n = 0$, for $n = 0, 1$, $N_n = 2^{n-2}$, for $n = 2, 3$, and $N_n = 2^{n-2} + (n-3)2^{n-4} = (n+1)2^{n-4}$, for $n \geq 4$. Finally, dividing by D_n, we have

$$\mu_n = 0, \;\; n = 0, 1, \qquad \mu_n = \tfrac{1}{4}, \;\; n = 2, 3, \qquad \mu_n = \frac{n+1}{16}, \;\; n \geq 4.$$

Note that when $n = 5$, this formula gives $\mu_5 = \frac{6}{16} = \frac{3}{8}$, in agreement with the data above.

Now we shall consider substrings that are to be avoided.

Example 12.1 *Determine the number of* $\{0,1\}$*-strings of length* n*, with no occurrences of "*0001111*" as a substring.*

SOLUTION. Let $\mathcal{S}$ be the set of $\{0,1\}$-strings with no occurrences of "0001111" as a substring. Then from the block decomposition we obtain

$$\mathcal{S} = \{1\}^* \left(\{0\}\{0\}^*\{1\}\{1\}^* \setminus \{0\}^*\{0001111\}\{1\}^*\right)^* \{0\}^*,$$

so the generating series for $\mathcal{S}$ with respect to length is

$$\begin{aligned} \Phi_{\mathcal{S}}(x) &= \frac{1}{1-x} \frac{1}{1-\left(\frac{x}{1-x}\frac{x}{1-x} - \frac{x^7}{(1-x)^2}\right)} \frac{1}{1-x} \\ &= \frac{1}{1-2x+x^7}, \end{aligned}$$

and the number of strings in $\mathcal{S}$ of length n is equal to the coefficient of x^n in $\Phi_{\mathcal{S}}(x)$, $n \geq 0$.

So far, as in the above example, we have used generating series in the following, direct way: we wish to enumerate a set $\mathcal{S}$, and decompose $\mathcal{S}$ into other sets, using the set operations of disjoint union and Cartesian product, where the generating series for these other sets are known. However, it is often advantageous to proceed more indirectly, as in the following example.

Example 12.2 *Determine the number of* $\{0,1\}$*-strings of length* n*, with no occurrences of "*0101111*" as a substring.*

SOLUTION. Let $\mathcal{A}$ be the set of $\{0,1\}$-strings with no occurrences of "0101111" as a substring. Then from the 1-decomposition, but with the string 0101111 playing the role of "1", we obtain the decomposition

$$\{0,1\}^* = \mathcal{A}\left(\{0101111\}\mathcal{A}\right)^*, \tag{11}$$

and claim that the elements of $\{0,1\}^*$ are uniquely created in this decomposition.

To justify our claim for (11) above, let $\alpha = 0101111$. Note there are no nonempty strings β, γ with $\text{length}(\beta) = \text{length}(\gamma)$ and $\text{length}(\beta) < \text{length}(\alpha)$, for which $\alpha\beta = \gamma\alpha$ (check this for the 6 choices of $\beta = 0, 01, 010, 0101, 01011, 010111$). This means that the occurrences of α in any $\{0,1\}$-string s cannot overlap (i.e., no symbol in the string can be contained in more than one occurrence of α). Thus we can write

$$s = s_0\alpha s_1\alpha \ldots \alpha s_k,$$

for some unique $k \geq 0$, where $s_0, s_1, \ldots, s_k$ are unique strings in $\mathcal{A}$.

It is straightforward to obtain the generating series for $\mathcal{A}$ from (11), by considering the generating series for both sides with respect to length. This gives

$$\frac{1}{1-2x} = \frac{\Phi_{\mathcal{A}}(x)}{1-x^7\Phi_{\mathcal{A}}(x)},$$

and, crossmultiplying and solving for $\Phi_{\mathcal{A}}(x)$, we obtain

$$\begin{aligned} 1 - x^7\Phi_{\mathcal{A}}(x) &= (1-2x)\Phi_{\mathcal{A}}(x) \\ (1 - 2x + x^7)\Phi_{\mathcal{A}}(x) &= 1 \\ \Phi_{\mathcal{A}}(x) &= \frac{1}{1-2x+x^7}. \end{aligned}$$

Note that this indirect method would also give the result of Example 12.1 immediately.

13 Lecture of February 11

Now we consider the generating series

$$A(x,y) = \sum_{s\in\{0,1\}^*} x^{\text{length}(s)} y^{\omega(s)},$$

where $\omega(s)$ is equal to the number of of occurrences of "0101111" as a substring in s. An expression for $A(x,y)$ now follows from (11), by considering the generating series for both sides with respect to length, and with respect to weight function ω, giving

$$A(x,y) = \frac{\Phi_{\mathcal{A}}(x)}{1 - yx^7\Phi_{\mathcal{A}}(x)} = \frac{\frac{1}{1-2x+x^7}}{1 - \frac{yx^7}{1-2x+x^7}} = \frac{1}{1-2x-(y-1)x^7}.$$

Now consider strings on the alphabet $\{1,\ldots,n\}$, for fixed positive integer n. For a string b in $\mathcal{S} = \{1,\ldots,n\}^*$, define the weight function $\omega_i(b)$ to be the number of i's that occur in b, for $i = 1,\ldots,n$. Then, for a set of strings $\mathcal{B} \subseteq \mathcal{S}$, define the generating series

$$\Phi_{\mathcal{B}}(x_1,\ldots,x_n) = \sum_{b\in\mathcal{B}} x_1^{\omega_1(b)}\ldots x_n^{\omega_n(b)}.$$

Now, of course

$$\Phi_{\mathcal{S}}(x_1,\ldots,x_n) = \sum_{k_1,\ldots,k_n\geq 0} c_{k_1,\ldots,k_n} x_1^{k_1}\ldots x_n^{k_n},$$

where $c_{k_1,\ldots,k_n}$ is the number of strings with k_1 1's, ... ,k_n n's.

Now, determining $\Phi_{\mathcal{S}}(x_1,\ldots,x_n)$ using the sum and product rules, we obtain

$$\Phi_{\mathcal{S}}(x_1,\ldots,x_n) = \frac{1}{1-\Phi_{\{1,\ldots,n\}}(x_1,\ldots,x_n)} = \frac{1}{1-(x_1+\ldots+x_n)}.$$

Expanding this series, we use the geometric series to get

$$\Phi_{\mathcal{S}}(x_1,\ldots,x_n) = \sum_{m\geq 0}(x_1+\ldots+x_n)^m = \sum_{m\geq 0}\sum_{\substack{k_1,\ldots,k_n\geq 0 \\ k_1+\ldots+k_n=m}} \frac{m!}{k_1!\ldots k_n!} x_1^{k_1}\ldots x_n^{k_n},$$

where the last equality follows from the *multinomial theorem*. (It is easily proved by applying the binomial theorem to expand $(x_1 + (x_2 + \ldots + x_n))^m$, then repeating.) This implies that

the number of strings in $\{1,\ldots,n\}$ with k_i i's, for $i = 1,\ldots,n$, where $k_1,\ldots,k_n \geq 0$ and $k_1 + \ldots + k_n = m$, is given by

$$\frac{m!}{k_1!\ldots k_n!}, \tag{12}$$

which is usually referred to as the *multinomial coefficient*. Of course, it is easy to check that (12) is the correct cardinality for this set of strings, as follows: there are m positions in the string (say they're called $1,\ldots,m$). There are $\binom{m}{k_1}$ ways to choose positions for the 1's. Then, for $i = 2,\ldots,n$, suppose that positions have been chosen for the 1's, ... , $i-1$'s; independently of which set of $k_1 + \cdots + k_{i-1}$ positions has been chosen for these, there are $\binom{m-k_1-\cdots-k_{i-1}}{k_i}$ ways to choose positions for the i's. Thus the total number of ways of choosing positions for all symbols is (with convention that $k_0 = 0$)

$$\prod_{i=1}^{n} \binom{m-k_1-\cdots-k_{i-1}}{k_i} = \prod_{i=1}^{n} \frac{(m-k_1-\cdots-k_{i-1})!}{k_i!(m-k_1-\cdots-k_i)!} = \frac{m!}{k_1!\ldots k_n!},$$

in agreement with (12).

Now let $\mathcal{D}$ be the set of strings (called "Smirnov" strings) on the alphabet $\{1,\ldots,n\}$ in which adjacent elements are always distinct (for example, 1342532413253435251 is in $\mathcal{D}$, but 134255324132 is not, because of the substring "55"). Consider an arbitrary string $s \in \mathcal{S}$. Suppose that, for all $i = 1,\ldots,n$, we replace each block of i's in s by a single i. Then clearly we obtain a unique string $d \in \mathcal{D}$ by this operation. Moreover, if we reverse this construction, then we uniquely create the strings in $\mathcal{S}$ if we replace each element i, $i = 1,\ldots,n$, by a block of i's of any positive length, in all possible ways.

The block replacement construction above gives

$$\Phi_{\mathcal{S}}(x_1,\ldots,x_n) = \Phi_{\mathcal{D}}(x_1 + x_1^2 + \ldots,\ldots,x_n + x_n^2 + \ldots) = \Phi_{\mathcal{D}}(\frac{x_1}{1-x_1},\ldots,\frac{x_n}{1-x_n}).$$

Now let $y_i = x_i/(1-x_i)$, for $i = 1,\ldots,n$. Then crossmultiplying by $1-x_i$ and solving for x_i, we obtain $x_i = y_i/(1+y_i)$, for $i = 1,\ldots,n$. This then gives

$$\Phi_{\mathcal{D}}(y_1,\ldots,y_n) = \Phi_{\mathcal{S}}(\frac{y_1}{1+y_1},\ldots,\frac{y_n}{1+y_n}) = \frac{1}{1-\frac{y_1}{1+y_1}-\ldots-\frac{y_n}{1+y_n}},$$

which is the required generating series. For a consistency check on this series, note that if $y_1 = \ldots = y_n = z$, then we obtain

$$\begin{aligned}\Phi_{\mathcal{D}}(z,\ldots,z) &= \frac{1}{1-\frac{nz}{1+z}} = \frac{1+z}{1+z-nz} \\ &= 1 + \frac{nz}{1-(n-1)z} = 1 + \sum_{k\geq 1} n(n-1)^{k-1}z^k,\end{aligned}$$

which simply states the obvious fact that the number of strings in $\mathcal{D}$ of length k is $n(n-1)^{k-1}$, for $k \geq 1$.

Another feature of the above series $\Phi_{\mathcal{D}}(y_1, \ldots, y_n)$ is that it works *non-commutatively*. For example, expanding, we have

$$\begin{aligned} \frac{1}{1 - \frac{y_1}{1+y_1} - \ldots - \frac{y_n}{1+y_n}} &= 1 + \sum_{k \geq 1} \left(\frac{y_1}{1+y_1} + \ldots + \frac{y_n}{1+y_n} \right)^k \\ &= 1 + \sum_{k \geq 1} \left(y_1 - y_1^2 + y_1^3 - \ldots + y_n - y_n^2 + y_n^3 - \ldots \right)^k. \end{aligned}$$

Now the terms of total degree 1 in the y_i's are

$$y_1 + \ldots + y_n,$$

which checks with the fact that the strings of length 1 in $\mathcal{D}$ are $1, \ldots, n$. The terms of total degree 2 in the y_i's are

$$(y_1 + \ldots + y_n)(y_1 + \ldots + y_n) - y_1^2 - \ldots - y_n^2 = \sum_{\substack{1 \leq i,j \leq n \\ i \neq j}} y_i y_j,$$

which correctly produces the strings of length 2 in $\mathcal{D}$.

For more direct decompositions for strings on the alphabet $\{1, \ldots, n\}$, for n an arbitrary positive integer, note that the 1-decomposition extends easily, to

$$\{1, \ldots, n\}^* = \{2, \ldots, n\}^* \left(\{1\} \{2, \ldots, n\}^* \right)^*,$$

and the block decomposition also extends easily, to

$$\{1, \ldots, n\}^* = \{2, \ldots, n\}^* \left(\{1\} \{1\}^* \left(\{2, \ldots, n\}^* \setminus \{\varepsilon\} \right) \right)^* \{1\}^*.$$

14 Lecture of February 13

There are many examples of non-commutative results in combinatorics. For example, suppose that in products involving x and y we use the rule $yx = qxy$, where q commutes with both x and y. Then the binomial theorem becomes

$$(x + y)^n = \sum_{k=0}^{n} \binom{n}{k}_q x^k y^{n-k}, \tag{13}$$

where

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

The polynomial $\binom{n}{k}_q$ is often called the *q-binomial coefficient*, or *Gaussian coefficient*. Another interpretation of the expansion (13) is that

$$\binom{n}{k}_q = \sum_{\pi \in \mathcal{P}_{k,n-k}} q^{\text{area}(\pi)},$$

where $\mathcal{P}_{k,n-k}$ is the set of lattice paths from $(0,0)$ to $(k,n-k)$, and area(π) is the area under the path π. From a completely different point of view, in which q is a power of a prime, $\binom{n}{k}_q$ is equal to the number of k-dimensional subspaces of an n-dimensional vector space over GF(q) (the finite field with q elements).

The series that we have been using are *formal power series*, not the power series of real variables that have been studied in calculus courses. A formal power series is given by $A(x) = \sum_{i\geq 0} a_i x^i$, where $a_i = [x^i]A(x)$, the *coefficient of* x^i, is a complex number, for $i \geq 0$. The basic rule for $A(x)$ is that a_i is determined finitely for each finite i. Let $B(x) = \sum_{i\geq 0} b_i x^i$. Then $A(x) = B(x)$ if and only if $a_i = b_i$ for all $i \geq 0$, and we define sum and product by

$$A(x) + B(x) = \sum_{i\geq 0}(a_i + b_i)x^i, \qquad A(x)B(x) = \sum_{i\geq 0}\left(\sum_{j=0}^{i} a_j b_{i-j}\right)x^i,$$

and a special case of product is the scalar product $c\,A(x) = \sum_{i\geq 0}(c\,a_i)x^i$, for a complex number c. We write $A(0) = a_0$, and unless A is a polynomial, this is the only "evaluation" we allow. If $b_0 = 0$, then we define the composition

$$A(B(x)) = \sum_{i\geq 0} a_i B(x)^i = \sum_{n\geq 0} \sum_{\substack{i\geq 0, j_1,\ldots,j_i\geq 1 \\ j_1+\ldots+j_i=n}} a_i b_{j_1}\ldots b_{j_i} x^n,$$

and note that the summations above are finite.

Now suppose $A(0) = 1$. Then if $B(x)$ is a *multiplicative inverse* of $A(x)$, we have (since multiplication of complex numbers is commutative, so is multiplication of $A(x)$ with $B(x)$, so there is no difference between a left-inverse and a right-inverse) $\sum_{i\geq 0} a_i x^i \sum_{j\geq 0} b_j x^j = 1$, and equating coefficients of x^n on both sides, for $n \geq 0$, we obtain

$$\begin{aligned} b_0 &= 1 \\ a_1 b_0 + b_1 &= 0 \\ a_2 b_0 + a_1 b_1 + b_2 &= 0, \end{aligned}$$

where the nth equation is $a_n b_0 + a_{n-1} b_1 + \ldots + b_0 = 0$, $n \geq 1$. But this gives $b_0 = 1$, and allows us to determine b_n uniquely in terms of $b_0, \ldots, b_{n-1}$, for each $n \geq 1$, so, by induction on n, $B(x)$ is unique. Applying this process to obtain the multiplicative inverse of $A(x) = 1 - x$, we obtain $b_n = 1$, $n \geq 0$, by induction on n, or $(1-x)^{-1} = \sum_{i\geq 0} x^i$. But substitution into this, for an arbitrary $A(x)$ with $A(0) = 1$, gives

$$A(x)^{-1} = (1 - (1 - A(x)))^{-1} = 1 + \sum_{i\geq 1} (1 - A(x))^i,$$

which is therefore the *unique* multiplicative inverse of $A(x)$.

We define *differentiation* and *integration* operators by

$$\frac{d}{dx}A(x) = \sum_{i\geq 1} i a_i x^{i-1}, \qquad I_x\, A(x) = \sum_{i\geq 0} \frac{a_i}{i+1} x^{i+1}.$$

Now note that we have uniqueness for solution of differential equations: if $\frac{d}{dx}A(x) = \frac{d}{dx}B(x)$ and $A(0) = B(0)$, then $A(x) = B(x)$.

15 Lecture of February 15

Now

$$\frac{d}{dx}(A(x)+B(x)) = \sum_{i\geq 1} i(a_i+b_i)x^{i-1} = \sum_{i\geq 1} ia_i x^{i-1} + \sum_{i\geq 1} ib_i x^{i-1} = \frac{d}{dx}A(x) + \frac{d}{dx}B(x),$$

so this differentiation operator satisfies the sum rule, and

$$\begin{aligned}\frac{d}{dx}(A(x)B(x)) &= \sum_{i\geq 1}\sum_{j=0}^{i} ia_j b_{i-j}x^{i-1} \\ &= \sum_{i\geq 1}\sum_{j=0}^{i}(j+i-j)a_j b_{i-j}x^{i-1} \\ &= \left(\frac{d}{dx}A(x)\right)B(x) + A(x)\left(\frac{d}{dx}B(x)\right),\end{aligned}$$

and differentiation satisfies the product rule. Induction on n then gives $\frac{d}{dx}B(x)^n = nB(x)^{n-1}\frac{d}{dx}B(x)$ for positive integers n, which allows us to prove the *chain rule*:

$$\frac{d}{dx}A(B(x)) = A'(B(x))\frac{d}{dx}B(x).$$

We now define three special series

$$\varepsilon(x) = \sum_{n\geq 0}\frac{1}{n!}x^n, \quad \lambda(x) = \sum_{n\geq 1}\frac{1}{n}x^n, \quad B_a(x) = \sum_{n\geq 0}\frac{a(a-1)\ldots(a-n+1)}{n!}x^n,$$

where a is a complex number parameter in $B_a(x)$. Our object is to show that $\varepsilon(x), \lambda(x), B_a(x)$ have the properties of the familiar functions $e^x, \ln(1-x)^{-1}, (1+x)^a$, respectively. (Except that we will NOT be able to consider, for example, $\varepsilon(\varepsilon(x))$, since it uses composition with a series with constant term 1.) First, note that $\frac{d}{dx}\varepsilon(x) = \varepsilon(x)$. Then, for example, we can prove that $\varepsilon(x)\varepsilon(-x) = 1$, since $\varepsilon(x)\varepsilon(-x)$ has constant term $\varepsilon(0)\varepsilon(-0) = 1$, and

$$\frac{d}{dx}(\varepsilon(x)\varepsilon(-x)) = \varepsilon(x)\varepsilon(-x) - \varepsilon(x)\varepsilon(-x) = 0,$$

where we have used the product rule and chain rule. The result follows by the uniqueness of solution of differential equations (since 1 also has constant term 1 and derivative 0). Also, we have $\frac{d}{dx}\lambda(x) = \sum_{n\geq 0} x^n = (1-x)^{-1}$, so

$$\frac{d}{dx}\left(\lambda(1-\varepsilon(-x))\right) = \left(\varepsilon(-x)\right)^{-1}\varepsilon(-x) = 1,$$

by the chain rule, and $\lambda(1-\varepsilon(-0)) = 0$, and we conclude that $\lambda(1-\varepsilon(-x)) = x$, by uniqueness of solution of differential equations. (The series $1-\varepsilon(-x)$ has constant term 0,

so the composition $\lambda(1-\varepsilon(-x))$ is valid.) Similarly, we prove that $\varepsilon(\lambda(x)) = (1-x)^{-1}$, using $\varepsilon(\lambda(0)) = 1$, and

$$\frac{d}{dx}\left((1-x)\varepsilon(\lambda(x))\right) = -\varepsilon(\lambda(x)) + (1-x)\varepsilon(\lambda(x))(1-x)^{-1} = 0,$$

using the product rule and chain rule. For the series $B_a(x)$, we have $\frac{d}{dx}B_a(x) = a\,B_{a-1}(x)$, and we omit further details of these computations.

Now, it is easy to verify that there are no zero divisors for formal power series, and this fact allows us to establish that nth roots are unique, at least with given constant term, as follows. Suppose $A(0) = B(0) = 1$, and $A(x)^n = B(x)^n$, for some positive integer n. Then we have

$$0 = A(x)^n - B(x)^n = (A(x) - B(x))(A(x)^{n-1} + A(x)^{n-2}B(x) + \ldots + B(x)^{n-1}).$$

Now the constant term in the second factor is $A(0)^{n-1} + A(0)^{n-2}B(0) + \ldots + B(0)^{n-1} = n \neq 0$, so we conclude that $A(x) - B(x) = 0$, since there are no zero divisors, which gives $A(x) = B(x)$, as required. But we can determine the nth root of $A(x)$ with $A(0) = 1$ by substitution in the binomial series $B_a(x)$, to obtain

$$A(x)^{\frac{1}{n}} = (1 + (A(x)-1))^{\frac{1}{n}} = 1 + \sum_{i\geq 1} \frac{\frac{1}{n}(\frac{1}{n}-1)\ldots(\frac{1}{n}-i+1)}{i!}\,(A(x)-1)^i\,,$$

which is therefore the *unique* nth root with constant term 1.

We introduce trigonometric series by defining

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots, \qquad \cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots,$$

and then proving the properties of these series from properties of the series $\varepsilon(x) = e^x$, by

$$\sin(x) = \frac{e^{ix} - e^{-ix}}{2i}, \qquad \cos(x) = \frac{e^{ix} + e^{-ix}}{2},$$

so, for example, we have

$$\begin{aligned}\sin(x)^2 + \cos(x)^2 &= \left(\frac{e^{ix} - e^{-ix}}{2i}\right)^2 + \left(\frac{e^{ix} + e^{-ix}}{2}\right)^2 \\ &= \tfrac{1}{4}\left(-(e^{2ix} - 2 + e^{-2ix}) + (e^{2ix} + 2 + e^{-2ix})\right) = 1.\end{aligned}$$

Then, noting that $\cos(x)$ has constant term 1, so it is invertible, we define

$$\tan(x) = \frac{\sin(x)}{\cos(x)} = x + \frac{x^3}{3} + \frac{2x^5}{15} + \ldots = x + \frac{2x^3}{3!} + \frac{16x^5}{5!} + \ldots,$$

$$\sec(x) = \frac{1}{\cos(x)} = 1 + \frac{x^2}{2} + \frac{5x^4}{24} + \ldots = 1 + \frac{x^2}{2!} + \frac{5x^4}{4!} + \ldots.$$

Finally, we are going to show that the sequences $1, 2, 16, \ldots$ and $1, 1, 5, \ldots$, which are the coefficients in $\tan(x)$ and $\sec(x)$, scaled by the factorial, count some interesting combinatorial objects.

Let a_{2k+1}, for $k \geq 0$, denote the number of permutations $\sigma_1 \ldots \sigma_{2k+1}$ of $\{1, \ldots, 2k+1\}$, for which $\sigma_1 < \sigma_2 > \sigma_3 < \ldots < \sigma_{2k} > \sigma_{2k+1}$. Similarly, let b_{2k}, for $k \geq 0$, denote the number of permutations $\sigma_1 \ldots \sigma_{2k}$ of $\{1, \ldots, 2k\}$, for which $\sigma_1 < \sigma_2 > \sigma_3 < \ldots > \sigma_{2k-1} < \sigma_{2k}$. These permutations are called *alternating* permutations. For example, $a_1 = 1$ (the permutation here is 1), $a_3 = 2$ (the permutations are 132, 231), $b_0 = 1$ (the empty permutation ε), $b_2 = 1$ (the permutation 12), and $b_4 = 5$ (the permutations 1423, 2413, 3412, 1324, 2314). To obtain a recurrence equation for these numbers, first consider an alternating permutation on $\{1, \ldots, 2k+1\}$, where $k \geq 1$. Then we have $\sigma_{2i+2} = 2k+1$ for some unique $i = 0, \ldots, k-1$. In this case, $\sigma_1 \ldots \sigma_{2i+1}$ is an alternating permutation on some $(2i+1)$-subset α of $\{1, \ldots, 2k\}$, and $\sigma_{2i+3} \ldots \sigma_{2k+1}$ is an alternating permutation on $\{1, \ldots, 2k\} \setminus \alpha$. Then there are $\binom{2k}{2i+1}$ choices for α, and for each such α, a_{2i+1} choices for $\sigma_1 \ldots \sigma_{2i+1}$, and $a_{2k-2i-1}$ choices for $\sigma_{2i+3} \ldots \sigma_{2k+1}$. Moreover, this is bijective, and we conclude that

$$a_{2k+1} = \sum_{i=0}^{k-1} \binom{2k}{2i+1} a_{2i+1} a_{2k-2i-1}, \quad k \geq 1, \tag{14}$$

with initial condition $a_1 = 1$. Similarly, we obtain

$$b_{2k} = \sum_{i=0}^{k-1} \binom{2k}{2i+1} a_{2i+1} b_{2k-2i-2} \quad k \geq 1, \tag{15}$$

with initial condition $b_0 = 1$. Now let

$$A(x) = \sum_{k \geq 0} a_{2k+1} \frac{x^{2k+1}}{(2k+1)!}, \qquad B(x) = \sum_{k \geq 0} b_{2k} \frac{x^{2k}}{(2k)!},$$

which are called the *exponential generating series* for the sequences $\{a_{2k+1}\}_{k \geq 0}$, $\{b_{2k}\}_{k \geq 0}$, respectively. Now, multiply both sides of (14) by $\frac{x^{2k}}{(2k)!}$, and sum for $k \geq 1$, to obtain

$$\sum_{k \geq 1} a_{2k+1} \frac{x^{2k}}{(2k)!} = \sum_{k \geq 1} \sum_{i=0}^{k-1} \frac{a_{2i+1}}{(2i+1)!} \frac{a_{2k-2i-1}}{(2k-2i-1)!} x^{2k}.$$

Then change indices in the double summation from i, k to i, j, where $j = k - 1 - i$. This gives ranges of summation $i \geq 0$ and $j \geq 0$, and we have $2k = 2(i+j+1) = 2i+1+2j+1$, so the above equation becomes

$$\sum_{k \geq 1} a_{2k+1} \frac{x^{2k}}{(2k)!} = \sum_{i \geq 0} \frac{a_{2i+1}}{(2i+1)!} x^{2i+1} \sum_{j \geq 0} \frac{a_{2j+1}}{(2j+1)!} x^{2j+1}.$$

Translating this in terms of $A(x)$, we obtain

$$\frac{d}{dx} \left(A(x) - a_1 x \right) = A(x)^2,$$

so

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

(In CO 330, exponential generating series are considered in greater detail. There we give a product rule that allows us to write down the above differential equation immediately, avoiding all the details of summation indices, etc.) To solve this differential equation, divide both sides by $1 + A(x)^2$ (this has constant term 1, so it is invertible), and integrate with respect to x, to obtain

$$\arctan(A(x)) = x + c,$$

where we determine from the initial condition $A(0) = 0$ that $c = 0$, and conclude that $A(x) = \tan(x)$.

Similarly, from (15), multiplying by $\frac{x^{2k}}{(2k)!}$, and summing for $k \geq 1$, we obtain

$$\frac{d}{dx}B(x) = A(x)B(x).$$

To solve this differential equation, divide both sides by $B(x)$ (this has constant term 1, so it is invertible), and integrate with respect to x, to obtain

$$\ln(B(x)) = \ln(\sec(x)) + c,$$

where we determine from the initial condition $B(0) = 1$ that $c = 0$, and conclude that $B(x) = \sec(x)$.

There are many other types of generating series used for various types of applications. For this reason, the generating series $\sum_{n\geq 0} a_n x^n$ that we have used in MATH 249 is often referred to as the *ordinary* generating series for the sequence $\{a_n\}_{n\geq 0}$. For example, in number theoretic applications, the *Dirichlet* generating series $\sum_{n\geq 1} a_n\, n^{-s}$ is often used (here s is the variable). The significance of each choice of generating series is usually to be found in their product. For example, we have

$$\sum_{i\geq 0} a_i x^i \sum_{j\geq 0} b_j x^j = \sum_{n\geq 0} \left(\sum_{\substack{i,j\geq 0 \\ i+j=n}} a_i b_j \right) x^n,$$

$$\sum_{i\geq 0} a_i \frac{x^i}{i!} \sum_{j\geq 0} b_j \frac{x^j}{j!} = \sum_{n\geq 0} \left(\sum_{\substack{i,j\geq 0 \\ i+j=n}} \binom{n}{i} a_i b_j \right) \frac{x^n}{n!},$$

$$\sum_{i\geq 1} a_i\, i^{-s} \sum_{j\geq 1} b_j\, j^{-s} = \sum_{n\geq 1} \left(\sum_{\substack{i,j\geq 1 \\ i\cdot j=n}} a_i b_j \right) n^{-s}.$$

As an extra, the following example gives the general solution technique for the enumeration of strings excluding arbitrary substrings (that may overlap with themselves).

Example Find the number of $\{0,1\}$-strings of length n, with no occurrences of 01101 as a substring.

To solve this, we create a combinatorial set using "marking". Let $\mathcal{T}$ be the set of $\{0,1\}$-strings with some subset of occurrences of 01101 as a substring marked (in any example, this is shown by circling the marked substrings). In this case, the circled occurrences are not necessarily disjoint – the connected components of interlocked circled substrings are elements of the countable set whose first three elements are illustrated in Figure 1. We call this set the set of *clusters*. For $s \in \mathcal{T}$, let length(s) equal the length of the underlying string, and

Figure 1: The set of clusters for 01101.

let circ(s) equal the number of circled substrings. Then define

$$\Psi(x,u) = \sum_{s\in\mathcal{T}} x^{\text{length}(s)} u^{\text{circ}(s)}.$$

Clearly we have

$$\mathcal{T} = \{\{0,1\} \cup \mathcal{C}\}^\star,$$

where $\mathcal{C}$ is the set of clusters, and the circled occurrences of the substring 01101 are as illustrated in Figure 1. Thus we have $\mathcal{C} = \{01101\}\{101\}^\star$, and

$$\Psi(x,u) = \frac{1}{1-2x-C(x,u)},$$

where $C(x,u)$, the *cluster generating function*, is given by

$$C(x,u) = \frac{ux^5}{1-ux^3}.$$

But, we also have

$$\Psi(x,u) = \sum_{a\in\{0,1\}^*} x^{\text{length}(a)}(1+u)^{\omega(a)},$$

where $\omega(a)$ is the number of times that 01101 appears as a substring in a. Thus the required number is given by

$$[x^n]\Psi(x,-1) = [x^n]\frac{1}{1-2x-C(x,-1)} = [x^n]\frac{1}{1-2x+\frac{x^5}{1+x^3}}.$$

(For k occurrences of 01101 as a substring, we use $[x^n u^k]\Psi(x,u-1)$.)

16 Lecture of February 25

Now, we begin the study of graphs. A *graph* G consists of a finite, nonempty set, denoted $V(G)$, together with a set, denoted $E(G)$, of unordered pairs of distinct elements of $V(G)$. The elements of $V(G)$ are called *vertices* of G, and the elements of $E(G)$ are called *edges* of G.

For example, one particular graph is given by

$$V(G) = \{1, 2, 3, 4\}, \quad E(G) = \{\{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 4\}, \{3, 4\}\}.$$

We say that vertex 1 and edge $\{1, 2\}$ are *incident*; we say that vertex 1 and vertex 2 are *joined, adjacent* or *neighbours* (and that vertex 2 is not adjacent to vertex 3). The *degree* of a vertex v, denoted by $\deg(v)$, is the number of edges with which v is incident, so here we have $\deg(1)=\deg(4)=3$, and $\deg(2)=\deg(3)=2$. The graph G is drawn on the left of Figure 2. In

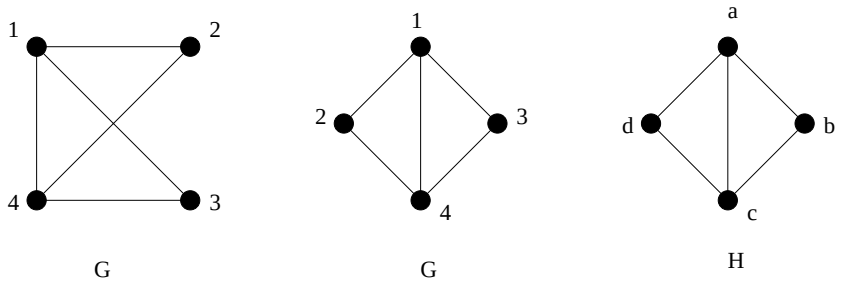

Figure 2: The graphs G and H.

this drawing we have represented the vertices as black circles, and drawn a curve between two circles if the corresponding vertices are adjacent (in general, these curves do not need to be straight lines). Another drawing of G is given in the middle of Figure 2. Note that, despite the similarity of this drawing of G, and the graph H displayed to the right of G, graphs G and H are NOT the same graphs, since they have different vertex-sets. One relationship between G and H that emphasizes the similarity of these drawings is to consider the function f from $V(G)$ to $V(H)$, specified by $f(1) = a, f(2) = d, f(3) = b, f(4) = c$. Then if vertex i is relabelled by $f(i)$, for all $i \in V(G)$, we obtain precisely graph H. In general, pairs of graphs like G and H are said to be isomorphic, and such a relabelling function f is called an isomorphism. These are defined precisely as follows:

We say that graph G is *isomorphic* to graph H is there is a bijection $f : V(G) \to V(H)$ that preserves adjacency (which means that for all unordered pairs u, v of vertices in G, u, v are adjacent in G if and only if $f(u), f(v)$ are adjacent in H. such a bijection f is called an *isomorphism* from G to H. Note that if f is an isomorphism from G to H, then f^{-1} is an isomorphism from H to G, so we usually use the language symmetrically, as in " graphs G and H are isomorphic".

Our first result about graphs is the "handshake theorem".

Theorem 16.1

$$\sum_{v \in V(G)} \deg(v) = 2\,|E(G)|$$

PROOF. Count the (vertex, edge) pairs that are incident. There are $\deg(v)$ such pairs for each vertex v, giving the lefthandside, and there are two such pairs for each edge, giving the righthandside.

Now, we define a class of graphs. For each $n \geq 0$, the n-cube Q_n is the graph whose vertices are the $\{0,1\}$-strings of length n, and two strings are adjacent if they differ in exactly one position. For example, The 0-cube Q_0 has a single vertex, the empty string ε, and no edges. The graphs Q_n, for $n = 1, 2, 3$, are given in Figure 3. Clearly there are $p = 2^n$ vertices in Q_n. How many edges q are there ? (We often use parameters p and q in this way, as generically denoting the numbers of vertices and edges in a graph.) Now, every vertex in Q_n has degree n, since there are n positions that can be changed (from 0 to 1, or from 1 to 0). Thus, we have

$$\sum_{v \in V(Q_n)} \deg(v) = \sum_{v \in V(Q_n)} n = n \cdot p = n2^n,$$

and from the handshake theorem this is equal to $2q$, so we conclude that $q = n2^{n-1}$.

In general, a graph in which all vertices have the same degree is called a *regular* graph. If in particular this degree is always k, then we also may say that the graph is k-regular (so Q_n is n-regular, $n \geq 0$).

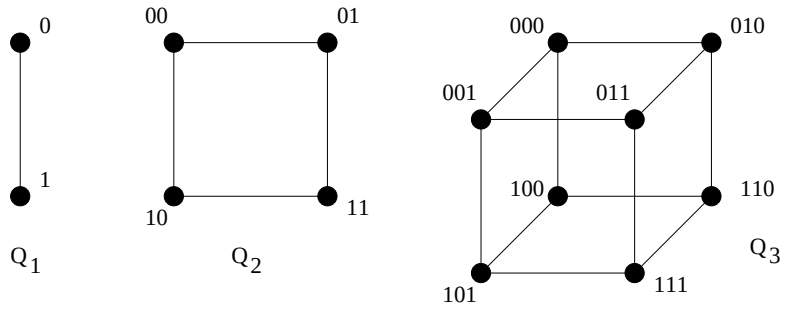

Figure 3: The cubes Q_n, for $n = 1, 2, 3$.

As another example of a parameterized family of graphs, fix nonnegative integers n, m, k. Define a graph to have vertices that are the m-subsets of a fixed n-set, and two m-subsets are adjacent if their intersection has cardinality k. For example, in the case $n = 5$, $m = 2$, $k = 0$, this graph is called the *Petersen* graph, and is given in Figure 4, as H.

Are any of the graphs G_1, G_2, G_3 also given in Figure 4 isomorphic to H ? Clearly G_1 is not isomorphic to H, since G_1 and H have different numbers of vertices. Also, G_2 is not isomorphic to H, since G_2 has two triangles (on vertices $2, 3, 4$, and on vertices $7, 8, 9$), but H has no triangles (since the vertices of such a triangle would have to consist, together, of six distinct elements of $\{1, 2, 3, 4, 5\}$ – impossible !). However, G_3 is isomorphic to H, with isomorphism f given by $f(a) = 12$, $f(b) = 34$, $f(c) = 15$, $f(d) = 23$, $f(e) = 14$, $f(f) = 25$,

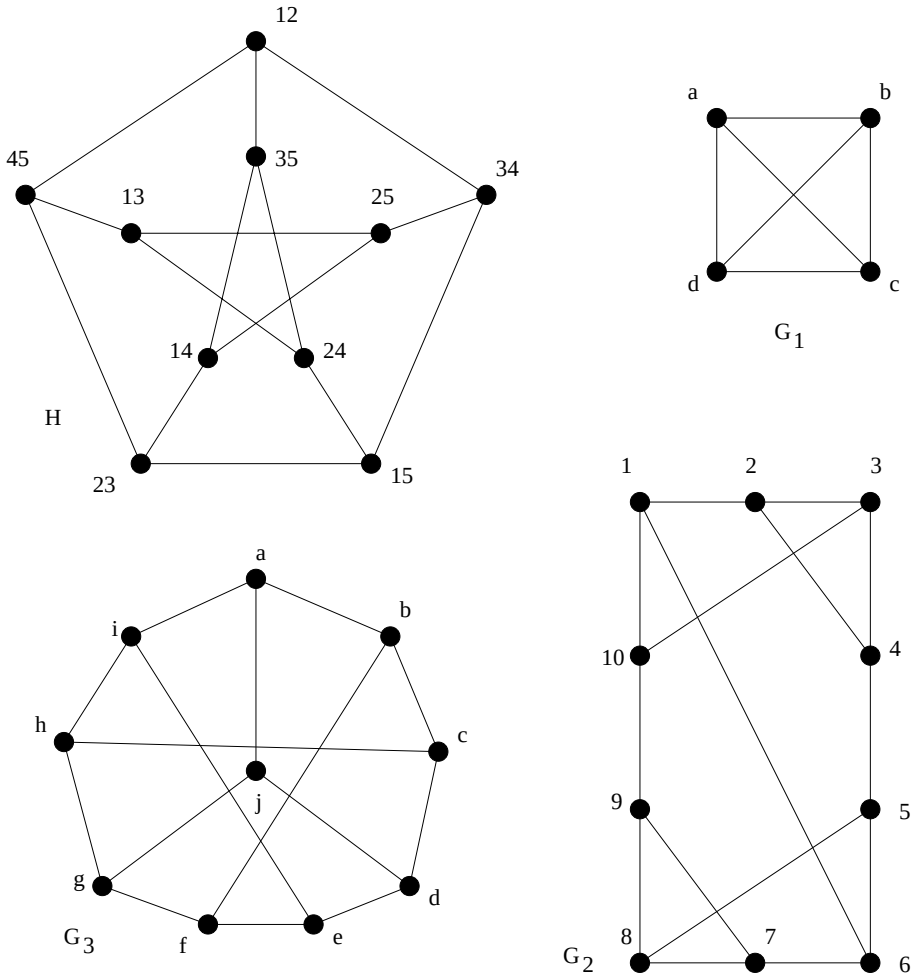

Figure 4: The Petersen graph H

$f(g) = 13$, $f(h) = 24$, $f(i) = 35$, $f(j) = 45$. Of course, this isomorphism needs to be checked: for example, vertex a is adjacent to vertices b, i, j in G_3, and indeed $f(a) = 12$ is adjacent to vertices $f(b) = 34, f(i) = 35, f(j) = 45$ in H.

17 Lecture of February 27

An isomorphism from a graph to itself is called an *automorphism* of the graph. If f is an automorphism of graph G, then for all pairs of distinct vertices $u, v \in V(G)$, we have $\{u, v\} \in E(G)$ if and only if $\{f(u), f(v)\} \in E(G)$. But f is a bijection on $V(G)$, so setting $u' = f(u)$ and $v' = f(v)$, we have $\{u', v'\} \in E(G)$ if and only if $\{f^{-1}(u'), f^{-1}(v')\} \in E(G)$, and we conclude that f^{-1} is also an automorphism of G. The identity function on $V(G)$ is always an automorphism of G. Define the product $f \cdot g$ of two automorphisms f, g by $(f \cdot g)(v) = f(g(v))$, for $v \in V(G)$. Then, with this product, the set of all automorphisms of a graph form a *group*, called the *automorphism group* of the graph, with the identity function as the identity element of this group. This group is a subgroup of the set of all $p!$

bijections on $V(G)$ (where $|V(G)| = p$). Among other consequences, this implies that the number of automorphisms of a graph divides $p!$. For example, how many automorphisms does the graph G in Figure 5 have ? The answer is 2, with the identity function giving one

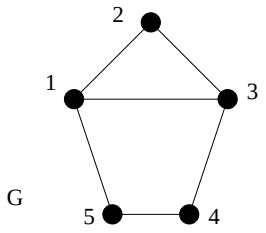

Figure 5: The graph G.

automorphism, and the function $f(1) = 3$, $f(2) = 2$, $f(3) = 1$, $f(4) = 5$, $f(5) = 4$ giving the other. (It is straightforward to check that f is an automorphism. To prove that there are no other automorphisms, note that vertices 1 and 3 have degree 3, but vertices $2, 4, 5$ have degree 2, so every automorphism must map $1, 3$ to $1, 3$ only, in either order; once this order is given, then $2, 4, 5$ are mapped to $2, 4, 5$ in one way only.)

What is the largest number of automorphisms that a graph on p vertices can have ? The answer is $p!$, for either the *complete* graph K_p on p vertices, in which every pair of vertices is adjacent, or the empty graph Z, with no edges. For $p = 5$, these are given in Figure 6.

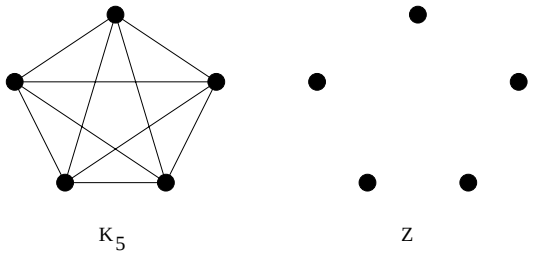

Figure 6: Graphs on 5 vertices with 5! automorphisms.

The relationship between K_5 and Z above is generalized by defining the *complement* of a graph. The complement $\widetilde{G}$ of a graph G is the graph with vertex set $V(G)$, and whose edges are the unordered pairs of vertices that are NOT in $E(G)$. For example, in Figure 6, $\widetilde{K_5} = Z$, and $\widetilde{Z} = K_5$. Now, in general, $\widetilde{G}$ and G have exactly the same automorphisms (this is easy to prove, since a pair of vertices is adjacent in G if and only if it is not adjacent in $\widetilde{G}$). One use of this result is to count automorphisms for a graph whose complement is much easier to handle. For example, consider the graph G given in Figure 7. The complement of G is H, given beside G in Figure 7. But it is easy to see that H has exactly $6 \cdot 4 \cdot 2 = 48$ automorphisms, since 1 can be mapped to any of the 6 vertices, but then 3 must be mapped to the vertex that is adjacent to the image of 1; then 2 can be mapped to any of the 4 remaining vertices, but then 5 must be mapped to the vertex that is adjacent to the image

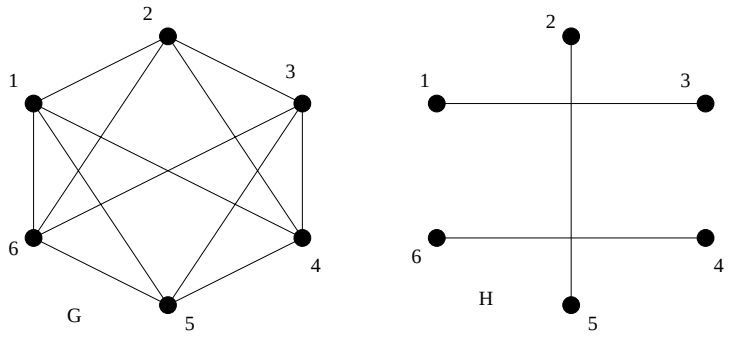

Figure 7: A graph G, and its complement H.

of 2; finally 4 can be mapped to any of the 2 remaining vertices, but then 6 must be mapped to the vertex that is adjacent to the image of 4.

A graph G is *vertex transitive* if, for every pair u, v of vertices in G, there exists an automorphism f such that $f(u) = v$. (Note: In general, the choice of f will depend on u, v.) A graph G is *edge transitive* if, for every pair $\{u, v\}, \{x, y\}$ of edges in G, there exists an automorphism f such that either $f(u) = x, f(v) = y$ or $f(u) = y, f(v) = x$. It follows immediately that a vertex transitive graph must be regular. Does vertex transitivity imply edge transitivity or vice-versa ? Consider the graphs in Figure 8. Note that graphs G_1 and

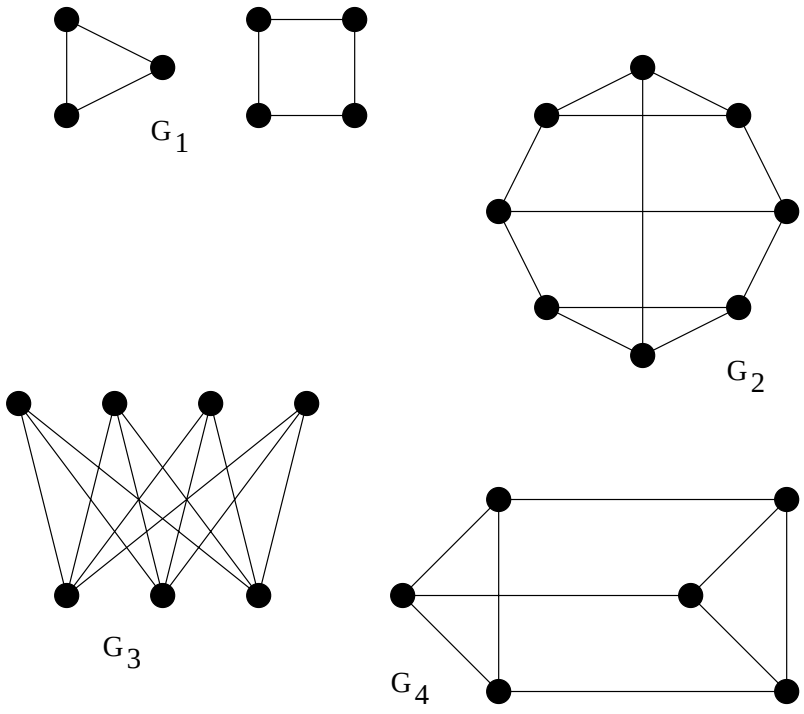

Figure 8: The graphs G_i, $i = 1, \ldots, 4$.

G_2 are both regular, but neither are vertex transitive. For G_1, no automorphism can map

any vertex on the triangle to any vertex on the square. For G_2, there are precisely two triangles, at the top and bottom, so no automorphism can map any of the 6 vertices on these triangles to any of the two remaining vertices.

Now, graph G_3 is edge transitive, but not vertex transitive (also, G_3 is not regular). We write G_3 as $K_{4,3}$, since it has a complete set of edges *between* the 4-set of vertices at the top and the 3-set of vertices at the bottom.

Finally, graph G_4 is vertex transitive, but it is not edge transitive, since the three horizontal edges are not contained in any triangle.

18 Lecture of February 29

A *subgraph* H of a graph G is a graph whose vertex set is a nonempty subset of $V(G)$, and edge set is a subset of the edges of G joining two vertices in $V(H)$. A *spanning* subgraph H of G has $V(H) = V(G)$. An *induced* subgraph of G has vertex set that is a nonempty subset of $V(G)$, and edge set consisting of *all* the edges of G joining two vertices in $V(H)$.

A *walk* in a graph G is a sequence $v_0e_1v_1 \ldots e_nv_n$, $n \geq 0$, in which $v_0, v_1, \ldots, v_n$ are vertices of G, and $e_1, \ldots, e_n$ are edges of G, with $e_i = \{v_{i-1}, v_i\}$ for $i = 1, \ldots, n$. This walk is *from* v_0 *to* v_n, and has *length* n. Note that we can reverse a walk, so that $v_ne_n \ldots v_1e_1v_0$ is then a walk from v_n to v_0, and we therefore often speak of a walk *between* a pair of vertices. A *path* is a walk in which the vertices are distinct.

Note that if there is a walk from vertex u to vertex v in a graph G, then there is a path from u to v in G, by the following argument: let the vertices on the walk be $v_0v_1 \ldots v_n$, where $v_0 = u$, $v_n = v$. Then if no vertex appears more than once, it is a path, and we are done. Otherwise, there exists $i < j$ with $v_i = v_j$, and in this case, $v_0v_1 \ldots v_iv_{j+1} \ldots v_n$ is a shorter walk from u to v. Continue until we have a path (n is finite, and G has a finite number of vertices, edges, so this is a finite procedure.)

Now, given a graph G, define the relation $\mathcal{P}_G$ on $V(G)$ by $u\mathcal{P}_Gv$ if there is a path in G from u to v.

Theorem 18.1 *For any graph G, $\mathcal{P}_G$ is an equivalence relation.*

PROOF. The relation is *reflexive*, since v is a path of length 0 from v to v, for all $v \in V(G)$ (so we have $v\mathcal{P}_Gv$). The relation is *symmetric*, since if $v_0v_1 \ldots v_n$ is a path from $u = v_0$ to $v = v_n$, then $v_n \ldots v_1v_0$ is a path from $v = v_n$ to $u = v_0$, for all $u, v \in V(G)$ (so we have $u\mathcal{P}_Gv$ implies $v\mathcal{P}_Gu$). The relation is *transitive*, since if there is a path $x_0 \ldots x_n$ from u to v in G (with $u = x_0$, $v = x_n$), and there is a path $y_0 \ldots y_m$ from v to w in G (with $v = y_0$, $w = y_m$), then $x_0 \ldots x_ny_1 \ldots y_m$ is a walk from u to w in G, but the result above then implies that there is a path from u to w (so we have $u\mathcal{P}_Gv$ and $v\mathcal{P}_Gw$ implies $u\mathcal{P}_Gw$).

Of course, an equivalence relation partitions the set on which it is defined into *equivalence classes* (which means that $u\mathcal{P}_Gv$ if and only if u and v are in the same equivalence class). We define a graph G to be *connected* if $\mathcal{P}_G$ has a single equivalence class. In general, the induced subgraph whose vertex set is an equivalence class is called a *(connected) component*

of G, so a connected graph has a single component. (Another definition of component is that it is a *maximal* connected subgraph.) For example, the graph G in Figure 9 has two components, one containing vertices $\{1, \ldots, 8\}$, and the other containing vertices $\{9, 10\}$.

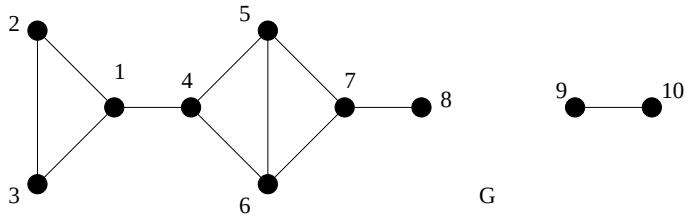

Figure 9: A graph G with two components.

For an edge e in a graph G, we define $G - e$ to be the spanning subgraph of G, with all edges of G except for e. An edge e in a connected graph G is defined to be a *bridge* if $G - e$ is not connected. If G is not connected, then an edge e in G is said to be bridge if e is a bridge of some component of G. For example, the bridges in the graph G of Figure 9 are $\{1, 4\}$, $\{7, 8\}$, and $\{9, 10\}$.

19 Lecture of March 3

If $e = \{x, y\}$ is a bridge in connected G, how many components does $G - e$ have ? The answer is that $G - e$ always has exactly two components, one containing x, and the other containing y. To prove this, let V_x be the set of vertices in the same component of $G - e$ as x. Let z be any vertex in $V(G)$ that is not in V_x (there must be at least one such z, since $G - e$ is not connected). Consider a path $v_0 \ldots v_n$ in G from $x = v_0$ to $z = v_n$ (there is such a path since G is connected). Now this path must contain edge e, since there is no path from x to z in $G - e$, and since a path contains no repeated vertices, we must have $v_0 = x$ and $v_1 = y$. Therefore, $v_1 \ldots v_n$ is a path from $y = v_1$ to $z = v_n$ in $G - e$, which implies that z is in the same component of $G - e$ as y for any such z. This means that the vertices not in V_x are all contained in a second component, which contains y.

Example 19.1 *Prove that there can be no bridge in a 4-regular graph.*

PROOF. If there is such a bridge $e = \{x, y\}$, then let G_x be the component of $G - e$ containing x (and not y, by the above result). Now, all vertices in G_x have degree 4, except for vertex x, which has degree 3. But this contradicts the handshake theorem, which says that G_x must have an even number of vertices of odd degree, so we conclude that G has no bridge.

For $n \geq 3$, a *cycle* (often called an n-cycle) in a graph G is a set of n distinct vertices $\{v_1, \ldots, v_n\}$ and n distinct edges $\{\{v_1, v_2\}, \ldots \{v_{n-1}, v_n\}, \{v_n, v_1\}\}$. For example, vertices $\{1, 2, 3\}$ and edges $\{\{1, 2\}, \{2, 3\}, \{3, 1\}\}$, form a 3-cycle (or *triangle*) in the graph G in Figure 9. Cycles give another, powerful, way of characterizing bridges: an edge e in a graph G is a bridge if and only if e is NOT contained in any cycle of G. To prove this, simply note

that a cycle containing edge $e = \{x, y\}$ consists of edge $\{x, y\}$, together with a path from x to y in $G - e$, in which case e is not a bridge. For example, this characterization can be checked for the three bridges in the graph G in Figure 9.

We begin with another lemma about cycles. Suppose that there are two different paths between some pair of distinct vertices in a graph G. Then G must contain a cycle. To prove this suppose the paths are $P_1 = u_0 \ldots u_m$, and $P_2 = v_0 \ldots v_n$, where $u_0 = v_0$, and $u_m = v_n$. Let $i \geq 0$ be the smallest value for which $u_{i+1} \neq v_{i+1}$ (so $u_0 = v_0, \ldots, u_i = v_i$). There must be such an i, since the paths are different. Let j be the smallest value, with $j > i$, for which u_j is contained in P_2, and suppose that $u_j = v_k$. There must be such a j, since u_m is contained in P_2. Then $u_i u_{i+1} \ldots u_j v_{k-1} \ldots v_{i+1}$ is a cycle in G, proving the result.

A *tree* is a connected graph with no cycles. Since it is connected, a tree must contain at least one path between every pair of vertices, but the above result implies that there cannot be more than one path between any pair of vertices (otherwise, we would have a cycle). Thus we conclude that there is a unique path between every pair of vertices in a tree. Also, since no edge in a tree is contained in a cycle, our cycle characterization of bridges implies that in a tree, every edge is a bridge.

Every tree on p vertices has $q = p - 1$ edges, for $p \geq 1$. We prove this by (strong) induction on p. For the base case, there is only one tree for $p = 1$, which has 0 edges (this is connected, with no cycles), so the result is true for $p = 1$. For the induction hypothesis, with $p > 1$, assume that the result holds for values smaller than p, and consider an arbitrary tree T on p vertices. Now T has at least one edge (otherwise, it is not connected, since it would have $p > 1$ components). Remove an arbitrary edge $e = \{x, y\}$ from T. We have proved above that e is a bridge, so $T - e$ has two components, with one containing vertex x, and the other containing vertex y. There can be no cycles in $T - e$, since T has no cycles, so both components of $T - e$ are trees. Let T_1 be the component of $T - e$ containing x, and T_2 be the component containing y. Let p_i be the number of vertices in T_i, for $i = 1, 2$. Then we have $p_1, p_2 \geq 1$, and $p_1 + p_2 = p$, so $p_1, p_2 \leq p - 1$, and we can thus apply the induction hypothesis to determine that the number of edges in T_i is $p_i - 1$, for $i = 1, 2$. Then the number of edges in T is given by

$$(p_1 - 1) + (p_2 - 1) + 1 = p_1 + p_2 - 1 = p - 1,$$

and we have proved that the result is true for T. The result has now been established for all $p \geq 1$, by mathematical induction.

A *spanning tree* of a graph is a spanning subgraph that is a tree.

Theorem 19.2 *A graph G is connected if and only if it has a spanning tree T.*

PROOF. If G has a spanning tree T, then T contains a unique path between every pair of vertices, and all such paths are contained in G, implying that G is connected.

For the converse, suppose that G is connected. If G has no cycle, then G is itself a tree, and so has a spanning tree. Otherwise, if G has a cycle, remove an arbitrary edge e on any cycle of G. Now, from the cycle characterization of bridges, e is not a bridge, so $G - e$ is a spanning subgraph of G that is connected, and has one fewer edge than G. Repeat this finite process, until there is no cycle, to obtain a spanning tree of G.

20 Lecture of March 5

We now consider various types of spanning subgraphs of a graph G with p vertices and q edges. The total number of spanning subgraphs is 2^q, since these correspond exactly to subsets of the q edges (every spanning subgraph contains all of the p vertices). How many of these spanning subgraphs have an even number of edges ? If $q \geq 1$, the answer is 2^{q-1}. For one proof, use the identity

$$\sum_{i \text{ even}} \binom{q}{i} = \sum_{i \text{ odd}} \binom{q}{i} = 2^{q-1},$$

which follows immediately from the binomial expansion of $0 = (1-1)^q$. For a second proof, set aside an arbitrary edge e, and consider any subset α of the remaining edges. We claim that there is a unique spanning subgraph with an even number of edges corresponding to α: if $|\alpha|$ is even, then α is the edge set of one of these spanning subgraphs; if $|\alpha|$ is odd, then $\alpha \cup \{e\}$ is the edge set of one of these spanning subgraphs, and this is a bijection.

An *even* graph is a graph in which every vertex has even degree (where 0 is included). An *even spanning subgraph* is a spanning subgraph which is even. For example, consider the graph G with $p = 6$ vertices and $q = 8$ edges given in Figure 10. There are 8 even spanning

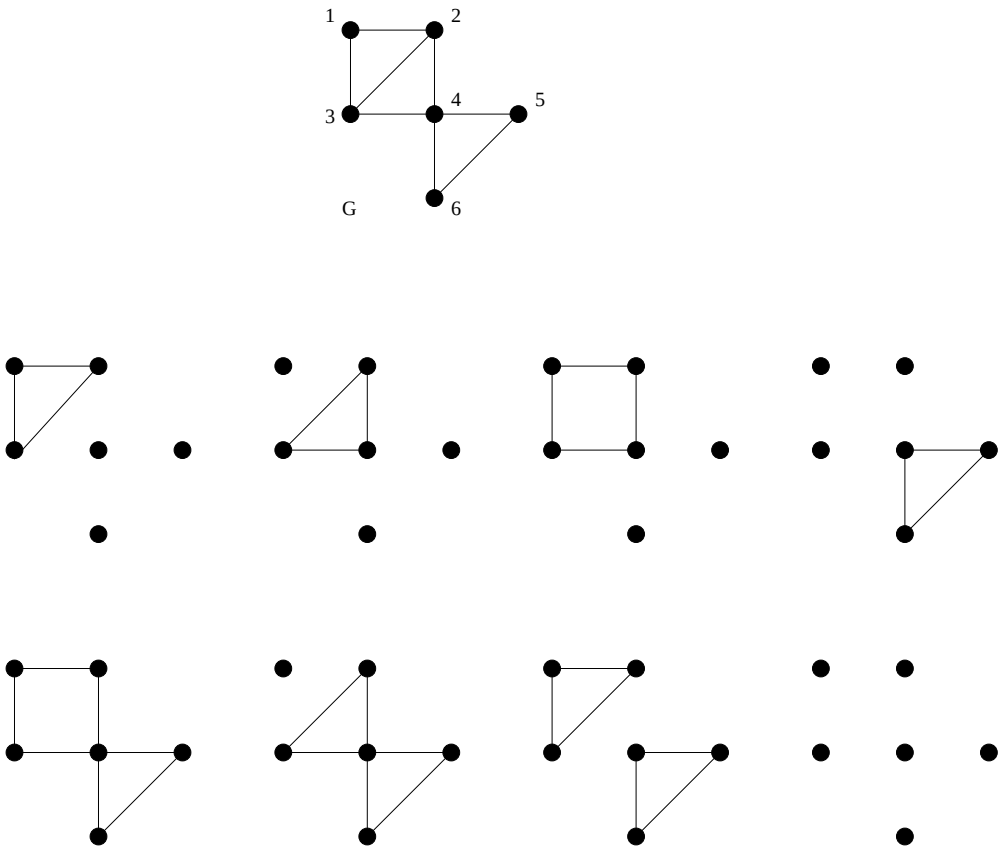

Figure 10: A graph G together with even spanning subgraphs.

subgraphs of G, also given in Figure 10, below G itself.

For any field F and positive integer q, the set F^q, consisting of q-tuples of elements of F, is a vector space of dimension q. This vector space is also given by span$\{(1,0,\ldots,0),(0,1,0,\ldots,0),\ldots,(0,\ldots,0,1)\}$, where 0 and 1 are the additive identity and multiplicative identity in the field, respectively. In this course, we consider the field $\mathrm{Z}_2 = \{0,1\}$ of integers modulo 2.

Consider a graph G with q edges, where we index the edges $e_1,\ldots,e_q$ in an arbitrary, but fixed, way. The *incidence vector* of a spanning subgraph of G is a $\{0,1\}$ q-tuple, with 1 in entry i if e_i is contained in the subgraph, and 0 otherwise, for each $i = 1,\ldots,q$. Using incidence vectors, we associate a unique element of Z_2^q with each spanning subgraph of G. The 0-vector $(0,\ldots,0)$ is the incidence vector of the spanning subgraph with no edges, and we denote this graph by Z. Now, to understand what the vector space operations on incidence vectors mean for the underlying spanning subgraphs themselves: for scalar multiples of a spanning subgraph H, we have $1\,H = H$, since scalar multiplication of the incidence vector by 1 doesn't change any of the entries; we also have $0\,H = Z$, since scalar multiplication of the incidence vector by 0 makes all of the entries 0. For spanning subgraphs H_1, H_2, the sum $H_1 + H_2$ is the spanning subgraph whose edge set is the symmetric difference of the edge sets of H_1 and H_2 – i.e., it has all the edges that are contained in *exactly* one of H_1 and H_2, which exactly mimics the sums in Z_2 that are carried out in each entry. Thus we refer to the set of spanning subgraphs of a graph as the *edge space* of the graph.

We now consider two subspaces of the edge space.

Theorem 20.1 *For any graph G with q edges, the set of spanning subgraphs with an even number of edges is a $(q-1)$-dimensional subspace of the edge space of G (the vector space of spanning subgraphs).*

SOLUTION. The set of incidence vectors $\mathbf{v}^{\mathrm{t}} = (v_1,\ldots,v_q)^{\mathrm{t}}$ (regarding them as column vectors) for these spanning subgraphs is exactly the *null space* of the $1 \times q$ matrix $A = (1\ 1\ \ldots 1)$. We recall from linear algebra, that this is always a vector space over the same field as the entries of A. (PROOF: If A is $m \times n$, then its null space consists of all $n \times 1$ vectors $\overline{x}$ such that $A\overline{x} = \overline{0}$, where $\overline{0}$ is $m \times 1$. Thus the null space contains $\overline{0}_{n\times 1}$, and is thus nonempty. It is closed under scalar multiplication by the field element c, since for $\overline{x}$ in the null space we have

$$Ac\overline{x} = cA\overline{x} = c\,\overline{0} = \overline{0}.$$

The null space is also closed under vector addition, since for $\overline{x},\overline{y}$ in the null space, we have

$$A(\overline{x}+\overline{y}) = A\overline{x} + A\overline{y} = \overline{0} + \overline{0} = \overline{0}.$$

Thus the null space of A is a subspace of F^n (as column vectors), and is thus a vector space over F.)

Since the matrix A has rank 1, then the null space has dimension $q-1$ (rank plus nullity equals number of columns). As in our linear algebra courses, we determine the null space by considering the linear equation

$$v_1 + v_2 + \ldots + v_q = 0,$$

and setting $v_2, \ldots, v_q$ as arbitrary parameters in Z_2. Thus we have $v_1 = -v_2 - \ldots - v_q = v_2 + \ldots + v_q$, so the null space is given by

$$\{(v_2 + \ldots + v_q, v_2, \ldots, v_q)^{\mathrm{t}} : v_2, \ldots, v_q \in \mathrm{Z}_2\}$$

$$= \mathrm{span}\{(1,1,0,\ldots,0)^{\mathrm{t}}, (1,0,1,0,\ldots,0)^{\mathrm{t}}, \ldots (1,0,\ldots,0,1)^{\mathrm{t}}\},$$

where, of course, these $q-1$ vectors are linearly independent. Note that the first entry in these vectors acts as a parity check.

We could also establish the above result more directly, instead of recognizing the set as the null space of a particular matrix A, by proving that the set is a nonempty subset of the edge space that is closed under scalar multiplication and vector addition. Indeed, we use this method of proof for the following result, concerning the second subspace of the edge space.

Theorem 20.2 *For any graph G with q edges, the set $C(G)$ of even spanning subgraphs is a subspace of the edge space of G.*

SOLUTION. The set $C(G)$ contains the subgraph Z, with no edges, since this subgraph has degree 0 at every vertex. Then for $H \in C(G)$, we have $0\,H = Z \in C(G)$, and $1\,H = H \in C(G)$, so $C(G)$ is closed under scalar multiplication. Now, for $H_1, H_2 \in C(G)$, we consider $H_1 + H_2$: for an arbitrary vertex $v \in G$, suppose that v has degree $2k$ in H_1, and degree $2m$ in H_2, and that exactly n of the edges incident with v in H_1 are also contained in H_2; then the degree of vertex v in $H_1 + H_2$ is given by $2k + 2m - 2n$, which is even, and we conclude that $H_1 + H_2 \in C(G)$, so $C(G)$ is closed under addition. Thus $C(G)$ is a subspace, giving the result.

To determine the dimension of the subspace $C(G)$, we consider first the case in which G is connected. Since G is connected, it must have a spanning tree, so let T be an arbitrary, but fixed spanning tree of G. We refer to the edges of G that are contained in T as *tree edges*, and the edges of G that are not contained in T as *nontree edges*.

For $H_1, H_2 \in C(G)$, suppose that H_1, H_2 have exactly the same sets of nontree edges. Then all edges of $H_1 + H_2$ are tree edges, so all components of $H_1 + H_2$ are trees (contained in T).

21 Lecture of March 7

In a tree on $p \geq 2$ vertices, there must be at least two vertices of degree 1. To prove this, consider a longest path $u_0 \ldots u_m$ in the tree. Since $p \geq 2$, we must have $m \geq 1$, since every edge would give a path of length 1, and a graph on p vertices and 0 edges has p components (and thus is not connected for $p \geq 2$). Now, u_0 is adjacent to u_1. If u_0 is adjacent to u_j for any $j \geq 2$, then $u_0 \ldots u_j$ is a cycle in the tree, which is impossible. If u_0 is adjacent to v for any vertex v in the tree that is not contained in the longest path, then $v u_0 \ldots u_m$ is a path in the tree that is longer than the longest path, which is impossible. Therefore, u_0 is adjacent only to vertex u_1 in the tree, and so has degree 1. The same argument proves that u_m has degree 1, so we have constructed two vertices of degree 1, proving the result.

Thus, returning to the argument at the end of the previous class, if any of the components of H_1+H_2 has two or more vertices, then it must have at least two vertices of degree 1, which would contradict the fact that $H_1+H_2 \in C(G)$. Thus we conclude that every component of H_1+H_2 is a single vertex, so $H_1+H_2 = Z$, or equivalently, $H_1 = H_2$. Thus we have proved that no two elements of $C(G)$ have exactly the same set of nontree edges.

Suppose that G has p vertices. Then T has $p-1$ edges, so there are exactly $q-p+1$ nontree edges, and we conclude that

$$|C(G)| \leq 2^{q-p+1},$$

or, equivalently, that dimension of $C(G)$ is less than or equal to $q-p+1$.

Now suppose that $e = \{u, v\}$ is a nontree edge of G. Then T contains a unique path between vertices u and v, so $T+e$ (which is the tree T, with edge e added) contains a unique cycle, formed by the path between u and v in T, together with edge e. We call this the *fundamental cycle* of e, and denote it by C_e. Index the edges of G so that the nontree edges are given by $e_1, \ldots, e_{q-p+1}$.

Theorem 21.1 *The fundamental cycles* $\{C_{e_1}, \ldots, C_{e_{q-p+1}}\}$ *form a linearly independent set in* $C(G)$.

PROOF. The 0 vector or additive identity in $C(G)$ is Z. Now consider which scalars $a_1, \ldots, a_{q-p+1}$ in Z_2 satisfy the equation

$$a_1 C_{e_1} + \ldots + a_{q-p+1} C_{e_{q-p+1}} = Z.$$

For each $i = 1, \ldots, q-p+1$, if $a_i = 1$, then the spanning subgraph on the lefthandside contains edge e_i, since e_i is contained in C_{e_i}, but e_i is not contained in C_{e_j} for any $j \neq i$. Since e_i is not contained in Z, we conclude that $a_i = 0$, for $i = 1, \ldots, q-p+1$. This gives the result.

Now, if $\dim C(G) = k$, then we have a basis $\{H_1, \ldots, H_k\}$ for $C(G)$, and the elements of $C(G)$ are then created uniquely in $\mathrm{span}\{H_1, \ldots, H_k\}$. This implies that $|C(G)| = 2^k$. In the above result, we have constructed a linearly independent set in $C(G)$ of size $q-p+1$, which implies that $\dim C(G) \geq q-p+1$, and so we conclude that

$$|C(G)| \geq 2^{q-p+1}.$$

From the two inequalities, we then get $|C(G)| = 2^{q-p+1}$, so $\dim C(G) = q-p+1$, and the linearly independent set $\{C_{e_1}, \ldots, C_{e_{q-p+1}}\}$ of fundamental cycles is a basis for $C(G)$.

For $H \in C(G)$, with $H \neq Z$, what is the minimum number of edges in H ? The answer is 3, since H must contain a cycle (otherwise, if H has no cycles, then all components of H are trees, and if any such component has an edge, it must have at least two vertices of degree 1, which is odd; but $H \neq Z$ must have at least one edge, so it contains a cycle). In terms of incidence vectors, this means that the incidence vector for H differs from the incidence vector for Z in at least 3 positions. Now consider $H_1, H_2 \in C(G)$, $H_1 \neq H_2$. Then, from arithmetic in Z_2 ($1+1 = 0+0 = 0$, $1+0 = 0+1 = 1$), we know that the incidence vectors for H_1 and H_2 differ in a position exactly when the incidence vector for

$H_1 + H_2$ has a 1 in that position. This implies that the number of positions in which the incidence vectors for H_1 and H_2 differ, is equal to the number of 1's in the incidence vector for $H_1 + H_2$. But $H_1 + H_2 \in C(G)$, so the incidence vector for $H_1 + H_2$ contains at least three 1's, and therefore the incidence vectors for H_1 and H_2 differ in at least 3 positions for all $H_1, H_2 \in C(G)$, $H_1 \neq H_2$.

For example, consider Figure 11, in which the graph G from Figure 10 appears, with the edges of a spanning tree T drawn as thick lines, the nontree edges as thin lines. For this graph, we have $q = 8$, $p = 6$, and the edges are indexed as marked in Figure 11. The spanning

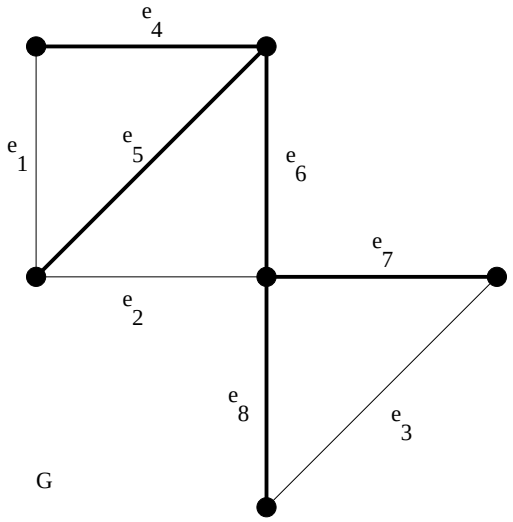

Figure 11: A graph G and spanning tree T.

subgraphs in $C(G)$ are given in Figure 10. With this indexing of edges, the incidence vectors of these spanning subgraphs are given as the rows of the following array.

e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8
1	0	0	1	1	0	0	0
0	1	0	0	1	1	0	0
1	1	0	1	0	1	0	0
0	0	1	0	0	0	1	1
1	1	1	1	0	1	1	1
0	1	1	0	1	1	1	1
1	0	1	1	1	0	1	1
0	0	0	0	0	0	0	0

Note that, indeed, every pair of distinct rows differs in at least three positions.

22 Lecture of March 10

The incidence vectors of $C(G)$ are said to form a *binary* $(q, q-p+1, 3)$-*code*, since they are binary vectors of length q forming a vector space of dimension $q-p+1$, in which the vectors

differ pairwise in at least 3 positions. If every cycle in G has length at least equal to t, then we get a $(q, q-p+1, t)$-code, since the vectors differ in at least t positions in this case. The importance of the pairwise difference in positions is that if information is encoded in terms of the vectors in the code, then errors in transmission/reception can be detected (if there are fewer than t errors), and even corrected (if there are at most $(t-1)/2$ errors). Coding theory is studied in CO 331.

Let $N_{m,n}$, $n \geq m$, denote the number of $m \times n$ $\{0,1\}$-matrices of rank m. Then we have $N_{1,n} = 2^n - 1$, since a row consisting entirely of 0's is the only way to avoid rank 1. Similarly, $N_{m,n} = (2^n - 2^{m-1})N_{m-1,n}$, since the first $m-1$ rows must be linearly independent (to give rank $m-1$), and then row m cannot be any of the 2^{m-1} linear combinations of the first $m-1$ rows. We conclude that

$$N_{m,n} = \prod_{k=0}^{m-1}(2^n - 2^k).$$

For example, this allows us to determine the probability that an $n \times n$ $\{0,1\}$-matrix is nonsingular (over Z_2), given by

$$\frac{\prod_{k=0}^{n-1}(2^n - 2^k)}{2^{(n^2)}} = \prod_{j=1}^{n}(1 - 2^{-j}).$$

It is straightforward to determine the number of m-dimensional subspaces of a vector space of dimension n over Z_2, since it is given by the ratio

$$\frac{N_{m,n}}{N_{m,m}} = \frac{(2^n-1)(2^{n-1}-1)\cdots(2^{n-m+1}-1)}{(2^n-1)(2^{m-1}-1)\cdots(2^1-1)} = \binom{n}{m}_2,$$

where $\binom{n}{k}_q$ is the Gaussian coefficient, defined near the beginning of the Lecture of February 13 (following (13)).

We now consider a spanning tree algorithm that, among other consequences, will allow us to find the shortest cycle in a graph.

Algorithm: Input a graph G, on p vertices. Initially, let subgraph D of G consist of an arbitrary vertex, denoted r, of G, with $pr(r) = \emptyset$. At every stage, find an edge $\{u, v\}$ of G, with $u \in D$, and $v \notin D$; add vertex v and edge $\{u, v\}$ to D, with $pr(v) = u$. Stop when there is no such edge.

Claim: At termination, if D has p vertices, then it is a spanning tree of G. Otherwise, if D has fewer than p vertices, then G is not connected (so it has no spanning tree).

PROOF. Initially, D has no edge and a single vertex, and at every stage we add one edge and one vertex, so at stage k, D has $k-1$ edges and k vertices. Also, initially D is connected, and at every stage we add an edge from a new vertex, v, to an existing vertex, u. Thus, by induction on the number of stages, D is always connected. But a connected graph with k vertices and $k-1$ edges must have a spanning tree (from Theorem 19.2), which must have all $k-1$ edges, and thus D is a tree at every stage. The first part of the claim, that D is a spanning tree of G if it has p vertices, follows immediately.

For the second part of the claim, suppose that D has fewer than p vertices at termination, but that G is connected. Then we have a vertex x of G, not in D. Consider a path $v_0 \ldots v_n$ in G, from $r = v_0$ (where r is the initial vertex of D), to $x = v_n$. (Such a path exists, since G is connected.) Now, we have $v_0 \in D$, $v_n \notin D$, so there exists $v_i \in D$, $v_{i+1} \notin D$, for some $i = 0, \ldots, n-1$. But the algorithm stopped because there is no edge such as $\{v_i, v_{i+1}\}$, so we have a contradiction, and conclude that G cannot be connected. This proves the second part of the claim.

23 Lecture of March 12

We call D a *search tree*, since the "*pr*" (for pointer, parent, predecessor) function allows us to find paths in D (and therefore in G): apply pr repeatedly to a vertex until we get r, to obtain a path from that vertex to the root vertex of D. If we apply pr exactly k times to do this, then we say that the vertex is at *level* k in D.

For *breadth first search*, we apply the algorithm above, but choose an edge $\{u, v\}$, with $u \in D$, $v \notin D$, for which u joined D first among all such edges. For *depth first search*, we apply the algorithm above, but choose an edge $\{u, v\}$, with $u \in D$, $v \notin D$, for which u joined D last among all such edges. We begin by proving the primary property of Breadth First Search.

Theorem 23.1 *In a Breadth First Search tree, the nontree edges join vertices that are at most one level apart in the tree.*

PROOF. We first prove that the vertices join the tree in weakly increasing order by level, using induction on the number of vertices in the tree. As base cases, the result is true if there are one or two vertices in the tree, since the first two vertices have levels 0 and 1, respectively. For the induction hypothesis, assume that the result holds for k vertices in the tree, for some $k \geq 2$, and denote these vertices, in order of joining the tree, as $v_1, \ldots, v_k$ (so the induction hypothesis is that $level(v_1) \leq \ldots \leq level(v_k)$). Now consider the $k+1$st vertex to join the tree, v_{k+1}. Suppose $pr(v_{k+1}) = v_i$ and $pr(v_k) = v_j$. Then we must have $j \leq i$, and the induction hypothesis gives $level(v_j) \leq level(v_i)$, so we obtain

$$level(v_k) = level(v_j) + 1 \leq level(v_i) + 1 = level(v_{k+1}),$$

and we have proved that the result holds for $k+1$ vertices in the tree. Therefore the result is true by mathematical induction.

Now suppose that the vertices in the Breadth First Search tree, in order of joining the tree, are $v_1, \ldots, v_m$. Suppose that there is a nontree edge joining v_i to v_j, where $i < j$. Then we have $pr(v_j) = v_k$, where $k < i$, since otherwise, if $k > i$, we would have had $pr(v_j) = v_i$. This gives

$$level(v_j) = level(v_k) + 1 \leq level(v_i) + 1,$$

from the result above, and so we have proved that the nontree edge $\{v_i, v_j\}$ joins vertices at most one level apart.

One consequence of the primary property of Breadth First Search is for length of shortest paths: The length of the shortest path between vertices u and v in a connected graph G is equal to $level(v)$ in any Breadth First Search tree of G rooted at u. (Or, equivalently, to $level(u)$ in any Breadth First Search tree of G rooted at v.) The proof is immediate, since there is a path of the given length, using tree edges only, and there can be no shorter path, since it would require an edge that joins vertices two or more levels apart.

Another consequence is for parity of cycles. Note that if a nontree edge joins two vertices at the same level in a Breadth First Search tree, then the fundamental cycle for that edge must have odd length (if the vertices are at the same level, k, and the paths from these vertices to the root first meet at a vertex at level m, then the path between these vertices in the tree has length $2(k-m)$, and together with the nontree edge, the fundamental cycle thus has length $2(k-m)+1$). If a nontree edge joins vertices one level apart, then the fundamental cycle for that edge must have even length (if the vertices are at levels k and $k+1$, and the paths from these vertices to the root first meet at a vertex at level m, then the path between these vertices in the tree has length $(k-m)+(k+1-m)$, and together with the nontree edge, the fundamental cycle thus has length $2(k+1-m)$).

A graph G is said to be *bipartite* if $V(G)$ can be partitioned into two nonempty subsets A, B, so that every edge of G joins a vertex of A to a vertex of B.

Theorem 23.2 *A graph is bipartite if and only if it has no odd cycles.*

PROOF. Consider a cycle $v_1, v_2, \ldots, v_k$ in a bipartite graph. Then, without loss of generality, suppose $v_1 \in A$ (otherwise, we can interchange the sets A and B). This forces $v_2 \in B$ (because of edge $\{v_1, v_2\}$), and then we have $v_3 \in A$. Continuing, we can prove by induction that $v_m \in A$ for m odd, and $v_m \in B$ for m even. But we also have $v_k \in B$, because of edge $\{v_k, v_1\}$, and thus conclude that k must be even. Thus we have proved that if G is bipartite, then it has no odd cycles.

For the converse, suppose that G has no odd cycles. Then if we carry out Breadth First Search on every component of G, we will never find a nontree edge joining vertices at the same level (otherwise, we would have an odd cycle, from the discussion above, which is impossible, from the first part of this result). But this means that the graph is bipartite, with bipartition given by A, the vertices at even levels in these trees, and B, the vertices at odd levels.

24 Lecture of March 14

Define the *girth* of a graph to be the length of the shortest cycle in the graph. If the graph has no cycles, then we define its girth to be ∞. By considering Breadth First Search, we are able to prove the following result.

Theorem 24.1 *For $k \geq 2$, a k-regular graph of girth 5 has at least k^2+1 vertices.*

PROOF. Carry out Breadth First Search on the graph, with root vertex an arbitrary vertex v. Then the k vertices adjacent to v, call them $u_1, \ldots, u_k$, all join the tree at level 1, in

that order. There can be no edges between u_i and u_j for any $i \neq j$, since otherwise $u_i u_j v$ would be a 3-cycle in the graph, which is impossible since the girth is 5. Thus the $k-1$ vertices other than v that are adjacent to u_1, call them $w_{11}, \ldots, w_{1\,k-1}$, all join the tree at level 2, with a pointer to u_1. But u_2 cannot be adjacent to any w_{1j}, since otherwise $u_2 w_{1j} u_1 v$ would be a 4-cycle, which is impossible since the girth is 5. By a similar argument, for every $i = 1, \ldots, k$, we establish that the $k-1$ vertices other than v that are adjacent to u_i, call them $w_{i1}, \ldots, w_{i\,k-1}$, all join the tree at level 2, with a pointer to u_i. But this means that the tree already has $1 + k + k(k-1) = k^2 + 1$ vertices, all of which are in the graph, and the result follows.

For which values of $k \geq 2$ is there a k-regular graph of girth 5 that exactly achieves this bound of $k^2 + 1$ vertices ? The answer is only $k = 2, 3, 7$ and (possibly) 57. I say "possibly" because it is still an open research problem as to whether such a graph exists – such a graph is referred to as the "Moore graph". The proof of the impossibility of other values for k is given below:

Let A be the *adjacency* matrix for such a graph. That is, call the vertices $v_1, \ldots, v_n$, in some arbitrary way, and then A is $n \times n$, where $n = k^2 + 1$, with (i, j)-entry equal to 1 if vertex v_i is adjacent to vertex v_j. Clearly A is a symmetric $0, 1$ matrix with diagonal entries equal to 0.

Theorem 24.2 *Let I_n be the $n \times n$ identity matrix, and J_n be the $n \times n$ matrix with all entries equal to* 1. *Then for $n = k^2 + 1$ and A given above, we have*

$$A^2 = (k-1)I_n + J_n - A.$$

PROOF. The (i, j)-entry of A^2 is given by

$$\left(A^2\right)_{i,j} = \sum_{m=1}^{n} A_{i,m} A_{m,j},$$

where $A_{i,j}$ is the (i, j)-entry of A. Then, for $i = j$, we have

$$\left(A^2\right)_{i,i} = \sum_{m=1}^{n} A_{i,m} = k,$$

for all $i = 1, \ldots, n$, since every vertex v_i has degree k. For v_i adjacent to v_j, we have $\left(A^2\right)_{i,j} = 0$, since if $A_{i,m} = A_{m,j} = 1$ for any m, then $v_i v_j v_m$ would be a 3-cycle in the graph, which is a contradiction. Finally, we treat the remaining case, that $i \neq j$ and v_i is not adjacent to v_j. In this case, consider a Breadth First Search tree rooted at v_i. Then v_j appears at level 2 in the tree, and $pr(v_j) = v_\ell$ for some unique v_ℓ at level 1 in the tree, which means that $A_{i,\ell} A_{\ell,j} = 1$. But there is no other v_m with $A_{i,m} A_{m,j} = 1$, since this would imply that $v_i v_\ell v_j v_m$ would be a 4-cycle in the graph, which is a contradiction. Thus, in this case we have $\left(A^2\right)_{i,j} = 1$. The result follows, since the matrix on the right hand side has entries $k, 0, 1$ in precisely the correct positions.

Now, let $\mathbf{1}_n$ be the $n \times 1$ vector of 1's. Then the fact that

$$A\mathbf{1}_n = k\mathbf{1}_n$$

follows immediately from the fact that the graph is k-regular. Thus A has $\mathbf{1}_n$ as an eigenvector, with k as the corresponding eigenvalue. Now consider another eigenvector $\mathbf{x}$, with corresponding eigenvalue λ. Then the fact that A is real and symmetric implies that $\mathbf{x}$ and $\mathbf{1}_n$ are orthogonal (they correspond to distinct eigenvalues). But this implies that $J_n\mathbf{x} = \mathbf{0}_n$, the $n \times 1$ vector of 0's. Thus from Theorem 24.2, we have

$$\lambda^2\mathbf{x} = A^2\mathbf{x} = ((k-1)I_n + J_n - A)\,\mathbf{x} = (k-1)\mathbf{x} + \mathbf{0}_n - \lambda\mathbf{x},$$

which implies that $(\lambda^2 + \lambda + 1 - k)\,\mathbf{x} = \mathbf{0}_n$, and thus that

$$\lambda^2 + \lambda + 1 - k = 0,$$

since $\mathbf{x} \neq \mathbf{0}_n$. The roots of this quadratic are $-\frac{1}{2} \pm \frac{D}{2}$, where $D = \sqrt{4k-3}$. Thus A has eigenvalues $\lambda_0 = k$, $\lambda_1 = -\frac{1}{2} + \frac{D}{2}$, $\lambda_2 = -\frac{1}{2} - \frac{D}{2}$, and we let m_0, m_1, m_2 be the corresponding algebraic multiplicities. Of course, we have $m_0 + m_1 + m_2 = n$. To obtain two other linear equations for m_0, m_1, m_2, use the fact that A is real symmetric, so it is diagonalizable. Thus we have $A = P\Lambda P^{-1}$, where Λ is a diagonal matrix with diagonal entries $\lambda_0, \lambda_1, \lambda_2$, with multiplicities m_0, m_1, m_2, respectively. Thus we obtain $A^\ell = P\Lambda^\ell P^{-1}$ for any positive integer ℓ (prove this, say, by induction on ℓ). Now for the trace of a square matrix (the sum of diagonal entries) we have traceBC = traceCB, where B is $s \times t$ and C is $t \times s$ (for a proof, consider $\sum_{i=1}^{s}\sum_{j=1}^{t} B_{i,j}C_{j,i} = \sum_{j=1}^{t}\sum_{i=1}^{s} C_{j,i}B_{i,j}$.) This gives trace$A^\ell$ = trace $\Lambda^\ell P^{-1}P$ = traceΛ^ℓ, so $\ell = 1, 2$ give us the equations $m_0\lambda_0 + m_1\lambda_1 + m_2\lambda_2 =$ trace A, and $m_0\lambda_0^2 + m_1\lambda_1^2 + m_2\lambda_2^2 =$ trace A^2. Now, A has all diagonal entries equal to 0, so trace $A = 0$, and from the proof of Theorem 24.2 above, A^2 has diagonal entries all equal to k, so trace$A^2 = kn$. Putting these together, we have the linear system

$$\begin{aligned} m_0 + m_1 + m_2 &= n \\ \lambda_0 m_0 + \lambda_1 m_1 + \lambda_2 m_2 &= 0 \\ \lambda_0^2 m_0 + \lambda_1^2 m_1 + \lambda_2^2 m_2 &= nk \end{aligned}$$

for m_0, m_1, m_2. But this system has determinant $(\lambda_2 - \lambda_1)(\lambda_2 - \lambda_0)(\lambda_1 - \lambda_0) \neq 0$ (called the *Vandermonde* determinant), and so has a unique solution, given by

$$m_0 = 1, \qquad m_1 = \tfrac{1}{2}k\left(k + \frac{k-2}{D}\right), \qquad m_2 = \tfrac{1}{2}k\left(k - \frac{k-2}{D}\right).$$

(To check this, you will need the facts that $\lambda_0 = k$, $n = k^2 + 1$, $\lambda_1\lambda_2 = 1 - k$, $\lambda_1 + \lambda_2 = -1$, and $\lambda_i^2 = -\lambda_i + k - 1$, for $i = 1, 2$.)

The proof concludes by using the fact that m_1, m_2 are nonnegative integers. There are two cases. **Case 1:** If $4k-3$ is not a perfect square, then D is irrational, which implies that m_1 (and m_2) are irrational unless $k(k-2) = 0$. But $k \geq 2$, so in this case we have $k = 2$ only. **Case 2:** If $4k-3$ is a perfect square, then we have $k = \frac{1}{4}(D^2+3)$, which gives

$$m_1 = \tfrac{1}{8}(D^2+3)\left(\tfrac{1}{4}(D^2+3) + \frac{\frac{1}{4}(D^2-5)}{D}\right),$$

and when we multiply this equation by $32D$ and rearrange slightly, we obtain

$$32Dm_1 - D(D^2+3)^2 - D^4 + 2D^2 = -15.$$

But D and m_1 are integers, and every term on the left in this expression is divisible by D, so we conclude that D divides 15. Thus possible values of D are $1, 3, 5, 15$, which correspond to the values $1, 3, 7, 57$ for k. We have ruled out $k = 1$ (such a graph has no cycles, so has girth ∞), so by combining the two cases, we have possible values for k as $2, 3, 7, 57$. For $k = 2, 3, 7$ such a graph is known to exist, and is unique up to automorphism – for $k = 2$ this is the 5-cycle, for $k = 3$ this is the Petersen graph (seen when we introduced graph isomorphisms), and for $k = 7$ this is called the Hoffman-Singleton graph. As stated above, it is still unknown whether such a graph exists for $k = 57$ (it would have 3250 vertices and 92625 edges).

25 Lecture of March 17

Now we change topics, and consider planarity in graphs. A graph is called *planar* if it can be drawn in the plane so that edges do not intersect (except at vertices). Such a drawing is called a *planar embedding* of the graph. For example, the 3-cube is planar, with a planar embedding given in Figure 12. A planar embedding partitions the plane into connected

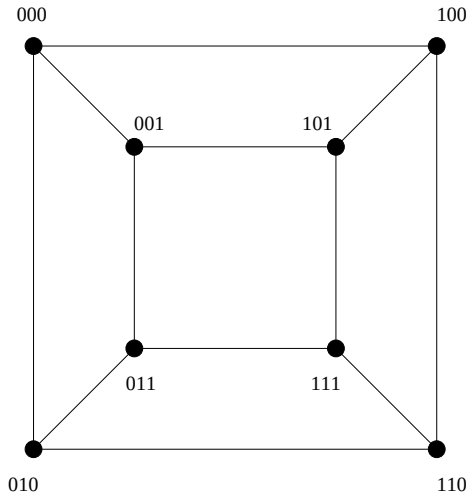

Figure 12: A planar embedding of the 3-cube.

regions called *faces*; one of these regions, called the *outer* face, is unbounded. For example, the planar embedding G given in Figure 13 has 3 faces, identified as f_1, f_2, f_3 in the diagram. The outer face is f_3. The vertices and edges incident with a face are called the *boundary* of the face. In a connected graph, if we traverse the boundary of a face in a fixed direction (e.g., clockwise), starting at any vertex incident with that face, then we encounter an alternating sequence of vertices and edges, which is a closed walk in the graph. This is called a *boundary walk* of the face. The number of edges in a boundary walk of a face is called the *degree* of the face. Note that a bridge of a planar embedding is incident with one face, and is contained in a boundary walk of that face twice, once for each side. Thus a bridge contributes 2 to the degree of the face with which it is incident. Otherwise, an edge that is not a bridge (and therefore must appear on a cycle) is incident with 2 different faces, and is contained in

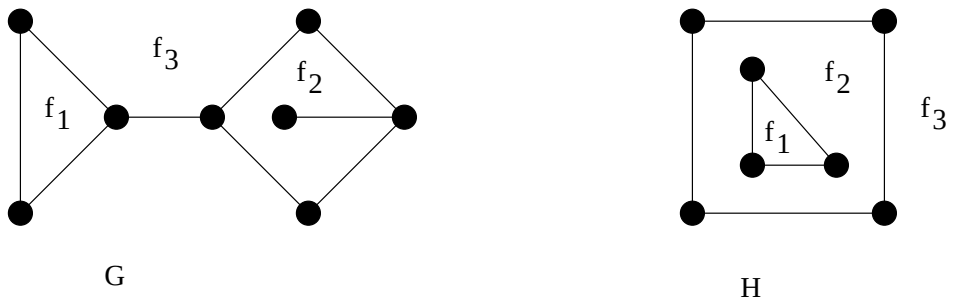

Figure 13: Two planar embeddings with 3 faces.

the boundary walks of both faces. For example, in Figure 13, for G we have $\deg(f_1) = 3$, $\deg(f_2) = 6$, $\deg(f_3) = 9$. In a planar embedding that is not connected, the boundary of some faces are the disjoint union of several boundary walks, and the degree of such a face is the number of edges contained in all boundary walks in the boundary. For example, in Figure 13, for H we have $\deg(f_1) = 3$ and $\deg(f_3) = 4$, and the boundary of f_2 consists of two disjoint boundary walks, giving $\deg(f_2) = 3 + 4 = 7$. We use s generically for the number of faces in a planar embedding. For face degrees, we have the following analogue of the handshake theorem.

Theorem 25.1 *In a planar embedding with faces $f_1, \ldots, f_s$, we have*

$$\sum_{i=1}^{s} \deg(f_i) = 2q.$$

PROOF. Each bridge contributes 2 to the degree of a single face, and each edge on a cycle contributes 1 to the degree of two different faces, and the result follows.

The next result gives *Euler's Formula*, which proves that all planar embeddings of a given planar graph have the same number of faces.

Theorem 25.2 *For a planar embedding with p vertices, q edges, s faces, and c components, we have*

$$p - q + s = 1 + c.$$

PROOF. For each fixed $p \geq 1$, we prove this by induction on $q \geq 0$, in the form $p-q+s-c = 1$. The base case is $q = 0$, for which we have $s = 1$ and $c = p$, and thus

$$p - q + s - c = p - 0 + 1 - p = 1,$$

so the result is true in the base case.

The induction hypothesis is to assume that the result holds for all planar embeddings with $q = k$ edges, for some $k \geq 0$. Now, consider an arbitrary planar embedding, P, with $q = k+1$ edges. Let p, s, c denote the numbers of vertices, faces, components in P, respectively. Let e be any edge of P, and let $P' = P - e$. Suppose that p', q', s', c' denote the numbers of

vertices, edges, faces, components in P', respectively. There are two cases. If e is a bridge of P, then we have $p = p'$, $q = q' + 1$, $s = s'$, and $c = c' - 1$, giving

$$p - q + s - c = p' - (q' + 1) + s' - (c' - 1) = p' - q' + s' - c',$$

and the result holds for this case by the induction hypothesis. Otherwise, if e is not a bridge, then we have $p = p'$, $q = q' + 1$, $s = s' + 1$, and $c = c'$, giving

$$p - q + s - c = p' - (q' + 1) + (s' + 1) - c' = p' - q' + s' - c',$$

and the result holds for this case also, again by the induction hypothesis. Therefore, the result is true by mathematical induction.

26 Lecture of March 19

Our first application of Euler's Formula is to put an upper bound on the number of edges in a planar embedding.

Theorem 26.1 *If each face in a planar embedding with p vertices and q edges has degree at least $d^\star$, then*

$$(d^\star - 2)q \leq d^\star(p - 2).$$

PROOF. Summing the face degrees, we get $2q \geq sd^\star$. But from Euler's Formula, we have $p - q + s \geq 2$, or $s \geq 2 - p + q$, so

$$2q \geq d^\star s \geq d^\star(2 - p + q).$$

Now, rearrange the inequality $2q \geq d^\star(2 - p + q)$, to get the result.

Next we state a lemma about face boundaries.

Lemma 26.2 *In a planar embedding with at least one cycle, the boundary of every face contains the edges of a cycle.*

These results allow us to prove that a number of graphs are nonplanar.

Theorem 26.3 *K_5, $K_{3,3}$, the 4-cube and the Petersen graph are all nonplanar.*

PROOF. For K_5, the complete graph on 5 vertices, we have $p = 5$, $q = 10$, and there are (many) cycles in K_5. Since every cycle must contain at least 3 edges then, from the Lemma above, if K_5 has a planar embedding, every face must have degree at least 3. Then, Theorem 26.1 must hold with $p = 5$, $q = 10$, and $d^\star = 3$. But we have

$$(d^\star - 2)q = q = 10 > 9 = 3(p - 2) = d^\star(p - 2),$$

so the inequality of Theorem 26.1 does not hold, and we conclude (by the contrapositive) that K_5 has no planar embedding. This means that K_5 is not planar.

For $K_{3,3}$, we have $p = 6$, $q = 9$, and there are (many) cycles. Also, $K_{3,3}$ is bipartite, so it has no odd cycles, and thus no 3-cycles. Therefore every cycle in $K_{3,3}$ must contain at least $d^\star = 4$ edges, so Theorem 26.1 must hold with the given values of $p, q, d^\star$. But we have

$$(d^\star - 2)q = 2q = 18 > 16 = 4(p-2) = d^\star(p-2),$$

so the inequality of Theorem 26.1 does not hold, and we conclude that $K_{3,3}$ is not planar.

For the 4-cube, we have $p = 2^4 = 16$, $q = 4 \cdot 2^3 = 32$, and there are (many) cycles. Also, all n-cubes are bipartite (every edge joins a vertex with an even number of 1's to a vertex with an odd number of 1's), so we can use $d^* = 4$, and Theorem 26.1 must hold with the given values of $p, q, d^\star$. But we have

$$(d^\star - 2)q = 2q = 64 > 56 = 4(p-2) = d^\star(p-2),$$

so the inequality of Theorem 26.1 does not hold, and we conclude that $K_{3,3}$ is not planar. Note that this inequality also allows us to deduce that there is no set of 3 edges that can be removed from the 4-cube to give a planar subgraph (since we would have $58 > 56$ for such a graph).

For the Petersen graph, we have $p = 10$, $q = 15$, and there are (many) cycles. Also, there are no 3-cycles, nor 4-cycles (it has girth 5). Thus every cycle must contain at least $d^\star = 5$ edges, so Theorem 26.1 must hold with the given values of $p, q, d^\star$. But we have

$$(d^\star - 2)q = 3q = 45 > 40 = 5(p-2) = d^\star(p-2),$$

so the inequality of Theorem 26.1 does not hold, and we conclude that the Petersen graph is not planar.

It is important to note that Theorem 26.1 is not if and only if. To prove that the converse doesn't always hold, consider the graph in Figure 14. This graph is clearly nonplanar, since

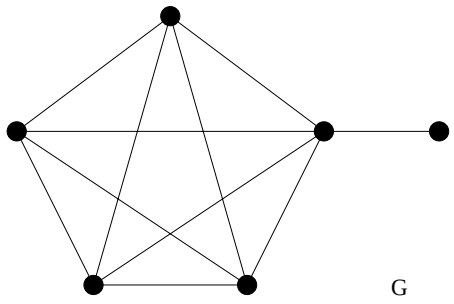

Figure 14: A nonplanar graph G.

it has K_5 as a subgraph. Also, this graph has 3-cycles, so we have $p = 6$, $q = 11$, $d^\star = 3$, and thus

$$(d^\star - 2)q = q = 11 \leq 12 = 3(p-2) = d^\star(p-2).$$

This shows that the inequality of Theorem 26.1 can hold for a nonplanar graph.

We conclude this section on nonplanarity with a useful inequality.

Theorem 26.4 *If a graph with $p \geq 3$ vertices and q edges has a planar embedding, then $q \leq 3p - 6$.*

PROOF. There are two cases. If the graph has a cycle, then the cycle must contain at least 3 edges, and so we can apply Theorem 26.1 with $d^\star = 3$, giving the result in this case. Otherwise, if the graph has no cycle, then the number of edges is at most that of a tree (which has $q = p - 1$ edges), so we have

$$q \leq p - 1 \leq p - 1 + (2p - 5) = 3p - 6,$$

since $2p - 5 \geq 0$ for $p \geq 3$, giving the result in this case also.

An *edge-subdivision* of a graph G is obtained by replacing each edge of G by a path with a nonnegative number of internal vertices, where each edge of G is treated independently. For example, in Figure 15, we give a graph G, together with one of its edge-subdivisions H. Now,

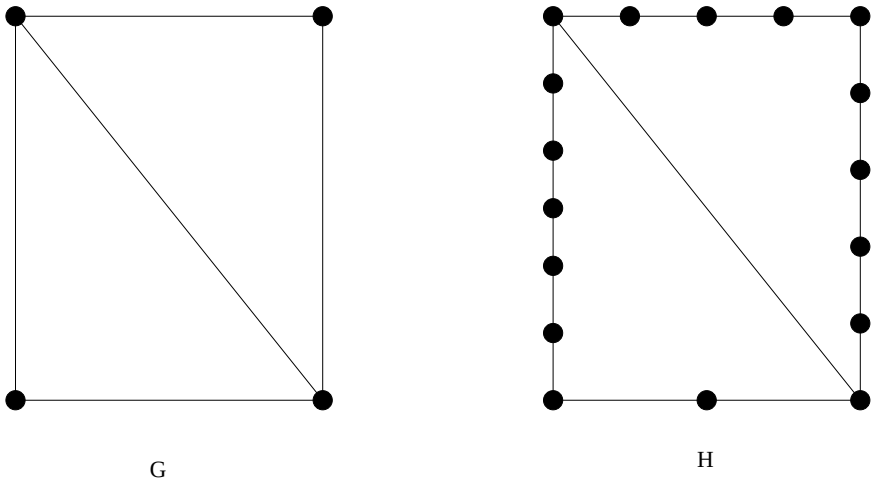

Figure 15: A graph G and an edge-subdivision H.

clearly, an edge-subdivision of a graph is planar if and only if the graph is planar. Therefore, if a graph has a subgraph that is an edge-subdivision of a nonplanar graph, then the graph is nonplanar. The following result, which we do not prove, is *Kuratowski's Theorem.* In view of the above statement, the surprising part of this result is the "only if".

Theorem 26.5 *A graph is nonplanar if and only if it has a subgraph that is an edge-subdivision of K_5 or $K_{3,3}$.*

To illustrate this result, consider the Petersen graph, which we have previously proved to be nonplanar. Kuratowski's Theorem implies that the Petersen graph must have a subgraph that is an edge-subdivision of K_5 or $K_{3,3}$. In Figure 16, we give the Petersen graph. If the middle vertex (drawn as an unfilled circle) and its three incident edges (drawn as dotted lines) are removed, then the result is a subgraph that is an edge subdivision of $K_{3,3}$. (To see this, consider $K_{3,3}$ as a 6-cycle with three pairwise crossing chords.)

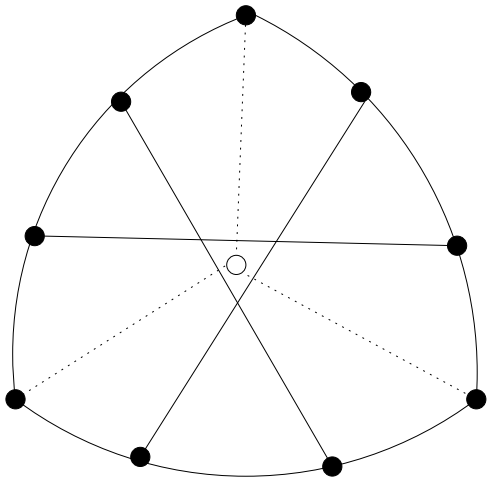

Figure 16: The Petersen graph and a subgraph.

27 Lecture of March 24

We can also consider embeddings of graphs, without edges crossing, in surfaces other than the sphere (equivalent to the plane for embeddings of finite graphs). For example, the *torus* is a sphere with a handle. The complete graph K_5 is not embeddable on the sphere, but can be embedded on the torus, as given in Figure 17. The meaning of the matching arrows on the top and bottom is that these segments are identified (like rolling up the diagram

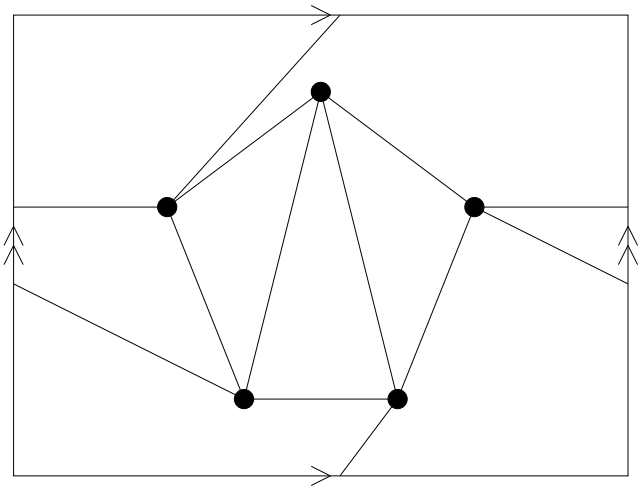

Figure 17: An embedding of K_5 on the torus.

horizontally, into a cylinder), and the matching arrows on the left and right is that these segments also are identified (to turn the cylinder into a "donut" shape).

In fact we can embed K_7 on the torus, as well as $K_{4,4}$ and the Petersen graph. Suitable embeddings are given in Figure 18. (Note that on of the vertices of K_7 is represented by four copies, one in each corner of the diagram.)

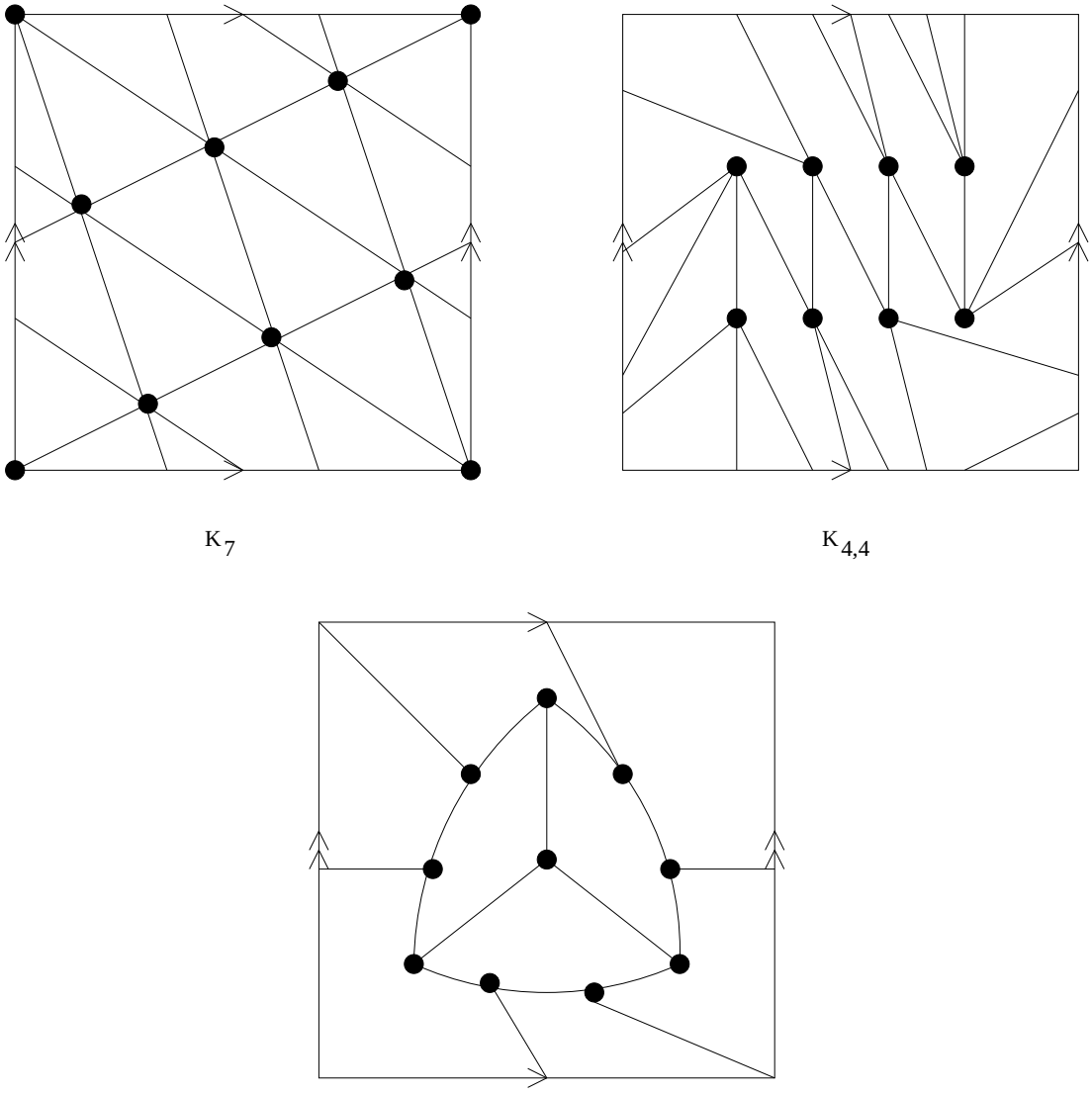

Figure 18: Embeddings on the torus for K_7, $K_{4,4}$, and the Petersen graph.

The embedding in Figure 17 has faces, just as for embeddings on the sphere, with boundary walks that can be obtained in the same way. For example, there are 5 faces for the embedding in Figure 17, four of degree 3, and one of degree 8. Of course, the sum of face degrees is equal to $2q$, for example, here we have $3 + 3 + 3 + 3 + 8 = 20 = 2 \cdot 10$. The embedding in Figure 17 is called a 2-*cell embedding*, meaning that all the faces are homeomorphic to discs. In a similar way, we can consider embeddings on T_g, the surface obtained by adding g handles to the sphere. We say that T_g has *genus* g, and a 2-cell embedding on T_g partitions the surface into s faces. For any nonnegative integer g, there is an extension of Euler's formula, giving

$$p - q + s = 2 - 2g$$

for embeddings of connected graphs. For example, for the embedding of K_5 in Figure 17, we have $p = 5$, $q = 10$, and $s = 5$, so $p - q + s = 0 = 2 - 2g$, with $g = 1$.

There is also a Kuratowski-style theorem for T_g: it is known that there is a finite list of graphs such that a graph can be embedded on T_g if and only if it does not have a subgraph that is an edge-subdivision of some graph on the list. However, even for $g = 1$, such a list is not known precisely, but it is known is that it must contain at least 1000 graphs. The surfaces T_g are called *orientable* surfaces. We can also consider 2-cell embeddings on nonorientable surfaces, like the projective plane, or the Klein bottle. There is a Kuratowski-style theorem for these, too; again there is always a finite list. The simplest case is for the projective plane, where there is a known list of 103 graphs.

Now, we showed in Figure 18 that the complete graph K_7 can be embedded on the torus. We now consider whether K_n can be embedded on the torus for any larger n, and then which complete graphs can be embedded on T_g for each $g \geq 1$. Let

$$M_g = \frac{7 + \sqrt{1 + 48g}}{2}, \quad N_g = \lfloor M_g \rfloor.$$

This gives, for example, $N_0 = 4$, $N_1 = 7$, $N_2 = 8$. In 1974, Ringel and Youngs proved that K_{N_g} has a 2-cell embedding in T_g, for $g \geq 0$. On the other hand, we have the following result.

Theorem 27.1 *If a connected graph G has a 2-cell embedding on T_g, then G must have a vertex of degree less than or equal to $N_g - 1$, for $g \geq 1$.*

28 Lecture of March 26

PROOF. From Euler's formula, we have $p - q + s = 2 - 2g$, and from the sum of face degrees, we have $2q \geq 3s$. Also, M_g is the solution to a quadratic equation, and we have $M_g(M_g - 7) = 12g - 12$. Now, to arrive at a contradiction, suppose otherwise, that every vertex of G has degree greater than $N_g - 1$, or equivalently, greater than $M_g - 1$. Then, from the sum of vertex degrees, we have $2q > (M_g - 1)p$. Also, we have $p > M_g$. Then, putting these results together, we obtain

$$2 - 2g = p - q + s \leq p - q + \tfrac{2}{3}q = p - \tfrac{1}{3}q < p - \frac{M_g - 1}{6}p = \frac{7 - M_g}{6}p \leq \frac{7 - M_g}{6}M_g = 2 - 2g,$$

giving the contradiction (for the last inequality, note that $M_g \geq 7$, since $g \geq 1$). The result follows.

Comparing the above result with the embeddings of Ringel and Youngs, we see that their construction is best possible: if $g \geq 1$, K_n cannot be embedded on T_g for any $n > N_g$, since *every* vertex of K_n has degree $n - 1 > N_g - 1$.

Moreover, for $g = 0$, we have already proved that K_5 is not planar, so Ringel and Youngs' construction is also best possible for $g = 0$.

We now give the analogue of Theorem 27.1 for the case $g = 0$.

Theorem 28.1 *In every planar graph, there is a vertex of degree at most* 5.

PROOF. Suppose otherwise, that we have a planar graph G in which every vertex has degree at least 6. Then, considering the sum of vertex degrees, we have $2q \geq 6p$, or $q \geq 3p$. But this contradicts Theorem 26.4, and the result follows.

Now, in order to discuss the four colour theorem, we consider colouring the vertices of a graph. A *k-colouring* of a graph G is a function $f : V(G) \to \{1, \ldots, k\}$, with the property that $f(u) \neq f(v)$ if u and v are adjacent. The elements of $\{1, \ldots, k\}$ are referred to as *colours*. If a graph has a k-colouring f, then we say that it is *k-colourable*, and refer to $f(v)$ as the colour of vertex v. For example, a 3-colouring of a graph H is given in Figure 19, where the number beside each vertex indicates its colour. Of course, a vertex with p vertices

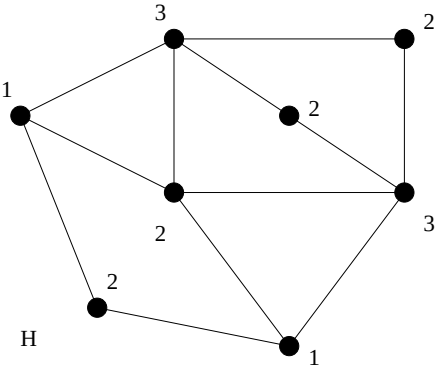

Figure 19: A 3-colouring of a graph H.

is always p-colourable, by assigning a different colour to every vertex. The complete graph K_p is not k-colourable for any $k < p$, since every pair of vertices is adjacent, and so must have a different colour. A graph is bipartite if and only if it is 2-colourable, with colour 1 for the vertices in A, and colour 2 for vertices in B.

Theorem 28.2 *Every connected graph with a* 2*-cell embedding on T_g is N_g-colourable for $g \geq 1$, and is* 6*-colourable for $g = 0$.*

PROOF. The proof is by strong induction on $g \geq 0, p \geq 1$. For base cases, using the fact that *every* graph on n vertices is n-colourable (use a different colour for each vertex), we note that the result is true for $g = 0$, $p \leq 6$, and $g \geq 1$, $p \leq N_g$.

For the induction hypothesis, assume that the result is true for $g < m$, $p \geq 1$ and for $g = m$, $p \leq k$, where $m \geq 0$, and $k \geq 6$ if $m = 0$, or $k \geq N_m$ if $m \geq 1$. Now consider an arbitrary connected embedding G on T_m with $k + 1$ vertices. Then Theorem 28.1 and Theorem 27.1 imply that G has a vertex v of degree at most 5 if $m = 0$, or at most $N_g - 1$ if $m \geq 1$. Then if we remove v and incident edges from G, we obtain a positive number of components, each of which has at most k vertices (since the total number of vertices among all components is k), and each of which can be embedded on T_n for some $n \leq m$ (the latter is a topological property whose proof we don't give here). Thus, each component can be N_m-coloured by the induction hypothesis (and the fact that $6 \leq N_1 \leq N_2 \leq \ldots$). But, the $N_m - 1$ (or 5 if $m = 0$) neighbours of v cannot receive, together, more than $N_m - 1$ colours (or 5 if $m = 0$), leaving a different colour for v in all cases, proving the result for G. The result follows by mathematical induction.

The above colouring theorem is best possible for $g \geq 1$, as an immediate consequence of Ringel and Youngs' construction, since it is not possible to use fewer than N_g colours to colour K_{N_g}. However, for $g = 0$, the existence of an embedding of K_4 on the plane, but not of K_5 means that the 6-colouring result for $g = 0$ might not be best possible, and perhaps 6 can be replaced by 5 or 4 for the best possible result. In fact, the correct answer is 4, and this is the statement of the celebrated Four Colour Theorem.

29 Lecture of March 28

The statement of the above Theorem as applying to *vertex* colouring may seem strange, since most popular accounts of the Four Colour Theorem refer to colouring the *faces* of a planar embedding. However, we now show that colouring the vertices of a planar embedding so that adjacent vertices are assigned different colours, is equivalent to colouring the faces of a planar embedding so that faces with a common edge in their boundaries are assigned different colours. To do this, we define the planar embedding G^*, called the *dual* of the planar embedding G. There is one vertex of G^* for each face of G, and we locate such a vertex in the interior of the corresponding face. Two vertices of G^* are adjacent when the corresponding faces of G are incident with a common edge of G. In the embedding G^*, the curve representing such an edge crosses the curve representing the corresponding edge in G. For example, in Figure 20 we give a planar embedding G, with vertices as circles, and edges as solid curves, and the planar embedding G^*, with vertices as boxes, and lines as dashed curves. Note that G has $p = 7$ vertices, $q = 10$ edges, and $s = 5$ faces, where G^* has $p^* = 5$ vertices, $q^* = 10$ edges, and $s^* = 7$ faces. Also, here we have $(G^*)^* = G$ (actually, there is some choice about where to embed the edges of G^* that are embedded in the outer face of G, and vice-versa; however, there is no choice if the embedding is in the surface of the sphere). For any planar embedding G and its dual G^*, these relationships always hold; in general, we have $p = s^*$, $q = q^*$, $s = p^*$, and $(G^*)^* = G$. In particular, a vertex of degree i in G becomes a face of degree i in G^*, and a face of degree i becomes a vertex of degree i in G^*. Moreover, two vertices are adjacent in G exactly when the corresponding faces in G^* have a common edge in their boundaries. But this means that a vertex colouring of G becomes a face colouring of G^*, which explains why the Four Colour Theorem is often stated for colouring faces.

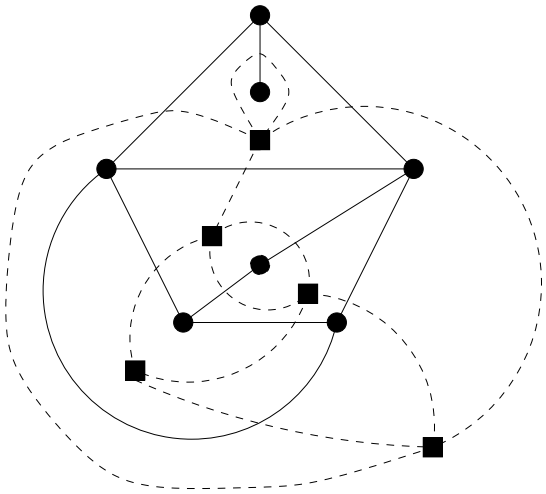

Figure 20: A planar embedding and its dual.

However, there are some additional aspects that arise in considering the dual. Note in Figure 20 that the bridge of G gives rise to a loop in G^* (an edge that joins a vertex to itself). Also, the vertex of degree 2 in G gives rise to a pair of edges between the same pair of vertices in G^* (this is called a multi-edge). Thus in taking the dual of a planar embedding, we can create a graph that has loops and multiple edges (and is thus is no longer a *simple* graph).

Now suppose that we have a planar embedding whose faces we wish to colour with k colours, so that faces that meet at an edge (because that edge is incident with the faces) have different colours. Suppose that this planar embedding has no vertices of degree 2 (they can always be removed without changing the underlying regions that are to be coloured), and no bridges (if there were a bridge, then the face of the planar embedding that is twice incident with that bridge could not be coloured without violating the "different colour" restriction). Then the dual of the planar embedding is a planar embedding without loops or multiple edges, and a k-colouring of the vertices of the dual will correspond exactly to a k-colouring of the faces of the original planar embedding.

We have proved above that every planar graph is k-colourable for $k = 6$. We now prove the case $k = 5$. In the proof, we use a graph construction known as *edge-contraction*. If $\{x, y\} \in E(G)$, and $z \notin V(G)$, then define H to be a graph with $V(H) = V(G) \cup \{z\} \setminus \{x, y\}$. The edges of G consist of all edges $\{a, b\}$ of G with $a, b \in V(G) \setminus \{x, y\}$, together with all $\{z, c\}$ such that $\{x, c\}$ or $\{y, c\}$ (or both) is an edge of G, with $c \in V(G) \setminus \{x, y\}$. We say that H has been obtained from G by edge-contraction of $\{x, y\}$. Note that if G is planar, then every edge-contraction of G is planar, a fact that we shall not prove here. This is often useful in the contrapositive: for example, we can obtain K_5 from the Petersen graph by 5 edge-contractions (using the "spokes" that join the outer pentagon to the inner pentagonal star), and thus conclude, from the nonplanarity of K_5, that the Petersen graph is nonplanar.

Be careful with edge-contraction, however. It is *not* the case that every edge-contraction of a nonplanar graph is nonplanar; for example, applying edge-contraction to any single edge of $K_{3,3}$ (which is of course nonplanar), we obtain a planar graph (it has 5 vertices and 8

edges so it can't possibly contain an edge-subdivision of K_5 or $K_{3,3}$ as a subgraph).

Theorem 29.1 *Every planar graph is* 5*-colourable.*

PROOF. Again, the proof is by induction on p, the number of vertices. As base cases, the result holds for planar graphs with $p \leq 5$, since *every* graph on at most 5 vertices is 5-colourable. For the (strong) induction hypothesis, assume that every planar graph on $p \leq k$ vertices is 5-colourable, where $k \geq 5$. Consider a planar graph G on $p = k+1$ vertices. Then Theorem 28.1 implies that G has a vertex of degree at most 5. There are two cases.

If G has a vertex v of degree at most 4, then $G - v$ has k vertices, and we can apply the induction hypothesis to obtain a 5-colouring of $G - v$. In G, v has at most 4 neighbours, so in the 5-colouring of $G - v$, they are coloured, together, with at most 4 colours. This means that there is a different colour available for v, to give a 5-colouring for G.

Otherwise, G has a vertex v of degree 5. Now, there must exist a pair a, b of neighbours of v that are not adjacent (else, the graph would contain K_5 as a subgraph, whose vertices are the neighbours of v; but this is a contradiction, since the planar graph G cannot contain any nonplanar subgraph). Now, let G' be the graph obtained from G by edge-contraction of $\{a, v\}$ and $\{b, v\}$, where the new vertex of G' is v'. Note that G' is planar, since G is planar. Also, G' has $k - 1$ vertices, so we can apply the induction hypothesis to obtain a 5-colouring of G'. Now, we produce a 5-colouring of G: for all vertices of G except a, b, v, use the colour assigned in the 5-colouring of G'; for vertices a, b, which are not adjacent, assign the same colour, namely the colour assigned to v' in the 5-colouring of G; now v has exactly 5 neighbours in G, and two of them have the same colour, so, together, they are coloured with at most 4 colours. This means that there is a different colour available for v, and we have successfully 5-coloured G.

Therefore, in both cases, G is 5-colourable, and the result follows by mathematical induction.

The statement of the celebrated Four Colour Theorem is "Every planar graph is 4-colourable." The published proofs use induction on the number of vertices, and consist of a case analysis based on a vertex v of small degree, just as in the above proof of the Five Colour Theorem. However, there are many cases (in the first proof, over 2000; in the most recent proof, about 600), and the constructions needed to produce a 4-colouring are often very complicated (the original proof is perhaps especially well known because it incorporated computer-aided manipulations to prove the cases).

30 Lecture of March 31

Now, we turn to matchings in a graph. A *matching* M in a graph G is a set of edges of G with the property that no vertex of G is incident with more than one edge of M. For example, $M_1 = \{\{1, 4\}, \{2, 6\}\}$ and $M_2 = \{\{1, 2\}, \{3, 4\}, \{5, 6\}\}$ are matchings of the graph G given in Figure 21. The size of a matching is the number of edges in the matching, so we write $|M_1| = 2$ and $|M_2| = 3$. The set of vertices incident with the edges of a matching M are said to be *saturated* by the matching. For example, the set of vertices saturated by M_1 is $\{1, 2, 4, 6\}$, and the set of vertices saturated by M_2 is $V(G)$. If G has p vertices, then no

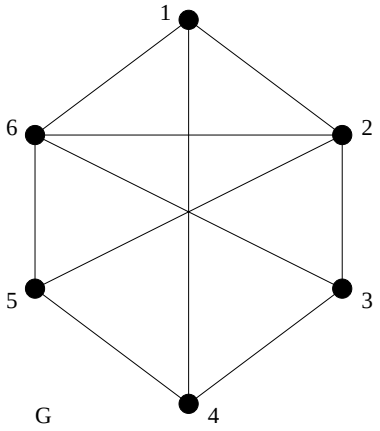

Figure 21: A graph G.

matching of G can saturate more than p vertices, so we have immediately the upper bound $|M| \leq p/2$ for the size of a matching M of G. If a matching M saturates all the vertices of G (so $|M| = p/2$) then we say that M is a *perfect* matching of G.

A *maximum* matching of G is a matching of maximum possible size. Note that M_1 above is a *maximal* matching of G, since it is not a proper subset of another matching (the two vertices, $3, 5$ that are not saturated by M_1 are not adjacent), but it is not a maximum matching, because M_2 has bigger size than M_1.

A *cover* C in a graph G is a set of vertices of G with the property that no edge of G is incident with less than one vertex of C. For example, $C_1 = \{1, 2, 3, 4\}$ and $C_2 = \{2, 4, 6\}$ are covers of the graph G given in Figure 21. The size of a cover is the number of vertices in the cover. The reason that we introduce covers in a discussion of matchings is the following result, which relates the size of matchings and covers.

Theorem 30.1 *For every matching M and cover C of a graph G, we have $|M| \leq |C|$.*

PROOF. Consider an arbitrary matching M of G, and suppose that M contains edges $\{u_1, v_1\}, \ldots \{u_k, v_k\}$. Then every cover C of G must contain u_i or v_i or both, for each $i = 1, \ldots, k$. But the u_i and v_i are all different, so it follows that $|C| \geq k = |M|$, and the result follows.

As an immediate corollary, we obtain that, if M is a matching and C is a cover with $|M| = |C|$, then M is a maximum matching, and C is a minimum cover (i.e., a matching of maximum size in the graph, and a cover of minimum size in the graph). Is it always possible to find a matching and a cover of equal size in every graph G ? The answer is "No" – in the complete graph K_3, every pair of edges is incident with a common vertex, so the maximum size of a matching is 1; however, each vertex is incident with only two of the three edges, so the minimum size of a cover is 2. Similarly, for any cycle of odd length $2m+1$, the maximum size of a matching is m, and the minimum size of a cover is $m + 1$. However, odd cycles are the only thing that prevents the above equality, as given by the following result, called *König's* Theorem.

Theorem 30.2 *In a bipartite graph, the size of the maximum matching is equal to the size of the minimum cover.*

Before we prove König's Theorem, we note that the converse is false: for example, the graph G in Figure 21 is not bipartite (it has a number of 3-cycles), but it has a matching M_2 and a cover C_2 of the same size. We also need to introduce more terminology. In a graph G, with matching M, an *alternating* path is a path in which the edges alternate between belonging to M, and not belonging to M. A path with no edges is permitted. An *augmenting* path is an alternating path of positive length that starts and ends at an unsaturated vertex. For example, for the graph G given in Figure 21, with matching given by $M_1 = \{\{1,4\},\{2,6\}\}$, some of the alternating paths are given by 1, 12, 14, 126, 2614, 54126, 541623, and of these, 541623 is an augmenting path.

Theorem 30.3 *If G has an augmenting path for a matching M, then M is not a maximum matching of G.*

PROOF. Since an augmenting path begins and ends at unsaturated vertices, and has positive length, the first and last steps of the path do not belong to M (since an unsaturated vertex of G is incident with no edge of M). Therefore, an augmenting path has odd length, and contains k edges of M, and $k+1$ edges not in M, for some $k \geq 0$. Now, construct a new matching M' as follows: M' consists of all edges of M not in the augmenting path, together with all edges of the augmenting path not in M. Thus, to obtain M' from M, we replace k edges by $k+1$ edges, and so we have $|M'| = |M|+1$. The fact that M' is a matching follows immediately, since the vertices on the augmenting path are incident with exactly one edge of M', and the remaining vertices are incident with at most one edge of M', since M is a matching. We have constructed a larger matching than M, giving the result.

In order to complete the terminology for a proof of König's Theorem, consider a bipartite graph G, with bipartition A, B, together with a matching M of G. The matching M doesn't need to be a maximum matching of G, and could even have no edges. Let X_0 be the set of unsaturated vertices in A. Let X be the set of vertices in A that are reachable by an alternating path from a vertex in X_0, and let Y be the set of vertices in B that are reachable by an alternating path from a vertex in X_0.

31 Lecture of April 2

Note that $X_0 \subseteq X$, since alternating paths with no edges are permitted. Also, let U be the set of unsaturated vertices in Y. For example, for the graph G in Figure 22, with sets A and B as drawn, and matching M consisting of the thick edges, we have $X_0 = \{3,5\}$. Then (say, by creating search trees rooted at vertices 3 and 5), we determine $X = \{1,3,4,5\}$ and $Y = \{a,b,e,f\}$, and hence $U = \{b,e\}$. In order to clearly illustrate the results that come next, we redraw the graph G of Figure 22, with the sets X and Y on the left, in Figure 23. In general, we have a number of results for the sets X, Y, U.

Theorem 31.1 *1. There is no edge of G joining a vertex of X to a vertex of $B \setminus Y$.*

2. The set $C = (A \setminus X) \cup Y$ is a cover of G.

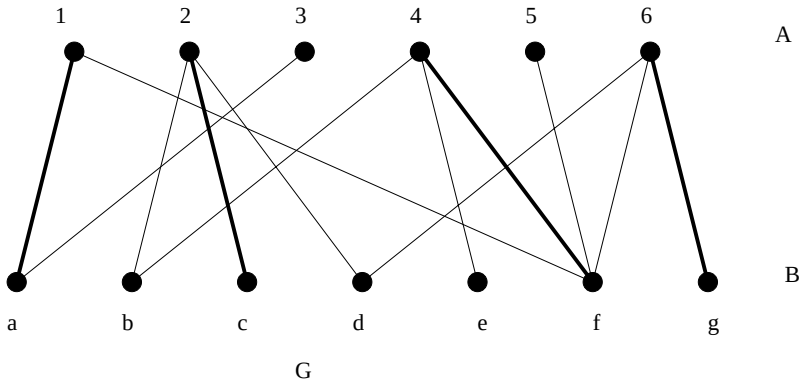

Figure 22: A bipartite graph G with a matching M.

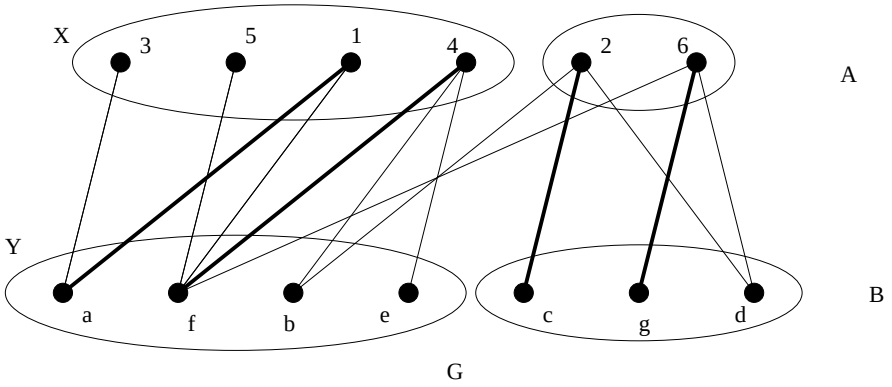

Figure 23: A redrawing of bipartite graph G with matching M.

3. *There is no edge of M joining a vertex of Y to a vertex in $A \setminus X$.*

4. $|M| = |C| - |U|$.

5. *If $|U| > 0$, then M is not a maximum matching of G; if $|U| = 0$, then M is a maximum matching of G.*

PROOF.

1. Suppose otherwise, to arrive at a contradiction. If there is a nonmatching edge of G joining a vertex x of X to a vertex v of $B \setminus Y$, then there is an alternating path from a vertex in X_0 to x. But this alternating path, together with the nonmatching edge, would give an alternating path from a vertex in X_0 to v, implying $v \in Y$, which is a contradiction. Otherwise, if there is an edge of M joining a vertex x of X to a vertex v of $B \setminus Y$, then there is an alternating path from a vertex in X_0 to x. But the last edge in this path must be an edge of M joining a vertex y of Y to x, so x is incident with two edges of M. This is a contradiction, since no vertex can be incident with more than at one edge in a matching.

2. This follows immediately from part 1 of this result, together with the fact that G is bipartite.

3. Suppose otherwise, to arrive at a contradiction. If there is an edge of M joining a vertex y of Y to a vertex v of $A \setminus X$, then there is an alternating path from a vertex in X_0 to y. But this alternating path, together with the edge of M, would give an alternating path from a vertex in X_0 to v, implying $v \in X$, which is a contradiction.

4. From parts 1 and 3 of this result, if follows that the edges of M are of two types. Type 1 join a vertex of X to a vertex of Y, and type 2 join a vertex of $A \setminus X$ to a vertex of $B \setminus Y$. The number of edges of type 1 is $|Y| - |U|$. The number of edges of type 2 is $|A \setminus X|$, and together these give

$$|M| = |Y| - |U| + |A \setminus X| = |C| - |U|,$$

   giving the result.

5. If $u \in U$, then there is an alternating path from a vertex in X_0 to a vertex in U, and thus this is an augmenting path. Theorem 30.3 then implies that M is not a maximum matching. Otherwise, if $|U| = 0$, then part 4 of the result implies that $|M| = |C|$, which means that M is a maximum matching, and C is a minimum cover, by the corollary stated below Theorem 30.1.

For example, for the graph G and matching M in Figure 22, we have $|U| > 0$, and this implies that M is not a maximum matching of G. Now, there is an augmenting path to each vertex in U, for example, $5f4b$ is one of these. If we "switch" the matching and non-matching edges on this path , as in the proof of Theorem 30.3, then we obtain the matching M', indicated by the thick edges in Figure 24. For this matching, we have $X_0 = \{3\}$, and

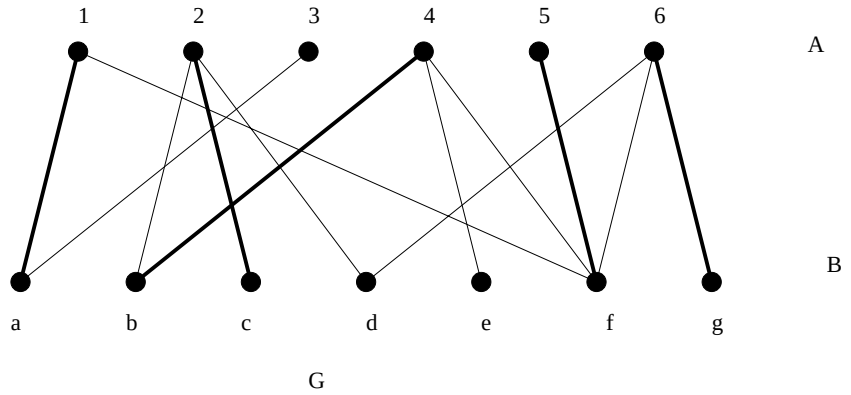

Figure 24: The bipartite graph G with matching M'.

determine $X = \{1, 3, 5\}$, $Y = \{a, f\}$, so $U = \emptyset$. Since $|U| = 0$, we conclude that M' is a maximum matching, and indeed can check that $(A \setminus X) \cup Y = \{2, 4, 6, a, f\}$ is a cover of G, of the same size as M'.

Proof of König's Theorem: For the bipartite graph G, with bipartition A, B, find a matching M (the empty matching will do; otherwise, for a bigger matching, you could consider the edges in any order, and select them for the matching if they join two previously unsaturated vertices). Then construct the sets X_0, X, Y, U and C. There are two cases: if $|U| = 0$, then M is a maximum matching, of the same size as the cover C; otherwise, if $|U| > 0$, then find an augmenting path to a vertex in U, and apply the "switching" method as in the proof of Theorem 30.3, to obtain a matching M' that is larger than M. Repeat this for G and the new matching M'. Since the matching increases in size at every stage, and no matching can have more than $p/2$ edges (where p is the finite number of edges in G), this must terminate finitely in the first case, with a matching and cover of the same size. This finishes the proof of König's Theorem.

As a corollary of König's Theorem, we obtain the following result, which is called *Hall's Theorem*. For a set of vertices D, it uses $N(D)$ for the union of the sets of neighbours of all the vertices in D.

Theorem 31.2 *In a bipartite graph with bipartition A, B, there is a matching that saturates all vertices of A if and only $|N(D)| \geq |D|$, for all $D \subseteq A$.*

32 Lecture of April 4

PROOF. For the "only if": suppose there is a matching that saturates all vertices of A. Then, for any $D \subseteq A$, this matching saturates all vertices of D. But these $|D|$ matching edges are incident with $|D|$ vertices of B, and all of these vertices belong to $N(D)$, so $|N(D)| \geq |D|$.

For the "if": we'll prove the contrapositive. Suppose there is no matching saturating all vertices of A. Then, since the graph is bipartite, the size of a maximum matching is less than $|A|$. Now, consider a minimum cover C, so König's Theorem implies that $|C| < |A|$. Let $\overline{C}$ be the complement of C in the vertex set of the graph. Partition the vertex set into four disjoint sets: $A \cap C$, $A \cap \overline{C}$, $B \cap C$, $B \cap \overline{C}$. The fact that C is a cover implies that there are no edges between $A \cap \overline{C}$ and $B \cap \overline{C}$, which in turn implies that $N(A \cap \overline{C}) \subseteq B \cap C$. Thus, we have

$$|N(A \cap \overline{C})| \leq |B \cap C| = |C| - |A \cap C| < |A| - |A \cap C| = |A \cap \overline{C}|,$$

so for the set $D = A \cap \overline{C} \subseteq A$ we have $|N(D)| < |D|$. The result follows.

We end with a problem about matchings.

Example 32.1 *Prove that if the opposite corners of a standard 8×8 chessboard are removed, then it is impossible to cover the remaining 62 squares with 31 1×2 dominoes.*

SOLUTION. Of course, the usual proof is to analyze the number of black and white squares, and we will do that here, but in a way that makes clear the connection to covers and matchings of a graph. Consider the graph whose vertices are the 64 squares of the chessboard. Two squares are adjacent if, together, they create a 1×2 domino (either vertically or horizontally). Thus this graph has 4 vertices of degree 2 (the corner squares), 24 vertices of

degree 3 (the squares on the outside, not on the corners), and 36 squares of degree 4 (the squares on the inside of the board). A covering of the chessboard with 1×2 dominoes is a perfect matching of the graph. Note that the graph is bipartite (with the sets A and B given by the 32 black and 32 white squares, respectively).

When the opposite corners are removed, this problem asks us to find a perfect matching of the graph with two vertices in the same class removed, so we must find a perfect matching of a bipartite graph whose bipartite classes have 30 and 32 vertices. But the 30 vertices in one class give a cover of the graph of size 30, and so Theorem 30.1 implies that no matching has size greater than 30. Thus it is impossible to find a perfect matching (of size $62/2 = 31$).