Math 188 Notes (Professor: Steven Sam)

Isabelle Mills

December 4, 2024

Lecture 1 Notes: 9/27/2024

Linear Recurrence Relations:

A sequence $(a_n)_{n\geq 0}$ satisfies a linear recurrence relation of order d if there exists $c_1, \ldots, c_d$ with $c_d \neq 0$ so that $a_n = c_1 a_{n-1} + c_2 a_{n-2} + \ldots + c_d a_{n-d}$ for all $n \geq d$.

(For $0 \leq n < d$, we usually explicitely specify a_n. Also, it seems like we are assuming all a_n and c_n are complex numbers right now)

To start this course, we're gonna discuss finding explicit (non-recursive) solutions.

Firstly, if $d = 1$, then this problem is easy. We can just plug in previous elements repeatedly to get that:

$$a_n = c_1 a_{n-1} = c_1^2 a_{n-2} = \ldots = c_1^n a_0$$

If $d = 2$, then plugging in previous elements doesn't help us really anymore. So how do we solve this problem now?

Theorem: Consider the characteristic polynomial $t^2 - c_1 t - c_2$ and let r_1, r_2 be the roots of that polynomial. If $r_1 \neq r_2$, then there exists α_1, α_2 such that $a_n = \alpha_1 r_1^n + \alpha_2 r_2^n$ for all $n \geq 0$.

To solve for α_1 and α_2, plug in different values of n into our equation. Since $r_1 \neq r_2$, we know the below linear system has a unique solution:

$$a_0 = \alpha_1 + \alpha_2$$
$$a_1 = \alpha_1 r_1 + \alpha_2 r_2$$

Now backing up, why does the above method work?

Approach 1: (Vector Spaces)

The set of sequences $(a_n)_{n\geq 0}$ form a vector space. Furthermore given any constants c_1 and c_2, we know that the set of sequences satisfying $a_n = c_1 a_{n-1} + c_2 a_{n-2}$ for all $n \geq 2$ is a subspace.

Proof:

Suppose (a_n) and (b_n) both satisfy that $a_n = c_1 a_{n-1} + c_2 a_{n-2}$ and $b_n = c_1 b_{n-1} + c_2 b_{n-2}$. Then given any constants γ and δ, we have that:

$$(\gamma a_n + \delta b_n) = c_1(\gamma a_{n-1} + \delta b_{n-1}) + c_2(\gamma a_{n-2} + \delta b_{n-2})$$

Hence, all linear combinations of any two sequences satisfying our linear recurrence relation also satisfies our linear recurrence relation.

Now what our above theorem is stating is that the sequences (r_1^n) and (r_2^n) span the subspace of solutions to our linear recurrence relation.

To see this, first note that (r_1^n) and (r_2^n) satisfy our recurrence relation.

If $n \geq 2$, then $r_i^n - c_1 r_i^{n-1} - c_2 r_i^{n-2} = r_i^{n-2}(r_i^2 - c_1 r_i - c_2) = r_i^{n-2}(0)$.

Hence, we know that $r_i^n = c_1 r_i^{n-1} + c_2 r_i^{n-2}$ for all $n \geq 2$.

Also, since we assumed $r_1 \neq r_2$, we know that (r_1^n) is linearly independent to (r_2^n). And finally, as mentioned before, we can solve a linear system of equations to find coffecients for a linear combination of (r_1^n) and (r_2^n) equal to any other sequence satisfying our recurrence relation.

Approach 2: (Formal Power Series)
Define the power series $A(x) = \sum_{n \geq 0} a_n x^n$. We call $A(x)$ a generating function of the sequence (a_n).

(We'll treat the formal power series more rigorously later...)

Now note that:

$$\begin{aligned} A(x) &= a_0 + a_1 x + \sum_{n \geq 2} a_n x^n \\ &= a_0 + a_1 x + \sum_{n \geq 2} (c_1 a_{n-1} + c_2 a_{n-2}) x^n \\ &= a_0 + a_1 x + c_1 \sum_{n \geq 2} a_{n-1} x^n + c_2 \sum_{n \geq 2} a_{n-2} x^n \\ &= a_0 + a_1 x + c_1 (A(x) - a_0) x + c_2 (A(x)) x^2 \end{aligned}$$

Isolating $A(x)$, we get the equation: $A(x) = \dfrac{a_0 + a_1 x - a_0 c_1 x}{1 - c_1 x - c_2 x^2}$.

Next, let's do fraction decomposition on our equation for $A(x)$.

Issue: We defined r_1 and r_2 as the roots of $t^2 - c_1 t - c_2 = (t - r_1)(t - r_2)$.

Trick: Plug in $t = \frac{1}{x}$. That way, we have that:

$$x^{-2} - c_1 x^{-1} - c_2 = (x^{-1} - r_1)(x^{-1} - r_2).$$

After that, multiply both sides of our equation by x^2 to get that:

$$1 - c_1 x - c_2 x^2 = (1 - r_1 x)(1 - r_2 x)$$

Since we're assuming $r_1 \neq r_2$, we know that for some constants α_1 and α_2, we have that:

$$A(x) = \frac{\alpha_1}{1 - r_1 x} + \frac{\alpha_2}{1 - r_2 x}$$

(If $r_1 = r_2$, then this step is where things will go differently.)

Now finally, we can rewrite $\frac{\alpha_1}{1 - r_1 x}$ as the geometric series $\alpha_1 \sum_{n \geq 0} (r_1 x)^n$. Doing likewise with $\frac{\alpha_2}{1 - r_2 x}$, we get that:

$$A(x) = \sum_{n \geq 0} a_n x^n = \alpha_1 \sum_{n \geq 0} (r_1 x)^n + \alpha_2 \sum_{n \geq 0} (r_2 x)^n = \sum_{n \geq 0} (\alpha_1 r_1^n + \alpha_2 r_2^n) x^n$$

Hence, we have for each n that $a_n = \alpha_1 r_1^n + \alpha_2 r_2^n$.

Lecture 2: 9/30/2024

Approach 3: (Matrices)

If $a_n = c_1 a_{n-1} + c_2 a_{n-2}$, then we can say that: $\begin{bmatrix} c_1 & c_2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_{n-1} \\ a_{n-2} \end{bmatrix} = \begin{bmatrix} a_n \\ a_{n-1} \end{bmatrix}$

Letting $\boldsymbol{C} = \begin{bmatrix} c_1 & c_2 \\ 1 & 0 \end{bmatrix}$, we thus know that: $\boldsymbol{C}^n \begin{bmatrix} a_1 \\ a_0 \end{bmatrix} = \begin{bmatrix} a_{n+1} \\ a_n \end{bmatrix}$

Notably, the characteristic polynomial of $\boldsymbol{C}$ is $t^2 - c_1 t - c_2$. So the eigenvalues of $\boldsymbol{C}$ are r_1 and r_2. Because we assumed r_1 and r_2 are distinct, we know $\boldsymbol{C}$ is diagonalizable. Hence there exists an invertible matrix $\boldsymbol{B}$ such that:

$$\boldsymbol{B} \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix} \boldsymbol{B}^{-1} = \boldsymbol{C}$$

Now set $\begin{bmatrix} x \\ y \end{bmatrix} = \boldsymbol{B}^{-1} \begin{bmatrix} a_1 \\ a_0 \end{bmatrix}$. Then we can see that:

$$\begin{bmatrix} a_{n+1} \\ a_n \end{bmatrix} = \boldsymbol{C}^n \begin{bmatrix} a_1 \\ a_0 \end{bmatrix} = \boldsymbol{B}\boldsymbol{D}^n \begin{bmatrix} x \\ y \end{bmatrix} = \boldsymbol{B} \begin{bmatrix} r_1^n x \\ r_2^n y \end{bmatrix} = \begin{bmatrix} b_{1,1} r_1^n x + b_{1,2} r_2^n y \\ b_{2,1} r_1^n x + b_{2,2} r_2^n y \end{bmatrix}$$

Setting $\alpha_1 = b_{2,1} x$ and $\alpha_2 = b_{2,2} y$, we have thus found constants α_1 and α_2 such that $a_n = \alpha_1 r_1^n + \alpha_2 r_2^n$.

Now some further questions to ask about recurrence relations are:

1. What if $r_1 = r_2$?
2. What if $d \geq 3$?
3. What if the recurrence relation is non-homogeneous or non-linear?

To start, let's answer question 1.

Theorem: Suppose r_1 and r_2 are the roots of $t^2 - c_1 t - c_2$ with $r_1 = r_2$. Then there exists α_1, α_2 such that $a_n = \alpha_1 r_1^n + \alpha_2 n r_1^n$ for all $n \geq 0$.

As was true when $r_1 \neq r_2$, you can solve for α_1 and α_2 by plugging in different values of n into the equation in order to get a linear system of equations.

To explain why this is, let's revisit two of our previous approaches.

The Formal Power Series Approach Revisited:

Before, we were able to show that $A(x) = \dfrac{a_0 + (a_1 - a_0 c_1)x}{(1 - r_1 x)(1 - r_2 x)}$ without assuming anything about r_1 and r_2.

But when we assume $r_1 = r_2$, we then get a different partial fraction decomposition for $A(x)$. Specifically, we have that there exists constants β_1, β_2 such that:

$$A(x) = \frac{\beta_1}{1 - r_1 x} + \frac{\beta_2}{(1 - r_1 x)^2}$$

Now we'll go into more rigor later. But for now, note that:

$$\tfrac{1}{(1-y)^2} = \left(\tfrac{1}{1-y}\right)' = \left(\sum_{n\geq 0} y^n\right)' = \sum_{n\geq 1} ny^{n-1} = \sum_{n\geq 0}(n+1)y^n$$

Comment from the future: as we'll cover two lectures from now, the definition of a derivative of a formal power series is different from the analysis definition we're familiar with.

Hence, we can write $A(x) = \sum\limits_{n\geq 0} a_n x^n = (\beta_1 + \beta_2)\sum\limits_{n\geq 0} r_1^n x^n + \beta_2 \sum\limits_{n\geq 0} n r_1^n x^n$.

Or in other words, setting $\alpha_1 = \beta_1 + \beta_2$ and $\alpha_2 = \beta_2$, we have that:

$$a_n = \alpha_1 r_1^n + \alpha_2 n r_1^n$$

The Matrix Approach Revisited:

If $r_1 = r_2$, then we must hav ethat the matrix $\boldsymbol{C}$ is not diagonalizable. For suppose it was, meaning there exists an invertible matrix $\boldsymbol{B}$ such that:

$$\boldsymbol{C} = \boldsymbol{B}\begin{bmatrix} r_1 & 0 \\ 0 & r_1 \end{bmatrix}\boldsymbol{B}^{-1}$$

Then we'd have to have that $\boldsymbol{C} = r_1\boldsymbol{B}\boldsymbol{B}^{-1} = \begin{bmatrix} r_1 & 0 \\ 0 & r_1 \end{bmatrix}$. But we know $\boldsymbol{C}$ isn't that.

Since we know $\boldsymbol{C}$ Is not diagonalizable, we will instead use the *Jordan-normal form* of $\boldsymbol{C}$. Specifically, we know there exists an invertible matrix $\boldsymbol{B}$ such that:

$$\boldsymbol{C} = \boldsymbol{B}\begin{bmatrix} r_1 & 1 \\ 0 & r_1 \end{bmatrix}\boldsymbol{B}^{-1}$$

Don't worry for the time being about how to prove the Jordan-normal form of a matrix always exists.

This tells us that $\boldsymbol{C}^n = \boldsymbol{B}\begin{bmatrix} r_1 & 1 \\ 0 & r_1 \end{bmatrix}^n\boldsymbol{B}^{-1}$.

Also, you can show by induction that $\begin{bmatrix} r_1 & 1 \\ 0 & r_1 \end{bmatrix}^n = \begin{bmatrix} r_1^n & nr_1^{n-1} \\ 0 & r_1^n \end{bmatrix}$.

So finally, defining $\begin{bmatrix} x \\ y \end{bmatrix}$ as before and expanding out the expression, you can get an explicit equation for a_n.

As for answering question 2, if $d \geq 3$, then our characteristic polynomial becomes $t^d - c_1 t^{d-1} - \ldots - c_d$. We'll assume this polynomial has distinct roots $r_1, \ldots, r_m$ with multiplicities $s_1, \ldots, s_m$ respectively.

Theorem: There exists constants $\alpha_1, \ldots, \alpha_d$ such that:

$$a_n = \sum_{i=1}^{s_1} \alpha_i n^{i-1} r_1^n + \ldots + \sum_{i=s_1+\ldots+s_{m-1}+1}^{s_1+\ldots+s_m} \alpha_i n^{i-1} r_m^n$$

As before, to solve for α_1 through α_d, you can plug in values of n and solve a linear system of equations.

The approaches to prove this are the same as when $d = 2$. However, there are just more terms floating around that need to be dealt with.

Special case: suppose the characteristic polynomial is $(t-1)^d$.

In that case, because the root of the polynomial r is 1, there exists $\alpha_1, \ldots, \alpha_d$ such that

$$a_n = \alpha_1 + n\alpha_2 + n^2\alpha_3 + \ldots + n^{d-1}\alpha_d.$$

In other words, the formula for a_n is a polynomial in n.

Another perspective on the characteristic polynomial:

Let V be the vector space of sequences $(a_n)_{n \geq 0}$, and define the translation operator $T : V \longrightarrow V$ such that $(a_n)_{n\geq 0} \mapsto (a_{n+1})_{n \geq 0}$. Now, given $\boldsymbol{a} \in V$ and the recurrence relation $a_n = c_1 a_{n-1} + \ldots + c_d a_{n-d}$ for all $n \geq d$, we have that $\boldsymbol{a}$ satisfies our recurrence relation if and only if:

$$T^d \boldsymbol{a} = c_1 T^{d-1}\boldsymbol{a} + c_2 T^{d-2}\boldsymbol{a} + \ldots + c_d \boldsymbol{A}$$

In other words, we must have that $\boldsymbol{a} \in \ker(T^d - c_1 T^{d-1} - \ldots - c_d)$.

If $r_1, \ldots, r_d$ are the roots of the characteristic polynomial $t^d - c_1 T^{d-1} - \ldots - c_d$, then we can rewrite this as:

$$(T - r_1) \cdots (T - r_d)\boldsymbol{a} = \boldsymbol{0}$$

Proposition: Given a sequence $\boldsymbol{a} = (a_n)_{n \geq 0}$, there exists a polynomial $p(n)$ of degree at most $d-1$ such that $a_n = p(n)$ if and only if $(T-1)^d \boldsymbol{a} = \boldsymbol{0}$.

We already saw in the special case above one direction of this statement. As for the other direction, suppose $p(n) = \alpha_d n^{d-1} + \alpha_{d-1} n^{d-2} + \ldots + \alpha_1$. Then $(T-1)$ applied to the sequence $(p(n))_{n \geq 0}$ is the sequence $(p(n+1) - p(n))_{n \geq 0}$ Importantly, $p(n+1)$ is also a polynomial of degree $d-1$ with α_d as the coefficient in front of n^{d-1}. So the difference is a polynomial of degree at most $d-2$.

Proceeding by induction, we know that $(T-1)^d(p(n))_{n\geq 0} = \mathbf{0}$.

Note that the operator $(T-1)$ can be thought of as the taking the "derivative" of a sequence $\boldsymbol{a}$. Going by that analogy, the previous proposition is saying that a sequence $\boldsymbol{a}$ is given by a polynomial if and only if a derivative of some order of the sequence is zero. Interestingly, the same is true of differential equations.

Lecture 3: 10/2/2024

To quickly address question 3, in general there is no unified approach to dealing with nonlinear recurrence relations. However, we can often solve non-homogeneous linear recurrence relations.

This will be addressed by the homework (see HW 1: Exercise (2)).

Formal Power Series:

A formal power series in the variable x is an expression of the form $A(x) = \sum_{n\geq 0} a_n x^n$ where a_n is a sequence of elements of a field.

Technically, we can go more general to a commutative ring (but we won't).

We call $A(x)$ the generating function of $(a_n)_{n\geq 0}$.

If $A(x)$ and $B(x)$ are the generating functions of $(a_n)_{n\geq 0}$ and $(b_n)_{n\geq 0}$ respectively, then:

- $A(x) = B(x)$ iff $a_n = b_n$ for all n.
- $A(x) + B(x) := \sum_{n\geq 0} (a_n + b_n)x^n$.
- $A(x)B(x) := \sum_{n\geq 0} c_n x^n$ where $c_n = \sum_{i=0}^{n} a_i b_{n-i}$.

(in other words this is the Cauchy Product of $A(x)$ and $B(x)$)

Note that polynomials and constants are special cases of formal power series with the sequence generating that function being eventually zero.

Also, sums and products of formal power series satisfy the commutative, associative, and distributive properties of a field.

The only one of those properties that's non-trivial to show is the associativity of products. One way that you can prove this property is to show that :

$$\sum_{i=0}^{n}\sum_{j=0}^{i} a_j b_{i-j} c_{n-i} = \sum_{(p,q,r)\in I} a_p b_q c_r = \sum_{i=0}^{n}\sum_{j=0}^{i} a_{n-i} b_j c_{i-j},$$

where $I = \{(p,q,r) \in \mathbb{Z}^3 \mid p+q+r = n \text{ and } p,q,r \geq 0\}$.

Plus, letting $0+0x+\ldots$ be the additive identity, then given any formal power series $A(x) = \sum_{n\geq 0} a_n x^n$, we have that:
$-A(x) = (-1+0x+\ldots)\sum_{n\geq 0} a_n x^n = \sum_{n\geq 0}(-a_n)x^n$ is the additive inverse of $A(x)$.

Finally, we clearly have that $1+0x+0x^2+\ldots$ is a multiplicative identity of the set of formal power series.

Therefore, to make this perfectly clear, the set of formal power series is not actually a set of functions. Rather, given a field (or commutative ring) F, the set of formal power series with coefficients in that ring is a commutative ring on the set F^ω of infinite sequences of elements of F using the $+$ and $\cdot$ operations defined above.

> In other words, it doesn't make sense to plug values of x into formal power series. Rather, the power series notation is just to make it clearer where the definitions of different operations are coming from.
>
> Also, since the set of formal power series on a commutative ring is itself a commutative ring, you can define the set of formal power series on the set of formal power series on that commutative ring. Spoiler: this is a way to define multivariate formal power series.

A formal power series $A(x)$ is invertible if there exists a formal power series $B(x)$ such that $A(x)B(x) = 1$. We write $B(x) = A(x)^{-1} = \frac{1}{A(x)}$, and call $B(x)$ the inverse of $A(x)$.

Example: If $A(x) = \sum_{n\geq 0} x^n$, then $A(x)$ is invertible with inverse $B(x) = 1-x$.

Proof:

$$\begin{aligned} A(x)B(x) &= (1+x+x^2+\ldots)(1-x) \\ &= 1+x-x+x^2-x^2+x^3-x^3+\ldots = 1 \end{aligned}$$

So $\sum_{n\geq 0} x^n = \frac{1}{1-x}$.

Theorem: $A(x) = \sum_{n\geq 0} a_n$ is invertible if and only if a_0 has a multiplicative inverse.

Proof:
If there exists $B(x)$ such that $A(x)B(x) = 1$, then we must have that:

$$\begin{aligned} &a_0 b_0 = 1 \\ &a_0 b_1 + a_1 b_0 = 0 \\ &a_0 b_2 + a_1 b_1 + a_2 b_0 = 0 \\ &\qquad\vdots \end{aligned}$$

If $\frac{1}{a_0}$ exists then we can inductively solve for each b_n. Specifically, $b_0 = \frac{1}{a_0}$ and $b_n = -\frac{1}{a_0}\sum_{i=1}^{n} a_i b_{n-i}$. Then $B(x) = \sum_{n\geq 0} b_n x^n$ satisfies that $A(x)B(x) = 1$.

If $\frac{1}{a_0}$ doesn't exist, then there is no choice of b_0 such that $A(x)(\sum_{n\geq 0} b_n x^n) = 1$. So $A(x)$ has no inverse.

Lecture 4: 10/4/2024

If $A(x)$ is a power series and $n \geq 0$, then $[x^n]A(x)$ refers to the coefficient a_n in front of x^n.

Let $A_0(x), A_1(x), \ldots$ be a sequence of formal power series. We say the sequence formally converges to $A(x)$ if:

$$\forall n \geq 0, \exists N \geq 0 \ s.t. \ i \geq N \implies [x^n]A_i(x) = [x^n]A(x)$$

We also write this as $\lim_{i\to\infty} A_i(x) = A(x)$

Note that this definition is different from the familiar definition of convergence in 140. For instance, the sequence $A_i(x) = \frac{1}{i+1}$ doesn't formally converge.

Lemma: Suppose $\lim_{i\to\infty} A_i(x) = A(x)$ and $\lim_{i\to\infty} B_i(x) = B(x)$. Then:

- $\lim_{i\to\infty}(A_i(x) + B_i(x)) = A(x) + B(x)$
- $\lim_{i\to\infty}(A_i(x)B_i(x)) = A(x)B(x)$

The proof for this is rather trivial. So do it yourself. :P

Continuing to let $A_0(x), A_1(x), \ldots$, be a sequence of formal power series, we define:

$$\sum_{i\geq 0} A_i(x) \coloneqq \lim_{i\to\infty}(\sum_{j=0}^{i} A_j(x))$$

$$\prod_{i\geq 0} A_i(x) \coloneqq \lim_{i\to\infty}(\prod_{j=0}^{i} A_j(x))$$

Lemma: (This is just reapplying the previous lemma for sequences and using the commutative property...)

- If $\sum_{i\geq 0} A_i(x)$ and $\sum_{i\geq 0} B_i(x)$ exist, then:
$$\sum_{i\geq 0}(A_i(x) + B_i(x)) = \sum_{i\geq 0} A_i(x) + \sum_{i\geq 0} B_i(x).$$
- If $\prod_{i\geq 0} A_i(x)$ and $\prod_{i\geq 0} B_i(x)$ exist, then:
$$\prod_{i\geq 0}(A_i(x)B_i(x)) = \left(\prod_{i\geq 0} A_i(x)\right)\left(\prod_{i\geq 0} B_i(x)\right).$$

Given a formal power series $A(x)$, we define:

$$\operatorname{mdeg} A(x) \coloneqq \inf(\{n \in \mathbb{Z}_+ \cup \{0\} \mid [x^n]A(x) \neq 0\} \cup \{\infty\}).$$

Proposition: Suppose $A_0(x), A_1(x), \ldots$ is a sequence of formal power series.

- $\sum_{i \geq 0} A_i(x)$ exists if and only if $\lim_{i \to \infty} \operatorname{mdeg} A_i(x) = \infty$.

  Proof: (The professor skipped this because he thinks it's boring.)

  ($\Longleftarrow$)

  Suppose $\lim_{j \to \infty} \operatorname{mdeg} A_j(x) = \infty$. Then for all $n \geq 0$, there exists $N \geq 0$ such that $\operatorname{mdeg} A_j(x) > n$ for all $j > N$. So:

  $$[x^n]\left(\sum_{j=0}^{i} A_j(x)\right) = [x^n]\left(\sum_{j=0}^{N} A_j(x)\right) \text{ for all } i > N.$$

  ($\Longrightarrow$)

  Suppose that $\lim_{j \to \infty} \operatorname{mdeg} A_j(x)$ either doesn't exist or doesn't equal infinity if it does exist. Then we know there must exist N such that $\operatorname{mdeg} A_j(x) < N$ for infinitely many $j \geq 0$. In turn, for some $n \in \{0, 1, \ldots, N-1\}$, there must be infinitely many $j \geq 0$ such that $\operatorname{mdeg} A_j(x) = n$. Thus, there does not exist $M \geq 0$ such that:

  $$[x^n]\left(\sum_{j=0}^{i} A_j(x)\right) \text{ is the same for all } i \geq M.$$

- Assume each A_i has no constant term. Then $\prod_{i \geq 0}(1 + A_i(x))$ exists if and only if $\lim_{i \to \infty} \operatorname{mdeg} A_i(x) = \infty$.

  Proof: (btw I'm having to figure this all out without any outside help)

  Lemma: Suppose $B(x)$ and $C(x)$ are formal power series such that $[x^0]B(x) = 1$ and $\operatorname{mdeg} C(x) = n$. Then $\operatorname{mdeg} B(x)C(x) = n$ with $[x^n](B(x)C(x)) = [x^n](C(x))$.

  Corollary 1: Given $B(x)$ and $C(x)$ defined as before, for all $0 \leq i < n$:

  $$[x^i](B(x)(1 + C(x))) = [x^i](B(x) + B(x)C(x)) = [x^i](B(x)).$$

  Corollary 2: $[x^0]\left(\prod_{j=0}^{i}(1 + A_j)\right) = 1$ for all $i \geq 0$.

  ($\Longleftarrow$)

  Suppose $\lim_{j \to \infty} \operatorname{mdeg} A_j(x) = \infty$. Then for any $n \geq 0$, there exists $N \geq 0$ such that $\operatorname{mdeg} A_j(x) > n$ for all $j > N$. So given any $i > N$, we can inductively show using the above lemma and corollaries that:

$$[x^n]\left(\prod_{j=0}^{i}(1+A_j(x))\right) = [x^n]\left(\prod_{j=0}^{i-1}(1+A_j(x))\right)$$
$$= \cdots = [x^n]\left(\prod_{j=0}^{N}(1+A_j(x))\right)$$

($\Longrightarrow$)
As before, we can show there must be infinitely many $i \geq 0$ such that $\text{mdeg } A_i(x) = n$ for some n. And for any such i, we have by the above lemma that:

$$\begin{aligned}
&[x^n]\left(\prod_{j=0}^{i}(1+A_j(x))\right)\\
&\quad= [x^n]\left(\prod_{j=0}^{i-1}(1+A_j(x)) + \left(\prod_{j=0}^{i-1}(1+A_j(x))\right)A_i(x)\right)\\
&\quad= [x^n]\left(\prod_{j=0}^{i-1}(1+A_j(x))\right) + [x^n]A_i(x)\\
&\quad\neq [x^n]\left(\prod_{j=0}^{i-1}(1+A_j(x))\right)
\end{aligned}$$

So there is no $N \geq 0$ such that:

$$[x^n]\left(\prod_{j=0}^{i}(1+A_j(x))\right) \text{ is the same for all } i \geq N.$$

Suppose $A(x)$ and $B(x)$ are formal power series such that $A(x)$ has no constant term and $B(x) = \sum_{n\geq 0} b_n x^n$. Then we define their composition:

$$(B \circ A)(x) = B(A(x)) := \sum_{n\geq 0} b_n A(x)^n$$

This is well defined because $\text{mdeg } A(x) \geq 1 \Longrightarrow \text{mdeg } A(x)^n \geq n$. Therefore, $\lim_{n\to\infty} b_n A(x)^n = \infty$, meaning we can apply the previous proposition.

Special Case: If $A(x) = 0$, then $(B \circ A)(x) = b_0$.

Proposition: If $A(x)$, $B(x)$, and $C(x)$ are power series generated by $(a_n)_{n\geq 0}$, $(b_n)_{n\geq 0}$, and $(c_n)_{n\geq 0}$ respectively such that $(B \circ A)(x)$ and $(C \circ A)(x)$ are defined, then:

- $((B + C) \circ A)(x) = (B \circ A)(x) + (C \circ A)(x)$

  Proof:
  By the second lemma on page 9, we know that:
  $\sum_{n\geq 0}(b_n+c_n)A(x)^n = \sum_{n\geq 0}(b_nA(x)^n+c_nA(x)^n) = \sum_{n\geq 0} b_nA(x)^n+\sum_{n\geq 0} c_nA(x)^n$

- $((BC) \circ A)(x) = (B \circ A)(x)(C \circ A)(x)$

Proof:
By the first lemma on page 9, we know that:

$$\begin{aligned}(B\circ A)(x)(C\circ A(x)) &= \left(\lim_{n\to\infty}\left(\sum_{i=0}^{n} b_i A(x)^i\right)\right)\left(\lim_{n\to\infty}\left(\sum_{i=0}^{n} c_i A(x)^i\right)\right)\\ &= \lim_{n\to\infty}\left(\left(\sum_{i=0}^{n} b_i A(x)^i\right)\left(\sum_{i=0}^{n} c_i A(x)^i\right)\right)\end{aligned}$$

Next note that given any $n \geq 0$, there exists a formal power series $R_n(x)$ with $\operatorname{mdeg} R_n(x) > n$ such that:

$$\left(\sum_{i=0}^{n} b_i A(x)^i\right)\left(\sum_{i=0}^{n} c_i A(x)^i\right) = \sum_{i=0}^{n}\left(\sum_{j=0}^{i} b_j c_{i-j}\right)A(x)^i + R_n(x)$$

Since $\lim\limits_{n\to\infty}\left(\sum_{i=0}^{n}\left(\sum_{j=0}^{i} b_j c_{i-j}\right)A(x)^i\right) = ((BC)\circ A)(x)$ and $\lim\limits_{n\to\infty}(R_n(x)) = 0$, we can thus apply the first lemma on page 9 again to get that:

$$\begin{aligned}&\lim_{n\to\infty}\left(\left(\sum_{i=0}^{n} b_i A(x)^i\right)\left(\sum_{i=0}^{n} c_i A(x)^i\right)\right)\\ &\qquad= \lim_{n\to\infty}\left(\sum_{i=0}^{n}\left(\sum_{j=0}^{i} b_j c_{i-j}\right)A(x)^i\right) + \lim_{n\to\infty}(R_n(x))\\ &\qquad= ((BC)\circ A)(x) + 0\end{aligned}$$

Suppose $A(x)$ is a formal power series. We define its derivative:

$$(DA)(x) = A'(x) := \sum_{n\geq 1} n a_n x^{n-1} = \sum_{n\geq 0}(n+1)a_{n+1}x^n$$

Note that for $n \in \mathbb{Z}^+$ and $a_n \in F$, we define na_n via repeated addition.

Proposition: The following rules hold for any two formal power series $A(x)$ and $B(x)$:

- Sum Rule: $(A+B)'(x) = A'(x) + B'(x)$

  This identity is hopefully obvious.
- Product Rule: $(AB)'(x) = A'(x)B(x) + A(x)B'(x)$

  The proof for this identity requires rearranging sums strategically.
- Power Rule: $(A^n)'(x) = nA^{n-1}(x)A'(x)$ if $n > 0$ and $(A^n)'(x) = 0$ if $n = 0$.

  To prove this, do induction on n using the product rule.

Also if $[x^0]A(x) = 0$, then:

- Chain Rule: $(B \circ A)'(x) = A'(x)B'(A(x))$

Proof: (seriously I'm doing this proof on my own...I order you to give me pity.)

Lemma: If $A_0(x), A_1(x), ...$ are a sequence of formal power series that converges to $A(x)$, then $\lim_{n\to\infty} A'_n(x) = A'(x)$.

The proof for this is rather trivial. However, this is notably different from the convergence of derivatives of sequences of functions in math 140.

Now suppose $B(x) = \sum_{n\geq 0} b_n x^n$, and for each n, set $B_n(x) = \sum_{i=0}^{n} b_i x^i$.

By definition, we know that:

$$\lim_{n\to\infty}((B_n \circ A)(x)) = \lim_{n\to\infty}\left(\sum_{i=0}^{n} b_i A(x)^i\right) = (B \circ A)(x).$$

Hence, applying the previous lemma, we know that:

$$\lim_{n\to\infty}((B_n \circ A)'(x)) = (B \circ A)'(x).$$

But now note that by the sum and power rules:

$$\begin{aligned}(B_n \circ A)'(x) &= \sum_{i=0}^{n} b_i (A^i)'(x) \\ &= \sum_{i=1}^{n} i b_i A(x)^{i-1} A'(x) = A'(x) \sum_{i=0}^{n-1} (i+1) b_{i+1} A(x)^i \\ &= A'(x)(B'_n \circ A)(x)\end{aligned}$$

Finally, by definition we know that:

$$\lim_{n\to\infty}((B'_n \circ A)(x)) = \lim_{n\to\infty}\left(\sum_{i=0}^{n-1} (i+1) b_{i+1} A(x)^i\right) = (B' \circ A)(x).$$

So by applying the first lemma on page 9, we know that:

$$\lim_{n\to\infty}((B_n \circ A)'(x)) = \lim_{n\to\infty}(A'(x)(B'_n \circ A)(x)) = A'(x)(B' \circ A)(x)$$

Meanwhile, if $A(x)$ is invertible, then:

- Multiplicative inverse rule: $(\frac{1}{A(x)})' = \frac{-A'(x)}{A(x)^2}$

To prove this, just apply product rule to the expression $A(x)(\frac{1}{A(x)}) = 1$.

Examples of proving identities:

1. Since $\frac{1}{1-x} = \sum_{n\geq 0} x^n$, we know $-\frac{-1}{(1-x)^2} = (\frac{1}{1-x})' = \left(\sum_{n\geq 0} x^n\right)' = \sum_{n\geq 0}(n+1)x^n$.
2. $\sum_{n\geq 0} n x^n = \sum_{n\geq 0}(n+1)x^n - \sum_{n\geq 0} x^n = \frac{1}{(1-x)^2} - \frac{1}{1-x}\left(\frac{1-x}{1-x}\right) = \frac{x}{(1+x)^2}$

Let F be the field or commutative ring that our formal power series are defined over.

Note that given $n, m \in \mathbb{Z}_+$, we define $n, m \in F$ by repeatedly adding $1 \in F$ to itself n and m times respectively. Then, we can define $nm = mn \in F$ by doing repeated addition of n or m with itself. With that in mind, defining $n! \in F$ in a similar fashion and assuming that there exists $\frac{1}{n!} \in F$, we have that:

$$[x^n]A(x) = \frac{(D^n A)(0)}{n!} = \frac{(D^n A)(0+0x+0x^2+\dots)}{n!}.$$

This is a random thought I had outside of lecture and wanted to write down:

Note: For the sake of clarity, I looked this up on wikipedia. If R is a commutative ring, then the set of formal power series in the variable x over R is written: $R[[x]]$.

Now given $A(x), B(x) \in R[[x]]$, we can define a metric $\rho(A(x), B(x)) = 2^{-n}$ where the nth coefficients of $A(x)$ and $B(x)$ are the first to differ, or if no such n exists, then we define $\rho(A(x), B(x)) = 0$. This somewhat trivially satisfies that:

- $\rho(A(x), B(x)) = 0 \iff A(x) = B(x)$.
- $\rho(A(x), B(x)) = \rho(B(x), A(x))$ for all $A(x), B(x) \in R[[x]]$.
- $\rho(A(x), B(x)) \leq \rho(A(x), C(x)) + \rho(C(x), B(x))$.

Also, we clearly have from of our definition of convergence that a sequence $(A_n(x))_{n\geq 0}$ in $R[[x]]$ converges if it is Cauchy. So this metric space is complete.

If I think of anything more to do with this, I'll add it to my notes.

Lecture 5: 10/7/2024

A formal power series in x, y is: $A(x,y) = \sum\limits_{n,m\geq 0} a_{m,n}x^m y^n$.

Suppose $A(x,y) = \sum\limits_{n,m\geq 0} a_{m,n}x^m y^n$ and $B(x,y) = \sum\limits_{n,m\geq 0} b_{m,n}x^m y^n$. Then define:

- $A(x,y) + B(x,y) \coloneqq \sum\limits_{n,m\geq 0} (a_{m,n} + b_{m,n})x^m y^n$.
- $A(x,y)B(x,y) \coloneqq \sum\limits_{m,n\geq 0} c_{m,n}x^m y^n$ where $c_{m,n} = \sum\limits_{i=0}^{m}\sum\limits_{j=0}^{n} a_{i,j}b_{m-i,n-j}$.

Like before, we know that addition and multiplication are commutative, associative, and distributive.

To prove this, note that if we think of $A(x,y)$ and $B(x,y)$ as being single variable formal power series in y whose coefficients are formal power series in x, then our old definition of sums and products match up with the new definitions above. So, we can just apply that those old definitions are commutative, associative, and distributive.

Also, $A(x,y)$ is invertible if and only if $a_{0,0}$ has a multiplicative inverse.

This is because a single variable power series only has a multiplicative inverse if its constant term has a multiplicative inverse.

Let $A_1(x,y), A_2(x,y), \ldots$ be a sequence of formal power series in x, y. We say that the sequence converges to $A(x,y)$ if:

$$\forall n, m \geq 0, \exists N \geq 0 \; s.t. \; i \geq N \implies [x^n y^m]A_i(x,y) = [x^n y^m]A(x,y)$$

Going back to the idea of thinking of two variable formal power series as a formal power series whose coefficients are formal power series, note that:

This definition of convergence is slightly different from applying our definition of formal convergence to a sequence of formal power series of formal power series, because we are only requiring that every coefficient power series in the sequence converge.

If that sequence converges, we write: $\lim\limits_{i\to\infty} A_i(x,y) = A(x,y)$.

Same as before, if $\lim\limits_{i\to\infty} A_i(x,y) = A(x,y)$ and $\lim\limits_{i\to\infty} B_i(x,y) = B(x,y)$. Then:

- $\lim\limits_{i\to\infty}(A_i(x,y) + B_i(x,y)) = A(x,y) + B(x,y)$
- $\lim\limits_{i\to\infty}(A_i(x,y)B_i(x,y)) = A(x,y)B(x,y)$

Using limits, we also define infinite sums and infinite products of two variable formal power series as you would expect.

Define $\operatorname{mdeg} A(x,y)$ to be the infimum of the set:

$$\{N \in \mathbb{Z}_{\geq 0} \mid \exists n, m \geq 0 \; s.t. \; n + m = N \text{ and } [x^n y^m]A(x) \neq 0\} \cup \{\infty\}$$

Then analogous propositions to those on page 10 and 11 hold, and their proofs are mostly identical.

If $A(x,y) = \sum\limits_{n,m\geq 0} a_{n,m}x^n y^m$, $B(x,y) = \sum\limits_{n,m\geq 0} b_{n,m}x^n y^m$, and $C(x,y) = \sum\limits_{n,m\geq 0} c_{n,m}x^n y^m$, are formal power series such that $a_{0,0} = b_{0,0} = 0$, then we define:

$$C(A(x,y), B(x,y)) = \lim_{N\to\infty} \sum_{i=0}^{N}\sum_{j=0}^{N} c_{i,j}A(x,y)^i B(x,y)^j$$

This will be well-defined because for all $N \geq 0$, since $a_{0,0} = b_{0,0} = 0$:

$$\operatorname{mdeg}\left(\sum_{i=0}^{N}\sum_{j=0}^{N} c_{i,j}A(x,y)^i B(x,y)^j - \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} c_{i,j}A(x,y)^i B(x,y)^j\right) \geq N$$

Note that by slightly modifying the reasoning on page 11 and page 12, we can show that:

- $(C+D)(A(x,y),B(x,y)) = C(A(x,y),B(x,y)) + D(A(x,y),B(x,y))$
- $(CD)(A(x,y),B(x,y)) = C(A(x,y),B(x,y))D(A(x,y),B(x,y))$

Given a formal power series $A(x,y)$, we define the partial derivatives D_xA and D_yA as you would expect.

If we once again consider $A(x,y)$ to be a formal power series in y whose coefficients are formal power series in x:

$$A(x,y) = \sum_{n\geq 0} A_n(x)y^n = A_0(x) + A_1(x)y + A_2(x)y^2 + \ldots,$$

then $D_yA(x,y) = \sum_{n\geq 0}(n+1)A_{n+1}(x)y^n = D\left(\sum_{n\geq 0} A_n(x)y^n\right)$.

It follows then that the following single-variable formal power series derivative rules hold for partial derivatives with respect to y.

- Sum Rule: $D_y(A+B)(x,y) = D_yA(x,y) + D_yB(x,y)$
- Product Rule: $D_y(AB)(x,y) = D_yA(x,y)B(x,y) + A(x,y)D_yB(x,y)$
- Power Rule: $D_y(A^n)(x,y) = nA^{n-1}(x,y)D_yA(x,y)$ if $n > 0$.

Obviously, analogous reasoning holds for partial derivatives with respect to x.

Of course there's no reason to stop at two variable formal power series. We can extend all the properties covered so far in this lecture to power series in any number of variables. Most often in this class though, we'll only need one or two variables.

Binomial Theorem:

In this section, we'll assume that the coefficients of all power series are complex numbers.

Lemma: Let $A(x)$ be a formal power series with $a_0 = 1$ and $d > 0$ be an integer. Then there exists a unique formal power series $B(x)$ with $b_0 = 1$ and $B(x)^d = A(x)$. So define $A(x)^{1/d} := B(x)$.

Proof:

We can inductively solve for the coefficients of $B(x) = \sum_{n\geq 0} b_nx^n$ so that $[x^n]B(x)^d = [x^n]A(x)$ for all $n > 0$.

For our base case, we must have that $a_1 = [x^1]B(x)^d = db_1$. Thus, we know $b_1 = \frac{a_1}{d}$. After that, we know for $n > 1$ that $a_n = [x^n]B(x)^d = db_n + f_{n,d}$ where $f_{n,d}$ is an expression only involving $b_1, \ldots, b_{n-1}$. So, assuming we know what $b_1, \ldots, b_{n-1}$ are, we can uniquely solve for b_n.

Thus, we've shown that given a formal power series $A(x)$ with $a_0 = 1$, we know $A(x)^c$ and $A(x)^{1/d}$ are well defined for all $c \in \mathbb{Z}$ and $d \in \mathbb{Z}_+$. The next obvious question is, can we define $A(x)^q$ for any nonzero rational number q?

Firstly, given integers c and d with $d \neq 0$, we need to sort out whether we want to define $A(x)^{c/d} = (A(x)^c)^{1/d}$ or $A(x)^{c/d} = (A(x)^{1/d})^c$ (both are defined because taking powers of $A(x)$ or $A(x)^{1/d}$ won't change the fact that it's constant term is 1). Luckily the choice doesn't actually matter because both are equivalent.

$((A(x)^{1/d})^c)^d = ((A(x)^{1/d})^d)^c = A(x)^c$. So $(A(x)^{1/d})^c = (A(x)^c)^{1/d}$.

The rest of the checks we need to do are in exercise 5.(a) in homework 2.

With that, we've now proven we can raise certain formal power series to rational powers. Although notably, we don't have an easy way right now of calculating those rational powers.

If $m \in \mathbb{Q}$ and $k > 0$ is an integer, then we define:

$\binom{m}{0} := 1$ and $\binom{m}{k} := \frac{m(m-1)\ldots(m-k+1)}{k!}$

Binomial Theorem: Suppose $m \in \mathbb{Q}$. Then $(1+x)^m = \sum\limits_{n \geq 0} \binom{m}{n} x^n$.

Before proving this theorem, here are some special uses of this theorem:

1. Suppose $A(x)$ is a formal power series with $a_0 = 1$. Then substituting x with $A(x) - 1$ in our above theorem, we get that:

   $$A(x)^m = (1 + (A(x) - 1))^m = \sum_{n \geq 0} \binom{m}{n} (A(x) - 1)^n.$$

   While this still is not the most wieldy formula ever, it does give us a way to calculate rational powers of formal power series without struggling to calculate a root.

2. Suppose $m \in \mathbb{Z}_{\geq 0}$. If $n > m$, then $\binom{m}{n} = 0$. So, $(1+x)^m = \sum\limits_{n=0}^{m} \binom{m}{n} x^n$

   Also, when m is an integer and $0 \leq k \leq m$, we have $\binom{m}{k} = \frac{m!}{(m-k)!k!}$

3. Note that $\binom{-1}{n} = \frac{(-1)(-2)\ldots(-n)}{n!} = (-1)^n$. Hence, applying the binomial theorem we get that: $(1+x)^{-1} = \sum\limits_{n \geq 0} (-x)^n$. This should make sense to you because:

   $$(1-x)^{-1} = (1 + (-x))^{-1} = \sum_{n \geq 0} (-(-x))^n = \sum_{n \geq 0} (x)^n.$$

4. Suppose d is a positive integer. Then:

$$\binom{-d}{n} = \frac{-d(-d-1)\ldots(-d-n+2)(-d-n+1)}{n!}$$
$$= (-1)^n \frac{(d+n-1)(d+n-2)\cdots(d-1)(d)}{n!}$$
$$= (-1)^n \frac{(d+n-1)!}{((d+n-1)-n)!n!} = (-1)^n \binom{d+n-1}{n}$$

It follows that $(1+x)^{-d} = \sum\limits_{n\geq 0} (-1)^n \binom{d+n-1}{n} x^n$.

In turn, we can conclude that: $(1-x)^{-d} = \sum\limits_{n\geq 0} \binom{d+n-1}{n} x^n$.

5. We define $\sqrt{1+x} := (1+x)^{1/2} = \sum\limits_{n\geq 0} \binom{1/2}{n} x^n$.

Now note that for $n \geq 2$,

$$\binom{1/2}{n} = \frac{(1/2)(-1/2)(-3/2)\cdots(\frac{1}{2}-n+1)}{n!} = \frac{(-1)^{n-1}}{2^n}\frac{(2n-3)(2n-5)\cdots(3)(1)}{n!}$$

To make this neater, we define the double factorial $k!! = k(k-2)(k-4)\cdots$. If k is even, then $k!!$ is the product of all even positive integers at most k. If k is odd, then $k!!$ is the product of all odd positive integers at most k.

Thus, $(1+x)^{1/2} = \sum\limits_{n\geq 0} \binom{1/2}{n} x^n = 1 + \frac{x}{2} + \sum\limits_{n\geq 2} \frac{(-1)^{n-1}(2n-3)!!}{2^n n!} x^n$

Proof of the Binomial Theorem:

Lemma: If $m \in \mathbb{Q}$ and $A(0) = 1$, then $D(A^m(x)) = mD(A(x))A(x)^{m-1}$.

Proof:

We know there exists integers p and q with $m = p/q$. Then:

$$D(A(x)^p) = pA(x)^{p-1}DA(x).$$

But also, $D(A(x)^p) = D((A(x)^m)^q) = q(A(x)^m)^{q-1}D(A(x)^m)$.

So combining these expressions, we get that:

$$D(A(x)^m) = \frac{pA(x)^{p-1}}{q(A(x)^m)^{q-1}}DA(x)$$
$$= m\frac{A(x)^{p-1}}{A(x)^p A(x)^{-m}}DA(x) = mA(x)^{m-1}DA(x)$$

Now recall that $[x^n](1+x)^m = \frac{D^n((1+x)^m)(0)}{n!}$.

Also, because of our above lemma, we know that:

$$D((1+x)^m) = m(1+x)^{m-1}$$
$$\Downarrow$$
$$D^2((1+x)^m) = m(m-1)(1+x^{m-2})$$
$$\Downarrow$$
$$\vdots$$
$$\Downarrow$$
$$D^n((1+x)^m) = m(m-1)\cdots(m-n+1)(1+x)^{m-n}$$

Thus:

$$[x^n](1+x)^m = \frac{D^n((1+x)^m)(0)}{n!} = \frac{m(m-1)\cdots(m-n+1)(1+(0))^{m-n}}{n!}$$
$$= \frac{m(m-1)\cdots(m-n+1)}{n!} = \binom{m}{n}.$$

One more note given at the end of class:

We will not prove it in this class, but if we have a quadratic equation: $A(x)t^2 + B(x)t + C(x) = 0$, where $A(x)$, $B(x)$, and $C(x)$ are formal power series, then there are at most 2 formal power series solutions to this quadratic.

Also, any solution formal power series t will satisfy that:

$$2A(x)t = -B(x) \pm \sqrt{B(x)^2 - 4A(x)C(x)}$$

Finally, this doesn't contradict the lemma on the bottom of page 16 because that lemma guarentees a unique root that has a constant term of 1. Because of this, $A^{1/2}(x)$ and $-A^{1/2}(x)$ are the only formal power series whose square can be $A(x)$.

Lecture 6: 10/9/2024

Choice problems: (where we finally start applying power series)

Example: Given $n \in \mathbb{Z}_{\geq 0}$, denote $[n] = \{1, \ldots, n\}$. Suppose we want to count the k-element subsets of $[n]$.

Trick: Consider the expansion of $(1+x)^n = (1+x)(1+x)\cdots(1+x)$.

Given a subset $S \subseteq [n]$, we can think of it as a term in the above expansion as follows:

At the ith step of multiplying x or 1, choose x if $i \in S$ and 1 if $i \notin S$.

Thus, we have that:

$$(1+x)^n = \sum_{S \subseteq [n]} x^{|S|} = \sum_{i=0}^{n} (\text{\# of subsets of size } i)x^i$$

Applying the binomial theorem, we thus know that $\binom{n}{i}$ equals the number of subsets of $[n]$ with i elements.

Corollary: The total number of subsets of $[n]$ of any size is:

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

Many identities can be proved by manipulating power series.

Example: Pascal's identity

Note that $(1+x)^n = (1+x)(1+x)^{n-1} = (1+x)^{n-1} + x(1+x)^{n-1}$. Thus, we have that:

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

By adding more variables, we can store more information about a problem.
Example: For $S \subseteq [n]$, define $\sum(S) = \sum_{i \in S} i$.

Consider the expansion of $(1+yx)(1+y^2x)\cdots(1+y^nx)$.

Any subset $S \subseteq [n]$, can be matched to a term in the above expression the same way as before. Specifically, S will be matched to $y^N x^M$ where $N = \sum(S)$ and $M = |S|$.

Thus, $(1+yx)(1+y^2x)\cdots(1+y^nx) = \sum a_{i,j} x^i y^j$ where $a_{i,j}$ is the number of subsets of $[n]$ with size i whose sum adds up to j.

An equivalent form of the binomial theorem can be gotten as follows:

Given $(1+x)^n = \sum_{i=0}^{n} \binom{n}{i} x^i$ where n is a positive integer, substitute $\frac{x}{y}$ for x.

Then the expression becomes $(1+\frac{x}{y})^n = \sum_{i=0}^{n} \binom{n}{i} x^i y^{-i}$. And, if we multiply this all by y^n, we thus get:

$$(y+x)^n = \sum_{i=0}^{n} x^i y^{n-i}$$

To further generalize this, let $k_1, \ldots, k_d \geq 0$ be integers and $n = k_1 + \ldots + k_d$. We define:

$$\binom{n}{k_1, \ldots, k_d} := \frac{n!}{k_1! \cdots k_d!}$$

Multinomial Theorem: Let $x_1, \ldots, x_d$ be variables. Then:

$$(x_1 + \ldots + x_d)^n = \sum_{\substack{k_1,\ldots,k_d \in \mathbb{Z}_{\geq 0} \\ k_1+\ldots+k_d=n}} \binom{n}{k_1,\ldots,k_d} x_1^{k_1} \cdots x_d^{k_d}$$

Proof:
If $d = 1$, then this theorem holds because $\binom{n}{n} x_1^n = x_1^n$.

Now assume the theorem is known for $d-1$ variable. Then substitute $x = x_1 + \ldots + x_{d-1}$ and $y = x_d$ into our rewrite of the binomial theorem above to get that:

$$\begin{aligned}((x_1 + \ldots + x_{d-1}) + x_d)^n &= \sum_{i=1}^{n} \binom{n}{i} (x_1 + \ldots + x_{d-1})^i x_d^{n-i} \\ &= \sum_{i=1}^{n} \binom{n}{i} \left(\sum_{\substack{k_1,\ldots,k_{d-1} \in \mathbb{Z}_{\geq 0} \\ k_1+\ldots+k_{d-1}=i}} \binom{i}{k_1,\ldots,k_{d-1}} x_1^{k_1} \cdots x_{d-1}^{k_{d-1}} \right) x_d^{n-i}\end{aligned}$$

If we set $k_d = n - i$, this expression becomes:

$$(x_1 + \ldots + x_{d-1} + x_d)^n = \sum_{\substack{k_1,\ldots,k_d \in \mathbb{Z}_{\geq 0} \\ k_1+\ldots+k_d=n}} \binom{n}{k_1,\ldots,k_d} x_1^{k_1} \cdots x_d^{k_d}$$

One choice problem we can model with the multinomial theorem is as follows:

Suppose we have d types of objects. Then $\binom{n}{k_1,\ldots,k_d}$ is the number of ways to arrange n objects such that exactly k_i of them have the ith type.

Consider the problem of picking multisets of $[n]$.

Consider the expansion of the formal power series $(1 + x + x^2 + \ldots)^n$. Given any multiset S of $[n]$, we can associate it with a term in that expansion as follows:

At the ith step of multiplying in a term, choose x^k where k is the number of times i appears in S.

Thus, $(1 + x + x^2 + \ldots)^n = \sum_{k \geq 0}$(# of multisets of $[n]$ of size k)x^k.

But note that $(1 + x + x^2 + \ldots) = \frac{1}{1-x}$. So, $(1 + x + x^2 + \ldots)^n = (1 - x)^{-n}$, meaning that:

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

Lecture 7: 10/11/2024

Rational Generating Functions:

A formal power series $F(x)$ is a rational function if there exists polynomials $P(x)$ and $Q(x)$ with $Q(x) \neq 0$ such that $Q(x)F(x) = P(x)$.

Note that $P(x)$ and $Q(x)$ are not unique because we can multiply and divide them by common polynomials and the relation still holds. However, a fact that's beyond the scope of this class is that because polynomials have unique factorizations, we know that any other $P'(x)$ and $Q'(x)$ will be off of $P(x)$ by $Q(x)$ only by some factor $R(x)$.

Define $\deg F(x) = \deg P(x) - \deg Q(x)$

Theorem: Suppose $F(x) = \sum_{n \geq 0} a_n x^n$.

Write $Q(x) = 1 + c_1 x + \ldots + c_r x^r = (1 - \gamma_1 x)^{m_1} \cdots (1 - \gamma_s x)^{m_s}$ where $\gamma_1^{-1}, \ldots, \gamma_s^{-1}$ are the roots of $Q(x)$. (We can do this because we know 0 isn't a root of $Q(x)$ on account of $Q(x)$ having a nonzero constant term). Then given $N \geq 0$ the following are equivalent:

(a) For all $n \geq N, \quad a_{n+r} + c_1 a_{n+r-1} + \ldots + c_r a_n = 0.$

(b) $Q(x)F(x)$ is a polynomial of degree less than $N+r$.

(c) There exists polynomials $f_1, \ldots, f_s$ with $\deg f_i < m_i$ such that $a_n = \sum_{i=1}^{s} f_i(n)\gamma_i^n$. for all $n \geq N$.

Proof:

(a) $\iff$ (b)

For any $n \geq N$, note that:

$$[x^{n+r}]Q(x)F(x) = a_{n+r} + c_1 a_{n+r-1} + \ldots + c_r a_n.$$

Thus, assuming (a), we'd have that $[x^{n+r}]Q(x)F(x) = 0$ for all $n \geq N$. So, $Q(x)F(x)$ is a polynomial of degree less than $N+r$.

Conversely, assuming (b), then we know that $[x^{n+r}]Q(x)F(x) = 0$ for all $n \geq N$. This is equivalent to (a).

(b) $\implies$ (c)

Assume that (b) holds and write $F(x) = \frac{P(x)}{Q(x)}$ where $\deg P(x) < N+r$. Then, doing polynomial long division, we can find polynomials $g(x)$ and $P_0(x)$ where $g(x)$ is a polynomial of degree less than N, $P_0(x)$ is a polynomial of degree less than r, and $F(x) = g(x) + \frac{P_0(x)}{Q(x)}$.

Doing partial fractions, we get that: $\frac{P_0(x)}{Q(x)} = \sum_{i=1}^{s} \frac{p_i(x)}{(1-\gamma_i x)^{m_i}}$ where $\deg p_i(x) < m_i$.

Now for the next bit of the proof, consider any term in $\frac{p_i(x)}{(1-\gamma_i x)^{m_i}}$ of the form $\frac{ax^d}{(1-\gamma x)^m}$.

By the binomial theorem, if d is an integer less than m, then:

$$\begin{aligned}\frac{ax^d}{(1-\gamma x)^m} &= ax^d \sum_{n\geq 0} \binom{m+n-1}{n}\gamma^n x^n \\ &= a\sum_{n\geq 0}\binom{m+n-1}{n}\gamma^n x^{n+d} = a\sum_{n\geq d}\binom{m+n-d-1}{n-d}\gamma^{n-d}x^n\end{aligned}$$

But note that when $n \geq d$, then:

$$\binom{m+n-d-1}{n-d} = \frac{1}{(m-1)!}(m+n-d-1)\cdots(n-d+2)(n-d+1)$$

is a polynomial in n of degree less than or equal to $m-1$.

Also, if $n \in \{d-1, d-2, \ldots, d-m+1\}$, then this polynomial equals 0. Thus, since $d \leq m-1$, we can say that:

$$\frac{ax^d}{(1-\gamma x)^m} = \sum_{n\geq 0} f(n)\gamma^n x^n$$

where $f(n) = \frac{a}{\gamma^d (m-1)!}(m+n-d-1)\cdots(n-d+2)(n-d+1)$ is a polynomial of degree at most $m-1$.

Doing this for every term in $\frac{p_i(x)}{(1-\gamma_i x)^{m_i}}$ and adding the polynomials together, we thus can show that $\frac{p_i(x)}{(1-\gamma_i x)^{m_i}} = \sum_{n\geq 0} f_i(n)\gamma^n x^n$ where $f_i(n)$ is a polynomial of degree less than m_i.

In turn, we have that $\frac{P_0(x)}{Q(x)} = \sum_{n\geq 0}\left(\sum_{i=1}^{s} f_i(n)\gamma_i^n\right) x^n$ where each f_i is a polynomial of degree less than m_i. Thus the conclusion of (c) holds for all $n \geq N$

since $[x^n]g(x) = 0$ for all $n \geq N$.

(c) $\Longrightarrow$ (b)

Assume there exists polynomials $f_1, \ldots, f_s$ with $\deg f_i < m_i$ such that $a_n = \sum_{i=1}^{s} f_i(n)\gamma_i^n$ for all $n \geq N$.

Then firstly, define $g_0(x) = F(x) - \sum_{n\geq 0}\left(\sum_{i=1}^{s} f_i(n)\gamma_i^n\right) x^n$. Note that g_0 will be a polynomial of degree less than N satisfying that:

$$F(x) = g_0(x) + \sum_{n\geq 0}\left(\sum_{i=1}^{s} f_i(n)\gamma_i^n\right) x^n$$

Next, observe that $a_n = \sum_{i=1}^{s} f_i(n)\gamma_i^n$ satisfies the relation:

$$a_{n+r} + c_1 a_{n+r-1} + \ldots + c_r a_n \text{ for all } n \geq 0.$$

This is because by manipulating $Q(x)$, we can see that the characteristic polynomial of the above recurrence relation has γ_1 throught γ_s as its roots with multiplicities m_1 through m_s respectively.

Since we already showed that (a) implies (b), we thus know that $Q(x)\sum_{n\geq 0}\left(\sum_{i=1}^{s} f_i(n)\gamma_i^n\right) x^n$ is a polynomial $P_0(x)$ of order less than r.

Finally, we now have the expression $F(x) = g_0(x) = \frac{P_0(x)}{Q(x)}$. Multiplying both sides by $Q(x)$, we get that $Q(x)F(x) = g_0(x)Q(x) + P_0(x)$ which is a polynomial of degree less than $N + r$.

Special Case: Let $Q(x) = (1-x)^r = \sum_{i=0}^{r}\binom{r}{i}(-1)^i x^i$. Then pick an integer $N \geq 0$. The following are equivalent:

(a) $\forall n \geq N,\ \sum_{i=0}^{r}(-1)^{r-i}\binom{r}{i}a_{n+i} = 0$

(b) $(1-x)^r F(x)$ is a polynomial of degree less than $N + r$.

(c) There exists a polynomial f with degree less than r such that $a_n = f(n)$ for all $n \geq N$.

Example: Given $a = (a_1, a_2, a_3), b = (b_1, b_2, b_3) \in (\mathbb{Z}_{\geq 0})^3$, let us write $a \leq b$ if $a_i \leq b_i$ for $i = 1, 2, 3$. Fixing a and letting $|b| = b_1 + b_2 + b_3$, we define:

$$F_a(x) = \sum_{b \geq a} x^{|b|} = \frac{x^{|a|}}{(1-x)^3}$$

The second equality comes from the fact that the generating function representing how we are choosing $b \geq a$ is:

$$(x^{a_1}(1 + x + x^2 + \ldots))(x^{a_2}(1 + x + x^2 + \ldots))(x^{a_3}(1 + x + x^2 + \ldots))$$

Applying the previous special case, if $N = |a| - 2$, then we've shown that $(1-x)^4 F_a(x)$ is a polynomial of degree less than $N + r$. Thus, there exists a quadratic in n which equals the number of $b \in (\mathbb{Z}_{\geq 0})^3$ satisfying that $b \geq a$ and $|b| = n$ given some $n \geq |a| - 2$.

Now suppose we are given another $a' = (a'_1, a'_2, a'_3) \in (\mathbb{Z}_{\geq 0})^3$, and fix $a'' = (\max(a_1, a'_1), \max(a_2, a'_2), \max(a_3, a'_3))$. Then:

$$\sum_{\substack{b \in (\mathbb{Z}_{\geq 0})^3 \\ b \geq a \text{ or } b \geq a'}} x^{|b|} = F_a(x) + F_{a'}(x) - F_{a''}(x)$$

Lecture 8: 10/14/2024

Catalan Numbers

We define the nth Catalan number C_n to be equal to the number of ways to list n pairs of parentheses such that they are balanced.

A string of pairs of parentheses are balanced if for all $i \leq 2n$, the number of open parentheses: "(", in the first i positions is greater than or equal to the number of closed parentheses: ")" in the first i positions.

Define $C(x) = \sum_{n \geq 0} C_n x^n$.

Lemma: If $n > 0$, then $C_n = \sum_{i=0}^{n-1} C_i C_{n-1-i}$.

Proof:
Every balanced set starts with "(". Consider it's matching ")". Between them is another set of balanced parentheses (which could be empty). Also, to the right of the ")" there is another set of balanced parentheses (which could be empty).

If the set of inside parentheses has i pairs, then the set of parentheses to the right will have $(n-1-i)$ pairs. And since we can choose any $i \in \{1, \ldots, n-1\}$ and also since both the inside and right sets can be chosen independently, it follows that:

$$C_n = \sum_{i=0}^{n-1} C_i C_{n-1-i}$$

Note that $[x^{n-1}]C(x)^2 = \sum_{i=0}^{n-1} C_i C_{n-1-i}$. Thus:

$$\begin{aligned} C(x) &= 1 + \sum_{n\geq 1}\left(\sum_{i=0}^{n-1} C_i C_{n-1-i}\right)x^n \\ &= 1 + x\sum_{n\geq 1}\left(\sum_{i=0}^{n-1} C_i C_{n-1-i}\right)x^{n-1} = 1 + xC(x)^2 \end{aligned}$$

Hence, we know $C(x)$ is a solution to the quadratic polynomial $xt^2 - t + 1 = 0$. Using the quadratic formula, we deduce that $C(x)$ is a solution to $2xt = 1 \pm \sqrt{1-4x}$.

Note that $2x$ doesn't have a multiplicative inverse. So we can't technically divide by it.

To figure out which solution $C(x)$ equals, let's plug in values of x. Note that when $x = 0$, the left side equals 0 while the right side equals 1 ± 1. Thus, it follows that:

$$2xC(x) = 1 - \sqrt{1-4x}$$

Next, we can apply the binomial theorem to get that:

$$1 - (1-4x)^{1/2} = 1 - \sum_{n\geq 0}\binom{1/2}{n}(-4x)^n = -\sum_{n\geq 1}\binom{1/2}{n}(-1)^n 4^n x^n.$$

Note that:

$$-(-1)^n 4^n \binom{1/2}{n} = -(-1)^n 4^n \frac{\frac{1}{2}\cdot\frac{-1}{2}\cdots\frac{-(2n-3)}{2}}{n!} = 2^{2n}\frac{\frac{1}{2}\frac{1}{2}\cdot\frac{3}{2}\cdots\frac{(2n-3)}{2}}{n!} = 2^n\frac{(2n-3)!!}{n!}$$

Then note that $(2n-2)!!(2n-3)!! = (2n-2)!$ and $2^{n-1}\frac{1}{(2(n-1))!!} = \frac{1}{(n-1)!}$. Therefore,

$$2^n\frac{(2n-3)!!}{n!} = 2\frac{(2n-2)!}{n!(n-1)!} = \frac{2}{n}\cdot\frac{(2n-2)!}{(n-1)!(n-1)!} = \frac{2}{n}\binom{2n-2}{n-1}$$

Hence, $2xC(x) = \sum_{n\geq 1}\frac{2}{n}\binom{2n-2}{n-1}x^n$. This implies:

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

So we conclude that:

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

Remarks:

1. In the last lecture, we showed that the coefficients of a formal power series eventually satisfy a linear reccurence relation if and only if the formal power series is a rational function. However, it's possible to show that $C(x)$ is not a rational formal power series. Therefore, the coeffiecients of $C(x)$ do not ever eventually satisfy a linear recurrence relation.

   A brief proof that $C(x)$ is irrational goes as follows:
   Suppose $Q(x)C(x) = P(x)$ where $Q(x)$ and $P(x)$ are polynomials. Then:

   $$Q(x) - Q(x)\sqrt{1-4x} = 2xP(x).$$

Rearranging terms and squaring everything, we thus get that:

$$(Q(x) - 2xP(x))^2 = Q^2(x)(1 - 4x)$$

But now we have a contradiction because the polynomial on the left side will have even degree and the polynomial on the right side will have odd degree.

2. Since $C(x)$ is the root of a polynomial, we say $C(x)$ is an algebraic function.

Other examples of usages of Catalan numbers:

(a) The number of ways to apply a binary operation $*$ to $n + 1$ arguments.

(b) The number of rooted binary trees with $n + 1$ leaves.

Lecture 9: 10/16/2024

Homework 1:

(1) Find a closed formula for the following recurrence relation:

$$a_0 = 1, \quad a_1 = 0, \quad a_2 = 2,$$
$$a_n = 5a_{n-1} - 8a_{n-2} + 4a_{n-3} \quad \text{for } n \geq 3$$

The characteristic polynomial of this relation is $t^3 - 5t^2 + 8t - 4$.

Because I wanted to trust that the professor wouldn't give us a messy polynomial, I used the rational root theorem to get a list of candidate roots to test. Those candidates are ± 1, ± 2, and ± 4.

After testing, I found that $(t-1)(t-2) = t^2 - 3t + 2$ is a factor of the characteristic polynomial. Doing polynomial long division, I then got that the other factor is $(t-2)$. So, our characteristic polynomial equals $(t-1)(t-2)^2$.

With that, we now know that $a_n = \beta_1 + \beta_2 2^n + \beta_3 n 2^n$. Plugging in $n = 0$, 1, and 2 respectively, we get the following system of equations:

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 2 \\ 1 & 4 & 8 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$$

To solve this, I wrote the following code:

```
C:\Users\jmill>py
Python 3.12.1 (tags/v3.12.1:2305ca5, Dec  7 2023, 22:03:25) [MSC v.1937 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license" for more information.
>>>
>>> import numpy as np
>>>
>>> M = np.array([[1, 1, 0], [1, 2, 2], [1, 4, 8]])
>>> y = np.array([[1], [0], [2]])
>>>
>>> print((np.linalg.inv(M))@y)
[[ 6.]
 [-5.]
 [ 2.]]
>>>
```

So, $a_n = 6 - (5 + 2n)2^n$.

(2) Let $(a_n)_{n\geq 0}$ and $(b_n)_{n\geq 0}$ be sequences. Assume that $(b_n)_{n\geq 0}$ satisfies a linear recurrence relation of order e. Then let $c_1, \ldots, c_d$ be scalars with $c_d \neq 0$ and assume that $(a_n)_{n\geq 0}$ satisfies:

$$a_n = c_1 a_{n-1} + \ldots + c_d a_{n-d} + b_n \text{ for all } n \geq d$$

Prove that $(a_n)_{n\geq 0}$ satisfies a linear recurrence relation of order $d + e$.

To start, let's show some smaller facts:

Observation 1: Suppose $a_n = c_1 a_{n-1} + \ldots + c_d a_{n-d} + p_m(n) r^n + F(n)$ where F is some arbitrary function, p_m is a polynomial of degree m and r is a nonzero constant. Then a_n satisfies a recurrence relation of degree $e = d + m + 1$ where $F'(n)$ is an arbitrary function $F(n)$ determined by m, r, and $F(n)$, and:

$$a_n = c'_1 a_{n-1} + \ldots + c'_e a_{n-e} + F'(n) \text{ for all } n \geq e.$$

Proof by induction:

Lemma: If $p_m(n)$ is a polynomial of degree $m > 0$, then $p_m(n) - p_m(n-1)$ is a polynomial of degree $m-1$.

This is because for all $N > 0$ and coefficients b, we have that: $bn^N - b(n-1)^N = bn^N - bn^N + q(n)$ where q is a polynomial of degree $N-1$. Meanwhile, the case where $N = 0$ is trivial.

Base Case: If $m = 0$, meaning $p_m(n) = b$ where b is a constant, then by taking the difference of a_n and ra_{n-1} for all $n \geq d+1$, we get that:

$$a_n = (c_1 + r)a_{n-1} + (c_2 - rc_1)a_{n-2} + \ldots + (c_d - rc_{d-1})a_{n-d} - rc_d a_{n-d-1} + F(n) - rF(n-1).$$

Setting $c'_1 = c_1 + 1, c'_2 = c_2 - rc_1, \ \ldots, \ c'_d = c_d - rc_{d-1}, c'_{d+1} = -rc_d$, and $F'(n) = F(n) - rF(n-1)$, we thus get that:

$$a_n = c'_1 a_{n-1} + \ldots + c'_{d+1}a_{n-d-1} + F'(n) \text{ for all } n \geq d+1.$$

Also $c'_{d+1} \neq 0$ because neither r nor c_d equal 0.

Induction on m: If $m > 0$, then by taking the difference of a_n and ra_{n-1} for all $n \geq d+1$, since $r^n(p(n)) - rr^n(p(n-1)) = r^n(p(n) - p(n-1))$, we get by our lemma above that:

$$a_n = (c_1 + r)a_{n-1} + (c_2 - rc_1)a_{n-2} + \ldots + (c_d - rc_{d-1})a_{n-d} - rc_d a_{n-d-1} + q(n)r^n + F'(n)$$

where q is a polynomial of degree $m-1$ and $F'(n) = F(n) - rF(n-1)$.

And same as before, $-rc_d \neq 0$ because $r \neq 0$ and $c_d \neq 0$.

But now we can conclude by induction that a_n satisfies a (possibly inhomogeneous) recurrence relation of order $e = (d+1) + ((m-1)+1) = d+m+1$ such that $F''(n)$ is some function determined by m, r and $F(n)$, and:

$$a_n = c''_1 a_{n-1} + \ldots + c''_e a_{n-e} + F''(n) \text{ for all } n \geq e.$$

One important observation from above is that if $F(n) = 0$, then $F'(n) = 0$. In some other situations, $F'(n)$ also behaves nicely.

Observation 2: Let $p_1(n), \ldots, p_k(n)$ be polynomials of degree $m_1, \ldots, m_k$ respectively. Also let $r_1, \ldots r_k$ be distinct nonzero constants. If:

$$a_n = c_1 a_{n-1} + \ldots + c_d a_{n-d} + p_1(n)r_1^n + F(n)$$

where $F(n) = \sum_{i=2}^{k} p_i(n)r_i^n$, then part 1 will make $F'(n) = \sum_{i=2}^{k} q_i(n)r_i^n$ where $q_2(n), \ldots, q_k(n)$ are also polynomials of degree $m_2, \ldots, m_k$ respectively.

To see why, note that:

$$F(n) - r_1F(n-1) = \sum_{i=2}^{k} p_i(n)r_i^n - \sum_{i=2}^{k} p_i(n-1)r_1r_i^{n-1}$$
$$= \sum_{i=2}^{k} (p_i(n) - \tfrac{r_1}{r_i}p_i(n-1))r_i^n$$

Because $r_i \neq r_1$, we know $\frac{r_1}{r_i} \neq 1$, meaning that the degree m term of $p_i(n) - \frac{r_1}{r_i}p_i(n-1)$ doesn't cancel. So $q_i(n) := p_i(n) - \frac{r_1}{r_i}p_i(n-1)$ is still a degree m polynomial. If the process in part 1 takes more steps, then we can just repeat this reasoning.

Combining observations 1 and 2 together, we can inductively show that if

$$a_n = c_1 a_{n-1} + \ldots + c_d a_{n-d} + \sum_{i=1}^{k} p_i(n) r_i^n$$

where $r_1, \ldots, r_k$ are distinct nonzero constants and $p_1(n), \ldots, p_k(n)$ are polynomials of degree $m_1, \ldots, m_k$ respectively, then letting $e = \sum_{i=1}^{k}(m_i + 1)$, there exists constants $c'_1, \ldots c'_{d+e}$ such that:

$$a_n = c'_1 a_{n-1} + \ldots + c'_{e+d} a_{n-d-e}$$

But now note that if $(b_n)_{n\geq 0}$ satisfies a linear recurrence relation of order e, we can write $b_n = \sum_{i=1}^{k} p_i(n) r_i^n$ for some polynomials $p_1(n), \ldots p_k(n)$ with degrees $m_1, \ldots, m_k$, as well as some distinct nonzero constants $r_1, \ldots, r_k$.

We know that each r_i is nonzero because the constant term in the characteristic polynomial for (b_n)'s recurrence relation must be nonzero.

Also, as we showed in class, $\sum_{i=1}^{k}(m_i + 1) = e =$ the order of the recurrence relation of $(b_n)_{n\geq 0}$.

So, we've shown that $a_n = c_1 a_{n-1} + \ldots + c_d a_{n-d} + b_n$ can be rewritten as a homogenous linear recurrence relation of order $(d + e)$.

(3) Let $(f_n)_{n\geq 0}$ be the Fibonacci numbers, and define $a_n = \sum_{i=0}^{n} f_i$.

(a) Find a linear recurrence relation of order 3 that $(a_n)_{n\geq 0}$ satisfies.

Note that $(a_n)_{n\geq 0}$ satisfies the relation $a_n = a_{n-1} + f_n$ for all a_n. Also, we showed in the first lecture that:

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

So we know that: $a_n = a_{n-1} + \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n$.

Firstly, taking the difference of a_n and $\frac{1+\sqrt{5}}{2}a_{n-1}$, after a lot of simplifying we have that:

$$\begin{aligned} a_n &= \left(1 + \tfrac{1+\sqrt{5}}{2}\right)a_{n-1} - \tfrac{1+\sqrt{5}}{2}a_{n-2} + \left(-\tfrac{1}{\sqrt{5}} + \tfrac{1+\sqrt{5}}{2\sqrt{5}} \cdot \tfrac{2}{1-\sqrt{5}}\right)\left(\tfrac{1-\sqrt{5}}{2}\right)^n \\ &= \tfrac{3+\sqrt{5}}{2}a_{n-1} - \tfrac{1+\sqrt{5}}{2}a_{n-2} + \left(\tfrac{1-\sqrt{5}}{2}\right)^{n-1} \end{aligned}$$

Secondly, we take the difference of a_n and $\frac{1-\sqrt{5}}{2}a_{n-1}$, and after a lot more simplifying get:

$$\begin{aligned} a_n &= \left(\tfrac{3+\sqrt{5}}{2} + \tfrac{1-\sqrt{5}}{2}\right) a_{n-1} - \left(\tfrac{1+\sqrt{5}}{2} + \tfrac{1-\sqrt{5}}{2} \cdot \tfrac{3+\sqrt{5}}{2}\right) a_{n-2} + \left(\tfrac{1-\sqrt{5}}{2} \cdot \tfrac{1+\sqrt{5}}{2}\right) a_{n-3} \\ &= 2a_{n-1} + 0a_{n-2} - a_{n-3} \end{aligned}$$

(b) Find a closed formula for a_n.

Method 1:

The characteristic polynomial of $a_n = 2a_{n-1} - a_{n-3}$ is $t^3 - 2t^2 + 1$. Just by looking at it, I can already see that $(t-1)$ is a factor of that polynomial. So after doing polynomial long division, we have that $(t-1)(t^2-t-1) = t^3 - 2t^2 + 1$.

By quadratic formula, the remaining roots are $\frac{1+\sqrt{5}}{2}$ and $\frac{1-\sqrt{5}}{2}$ (a.k.a the same roots as with the Fibonacci recurrence relation).

Finally, we get a system of linear equations:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & \left(\frac{1+\sqrt{5}}{2}\right)^2 & \left(\frac{1-\sqrt{5}}{2}\right)^2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$$

To solve this, I finally started learning sympy:

```
C:\Users\jmill\Desktop\py test> virt_math_test\Scripts\activate.bat

(virt_math_test) C:\Users\jmill\Desktop\py test>py
Python 3.12.1 (tags/v3.12.1:2305ca5, Dec  7 2023, 22:03:25) [MSC v.1937 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license" for more information.
>>>
>>> from sympy.matrices import Matrix
>>> from sympy import sqrt
>>> from sympy import pprint
>>> from sympy import simplify
>>>
>>> M = Matrix([[1, 1, 1], [1, ((1 + sqrt(5)) / 2), ((1 - sqrt(5)) / 2)], [1, ((1 + sqrt(5)) / 2
)**2, ((1 - sqrt(5)) / 2)**2] ])
>>> y = Matrix([[0], [1], [2]])
>>>
>>> x = simplify(simplify(M ** -1) * y)
>>>
>>> pprint(x)
[    -1    ]
|          |
|1   3·√5|
|- + ----|
|2    10 |
|          |
|1   3·√5|
|- - ----|
[2    10 ]
>>>
>>> simplify(M * x) == y
True
>>>
>>> exit()

(virt_math_test) C:\Users\jmill\Desktop\py test> virt_math_test\Scripts\deactivate.bat
C:\Users\jmill\Desktop\py test>
```

So, assuming I've not made a silly error somewhere, we should have that:

$$a_n = -1 + \left(\tfrac{5+3\sqrt{5}}{10}\right)\left(\tfrac{1+\sqrt{5}}{2}\right)^n + \left(\tfrac{5-3\sqrt{5}}{10}\right)\left(\tfrac{1-\sqrt{5}}{2}\right)^n$$

Method 2: (The hinted route)

Note that $\frac{1-r^{n+1}}{1-r} = \sum\limits_{i=0}^{n} r^i$. Using this fact, we can see that:

$$\begin{aligned} a_n = \sum_{i=0}^{n} f_i &= \tfrac{1}{\sqrt{5}} \sum_{i=0}^{n} \left(\tfrac{1+\sqrt{5}}{2}\right)^i - \tfrac{1}{\sqrt{5}} \sum_{i=0}^{n} \left(\tfrac{1-\sqrt{5}}{2}\right)^i \\ &= \tfrac{1}{\sqrt{5}} \cdot \tfrac{1}{1-\frac{1+\sqrt{5}}{2}} \left(1 - (\tfrac{1+\sqrt{5}}{2})^{n+1}\right) - \tfrac{1}{\sqrt{5}} \cdot \tfrac{1}{1-\frac{1-\sqrt{5}}{2}} \left(1 - (\tfrac{1-\sqrt{5}}{2})^{n+1}\right) \end{aligned}$$

Now technically we are done since that is a closed formula for a_n. However, it looks ugly. So I'm going to learn more sympy so it can symplify this:

```
>>>
>>> from sympy import Symbol
>>> from sympy import sqrt
>>> from sympy import simplify
>>> from sympy import expand
>>> from sympy import pprint
>>>
>>> n = Symbol('n')
>>> p1 = (1 / sqrt(5)) * (1 / (1 - ((1+sqrt(5))/2) )) * (1 - (((1 + sqrt(5))/2) ** (n + 1)))
>>> p2 = (1 / sqrt(5)) * (1 / (1 - ((1-sqrt(5))/2) )) * (1 - (((1 - sqrt(5))/2) ** (n + 1)))
>>>
>>> pprint(simplify(expand(p1 - p2)))
 -n ⎛        n                 n                n                n                  n⎞
2  ⋅⎝- 10⋅2  - 3⋅√5⋅(1 - √5)  + 5⋅(1 - √5)  + 5⋅(1 + √5)  + 3⋅√5⋅(1 + √5) ⎠
─────────────────────────────────────────────────────────────────────────────
                                     10
>>>
```

Hence we get the same answer as before:

$$a_n = -1 + \left(\tfrac{5+3\sqrt{5}}{10}\right)\left(\tfrac{1+\sqrt{5}}{2}\right)^n + \left(\tfrac{5-3\sqrt{5}}{10}\right)\left(\tfrac{1-\sqrt{5}}{2}\right)^n$$

(4)

(a) Suppose that $(a_n)_{n\geq 0}$ and $(a'_n)_{n\geq 0}$ both satisfy the same linear recurrence relation of order d and that they agree in d consecutive places, i.e. there exists k such that $a_k = a'_k$, $a_{k+1} = a'_{k+1}$, ..., $a_{k+d-1} = a'_{k+d-1}$. Then these sequences are the same.

This is because for all integers $N \geq 0$ and sequences $(b_n)_{n\in\mathbb{Z}}$ satisfying a linear recurrence relation: $b_n = c_1 b_{n-1} + \ldots + c_d b_{n-d}$ for all integers $n \geq d$, we have that:

$$\begin{gathered} N + d \geq d \\ \Downarrow \\ b_{N+d} = c_1 b_{N+d-1} + \ldots + c_{d-1} b_{N+1} + c_d b_N \\ \Updownarrow \\ b_N = \tfrac{1}{c_d} b_{N+d} - \tfrac{c_1}{c_d} b_{N+d-1} - \ldots - \tfrac{c_{d-1}}{c_d} b_{N+1} \end{gathered}$$

Now suppose we know d consecutive elements of $(b_n)_{n\geq 0}$ (say b_k, b_{k+1}, ..., and b_{k+d-1} where k is a nonnegative integer). Then we can uniquely solve for b_n by inductively applying our recurrence relation when $n > k+d-1$. Plus, we can uniquely solve for b_n by inductively using the identity we found on the previous page when $n < k$. So $(b_n)_{n\geq 0}$ is uniquely determined by its consecutive elements b_k, b_{k+1}, ..., and b_{k+d-1}.

Since $(a_n)_{n\geq 0}$ and $(a'_n)_{n\geq 0}$ are uniquely determined by the same consecutive elements, we thus know that the sequences are the same.

(b) Suppose that $(a_n)_{n\geq 0}$ satisfies the linear recurrence relation of order d: $a_n = c_1 a_{n-1} + \ldots + c_d a_{n-d}$ for all $n \geq d$. Then there is a unique sequence $(b_n)_{n\in\mathbb{Z}}$ such that $b_n = a_n$ for $n \geq 0$ and $b_n = c_1 b_{n-1} + \ldots + c_d b_{n-d}$ for all $n \in \mathbb{Z}$.

If $N < 0$, we can still inductively apply the identity we found on the last page: $b_N = \frac{1}{c_d} b_{N+d} - \frac{c_1}{c_d} b_{N+d-1} - \ldots - \frac{c_{d-1}}{c_d} b_{N+1}$ in order to uniquely solve for b_N such that $b_{N+d} = c_1 b_{N+d-1} + \ldots + c_d b_N$.

Hence, given a sequence $(a_n)_{n\geq 0}$ satisfying a linear recurrence relation of order d: $a_n = c_1 a_{n-1} + \ldots + c_d a_{n-d}$, we define $b_n = a_n$ when $n \geq 0$. Meanwhile, when $n < 0$, we inductively define b_n as:

$$b_n = \frac{1}{c_d} b_{n+d} - \frac{c_1}{c_d} b_{n+d-1} - \ldots - \frac{c_{d-1}}{c_d} b_{n+1}.$$

Then $(b_n)_{n\in\mathbb{Z}}$ is the unique sequence satisfying the problem's requirements.

(c) Consider the Fibonacci sequence $(f_n)_{n\geq 0}$. How does the negatively indexed Fibonacci sequence relate to the usual one?

If $f_{n+2} = f_{n+1} + f_n$, then we know $f_n = -f_{n+1} + f_{n+2}$. Defining $g_n = f_{-n}$, we thus get the recurrence relation: $g_n = -g_{n-1} + g_{n-2}$, and its characteristic polynomial is $t^2 + t - 1$.

Now let r_1 and r_2 be the roots of $t^2 - t - 1$, the characteristic polynomial of f_n. Then subbing in $t = (-s)$, we get that:

$$t^2 - t - 1 = (t - r_1)(t - r_2) \implies s^2 + s - 1 = (s + r_1)(s + r_2)$$

So, there exists constants α_1, α_2 such that:

$$g_n = \alpha_1 (-1)^n \left(\frac{1+\sqrt{5}}{2}\right)^n + \alpha_2 (-1)^n \left(\frac{1-\sqrt{5}}{2}\right)^n.$$

Since $g_1 = f_{-1} = -f_0 + f_1 = 1$ and $g_0 = f_0 = 0$, by plugging in $n = 0$ and $n = 1$, we get the following matrix equation:

$$\begin{bmatrix} 1 & 1 \\ -\frac{1+\sqrt{5}}{2} & -\frac{1-\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

After solving that, we get that $\alpha_1 = \frac{-1}{\sqrt{5}}$ and $\alpha_2 = \frac{1}{\sqrt{5}}$ So in conclusion:

$$f_{-n} = g_n = \tfrac{1}{\sqrt{5}}(-1)^{n+1}\left(\tfrac{1+\sqrt{5}}{2}\right)^n - \tfrac{1}{\sqrt{5}}(-1)^{n+1}\left(\tfrac{1-\sqrt{5}}{2}\right)^n = (-1)^{n+1}f_n$$

I could have probably proven the identity $f_{-n} = (-1)^{n+1}f_n$ much faster by just doing induction and assuming $f_{-N} = (-1)^{N+1}f_N$ for all $N \in \{0, 1, \ldots, n-1\}$.

(5) Let p be a prime number and let $(a_n)_{n\geq 0}$ be a sequence such that $a_n \in \mathbb{Z}_p$ and which satisfies a homogeneous linear recurrence relation. Prove that the sequence is periodic, i.e. there exists N such that $a_n = a_{n+N}$ for all $n \geq 0$.

Consider the set $\{(a_n, a_{n+1}, \ldots, a_{n+d-1}) \mid n \in \mathbb{Z}$ and $n \geq 0\}$. Then note that it is a subset of $(\mathbb{Z}_p)^d$ which is finite with p^d elements. Hence, given some $(\beta_0, \ldots, \beta_d)$ in that set, we know there exists infinitely many $n \geq 0$ such that:

$$(a_n, a_{n+1}, \ldots, a_{n+d-1}) = (\beta_0, \beta_1, \ldots, \beta_{d-1}).$$

Let S be the set of all such n. Then pick n_1 and n_2 to be the least and second least elements of S respectively. We claim $(a_n)_{n\geq 0}$ is periodic with period $N = n_2 - n_1$.

Firstly, we'll prove by induction that $a_n = a_{N+n}$ when $n \geq n_1$. For our base case, we know from how we picked n_1 and n_2 that $a_n = a_{N+n}$ when $n_1 \leq n < n_1 + d$. Meanwhile, suppose that given some $k \geq d$ we have that $a_n = a_{N+n}$ for all $n_1 \leq n < n_1 + k$. Then using our recurrence relation, when we solve for a_k and a_{k+N}, we will get that they are equal. Hence, we know by induction that $a_n = a_{n+N}$ for all $n \geq n_1$.

Next, consider that if $a_n = c_1 a_{n-1} + \ldots + c_d a_{n-d}$ for all $n \geq d$, then for all $n \geq 0$ we have that: $a_n = \frac{1}{c_d}a_{n+d} - \frac{c_1}{c_d}a_{n+d-1} - \ldots - \frac{c_{d-1}}{c_d}a_{n+1}$.

This expression is still well-defined in $\mathbb{Z}_p$ because $c_d \neq 0$ and all nonzero elements in $\mathbb{Z}_p$ have a multiplicative inverse since p is prime.

Thus, we can proceed by induction to show that $a_n = a_{N+n}$ for all $n \geq 0$. Our base case is that if $n \geq n_1$, we know from before that $a_n = a_{n+N}$. Meanwhile, suppose that given some $k < d$ we have that $a_n = a_{n+N}$ for all $n > k$. Then when we use the expression above to calculate a_k and a_{k+N}, we will get that they equal each other. Hence, we know by induction that $a_n = a_{n+N}$ for all nonnegative integers n.

Homework 2:

(1) Let $A(x)$ be a formal power series with $A(0) = 0$.

(a) Show that there exists a formal power series $B(x)$ with $B(0) = 0$ such that $A(B(x)) = x$ if and only if $[x^1]A(x) \neq 0$.

Let us write $A(x) = \sum_{n\geq 1} a_n x^n$ and $B(x) = \sum_{n\geq 1} b_n x^n$.

We start indexing at 1 because we know from the problem statement that $a_0 = 0$, and $A(B(x))$ is not defined unless $b_0 = 0$.

Note that $[x^1]A(B(x)) = a_1 b_1$. So if $a_1 = [x^1]A(x) = 0$, then we can't solve for b_1 such that $a_1 b_1 = 1$. On the other hand, if $a_1 \neq 0$ (a.k.a it has a multiplicative inverse), then we can uniquely fix $b_1 = {}^1/_{a_1}$.

After that, note that for any $n \in \mathbb{Z}_{\geq 2}$, we have that:

$$[x^n]A(B(x)) = a_1 b_n + f_{a_2,\ldots,a_n,b_1,\ldots,b_{n-1}}$$

where the end term is some expression of given coefficients and coefficients which we can calculate by induction. So, forcing $[x^n]A(B(x)) = 0$, we can uniquely determine b_n to be:

$$b_n = -\tfrac{1}{a_1} f_{a_2,\ldots,a_n,b_1,\ldots,b_{n-1}}$$

To better explain why the coefficients of $A(B(x))$ take on the form above, note that $\operatorname{mdeg} A(x) = 1 \implies \operatorname{mdeg} A(x)^n \geq n$. Also, if $B(x)$ with $B(0) = 0$ is raised to a power $m \geq 2$, then the first coefficient which $[x^n]B(x)$ will affect in $[x^n]B(x)^m$ is $[x^{m+n-1}]B(x)^m$.

It follows that by induction, there exists a unique formal power series $B(x)$ such that $A(B(x)) = x$.

(b) Assuming $[x^1]A(x) \neq 0$, show that $B(x)$ is unique and also satisfies $B(A(x)) = x$. You may use without proof that composition of formal power series is associative.

We know from before that $B(x)$ is unique. Meanwhile, note that if $(A \circ B)(x) = x$ and composition is associative, then:

$$(A \circ B)(A(x)) = A(x) \implies A((B \circ A)(x)) = A(x) \implies (B \circ A)(x) = x$$

Actually, to be fully rigorous I still need to show that:

$$A(B(x)) = A(C(x)) \implies B(x) = C(x)$$

I'll write this after question 5 since I didn't notice I needed to do this until rather late.

(2)

(a) Prove that $\sum\limits_{m,n\geq 0} \min(m,n)x^m y^n = \frac{xy}{(1-x)(1-y)(1-xy)}$.

Firstly, note that $\sum\limits_{m,n\geq 0} \min(m,n)x^m y^n = \sum\limits_{m\geq 0} \left(\sum\limits_{n=0}^{m} ny^n + \sum\limits_{n>m} my^n \right) x^m$.

Also, we have that:

- $\sum\limits_{n>m} my^n = my^{m+1} \sum\limits_{n\geq 0} y^n = \frac{my^{m+1}}{1-y}$.
- $$\begin{aligned}\sum_{n=0}^{m} ny^n &= y \sum_{n=0}^{m-1} (n+1)y^n \\ &= yD\left(\sum_{n=0}^{m} y^n\right) = yD\left(\frac{1-y^{m+1}}{1-y}\right) \\ &= y\left(\frac{-(m+1)y^m}{1-y} + \frac{1-y^{m+1}}{(1-y)^2}\right) = \frac{-my^{m+1}-y^{m+1}+my^{m+2}+y}{(1-y)^2}\end{aligned}$$

Thus, $\sum\limits_{n=0}^{m} ny^n + \sum\limits_{n>m} my^n = \frac{y-y^{m+1}}{(1-y)^2}$.

And therefore:

$$\begin{aligned}\sum_{m,n\geq 0} \min(m,n)x^m y^n &= \frac{y}{(1-y)^2} \sum_{m\geq 0} (1-y^m)x^m \\ &= \frac{y}{(1-y)^2} \sum_{m\geq 0} x^m - \frac{y}{(1-y)^2} \sum_{m\geq 0} x^m y^m \\ &= \frac{y}{(1-y)^2} \cdot \frac{1}{1-x} - \frac{y}{(1-y)^2} \cdot \frac{1}{1-xy} \\ &= \frac{-xy^2+xy}{(1-y)^2(1-x)(1-xy)} = \frac{xy}{(1-y)(1-x)(1-xy)}\end{aligned}$$

(b) Let $a_{m,n}$ be the number of paths in $\mathbb{R}^2$ from $(0,0)$ to (m,n) using steps of the form $\overrightarrow{(1,0)}$, $\overrightarrow{(0,1)}$, and $\overrightarrow{(2,1)}$. Prove that:

$$\sum_{m,n\geq 0} a_{m,n} x^m y^n = \frac{1}{1-x-y-x^2y}$$

Note that for $m \geq 2$ and $n \geq 1$, we have that:

$$a_{m,n} = a_{m-1,n} + a_{m,n-1} + a_{m-2,n-1}$$

Also, we can somewhat trivially calculate that $a_{m,0} = 1 = a_{0,n}$ if $m,n > 0$, that $a_{1,n} = n+1$ if $n \geq 0$, and that $a_{0,0} = 1$ (the empty path). Thus:

- $m \geq 1 \implies [x^m](1-x-y-x^2y)A(x) = a_{m,0} - a_{m-1,0} = 0$
- $n \geq 1 \implies [y^n](1-x-y-x^2y)A(x) = a_{0,n} - a_{0,n-1} = 0$
- $n \geq 1 \implies [xy^n](1-x-y-x^2y)A(x) = a_{1,n} - a_{0,n} - a_{1,n-1} = 0$

- $[x^0y^0](1-x-y-x^2y)A(x) = a_{0,0} = 1$.

Thus, $(1-x-y-x^2y)A(x) = 1$. And so $A(x) = \frac{1}{1-x-y-x^2y}$.

(3) Let $n \geq 2$ be an integer. Evaluate the following sums:

(a) $\sum_{i=0}^{n} i^2\binom{n}{i}$

Note that by taking derivatives of the binomial theorem, we get that:

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

So, since $i\binom{n}{i}x^i = i^2\binom{n}{i}x^i$ when $i=0$ or $i=1$, we can multiply everything by x^2 rearrange terms, and then add the missing terms to both sides to get that:

$$\begin{aligned}\sum_{i=0}^{n} i^2\binom{n}{i}x^i &= n(n-1)x^2(x+1)^{n-2} + x\sum_{i=0}^{n} i\binom{n}{i}x^{i-1} \\ &= n(n-1)x^2(x+1)^{n-2} + nx(x+1)^{n-1}\end{aligned}$$

Subbing in $x=1$, we thus get that:

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

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

Starting back off with the expression for $\sum_{i=0}^{n} i^2\binom{n}{i}x^i$ we found last time, we instead sub in $x=-1$ to get that:

$$\sum_{i=0}^{n} i^2\binom{n}{i}(-1)^i x^i = n(n-1)x^2(1-x)^{n-2} - nx(1-x)^{n-1}$$

Interestingly, if $n>2$, then when we plug in $x=1$, everything cancels and we get $\sum_{i=0}^{n} i^2\binom{n}{i}(-1)^i = 0$. Meanwhile, if $n=2$, then $(1-x)^{n-2}$ doesn't equal zero and we get that:

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

(c) $\sum\limits_{\substack{0 \le i \le n \\ i \text{ is even}}} i^2 \binom{n}{i}$

Combining the two previous sums, we get that:

$$\sum_{\substack{0 \le i \le n \\ i \text{ is even}}} i^2 \binom{n}{i} = \frac{1}{2}\sum_{i=0}^{n} i^2 \binom{n}{i} + \frac{1}{2}\sum_{i=0}^{n} i^2 \binom{n}{i}(-1)^i = \begin{cases} (n^2+n)2^{n-3} & \text{if } n > 2 \\ 4 & \text{if } n = 2 \end{cases}$$

(d) $\sum\limits_{\substack{0 \le i \le n \\ i \text{ is odd}}} i^2 \binom{n}{i}$

Combining the sum in part (a) and the sum in part (c), we get that:

$$\sum_{\substack{0 \le i \le n \\ i \text{ is odd}}} i^2 \binom{n}{i} = \sum_{i=0}^{n} i^2 \binom{n}{i} - \sum_{\substack{0 \le i \le n \\ i \text{ is even}}} i^2 \binom{n}{i} = \begin{cases} (n^2+n)2^{n-3} & \text{if } n > 2 \\ 2 & \text{if } n = 2 \end{cases}$$

(4)

(a) How many ways can we rearrange the letters of the word "SASSAFRAS"?

Something I'd briefly like to note is that an obvious upper bound for this number is $9! = 362880$. Now while this bound does overcount (by quite a lot) since it isn't taking into account repeating letters, it is a small enough that you could write a snippet of code to manually count up this quantity. In fact, I did write some code to do that and preemptively got back that the answer is 2520.

```
S = set()

bank = ['S', 'A', 'S', 'S', 'A', 'F', 'R', 'A', 'S']
for i1 in range(9):
 for i2 in range(8):
  for i3 in range(7):
   for i4 in range(6):
    for i5 in range(5):
     for i6 in range(4):
      for i7 in range(3):
       for i8 in range(2):
        temp = bank.copy()
        stringTemp = [temp.pop(i1), temp.pop(i2), temp.pop(i3),
        temp.pop(i4), temp.pop(i5), temp.pop(i6), temp.pop(i7),
        temp.pop(i8), temp.pop()]
        S.add(''.join(stringTemp))

print(len(S))
```

Now, consider the expansion of the polynomial: $(s+a+f+r)^9$. Given any string of nine letters: 's', 'a', 'f', 'r', we can associate that string with the term in the expansion gained by choosing to multiply in the ith letter of the string at step i.

It follows that $[s^4a^3fr](s+a+f+r)^9$ equals the number terms in the expansion associated with strings which rearrange the letters of "SASSAFRAS". And, we know that:

$$[s^4a^3fr](s+a+f+r)^9 = \binom{9}{4,3,1,1} = \frac{9!}{4!3!} = \frac{9\cdot 8\cdot 7\cdot 6\cdot 5}{3\cdot 2} = 9\cdot 8\cdot 7\cdot 5 = 2520$$

(b) How many ways can this be done if we also require that the 3 A's all appear consecutively?

After modifying my code from before, I preemptively know that the right answer is 210. So, whatever we do next, our final answer should match that.

Now forcing all three A's to be next to each other means that whenever we choose A from among our bank of possible letters, we have to add A three times. The generating function that models this pattern of choosing letters is:

$$(s+(a^3)+f+r)^7 = \sum \binom{7}{k_1,k_2,k_3,k_4} s^{k_1}(a^3)^{k_2} f^{k_3} r^{k_4}$$

Importantly, $[s^4a^3fr](s+a^3+f+r)^7$ will still give the number of ways to arrange the letters of "SASSAFRAS". However, now:

$$[s^4a^3fr](s+a^3+f+r)^7 = \binom{7}{4,1,1,1} = 7\cdot 6\cdot 5 = 210$$

(5)

(a) Let a, b be rational numbers. Show that for any formal power series $A(x)$ with $A(0)=1$, we have $A(x)^a A(x)^b = A(x)^{a+b}$.

To start, here's the proof that if n, m, c are integers with $c \neq 0$, then $A(x)^{\frac{m}{n}} = A(x)^{\frac{cm}{cn}}$:

$$(A(x)^{1/cn})^{cm} = (((A(x)^{1/cn})^{cm})^n)^{1/n} = (((A(x)^{1/cn})^{cn})^m)^{1/n} = (A(x)^m)^{1/n}$$

Now, suppose $a = \frac{m}{n}$ and $b = \frac{p}{q}$ where $m, n, p, q \in \mathbb{Z}$. Then:

$$\begin{aligned} A(x)^a A(x)^b &= (A(x)^{1/n})^m (A(x)^{1/q})^p \\ &= (A(x)^{1/nq})^{mq}(A(x)^{1/qn})^{pn} = (A(x)^{1/nq})^{mq+pn} = A(x)^{a+b} \end{aligned}$$

(b) Deduce from (a) that $\binom{a+b}{n} = \sum_{i=0}^{n} \binom{a}{i}\binom{b}{n-i}$ for all nonnegative integers n.

Note that:

$$\sum_{n\geq 0} \binom{a+b}{n} x^n = (1+x)^{a+b} = (1+x)^a(1+x)^b = \left(\sum_{i\geq 0}\binom{a}{i}x^i\right)\left(\sum_{j\geq 0}\binom{b}{j}x^j\right)$$

By applying the definition of Cauchy products, we thus get the desired identity.

Adendum to problem 1:

Suppose $A(x) = \sum_{n\geq 0} a_n x^n$, $B(x) = \sum_{n\geq 0} b_n x^n$, and $C(x) = \sum_{n\geq 0} c_n x^n$ are formal power series with $a_0 = b_0 = c_0 = 0$ satisfying that $A(B(x)) = A(C(x))$.

Note that we must have that $a_1 b_1 = a_1 c_1$. So, supposing a_1 has a multiplicative inverse, then $b_1 = c_1$.

After that, taking $[x^n]A(B(x))$ and $[x^n]A(C(x))$ for $n \geq 2$, we get that:

$$f(a_1, \ldots, a_n, b_1, \ldots, b_{n-1}) + a_1 b_n = a_1 c_n + f(a_1, \ldots, a_n, c_1, \ldots, c_{n-1})$$

where importantly f is the same function on both sides (it's the result of expanding the same structure of terms).

By induction, we can conclude that $b_j = c_j$ for all $j \in \{1, \ldots, n-1\}$. So it follows that $f(a_1, \ldots, a_n, b_1, \ldots, b_{n-1}) = f(a_1, \ldots, a_n, c_1, \ldots, c_{n-1})$. Hence, we have $a_1 b_n = a_1 c_n$. And by the same reasoning as before, this tells us that $b_n = c_n$.

Homework 5:

(1) Pick integers satisfying $1 \le k_1 < \cdots < k_r \le n$. Let X be the set of subspaces $W_1, \ldots, W_r$ of $\mathbf{F}_q^n$ such that $\dim W_i = k_i$ for all i and $W_i \subseteq W_{i+1}$. Find a formula for $|X|$.

I claim that $|X| = \begin{bmatrix} n \\ n-k_r,\, \ldots,\, k_2-k_1,\, k_1 \end{bmatrix}_q$.

To prove this, we proceed by induction. Firstly, when $r = 1$, then our problem just becomes counting the number of k_1 dimensioned subspaces of $\mathbf{F}_q^n$. We already showed in class that this quantity is given by:

$$\frac{[n]_q!}{[k_1]_q![n-k_1]_q!} = \begin{bmatrix} n \\ n-k_1, k_1 \end{bmatrix}_q.$$

Next, suppose $r \ge 2$. Then there are $\frac{[n]_q!}{[k_r]_q![n-k_r]_q!} = \begin{bmatrix} n \\ n-k_r, k_r \end{bmatrix}_q$ many different k_r-dimensioned subspaces of $\mathbf{F}_q^n$. Letting W_r be any one of those subspaces, we then must have that $1 \le k_1 < \cdot < k_{r-1} < k_r$ and $W_1, \ldots, W_{r-1}$ are subspaces of W_r such that $\dim W_i = k_i$ for all i and $W_i \subseteq W_{i+1}$. Since $W_r \cong F_q^{k_r}$, we can apply our inductive hypothesis to say that there are $\begin{bmatrix} k_r \\ k_r-k_{r-1},\, \ldots,\, k_2-k_1,\, k_1 \end{bmatrix}_q$ many choices of $W_1, \ldots, W_{r-1}$ we can make.

Thus:

$$\begin{aligned} |X| &= \begin{bmatrix} n \\ n-k_r, k_r \end{bmatrix}_q \begin{bmatrix} k_r \\ k_r-k_{r-1},\, \ldots,\, k_2-k_1,\, k_1 \end{bmatrix}_q \\ &= \frac{[n]_q!}{[n-k_r]_q![k_r]_q!} \cdot \frac{[k_r]_q!}{[k_r-k_{r-1}]_q! \cdots [k_2-k_1]_q![k_1]_q!} = \frac{[n]_q!}{[n-k_r]-q![k_r-k_{r-1}]_q! \cdots [k_2-k_1]_q![k_1]_q!} \\ &= \begin{bmatrix} n \\ n-k_r,\, \ldots,\, k_2-k_1,\, k_1 \end{bmatrix}_q. \end{aligned}$$

(2)

(a) Use the following q-analogue of Pascal's identity (you don't need to prove it):

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = q^k \begin{bmatrix} n-1 \\ k \end{bmatrix}_q + \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q \text{ for } n \ge k > 0$$

...to show that if d is a nonnegative integer, then:

$$\sum_{n \ge 0} \begin{bmatrix} n+d \\ n \end{bmatrix}_q x^n = \prod_{i=0}^{d} (1 - q^i x)^{-1} = \frac{1}{(1-x)(1-qx)\cdots(1-q^d x)}$$

We shall do induction on d. If $d = 0$, then $\begin{bmatrix} n+d \\ n \end{bmatrix}_q = \begin{bmatrix} n \\ n \end{bmatrix}_q = 1$ for all n. So we have that:

$$\sum_{n \ge 0} \begin{bmatrix} n+d \\ n \end{bmatrix}_q x^n = \sum_{n \ge 0} x^n = \frac{1}{1-x} = \prod_{i=0}^{d} (1 - q^i x)^{-1}$$

Now let $d \ge 1$. Then note that:

$$\begin{aligned}
\sum_{n\geq 0} \begin{bmatrix} n+d \\ n \end{bmatrix}_q x^n &= 1 + \sum_{n\geq 1} \begin{bmatrix} n+d \\ n \end{bmatrix}_q x^n \\
&= 1 + \sum_{n\geq 1} \begin{bmatrix} n+d-1 \\ n \end{bmatrix}_q (qx)^n + \sum_{n\geq 1} \begin{bmatrix} n+d-1 \\ n-1 \end{bmatrix}_q x^n \\
&= \sum_{n\geq 0} \begin{bmatrix} n+d-1 \\ n \end{bmatrix}_q (qx)^n + \sum_{n\geq 0} \begin{bmatrix} n+d \\ n \end{bmatrix}_q x^{n+1} \\
&= \prod_{i=0}^{d-1} (1-q^i(qx))^{-1} + x \sum_{n\geq 0} \begin{bmatrix} n+d \\ n \end{bmatrix}_q x^n
\end{aligned}$$

So, $(1-x)\sum_{n\geq 0} \binom{n+d}{n}_q x^n = \prod_{i=1}^{d} (q - q^i x)^{-1}$.

Dividing both sides by $(1-x)$, we get: $\sum_{n\geq 0} \binom{n+d}{n}_q x^n = \prod_{i=0}^{d} (q - q^i x)^{-1}$.

(b) Give an explanation for why the coefficient of x^n on the right side is the sum $\sum_\lambda q^{|\lambda|}$ over all integer partitions λ whose Young diagram fits in the $n \times d$ rectangle.

We already established in class that $\begin{bmatrix} n+d \\ n \end{bmatrix}_q$ gives the number of n-dimensional subspaces of $\mathbf{F}_q^{n+d}$. Additionally, it was also established that there is a bijective correspondance between n-dimensional subspaces of $\mathbf{F}_q^{n+d}$ and $n \times (n+d)$ full rank rref-matrices.

Next, the set of $n \times (n+d)$ full rank rref-matrices can be split $\binom{n+d}{n}$ types where the type of a matrix is determined by the indices $s_1 < \cdots < s_n$ of its pivot columns.

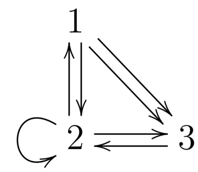
Given a type with pivot column indices $s_1 < \cdots < s_n$, the ith row will have $(n+d) - s_i - (n-i) = d - s_i + i$ components not forced to be anything. $i \leq s_i \leq d+i$. So, $0 \leq d - (s_i - i) \leq d$. Also $s_i - i$ monotonically increases as i increases. So, we can associate the type of matrix with a Young diagram whose ith row is the number of undetermined components in the ith row of that type of matrix. This Young diagram fits in an $n \times d$ box whose size is the number of undetermined components in the original matrix type.

Also, this association betweeen types of full rank rref matrices and young diagrams fitting in an $n \times d$ rectangle. We know this because we can invert it.

Given a young diagram with $r_1, \ldots, r_n$ being the number of boxes in row i, define $s_i = d - r_i + i$. Then $s_1 < \cdots < s_n$ determines a type of rref-matrix whose pivot rows are $s_1, \ldots, s_n$.

There are $q^{|\lambda|}$ many rref-matrices of the type associated with the young diagram given by λ since that type of matrix has $|\lambda|$ many arbitrary components.

So $\begin{bmatrix} n+d \\ n \end{bmatrix}_q = \sum_\lambda q^{|\lambda|}$ over all integer partitions λ whose Young diagram fits in the $n \times d$ rectangle.



(3) Let A be the adjacency matrix of the following directed graph:
Compute the rational function $P_{A;1,3}(x)$.

Note that $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$.

So:

$$P_{A;1,3}(x) = (-1)^{1+3} \frac{\det((\mathrm{Id}-xA);3,1)}{\det(\mathrm{Id}-xA)}$$

$$= \frac{\det(\begin{bmatrix} -x & -2x \\ 1-x & -x \end{bmatrix})}{\det(\begin{bmatrix} 1 & -x & -2x \\ -x & 1-x & -x \\ 0 & -x & 1 \end{bmatrix})} = \frac{x^2+2x-2x^2}{1(1-x-x^2)+x(-x)-2x(x^2)} = \frac{x-x^2}{1-x-2x^2+2x^3}$$

(4) Construct a directed graph for which walks between certain vertices can be interpreted as binary strings in which no symbol ever appears 3 times in a row.

Let $V = \{\text{binary strings of length } 2\} = \{00, 10, 01, 11\}$. Then given the words $v = a_1a_2, u = b_1b_2 \in V$, we say there is an edge from v to u if $a_2 = b_1$ and a_1, a_2, b_1, b_2 are not all equal. In other words, (v, w) is an edge if v and w are overlapping substrings of a word fulfilling our requirements.

This graph will look like:

$$\begin{tikzcd} 00 \arrow[d] & 10 \arrow[l] \arrow[dl, shift left] \\ 01 \arrow[r] \arrow[ur, shift left] & 11 \arrow[u] \end{tikzcd}$$

There is a bijective correspondance between walks of length k in the above graph and words of length $k+2$. Specifically suppose $a_1 \cdots a_k a_{k+1} a_{k+2}$ is a binary word of length $k+2$ where no symbol appears 3 times in a rows. For $0 \leq i \leq k$, define $v_i = a_{i+1}a_{i+2}$. Then $v_0, v_1, \cdots, v_k$ is a walk of length k in our graph.

To invert this process, if $v_0 = a_0b_0, v_1 = a_1b_1, \ldots, v_k = a_kb_k$ and $v_0, v_1, \ldots, v_n$ is a walk in our graph, then the word $a_0b_0b_1 \cdots b_k$ will be a binary word of length $k+2$ with no symbol appearing 3 times in a row.

(5) Let A be a finite alphabet. We say that $\boldsymbol{w} = w_1 \cdots w_k$ is a subword of $\boldsymbol{v} = v_1 \cdots v_n$ if there exists indices $1 \le i_1 < i_2 < \cdots < i_k \le n$ such that $v_{i_j} = w_j$ for all $j = 1, \dots, n$. Define $a_{\boldsymbol{w}}(n)$ to be the number of words of length n that have $\boldsymbol{w}$ as a subword. Prove that $\sum\limits_{n\ge 0} a_{\boldsymbol{w}}(n)x^n$ is a rational function of x.

We shall proceed by induction on the length of $\boldsymbol{w}$.

To start, assume $\boldsymbol{w} = w_1$ is just a single symbol. Then there are $|A|^n$ many word of length n and $(|A|-1)^n$ of those don't have $\boldsymbol{w}$ as a substring (meaning that they don't contain the symbol w_1 and so are words of the alphabet $A - \{w_1\}$). So:

$$\sum_{n\ge 0} a_{\boldsymbol{w}}(n)x^n = \sum_{n\ge 0}(|A|x)^n - \sum_{n\ge 0}((|A|-1)x)^n$$
$$= \tfrac{1}{1-|A|x} + \tfrac{1}{1-|A|x+x} = \tfrac{2-2|A|x+x}{(1-|A|x)(1-|A|x+x)} = \tfrac{2+(1-2|A|)x}{1+(1-2|A|)x+(|A|^2-|A|)x^2}$$

Now assume that $\sum\limits_{n\ge 0} a_{\boldsymbol{w}'}(n)x^n$ is a rational function of x for all words $\boldsymbol{w}'$ of length less than k.

Note that we can split the words $\boldsymbol{v}$ of length $n > 1$ with $\boldsymbol{w}$ as a subword into two types.

- Type 1: Suppose w_1 is not the first character of $\boldsymbol{v}$.

  Then we must have that $\boldsymbol{w}$ is a substring of $v_2 \cdots v_n$. Also, v_1 can be any symbol in $A - \{w_1\}$. Concatting these substrings since they can be chosen independently, there are $(|A|-1)a_{\boldsymbol{w}}(n-1)$ many words of this type.
- Type 2: Suppose $w_1 = v_1$.

  Let $\boldsymbol{w}' = w_2 \cdots w_k$. Then we know that $\boldsymbol{w}'$ is a substring $v_2 \cdots v_n$. So, the number of words of this type is $a_{\boldsymbol{w}'}(n-1)$.

Clearly, if $n < 2 \le k$, then $a_{\boldsymbol{w}}(n) = 0$. Also $a_{\boldsymbol{w}'}(0) = 0$. So:

$$\sum_{n\ge 0} a_{\boldsymbol{w}}(n)x^n = \sum_{n\ge 2} a_{\boldsymbol{w}}(n)x^n = \sum_{n\ge 2}(|A|-1)a_{\boldsymbol{w}}(n-1)x^n + \sum_{n\ge 2} a_{\boldsymbol{w}'}(n-1)x^n$$
$$= (|A|-1)x\sum_{n\ge 1} a_{\boldsymbol{w}}(n)x^n + x\sum_{n\ge 1} a_{\boldsymbol{w}'}(n)x^n$$
$$= (|A|-1)x\sum_{n\ge 0} a_{\boldsymbol{w}}(n)x^n + x\sum_{n\ge 0} a_{\boldsymbol{w}'}(n)x^n$$

By induction, we know that there are polynomials $P(x)$ and $Q(x)$ such that:

$$\sum_{n\ge 0} a_{\boldsymbol{w}'}(n)x^n = \tfrac{P(x)}{Q(x)}.$$

So, $(1-(|A|-1)x)\sum\limits_{n\ge 0} a_{\boldsymbol{w}}(n)x^n = \frac{xP(x)}{Q(x)}$.

Hence, $\sum\limits_{n\geq 0} a_{\boldsymbol{w}}(n)x^n$ is the rational function $\frac{xP(x)}{(1-(|A|-1)x)Q(x)}$.

Now while that was all I was required to do, for the sake of fun I'm gonna also find an explicit formula for $\sum\limits_{n\geq 0} a_{\boldsymbol{w}}(n)x^n$.

Define $\boldsymbol{w}^{(i)} = w_i \cdots w_k$.

From before, we know that whenever $\boldsymbol{w}$ has length greater than 2, we have that:

$$\sum_{n\geq 0} a_{\boldsymbol{w}}(n)x^n = \frac{x}{1-(|A|-1)x} \cdot \sum_{n\geq 0} a_{\boldsymbol{w}'}(n)x^n$$

So:

$$\begin{aligned}\sum_{n\geq 0} a_{\boldsymbol{w}}(n)x^n &= \frac{x}{1-(|A|-1)x} \cdot \sum_{n\geq 0} a_{\boldsymbol{w}^{(2)}}(n)x^n \\ &= \left(\frac{x}{1-(|A|-1)x}\right)^2 \cdot \sum_{n\geq 0} a_{\boldsymbol{w}^{(3)}}(n)x^n = \cdots = \left(\frac{x}{1-(|A|-1)x}\right)^{k-1} \sum_{n\geq 0} a_{\boldsymbol{w}^{(k)}}(n)x^n\end{aligned}$$

Since $\boldsymbol{w}^{(k)}$ is a word of length 1, we know that $\sum\limits_{n\geq 0} a_{\boldsymbol{w}^{(k)}}(n)x^n = \frac{2+(1-2|A|)x}{1+(1-2|A|)x+(|A|^2-|A|)x^2}$.

So $\sum\limits_{n\geq 0} a_{\boldsymbol{w}}(n)x^n = \frac{x^{k-1}}{(1-(|A|-1)x)^{k-1}} \cdot \frac{2+(1-2|A|)x}{1+(1-2|A|)x+(|A|^2-|A|)x^2}$.

We can simplify this further. Note that:

$$(1-2|A|)^2 - 4(|A|^2 - |A|) = 1 - 4|A| + 4|A|^2 - 4|A|^2 + 4|A| = 1.$$

So by quadratic formula there is a scalar c such that:

$$\begin{aligned}1 + (1-2|A|)x + (|A|^2 - |A|)x^2 &= c\left(x - \frac{2|A|-2}{2|A|^2-2|A|}\right)\left(x - \frac{2|A|}{2|A|^2-2|A|}\right) \\ &= c\left(x - \frac{1}{|A|}\right)\left(x - \frac{1}{|A|-1}\right)\end{aligned}$$

By matching the x^2 terms, we know that $c = (|A|^2 - |A|)$. Thus, rearranging terms we get that:

$$\begin{aligned}1 + (1-2|A|)x + (|A|^2 - |A|)x^2 &= (|A|x - 1)((|A|-1)x - 1) \\ &= (1 - |A|x)(1 - (|A|-1)x).\end{aligned}$$

Therefore $\sum\limits_{n\geq 0} a_{\boldsymbol{w}}(n)x^n = \frac{2x^{k-1}+(1-2|A|)x^k}{(1-(|A|-1)x)^k(1-|A|x)}$.

By our theorem on rational generating functions, we know there are constants $c_1, \ldots, c_{k+1}$ such that if $N + k + 1 > k$, then for all $n \geq N$ we have that:

$$a_{\boldsymbol{w}}(n) = (c_1 + c_2 n + \ldots + c_k n^{k-1})(|A|-1)^n + c_{k+1}|A|^n$$

$N = 0$ works for this. So for all $n \geq 0$ we have that:

$$a_{\boldsymbol{w}}(n) = (c_1 + c_2 n + \ldots + c_k n^{k-1})(|A|-1)^n + c_{k+1}|A|^n$$

To solve for each constant, use the initial conditions $a_{\boldsymbol{w}}(n) = 0$ when $n < k$ and $a_{\boldsymbol{w}}(k) = 1$. Unfortunately, I've run out of ideas for simplifying this more and it's 11:49PM.

Here's a different approach. Given any word $\boldsymbol{v} = v_1 \cdots v_n$ containing $\boldsymbol{w} = w_1 \cdots w_k$ as a subword, we can injectively map it to a sequence $(\boldsymbol{u}_0, \boldsymbol{u}_1 \ldots, \boldsymbol{u}_{k-1}, \boldsymbol{u}_k)$ of $k+1$ many words via the following procedure:

- $\boldsymbol{u}_0$ is the sequence of characters before the first instance of w_1 in $\boldsymbol{v}$.
- For $i < n$, $\boldsymbol{u}_i$ is the sequence of characters in $\boldsymbol{v}$ following $\boldsymbol{u}_1 w_1 \boldsymbol{u}_2 w_2 \cdots \boldsymbol{u}_{i-1} w_i$ but before the first instance of w_{i+1} in $\boldsymbol{v}$ after $\boldsymbol{u}_1 w_1 \boldsymbol{u}_2 w_2 \cdots \boldsymbol{u}_{i-1} w_i$.
- $\boldsymbol{u}_n$ is the rest of $\boldsymbol{v}$ after $\boldsymbol{u}_0 w_1 \boldsymbol{u}_1 w_2 \cdots \boldsymbol{u}_{k-1} w_k$.

Note that the $(\boldsymbol{u}_0, \boldsymbol{u}_1 \ldots, \boldsymbol{u}_{k-1}, \boldsymbol{u}_k)$ we get above has the three properties that:

- $\boldsymbol{u}_i$ does not contain w_{i+1} for all $0 \le i < k$.
- $\boldsymbol{u}_k$ can contain all symbols of A
- The lengths of $\boldsymbol{u}_0$, ..., $\boldsymbol{u}_k$ add up to $n - k$.

Importantly, given any sequence $(\boldsymbol{u}'_0, \boldsymbol{u}'_1 \ldots, \boldsymbol{u}'_{k-1}, \boldsymbol{u}'_k)$ also satisfying the above three properties, we know that the word $\boldsymbol{u}'_0 w_1 \boldsymbol{u}'_1 w_2 \cdots \boldsymbol{u}'_{k-1} w_k \boldsymbol{u}'_k$ will get mapped to $(\boldsymbol{u}'_0, \boldsymbol{u}'_1 \ldots, \boldsymbol{u}'_{k-1}, \boldsymbol{u}'_k)$. So our previous map is also surjective onto sequences with the above three properties.

So, we have now found a bijective correspondance between words $\boldsymbol{v}$ containing $\boldsymbol{w}$ and sequences of smaller words satisfying certain properties. We now move onto counting these sequences.

The collection of sequences will split into types determined by the lengths $r_0, r_1, \ldots, r_{k-1}, r_k$ of the words $\boldsymbol{u}_0, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_{k-1}, \boldsymbol{u}_k$ in a given sequence. By our third known property, the number of types is given by the number of weak compositions of $n - k$ into $k + 1$ parts.

Next, given a type determined by the lengths $r_0, r_1, \ldots, r_{k-1}, r_k$, we have that the number of sequences of that type is:

$$(|A| - 1)^{r_0}(|A| - 1)^{r_1} \cdots (|A| - 1)^{r_{k-1}}|A|^{r_k} = (|A| - 1)^{n-k-r_k}|A|^{r_k}$$

Also, the number of types with $r_k = i$ is the equal to the number of weak-compositions of $n - k - i$ into k parts. I.e. there are $\binom{n-k-i+k-1}{k-1} = \binom{n-i-1}{k-1}$ many types with $r_k = i$.

Finally, $0 \le r_k \le n - k$. So, we get that $a_{\boldsymbol{w}}(n) = \sum\limits_{i=0}^{n-k} \binom{n-i-1}{k-1}(|A| - 1)^{n-k-i}|A|^i$.

Homework 6:

(1) Let $b_{n,k}$ be the number of set partitions of $[n]$ with k blocks such that every block has an even (and positive) number of elements and let b_n be the same, but with no restrictions on the number of blocks.

(a) Find a formula for the EGF $B_k(x) = \sum_{n\geq 0} b_{n,k}\frac{x^n}{n!}$.

Define the structure α such that $\alpha(S) = 1$ if $|S|$ is positive and even, and $\alpha(S) = 0$ otherwise. That way, α selects sets which would be allowed to be blocks in a larger set partition. Also, we have:

$$\sum_{n\geq 0} |\alpha([n])|\frac{x^n}{n!} = \sum_{n\geq 1} \frac{x^{2n}}{(2n)!} = \sum_{n\geq 1} \frac{x^n}{2(n!)} + \sum_{n\geq 1} \frac{(-x)^n}{2(n!)}$$
$$= \frac{1}{2}\left(\sum_{n\geq 0} \frac{x^n}{n!} + \sum_{n\geq 0} \frac{(-x)^n}{n!}\right) - 2\frac{1}{2} = \frac{1}{2}(e^x + e^{-x}) - 1$$

Next, $|\alpha^k(S)|$ will give the number of ordered set partitions of S into k blocks with even positive sizes. And importantly:

$$\sum_{n\geq 0} |\alpha^k([n])|\frac{x^n}{n!} = \left(\sum_{n\geq 0} |\alpha([n])|\frac{x^n}{n!}\right)^k = \left(\frac{1}{2}(e^x + e^{-x}) - 1\right)^k$$

Now $|\alpha^k([n])| = k!b_{n,k}$. So $B_k(x) = \frac{1}{k!}\sum_{n\geq 0} |\alpha^k([n])|\frac{x^n}{n!} = \frac{1}{k!}(\frac{1}{2}(e^x + e^{-x}) - 1)^k$.

I hate how unsimplified this is but I've spent an hour and a half unable to simplify it further.

(b) Find a formula for the EGF $B(x) = \sum_{n\geq 0} b_n\frac{x^n}{n!}$.

Keeping α as defined in part (a), $b_n = |e^{\alpha}([n])|$.

Then as $\alpha(\emptyset) = 0$, we have that:

$$\sum_{n\geq 0} b_n\frac{x^n}{n!} = E_{e^\alpha}(x) = \exp(E_\alpha(x)) = \sum_{k\geq 0} \frac{(\frac{1}{2}(e^x+e^{-x})-1)^k}{k!}.$$

(2) For $n > 0$, let h_n be the number of permuations of size n that only have cycles of even length and define $h_0 = 1$. (go to next page...)

(a) Find a structure α so that $h_n = |e^\alpha([n])|$

Let $\alpha(S)$ equal the number of ways to cyclically order S if $|S|$ is even and positive and otherwise equal the emptyset.

Then $|\alpha(S)| = (|S|-1)!$ if $|S|$ is even and nonzero, and otherwise $|\alpha(S)| = 0$. Hence, we can say that $|e^\alpha([n])|$ is well defined and equals the number of permutations of $[n]$ with only even cycles.

(b) Apply the derivative identity (Proposition 5.2.6) to get a recurrence relation for h_n. What initial conditions do you need to add to determine the sequence.

Now define:

$$A(x) = \sum_{n\geq 0} a_n \frac{x^n}{n!} := E_\alpha(x)$$
$$H(x) = \sum_{n\geq 0} h_n \frac{x^n}{n!} := E_{e^\alpha}(x) = \exp(E_\alpha(x)) = \exp(A(x))$$

Then note $A(x) = \sum_{n\geq 0} |\alpha([n])| \frac{x^n}{n!} = \sum_{n\geq 1} (2n-1)! \frac{x^{2n}}{(2n)!} = \sum_{n\geq 1} \frac{1}{2n} x^{2n}$.

It follows that $A'(x) = \sum_{n\geq 1} x^{2n-1} = \sum_{n\geq 0} \frac{1}{2}(1-(-1)^n)x^n$

Then because $H'(x) = H(x)A'(x)$, we have $\frac{h_{n+1}}{n!} = \sum_{i=0}^{n} \frac{h_i}{i!} \cdot \frac{1-(-1)^{n-i}}{2}$ and $\frac{h_{n+3}}{(n+2)!} = \sum_{i=0}^{n+2} \frac{h_i}{i!} \cdot \frac{1-(-1)^{n+2-i}}{2} = \sum_{i=0}^{n+2} \frac{h_i}{i!} \cdot \frac{1-(-1)^{n-i}}{2}$ for $n \geq 0$.

So $\frac{h_{n+3}}{(n+2)!} - \frac{h_{n+1}}{n!} = \frac{h_{n+2}}{(n+2)!} \cdot \frac{1-1}{2} + \frac{h_{n+1}}{(n+1)!} \cdot \frac{1+1}{2} = \frac{h_{n+1}}{(n+1)!}$.

After some rearranging, this becomes:

$$h_{n+3} = ((n+2)(n+1) + (n+2))h_{n+1} = (n+2)^2 h_{n+1}$$

And we can shift indices to say that for $n \geq 3$, we have $h_n = (n-1)^2 h_{n-2}$.

And now we have the desired recursive formula. This is enough to recursively calculate out all h_n since we know that $h_0 = 1$, $h_1 = 0$, and $h_2 = 1$. Also, quick inspection shows that this recurrence relation thus works for $n = 2$.

Also just for kicks and giggles, it's obvious $h_{2n+1} = 0$ for all n and we can show by induction that $h_{2n} = ((2n-1)!!)^2$ when $n > 0$.

(3) In this problem, we will consider labeled simple graphs such that every vertex has degree 2 (i.e., is contained in exactly 2 edges). For $n \geq 1$, let h_n be the number of such graphs with vertex set $[n]$ and set $h_0 = 1$.

You may use this fact without proving it: every graph satisfying this condition is a disjoint union of cycle graphs, each one consisting of 3 or more vertices.

(a) Find a structure α so that $h_n = |e^\alpha([n])|$.

Let $\alpha(S)$ be the set of all cycles one can draw on a labeled graph whose vertex set is S such that the cycle contains all vertices (if $S = \emptyset$, then define $\alpha(S) = \emptyset$).

Then $|\alpha(S)| = 0$ when $|S| \leq 2$. Meanwhile, for $|S| > 2$, every cyclic ordering of S will give an element of $\alpha(S)$ and each graph will be associated with two cyclic orderings (a reversed and nonreversed ordering). So $|\alpha(S)| = \frac{1}{2}(|S|-1)!$ when $|S| \geq 3$.

Hence, $|e^\alpha(S)|$ is well defined and by the provided fact, $h_n = |e^\alpha([n])|$.

(b) Apply the derivative identity (Proposition 5.2.6) to get a recurrence relation for h_n. As in the previous problem, give relevant initial conditions.

As in problem (2), define the F.P.S. $A(x)$ and $H(x)$, although this time using the α defined in this problem's part (a).

Then note $A(x) = \sum\limits_{n\geq 0} |\alpha([n])| \frac{x^n}{n!} = \sum\limits_{n\geq 3} (n-1)! \frac{x^n}{2n!} = \sum\limits_{n\geq 3} \frac{x^n}{2n}$.

In turn, $A'(x) = \sum\limits_{n\geq 3} \frac{1}{2} x^{n-1} = \sum\limits_{n\geq 2} \frac{1}{2} x^n$.

So, applying the $H'(x) = H(x)A'(x)$ identity, we get that $\frac{h_{n+1}}{n!} = \sum\limits_{i=0}^{n-2} \frac{h_i}{i!} \cdot \frac{1}{2}$ when $n \geq 2$.

And thus $\frac{h_{n+2}}{(n+1)!} - \frac{h_{n+1}}{n!} = \sum\limits_{i=0}^{n-1} \frac{h_i}{i!} - \sum\limits_{i=0}^{n-2} \frac{h_i}{i!} = \frac{h_{n-1}}{2(n-1)!}$ Or in other words:

$$h_n = (n-1)h_{n-1} + \tfrac{1}{2}(n-1)(n-2)h_{n-3} \text{ for } n \geq 4.$$

I have no idea how'd you simplify that further. Luckily I'm not required to because this a recurrence relation.

The relevant initial conditions are $h_0 = 1$, $h_1 = 0$, $h_2 = 0$, $h_3 = 1$. Also, based on those conditions we can recurrence relation also holds when $n = 3$.

(4) Let $n \geq 2$ be an integer. We have n married couples ($2n$ people in total).

(a) How many ways can we have the $2n$ people stand in a line so that no person is standing next to their spouse? (go to next page...)

There are $(2n)!$ possible arrangements. Also, for $i \in \{1, \ldots, n\}$ define A_i to be the set of all arrangements with the ith couple next to each other. Thus, the quantity we want is:

$$|A_1^{\mathsf{C}} \cap \cdots \cap A_n^{\mathsf{C}}| = (2n)! - |A_1 \cup \cdots \cup A_n|$$

By inclusion-exclusion formula, we thus need to calculate:

$$(2n)! - \sum_{j=1}^{n} (-1)^{j-1} \sum_{1 \leq i_1 < \cdots < i_j \leq n} |A_{i_1} \cap \cdots \cap A_{i_j}|$$

Now by symmetry, no matter which j-many couples we pick, we'll get the same size intersection. So since there are $\binom{n}{j}$ to choose j couples, we can simplifying our expression to:

$$(2n)! + \sum_{j=1}^{n} (-1)^j \binom{n}{j} |\bigcap_{i=1}^{j} A_i|$$

And now to calculate $|\bigcap_{i=1}^{j} A_i|$:

- Each couple forced together will take up a block of 2 spots, and in each block we have a choice of who will be on the lower index spot. This gives 2^j many choices.
- If the blocks are indistinct, then we have the same number of ways to position those indistinct blocks in a line as there are weak compositions of $2n - 2j$ with $j + 1$ parts. This gives:
$$\binom{2n-2j+j+1-1}{2n-2j} = \binom{2n-j}{2n-2j} = \frac{(2n-j)!}{(2n-2j)!j!} \text{ choices.}$$
- We can order the blocks in $j!$ many ways.
- There are $(2n - 2j)$ many people left over who can get placed in any remaining spot. This gives $(2n - 2j)!$ choices.

Thus we get that: $|\bigcap_{i=1}^{j} A_i| = 2^j j!(2n-2j)! \frac{(2n-j)!}{(2n-2j)!j!} = 2^j (2n-j)!$

So our final total is:

$$(2n)! + \sum_{j=1}^{n} (-1)^j \binom{n}{j} 2^j (2n-j)!$$

(b) Same as (a) but replace "line" by "circle".

Imagining all the people sat around a table, I'm going to consider two arrangements equal if the former is just the latter rotated around several seats. Note that I'm assuming mirrored seating arrangements are not equivalent.

Now note that arranging people in a circle with a marked first position is equivalent to arranging people on a line while now caring if a couple occupies the first and last positions. To get the number of arrangement classes (i.e. where we don't have a marked first position in the circle), we can can count the number of arrangements with a marked first position and then divide by $2n$ (i.e. the number of ways to rotate an arrangement with a marked first position).

Other than that the total number of equivalent arrangements is $(2n-1)!$, we proceed identically as in part (a) until we get to calculating $|\bigcap_{i=1}^{j} A_i|$.

The arrangements of n couples around a circle with a marked first position will split into two cases.

1. There is no couple occupying the first and last position in the circle. Then the choices we make are identical to that of part (a). There are $2^j(2n-j)!$ arrangements in this case.
2. There is a couple occupying the first and last position in the circle.
   - Then we still have the j couples we're forcing together occupying j blocks. And in each block we can choose which person is first. This gives 2^j choices.
   - If the blocks are indistinct, then we have the same number of ways of positioning these indistinct blocks as there are weak compositions of $2n-2j$ with j parts. This gives: $\binom{2n-2j+j-1}{2n-2j} = \binom{2n-j}{2n-2j} = \frac{(2n-j-1)!}{(2n-2j)!(j-1)!}$ choices.
   - We can assign the j couples to the j blocks in $j!$ many ways.
   - There are $(2n-2j)$ people left over who can be positioned however.

   This gives $j!(2n-2j)!2^j \frac{(2n-j-1)!}{(2n-2j)!(j-1)!} = j2^j(2n-j-1)!$ many arrangements in this case.

So our number of equivalence classes of arrangements is:

$$\begin{aligned}|\bigcap_{i=1}^{j} A_i| &= \frac{1}{2n}\left(2^j(2n-j)! + j2^j(2n-j-1)!\right)\\ &= \frac{1}{2n}\cdot 2^j\cdot(2n-j-1)!\cdot(2n-j+j) = 2^j(2n-j-1)!\end{aligned}$$

By inclusion-exclusion, our desired total is thus:

$$(2n-1)! + \sum_{j=1}^{n}(-1)^j\binom{n}{j}2^j(2n-j-1)!$$

(5) Let q be a prime power. For this problem, we consider polynomials with coefficients in $\boldsymbol{F}_q$. A polynomial $f(x)$ of degree d is monic if the coefficient of x^d is 1. It is irreducible if it is not possible to factor it as $g_1(x)g_2(x)$ where both g_1, g_2 have degree $< d$. For this problem, you can assume without proof that all polynomials satisfy *unique factorization*: every monic polynomial can be written uniquely (up to rearranging the order of the factors) as a product of monic irreducible polynomials.

(a) Let N_n be the number of monic irreducible polynomials of degree n with coefficients in $\boldsymbol{F}_q$. Prove that:

$$(1 - qx)^{-1} = \prod_{d \geq 1} (1 - x^d)^{-N_d}$$

(b) By taking the logarithmic derivative of (a), we get the following identity:

$$\frac{1}{1-qx} = \sum_{d \geq 1} N_d \frac{dx^{d-1}}{1-x^d}$$

(The logarithmic derivative of A is A'/A and takes products to sums. You don't need to derive this.) Use this to prove that $q^n = \sum_{d \mid n} dN_d$.

(c) Use Möbius inversion to get a formula for N_n.