Algebraic Combinatorics

Jacob Steinhardt

1 Introduction

This is supposed to be (if not now, then eventually) a general compendium of techniques in algebraic combinatorics. I start with the most basic technique, generating functions, which requires only a knowledge of polynomials and some degree of calculus and power series. Then I move on to more advanced topics: for example, Combinatorial Nullstellensatz, a technique that I like to think of as a more sophisticated form of pigeonhole and which uses facts about polynomials over an arbitrary field (that is, a place where you can add, subtract, multiply, and divide). There is also the plethora of information one gets by thinking about a problem from the perspective of linear algebra. I hope to eventually add in sections on even more advanced techniques, such as spectral graph theory.

One can find more information about generating functions from Herbert Wilf's excellent book, *generatingfunctionology*. It is freely available online at `http://www.math.upenn.edu/~wilf/DownloadGF.html`. For those interested in the Combinatorial Nullstellensatz, I recommend reading Noga Alon's original survey paper, available at `http://www.math.tau.ac.il/~nogaa/PDFS/null2.pdf`.

2 The Basics

2.1 Binomial Coefficients

$\binom{n}{k}$ is called a binomial coefficient, and traditionally represents the number of ways to choose k out of n objects. It evaluates to $\frac{n!}{k!(n-k)!}$. One can see this as follows: there are $n!$ ways to arrange the objects, and you can choose the objects by taking the first k in a given arrangement. But the order of the first k and the last $n-k$ in that arrangement won't affect which k objects you chose, so we have counted each possibly choice $k!(n-k)!$ times, and so we should divide by $(n-k)!$.

Note also that $\binom{x}{k} = \frac{x(x-1)(x-2)...(x-k)}{k!}$. It is thus a polynomial of degree k in x, and makes sense for any value of x, not just integers[1].

The binomial coefficient is possibly the single most structured object in mathematics. Whatever you do with it, you will probably get something nice. Even the multiplicative inverses of binomial coefficients are workable, and adding consecutive values of functions involving binomial coefficients will often do something helpful.

[1] It is possible to define the coefficient even more generally as $\binom{x}{y} = \frac{\Gamma(x+1)}{\Gamma(y+1)\Gamma(x-y+1)}$, where $\Gamma(x) = \int_0^\infty t^{x-1}e^{-x}dt$, but this is unnecessary for the purposes of this lecture.

2.2 Extended Binomial Theorem

$$\sum_{i=0}^{\infty} x^i \binom{i}{y} = (1+x)^y$$

This holds for all y and for all x where the sum converges. In particular, we are fine when $|x| < 1$ and also whenever y is a positive integer (as then the left-hand-side is just a polynomial in x). It is more important, however, to think of this as an expression that holds *symbolically.* That is, we can replace the left-hand-side with the right-hand-side (or vice versa) without changing the truth of any statements (this is actually slightly false, but true enough for most purposes, and true enough that you shouldn't worry about it not being true).

2.3 Formal Power Series

Given any field F, (a set closed under addition and multiplication that is associative, commutative, distributive, and has identities and inverses), we can take all polynomials over that field. The result is often referred to as $F[x]$. The most common examples of a field are the reals ($\mathbb{R}$) and the complex numbers ($\mathbb{C}$), but other common fields are the rationals ($\mathbb{Q}$) as well as any prime mod ($\mathbb{Z}_p$). Basically, something is a field if you can divide by anything that isn't zero (so in particular, the integers are *not* a field since you can't divide 1 by 2 and still get an integer). If this confuses you, just assume that the field is the real numbers or something else you're familiar with.

We can associate any polynomial with its coefficients, that is, we can associate $(a_0, a_1, a_2, a_3, \ldots)$ with $a_0+a_1x+a_2x^2+a_3x^3+\ldots$. In a sense, they should be the same, as they both act as vectors over F (again, don't worry if you don't know what vectors/vector spaces are). However, the latter also has a built-in multiplication, namely $(a_0+a_1x+a_2x^2+\ldots)(b_0+b_1x+b_2x^2+\ldots) = (a_0b_0 + (a_1b_0 + a_0b_1)x + (a_2b_0 + a_1b_1 + a_0b_2)x^2 + \ldots)$. This is called a *convolution.* When the sequence a_i is eventually zero, we get a polynomial in the normal sense of the word. However, if we drop this stipulation and allow a_i to be arbitrary, then we get what is called a formal power series. This is denoted $F[[x]]$, as opposed to $F[x]$ as in the case of finite polynomials. The nice thing about $F[[x]]$ is that every formal power series with non-zero constant term has a unique multiplicative inverse, which is essentially found by solving for each of the coefficients in order (for those of you familiar with some linear algebra, note that if we were to try to write out a system of equations to solve for the inverse, it would be in lower triangular form, so that this makes plenty of sense if we ignore the fact that the resulting matrix is infinite).

The main idea of generating functions is to encode information about the sequence a_i into a polynomial/formal power series so that we can more easily manipulate it. Basically all normal algebraic manipulations are still valid for generating functions, so you shouldn't be afraid to use them.

3 Examples: Generating Functions for Sequences

Now that you know what generating functions are, it would be useful to know what to do with them. Here are a few examples:

Example one: *Find an explicit formula for the nth Fibonacci number, F_n.*

Solution: Let ϕ and $\bar{\phi}$ be the positive and negative roots of $x^2 - x - 1 = 0$. Let $F(x) = F_0 + F_1 x + F_2 x^2 + \ldots = 1 + x + 2x^2 + \ldots$. By definition, $F_n = F_{n-1} + F_{n-2}$, or $F_n - F_{n-1} - F_{n-2} = 0$. Thus, if we multiply $F(x) by (1 - x - x^2)$, all terms except the first term should telescope, yielding $F(x)(1 - x - x^2) = 1$. Then $F(x) = \frac{1}{1-x-x^2} = \frac{1}{(1-\phi x)(1-\bar{\phi} x)} = \frac{1}{\sqrt{5}}(\frac{\phi}{1-\phi x} - \frac{\bar{\phi}}{1-\bar{\phi} x})$ by partial fraction decomposition. Then, re-expressing the denominators as power series, we arrive at the following equivalent formulation of F:

$$F(x) = \sum_{n=0}^{\infty} (\frac{\phi^{n+1} - \bar{\phi}^{n+1}}{\sqrt{5}}) x^n$$

from which it follows that $F_n = \frac{\phi^{n+1} - \bar{\phi}^{n+1}}{\sqrt{5}}$.

Example two: *Find an explicit formula for the nth Catalan number, C_n.*

Solution: Let $C(x) = C_0 + C_1 x + C_2 x^2 + \ldots$. By definition, $C_n = C_{n-1} C_0 + C_{n-2} C_1 + \ldots$, so that if we define $D(x) = C(x)^2 = D_0 + D_1 x + \ldots$, then $C_n = D_{n-1}$. As $C_0 = 1$, we have $C(x) = xC(x)^2 + 1$, from which it follows that $C(x) = \frac{1-\sqrt{1-4x}}{2x}$ by the quadratic formula. (We can verify that the radical must be subtracted and not added in a number of ways. The easiest is to expand the first few terms of both and check to see which one yields positive coefficients). It then follows from the Extended Binomial Theorem that

$$C_n = 2(-4)^n \binom{\frac{1}{2}}{n+1} = \frac{1}{n+1} \binom{2n}{n}$$

4 Sums

Some problems will ask you to express a sum in closed form. In many cases, the goal is to find a suitable polynomial whose expansion is related to the sum in question. Here are a few tips for dealing with sums:

1. Be on the lookout for ways to turn an ugly-looking sum into the convolution of two nicer sequences. The main way to do this is to try to rearrange the sum so that in each term sums of variables are preserved. To give the canonical example, $\sum_{k=0}^{n} \binom{n}{k}^2$ is much easier to work with as $\sum_{k=0}^{n} \binom{n}{k} \binom{n}{n-k}$. Note that here the sums of the tops of the binomial coefficients is always $2n$ and the sums of the bottoms is always n. Once you have equal sums you can try to find a convolution that will help you out.

2. Try pairing terms in the sum together to simplify it, or adding it to itself, perhaps in reverse order. In particular, the sum or difference of adjacent terms, or functions of adjacent terms, in Pascal's triangle will usually simplify nicely.

3. Try finding a relation between consecutive elements in the sum and then multiplying by something to telescope it (see the Fibonacci example).

4. Try multiplying the kth term in the sum by x^k, simplifying in terms of x, then evaluating at $x = 1$. Alternately, if there is already some number raised to the kth power tacked onto each term, you might try replacing it with x, simplifying, then evaluating when x is the number in question.

5 Examples: Simplifying Sums

Example one: Show that $\binom{n}{k} = \binom{n}{n-k}$.

Solution: This follows directly from the fact that $(1+x)^n = (x+1)^n$ and the binomial theorem.

Example two: Show that $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$.

Solution: This follows from noting that $(1+x)^n = (1+x)(1+x)^{n-1}$ and comparing the x^k coefficient in both binomial expansions.

Example three: Simplify $\displaystyle\sum_{k=0}^{n} \binom{n}{k}$.

Method one: This is simply the expansion of $(1+1)^n = 2^n$.

Method two: This is the number of ways of choosing k out of n elements from a set, summed across k, so it is the number of ways of choosing some arbitrary number of elements from a set of size n. We can either choose or not choose each element, yielding two possibilities for each of n elements, so in total there are 2^n possibilities.

Method three: Use the identity that the sum is equal to 2^n as an inductive hypothesis. Then $\displaystyle\sum_{k=0}^{n} \binom{n}{k} = 2 + \sum_{k=1}^{n-1} \binom{n-1}{k} + \binom{n-1}{k-1} = 2 + 2 * (2^{n-1} - 1) = 2^n$, where the 3rd to last step is Pascal's identity, and the 2nd to last step uses the inductive hypothesis.

Example four: Simplify $\displaystyle\sum_{k=0}^{n} k\binom{n}{k}$.

Method one: As $\binom{n}{k} = \binom{n}{n-k}$, $k\binom{n}{k} + (n-k)\binom{n}{n-k} = \frac{n}{2}\binom{n}{k} + \frac{n}{2}\binom{n}{k}$, from which it immediately follows from the previous example that the sum in question simplifies to $n2^{n-1}$.

Method two: $(1+x)^n = \displaystyle\sum_{k=0}^{n} \binom{n}{k} x^k$, so $\frac{d}{dx}(1+x)^n = \displaystyle\sum_{k=0}^{n} k\binom{n}{k} x^{k-1}$. Then evaluating $\frac{d}{dx}(1+x)^n = n(1+x)^{n-1}$ at $x = 1$ immediately yields the identity.

Method three: We can use the following counting argument: rewriting $k = \binom{k}{1}$, we have the term $\binom{n}{k}\binom{k}{1}$. We can encode this algebraic quantity into a counting argument, by saying that we choose k out of n elements of a set, then mark one of them as a "special" element. As we sum across k, we can just say we are choosing sum arbitrary number of elements. However, we could reverse the order of this process by choosing the "special" element first, then choosing some arbitrary number of remaining elements. There are $\binom{n}{1}$ ways of choosing the "special" element, and as there are $n-1$ remaining elements, there are 2^{n-1} ways to choose some number of them. Thus the sum is equivalent to $n2^{n-1}$.

Excercise: Prove the above identity by induction.

Example five: Prove that $\binom{n+1}{k+1} = \displaystyle\sum_{m=k}^{n} \binom{m}{k}$.

Method one: Consider choosing $k+1$ elements of the set $S = \{1, 2, \ldots, n+1\}$ in order. Then, supposing that the first element chosen is $n+1-m$, there are $\binom{m}{k}$ ways to pick the remaining elements. If we sum across m, we have the above sum, but it is also equal to the number of ways of picking $k+1$ elements out of S, which is $\binom{n+1}{k+1}$.

Method two: Using the identity $1 + x\sum_{i=0}^{n-1}(1+x)^i = (1+x)^n$, and comparing the coefficient of x^k in both expansions, the identity immediately follows.

Excercise: Prove the above identity by induction.

Example six: Simplify $\displaystyle\sum_{k=0}^{n} \binom{n}{k}^2$.

Method one: If we rewrite $\binom{n}{k}^2 = \binom{n}{k}\binom{n}{n-k}$, the sum of the tops of the binomial coefficients is always $2n$, and the sum of the bottoms of the coefficients is always n. We now search for an expansion that would lead to this. If we were to place x^k after $\binom{n}{k}$ and x^{n-k} after $\binom{n}{n-k}$, we would end up with an overall term of x^n, suggesting that the sum consists of parts of the convolution for the coefficient of x^n in $(1+x)^n(1+x)^n$. In fact, it is not hard to see that it consists of all terms in this convolution, so that it is identically the coefficient of x^n in $(1+x)^n(1+x)^n = (1+x)^{2n}$, from which it becomes obvious that the sum is equal to $\binom{2n}{n}$ from the Binomial Theorem.

Method two: Rewrite $\binom{n}{k}^2$ as before and consider the following counting argument: the number of ways to choose some number of elements from $\{1, 2, \ldots, n\}$ and some other number of elmenets from $\{n+1, \ldots, 2n\}$ such that the total number of elements if n is the same as the number of ways to choose n elements from $\{1, 2, \ldots, 2n\}$. It then follows that the sum is equal to $\binom{2n}{n}$.

Example seven: Express the following in closed form:

$$\sum_{a=0}^{\lfloor \frac{k}{2} \rfloor} \binom{k}{a}\binom{k-a}{a}4^{k-a}$$

We can't directly preserve any sums. However, there is a sum of $2a$ in the bottom of the binomial coefficients. We can put a $2a$ somewhere else, namely in the exponent, by changing the 4 to a 2. Then we move a factor of 2^k to the outside, so that the sum becomes

$$2^k \sum_{a=0}^{\lfloor \frac{k}{2} \rfloor} \binom{k}{a}\binom{k-a}{a}2^{k-2a}$$

Now we start exploring possible expansions that could lead to this. $\binom{k}{a}\binom{k-a}{a}$ is equivalent to the trinomial coefficient $\binom{k}{a,a}$, indicating that we should have something of the form $(a+b+c)^k$. In addition, the values of the first two seem not to matter. Since they are both raised to an equal power (a) in any given term, they should be multiplicative inverses. Finally, the last term should yield the 2^{k-2a} part of the sum, so it should be 2. We are then looking for the coefficient of the constant term in $2^k(x+2+\frac{1}{x})^k$ If we let $x=y^2$, then $2^k(y^2+2+\frac{1}{y^2})^k$ factors to $2^k(y+\frac{1}{y})^{2k}$. The coefficient of the constant term in this expansion is $2^k\binom{2k}{k}$, and we are done.

Example eight: Prove that for every positive integer n,

$$\sum_{k=1}^{n} \frac{1}{k\binom{n}{k}} = \frac{1}{2^{n-1}} \sum_{k=1,\ k\ odd}^{n} \frac{\binom{n}{k}}{k}$$

Let A_n be the left-hand side, and B_n be the right-hand side. Because both sums are incredibly ugly to work with alone, or in terms of each other, we will deal with A_{n+1} and B_{n+1} in terms of A_n and B_n. For A_{n+1}, we have:

$$A_{n+1} = \sum_{k=1}^{n+1} \frac{1}{k\binom{n+1}{k}} = \frac{1}{n+1}(1+\sum_{k=1}^{n} \frac{n-k+1}{k\binom{n}{k}})$$

Breaking the sum apart, we get

$$\frac{1}{n+1} + \sum_{k=1}^{n} \frac{1}{k\binom{n}{k}} - \frac{1}{n+1}\sum_{k=1}^{n} \frac{1}{\binom{n}{k}} = \sum_{k=1}^{n} \frac{1}{k\binom{n}{k}} - \sum_{k=1}^{n} \frac{1}{(k+1)\binom{n+1}{k+1}} + \frac{1}{n+1}$$

from which it follows that

$$\sum_{k=1}^{n+1} \frac{1}{k\binom{n+1}{k}} + \sum_{k=2}^{n+1} \frac{1}{\binom{n+1}{k}} = \sum_{k=1}^{n} \frac{1}{k\binom{n}{k}} + \frac{1}{n+1}$$

or, equivalently,

$$2(\sum_{k=1}^{n+1}\frac{1}{k\binom{n+1}{k}}) = \sum_{k=1}^{n}\frac{1}{k\binom{n}{k}} + \frac{2}{n+1}$$

It follows that $A_{n+1} = \frac{A_n}{2} + \frac{1}{n+1}$. For B_{n+1}, we have two cases: n is even and n is odd. If n is odd, then

$$2^n B_{n+1} - 2^{n-1}B_n = \sum_{k=1,kodd}^{n}\frac{\binom{n+1}{k}-\binom{n}{k}}{k} = \sum_{k=1,kodd}^{n}\frac{\binom{n}{k-1}}{k}$$

And if n is even, the left-most expressiong becomes $\displaystyle\sum_{k=1,kodd}^{n}\frac{\binom{n}{k-1}}{k} + \frac{1}{n+1}$. Now let $S_n = \sum_{k=1,kodd}^{n}\frac{\binom{n}{k-1}}{k}$, we will prove by induction that $S_n = \frac{2^n}{n+1}$ if n is odd, and $\frac{2^n-1}{n+1}$ if n is even. The base cases are easily verified. The inductive step then immediately follows from the identity

$$(n+2)S_{n+1} - (n+1)S_n = \sum_{k=1,kodd}^{n+1}\binom{n+2}{k} - \sum_{k=1,kodd}^{n}\binom{n+1}{k} = 2^n \pm 1$$

from which it follows that $B_{n+1} = \frac{B_n}{2} + \frac{1}{n+1}$. As A_n and B_n are equal when $n = 1$, and they grow in the same manner, they are equal for all values of n, the desired result.

Comment: This problem is hard to solve with the above approach. Even with the idea of showing that they grow in the same manner, which is motivated by the fact that the two sums are too ugly to work with in basically any other way, there is a lot of difficult algebra involved. The algebra behind the left-hand side was motivated by trying to express A_{n+1} in terms of A_n. We started with the left-hand side because it was the less ugly of the two sums. Then, once we had a recurrence relation for A_n, our algebra for the right-hand side was aided because we knew what numbers we had to eventually get. We also used the golden rule that adding or subtracting numbers in Pascal's triangle yields nice results. Finally, we already knew that the sum of odd binomial coefficients would simplify. Currently, it may seem that, despite the difficulties arising in the approach we chose, there was not much other choice. It turns out, though, that a trick with integrals can completely destroy this problem! See the exercises if you are curious.

6 Problems on Sequences and Sums

1. Prove that $\binom{-n}{m} = (-1)^n\binom{n+m-1}{m}$.

2. (Vandermonde's Identity) Prove that

$$\sum_{i=0}^{c}\binom{a}{i}\binom{b}{c-i} = \binom{a+b}{c}$$

3. Suppose we have a Catalan-like sequence D satisfying $d_n = \sum_{i=0}^{n-2} d_i d_{n-i-2}$ for $n \geq 2$. Suppose $|p| < \frac{1}{4(a+b)}$. Find a closed form expression for $\sum_{n=0}^{\infty} d_n p^n$ in terms of d_0, d_1, and p.

4. (MOP 2006) Let $f(x) = \sum_{k=0}^{n} \binom{n}{k}^2 (1+x)^{2n-2k}(1-x)^{2k}$. Show that the coefficient of x^{2m-1} is zero for all positive integers m.

5. Suppose we have another sequence E satisfying $e_0 = \frac{1+\sqrt{3}}{2}$ and $e_n = \sum_{i=0}^{n} \sum_{j=0}^{n-i} e_i e_j$ for $n > 0$. Find a closed form expression for e_n.

6. Express $\sum_{k=0}^{n} \binom{n}{k}\binom{k}{a}$ in closed form.

7. (MOP 2006) Express $\binom{n}{0}^2 - \binom{n}{1}^2 + \binom{n}{2}^2 - \ldots + (-1)^n \binom{n}{n}^2$ in closed form.

8. Prove that

$$\sum_{i=0}^{c} \binom{i}{a}\binom{c-i}{b} = \binom{c+1}{a+b+1}$$

9. Prove that

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

and

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

This can be useful to make a k match up better with a $2k$ somewhere else.

10. Prove that

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

is 1 if $k = n-1$ and 0 otherwise.

11. (MOP 2007) Prove that

$$\sum_{i=0}^{n} \binom{n}{2i+1}\binom{i}{m} = 2^{n-2m-1}\binom{n-m-1}{m}$$

12. (MOP 2007) Let $u(i)$ denote the sum of the digits of the base-2 representation of i. For $k \geq mn$, find

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

13. Evaluate

$$(n+m+1)\int_0^1 x^n(1-x)^m dx$$

14. Evaluate

$$\frac{1}{\binom{2n}{1}} - \frac{1}{\binom{2n}{2}} + \frac{1}{\binom{2n}{3}} - \frac{1}{\binom{2n}{4}} + \ldots + \frac{1}{\binom{2n}{2n-1}}$$

15. Find a more direct approach to example eight in the section on combinatorial sums.

16. Show that

$$\sum_{k=0}^{n} \binom{2k}{k}\binom{2n-2k}{n-k} = 4^n$$

17. (MOP 2006) Let l be an even positive integer. Express

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

in closed form.

18. Express $\displaystyle\sum_{k=0}^{n} \frac{\binom{2k}{k}}{2k-1}\frac{\binom{2n-2k}{n-k}}{2n-2k-1}$ in closed form.

19. (MOP) Find

$$1 + \frac{3}{4} + \frac{3\cdot 5}{4\cdot 8} + \frac{3\cdot 5\cdot 7}{4\cdot 8\cdot 12} + \frac{3\cdot 5\cdot 7\cdot 9}{4\cdot 8\cdot 12\cdot 16} + \ldots$$

7 From Combinatorial to Algebraic

You have seen plenty of examples of how to solve algebraic problems by using ideas from generating functions. By now, if you have worked the examples together with some of the exercises, you have hopefully gained some intuition for manipulating generating functions. Now, finally, we will see the true magic of generating functions: how they can be used to kill not only algebraic but combinatorial problems. This is the reason I love algebraic combinatorics, because a seemingly difficult problem suddenly becomes easy when converted into algebra.

The idea is the same basic idea as in all those annoying word problems you had to do in Algebra I. It's tough to deal with all of the complexities in a real-world problem, so write down all of the essential information algebraically and then just grind through some algebra, get your answer, and convert it back to something sensible for the original question. Now, instead of a "real-world" problem, you have a combinatorics problem, but the same idea applies. Now, however, things can be complicated a bit since it might not be obvious how exactly to deal with something algebraically. The trick is to isolate the quantities you care about, look at how they interact, and then try to think about what algebraic structure interacts in the same way. If nothing interacts in the same way, then try to find something that behaves similarly and work from there. Of course, you might just be out of luck, but this is an issue with (almost) any technique.

The following examples may be illuminating:

Example one: (Putnam) How many ways can one number the sides of two (six-sided) dice in such a way that the likelihood of rolling any particular sum is the same as if the dice were numbered in the ordinary fashion? All numberings should be in positive integers, but the two dice might be numbered differently and we allow numbers to repeat on the same die.

Solution: What are the quantities we care about? We care about the numbers on each of the die. Now, there is a constraint on them, namely that if $d_1, \ldots, d_6$ are the numbers on the first die and $e_1, \ldots, e_6$ are the numbers on the second die, then the number of pairs (i, j) such that $d_i + e_j = k$ is equal to $1, 2, 3, 4, 5, 6, 5, 4, 3, 2, 1$ for $k = 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12$. When we see a constraint that has to do with all things summing to a fixed quantity, we should think of convolution. In fact, this constraint is the same as asking that the polynomials $D = \sum_{i=1}^{6} x^{d_i}$ and $E = \sum_{i=1}^{6} x^{e_i}$ satisfy $DE = x^2 + 2x^3 + \ldots + x^{12}$, or, equivalently, $DE = (x + x^2 + \ldots + x^6)^2$. We can factor the right-hand-side as $x^2(x+1)^2(x^2+x+1)^2(x^2-x+1)^2$. The constraint that the numbering consist of positive integers implies that x divides both D and E, so we know that the factors of x must distribute evenly across D and E. The constraint that the die be six-sided means that $D(1) = E(1) = 6$, so the factors of $x+1$ and x^2+x+1 must also distribute evenly. This leaves only the $x^2 - x + 1$ terms to distribute. If we distribute them evenly, we get the usual labeling. If we write $D = x(1+x)(1+x+x^2)(1-x+x^2)^2$ and $E = x(1+x)(1+x+x^2)$, we get a new labeling, which you can check for yourself is valid.

Example two: *Find the generating function for $\pi(n)$, the number of partitions of n.*

By definition, the number of partitions of n is the number of sets of positive integers with sum equal to n. For any positive integer k, a partition of n can involve using k 1 time, 2 times, 3 times, etc. Thus, we find that the number of paritions of n is equal to the coefficient of x^n in the product

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

8 Problems

1. Find the number of subsets of $\{1, 2, \dots, p\}$ with sum of elements divisible by p.

2. (Lucas' Theorem) If $n = n_0 + n_1 p + n_2 p^2 + n_3 p^3 + \dots$, where $0 \le n_i < p$, and m is defined similarly, then

$$\prod_i \binom{m_i}{n_i} \equiv \binom{m}{n} \pmod{p}$$

3. (Gauss-Lucas Theorem) If $P(x)$ is a polynomial, then the roots of $P'(x)$ lie in the convex hull of the roots of $P(x)$.

4. (MOP 2007) Suppose that there exist two distinct sets $A = \{a_1, \dots, a_n\}$ and $B = \{b_1, \dots, b_n\}$ such that for any r, the number of pairs (i, j), $i < j$ for which $a_i + a_j = r$ is the same as the number of such pairs for B. Prove that n is a power of two.

5. Let $A = \{a_1, a_2, \dots, a_n\}$ and $B = \{b_1, b_2, \dots, b_n\}$ be two subsets of $\{1, 2, \dots, 2n\}$ such that $A \cup B = \{1, 2, \dots, 2n\}$. Prove that for any r, the number of pairs (i, j) with $a_i + a_j \equiv r \pmod{2n}$ is equivalent to that for B. (Note that here the $i < j$ stipulation was dropped.)

6. Write an analogous statement to the above problem when there are three sets whose union is the first $3n$ natural numbers.

7. (Euler's Pentagonal Numbers Theorem) Let $\Pi(x)$ be the generating function of the partition sequence, and let $f(x)$ be the generating function defined as

$$f(x) = 1 + \sum_{k=1}^{\infty}(-1)^k(x^{k(3k-1)/2} + x^{k(3k+1)/2})$$

Prove that $\Pi(x)f(x) = 1$.

9 Combinatorial Nullstellensatz

9.1 Statement and Proof

Combinatorial Nullstellensatz is a technique involving polynomials that oftentimes allows one to convert combinatorial information very precisely into a tractable algebraic problem. The precise statement of Combinatorial Nullstellensatz is as follows:

Theorem 9.1 *Let F be a field. Let $S_1, \ldots, S_n$ be subsets of f. Suppose that $p = p(x_1, \ldots, x_n)$ is a polynomial in n variables, and that it has a highest-degree term $x_1^{t_1} \cdots x_n^{t_n}$ with non-zero coefficient. If $|S_i| > t_i$ for each i, there exists an s_i in each S_i such that $p(s_1, \ldots, s_n) \neq 0$.*

To understand the statement, we need to know what a field is. A field is simply a place where you can add, subtract, multiply, and divide (by any non-zero number). The examples you are familiar with are the rationals ($\mathbb{Q}$), the reals ($\mathbb{R}$), the complex numbers ($\mathbb{C}$), and the integers mod p ($\mathbb{Z}_p$). Before we go about proving Theorem 9.1, we need the following two lemmas, which are already familiar to you in the case of $\mathbb{C}$.

Lemma 9.2 *If p is a polynomial in one variable and $p(r) = 0$, then we can write p as $p = (x - r)p'$ for some polynomial p'.*

Proof Write $p = q(x - r) + c$, where c is a constant (we can do this by using polynomial division). If $p(r) = 0$, then we have $p(r) = q(r - r) + c \implies 0 = c$, so in fact $p = q(x - r)$ for some polynomial q, and setting $p' = q$ gives us the desired result.

Lemma 9.3 *Let p be a polynomial (in one variable) of degree n. If $r_1, \ldots, r_{n+1}$ are distinct and $p(r_1) = \cdots = p(r_{n+1}) = 0$, then $p = 0$. In other words, a non-zero polynomial of degree n has at most n distinct roots.*

Proof Let $p_0 = p$. We know that $p_0(r_1) = 0$, so by Lemma 9.2 we can write $p_0 = (x - r_1)p_1$ for some polynomial p_1. Then $p_0(r_2) = (r_2 - r_1)p_1(r_2)$, so $(r_2 - r_1)p_1(r_2) = 0$. $r_2 \neq r_1$, so $p_1(r_2) = 0$ and we can write $p_1 = (x - r_2)p_2$, and $p = (x - r_1)(x - r_2)p_2$. We can continue this way until we get $p = (x - r_1) \cdots (x - r_{n+1})p_{n+1}$. Then we see that p_{n+1} must equal zero since p has degree n whereas $(x - r_1) \cdots (x - r_{n+1})$ has degree $n + 1$, and the product of non-zero polynomials of degrees k and l has degree $k + l$. But then $p = (x - r_1) \cdots (x - r_{n+1}) \cdot 0$, so that $p = 0$ as well, as was to be shown.

We need one additional lemma that generalizes Lemma 9.3. It is as follows:

Lemma 9.4 *Let p be a polynomial in n variables with degree t_i in the ith variable. Suppose we have sets $S_1, \ldots, S_n$ with $|S_i| > t_i$ and such that $p(s_1, \ldots, s_n) = 0$ for all $s_i \in S_i$. Then $p = 0$.*

Proof We proceed by induction on n. The base case $n = 1$ is equivalent to Lemma 9.3. For $n > 1$, write

$$p(x_1, \ldots, x_n) = \sum_{j=1}^{t_n} p_j(x_1, \ldots, x_{n-1})x_n^j$$

In other words, we are writing p as a polynomial in x^n and grouping the terms together as such. Fix $s_1, \ldots, s_{n-1}$. Let $p_{s_1,\ldots,s_{n-1}}(x_n) = p(s_1, \ldots, s_{n-1}, x_n)$. Then

$$p_{s_1,\ldots,s_{n-1}}(x_n) = \sum_{j=1}^{t_n} p_j(s_1, \ldots, s_{n-1})x_n^j$$

With $s_1, \ldots, s_{n-1}$ fixed, $p_{s_1,\ldots,s_{n-1}}$ truly is just a polynomial in x_n. On the other hand, we know that $p_{s_1,\ldots,s_{n-1}}(x_n) = 0$ whenever $x_n \in S_n$. But $p_{s_1,\ldots,s_{n-1}}$ has degree less than $|S_n|$, so consequently $p_{s_1,\ldots,s_{n-1}} = 0$. This, then, implies that $p_j(s_1, \ldots, s_{n-1}) = 0$. But this logic applies to all $s_i \in S_i$, $i = 1, \ldots, n-1$. We can thus apply the inductive hypothesis and get that $p_j = 0$ for each j. But then this implies that $p = 0$, and so the induction is complete.

We are now ready for the proof of Combinatorial Nullstellensatz.

Proof We will argue by contradiction. Suppose that we have a polynomial satisfying the conditions, and that $p(s_1, \ldots, s_n) = 0$ whenever $s_i \in S_i$. Let g_i be defined as $\prod_{s \in S_i}(x_i - s)$. By replacing a monomial $x^{|S_i|}m$ with $(x^{|S_i|} - g_i)m$, we do not alter the value of f at any of the tuples $(s_1, \ldots, s_n)$, but we have replaced one monomial of degree at least $|S_i|$ in x_i with only lower-degree terms in x_i and without altering the degree any any of the x_j for $j \neq i$. Furthermore, we have done this by subtracting a polynomial of the form $g_i m$, where $\deg(m) + \deg(g_i) \leq \deg(f)$. Thus we can continue this process until we get a polynomial $\bar{f}$ whose degree in each x_i is less than $|S_i|$, and such that $f - \bar{f} = \sum_i g_i h_i$ for some polynomials h_i with $\deg(h_i) + \deg(g_i) \leq deg(f)$. On the other hand, $\bar{f}(s_1, \ldots, s_n) = 0$ for each $s_i \in S_i$, and as we have noted it has degree less than $|S_i|$ in x_i. Therefore, by Lemma 9.4, $\bar{f} = 0$ and so $f = \sum_i g_i h_i$.

On the other hand, consider any given $g_i h_i$. We are working under the assumption that $x_1^{t_1} \cdots x_n^{t_n}$ is a highest-order term in f, so there must in particular be some $g_i h_i$ whose $x_1^{t_1} \cdots x_n^{t_n}$ coefficient is non-zero. But we know that $\deg(g_i) + \deg(h_i) \leq \deg(f)$, so the only way to get a highest-order term in f is to multiply a highest-order term in g_i by a highest-order term in h_i. The only highest-order term in g_i is divisible by $x_i^{|S_i|}$, and $|S_i| > t_i$, so the coefficient of $x_1^{t_1} \cdots x_n^{t_n}$ must actually be zero. This is a contradiction, so our original assumption must have been false, and there is some $s_1, \ldots, s_n$ with $p(s_1, \ldots, s_n) \neq 0$.

9.2 Applications

Now that we've proved Combinatorial Nullstellensatz, here are some exciting applications. The first is known as the Cauchy-Davenport Theorem, and is a result about prime mods that I have found very useful.

Theorem 9.5 *Given two sets A and B, define $A + B$ to be the set $\{a + b \mid a \in A, b \in B\}$. If A and B are two subsets of $\mathbb{Z}_p$, then $|A + B| \geq \min |A| + |B| - 1, p$.*

Proof We consider two cases, depending on whether or not $|A|+|B|-1<p$. Suppose first that $|A|+|B|-1 \geq p$. Then for any $x \in \mathbb{Z}_p$, the sets $|A|$ and $s-|B|=\{s-b \mid b \in B\}$ must intersect by the pigeonhole principle (as the sum of their sizes exceeds p). Thus there exists some a, b such that $a=s-b$, and so $a+b=s$. This shows that $|A+B|=p$.

On the other hand, suppose that $|A|+|B|-1<p$. Suppose for the sake of contradiction that $|A+B|<|A|+|B|-1$, and let C be a set of size $|A|+|B|-2$ containins $A+B$. Consider the polynomial $p(x, y)=\prod_{c \in C}(x+y-c)$. Note that $p(a, b)=0$ for all $a \in A$, $b \in B$. The degree of p is $|A|+|B|-2$, so $x^{|A|-1} y^{|B|-1}$ is a highest-degree term. Its coefficient is, by the binomial theorem, $\binom{|A|+|B|-2}{|A|-1}$, which is not zero $(\bmod\ p)$ since $|A|+|B|-2<p$ by assumption. But then by Combinatorial Nullstellensatz (Theorem 9.1) there exists $a \in A$, $b \in B$ so that $p(a, b) \neq 0$, contradicting our previous observation. Therefore, our assumption that $|A+B|<|A|+|B|-1$ must be false. This completes the proof.

Remark Note that this bound is tight, as is seen by considering sets of the form $\{0,1, \ldots, k\}$.

The next problem is problem 6 from IMO 2007. I'm not too fond of this problem, but it at least shows a very direct application of Combinatorial Nullstellensatz that kills an otherwise difficult problem.

Problem Let n be a positive integer. Consider

$$S=\{(x, y, z) \mid x, y, z \in\{0,1, \ldots, n\}, x+y+z>0\}$$

as a set of $(n+1)^3-1$ points in three-dimensional space. Determine the smallest number of planes, the union of which contains S but does not include $(0,0,0)$.

Solution There are first of all the obvious $3n$ parallel planes of the form $x+y+z=k$, $k=1, \ldots, 3n$. We will show that $3n$ is actually the optimal number of planes. To do this, suppose that we have planes $P_1, \ldots, P_k$ satisfying the conditions. This means that we have planes of the form $a_i x+b_i y+c_i z=d_i$, $d_i \neq 0$. Now the obvious polynomial to use would be

$$p_0(x, y, z)=\prod_{i=1}^k\left(a_i x+b_i y+c_i z-d_i\right)$$

This is zero precisely when (x, y, z) is in the union of the k planes. However, this won't quite work because it doesn't actually address the fact that the point $(0,0,0)$ is not allowed to be in one of the hyperplanes. In fact, even if we could somehow apply Combinatorial Nullstellensatz (say with the possible values of x, y, and z being $\{0, \ldots, n\}$), then we wouldn't get anywhere because $p_0(0,0,0)$ is supposed to be non-zero. Let's fix this by taking a new polynomial p that *must* be zero at the point $0,0,0$:

$$p(x, y, z)=p_0(x, y, z)-p_0(0,0,0) \prod_{i=1}^n \frac{(x-i)(y-i)(z-i)}{-i^3}$$

This polynomial has been modified to be zero at $(0,0,0)$, but is equal to $p_0(x, y, z)$ for any $x, y, z \in\{0, \ldots, n\}$ other than $(x, y, z)=(0,0,0)$. Now we are in business, as if $k<3n$ then the highest-degree term in this polynomial is $3n$ and comes from the $x^n y^n z^n$ term, which has coefficient $\frac{p_0(0,0,0)}{(-1)^n(n!)^3}$, which is non-zero provided that $p_0(0,0,0)$ is non-zero (which must be true since $(0,0,0)$ isn't supposed to be in the union of the planes). Thus there is a point (x,y,z), $x,y,z \in \{0,\ldots,n\}$ such that $p(x,y,z) \neq 0$. We know that $(x,y,z) \neq (0,0,0)$, as that was the point of our construction, so there is some other point not covered by our planes, which is a contradiction. It follows that k must be at least $3n$, so at least $3n$ planes are needed to cover all the points but $(0,0,0)$, as claimed.

Another application is the following theorem in graph theory, which seems pretty cool to me. The solution uses the same idea of modifying the polynomial to rule out a "bad" solution.

Theorem 9.6 *If X is a graph with average degree greater than $2p-2$ and maximum degree at most $2p-1$, then X contains a p-regular subgraph.*

Corollary 9.7 *Every $(2p-1)$-regular graph contains a p-regular subgraph.*

Proof For each edge e, let x_e be a variable corresponding to the edge. A subgraph is just a subset of the edges and vertices, so we will let $x_e = 1$ represent that the edge e is used and $x_e = 0$ represent that it is not used (we'll worry about the vertices later). Then let

$$p((x_e)_{e\in E}) = \prod_{v\in V}(1 - (\sum_{e\sim v} x_e)^{p-1}) - \prod_{e\in E}(1 - x_e)$$

Here $e \sim v$ means that e is an edge incident on the vertex v. The left-hand term in p has been constructed to be 1 if and only if every vertex has degree 0 mod p in the subgraph. Note that this is exactly the conditions under which we have found a p-regular subgraph, as then we can just use all the vertices that have degree p (as the condition that the maximum degree is $2p-1$ means that a vertex has degree 0 mod p if and only if it has degree 0 or p). The only problem that can arrise, then, is if every vertex has degree 0 and none have degree p. This only happens if all the x_e are equal to zero, which the right-hand term prevents, as it is 1 if all the edges are unused and 0 otherwise.

Now, we are ready to apply Combinatorial Nullstellensatz. We are picking each variable x_e from among the values 0 and 1, so to apply Combinatorial Nullstellensatz we need to make sure that the $\prod_{e\in E} x_e$ term is non-zero as well as of highest-degree. Each term in the left-hand product has degree $p-1$ in the x_e, and there are $|V|$ of them, so that the degree of the left-hand term is (at most) $|V|(p-1)$. On the other hand, the average degree is greater than $2p-2$, so the total number of edges is greater than $\frac{1}{2}|V|(2p-2) = |V|(p-1)$, so the $\prod_{e\in E} x_e$ term from the right-hand product is of highest degree, and its coefficient is $-(-1)^{|E|}$, which is non-zero. Thus we can apply Combinatorial Nullstellensatz to find some values for the x_e among $\{0,1\}$ such that $p((x_e)_{e\in E}) \neq 0$, which is exactly what we wanted. The theorem follows.

9.3 Exercises

1. Given a set of $2p-1$ integers, show that there exists a subset of p of them whose sum is divisible by p. Use this to show that, given any set of $2n-1$ integers, there exists a subset of n of them whose sum is divisible by n (hint: use induction on the number of prime factors of n).

2. Let R be the unit hypercube in n-dimensional space (that is, its vertex set is $\{0,1\}^n$). What is the minimal number of hyperplanes ($(n-1)$-dimensional planes) whose union contains all the vertices except the all-0's vector?

3. Given a matrix $A = (a_{ij})$, the *permanent* $Per(A)$ is $\sum_\pi \prod_{i=1}^n a_{i\pi(i)}$, where π ranges over all possible permutations. It is basically the determinant without any of the negative signs. Prove the following: If A is a matrix with $Per(A) \neq 0$, then for any vector b and any family of 2-element sets $S_1, \ldots, S_n$, there is a vector x whose ith coordinate is in S_i and such that Ax differs from b in all of its coordinates.

10 Linear Algebra

10.1 Notation

$V(X)$ = vertices of X
$E(X)$ = edges of X
$A(X)$ = adjacency matrix of X
$d(v)$ = degree of v = number of vertices adjacent to v
$Null(A)$ = nullspace of $A = \{h \mid Ah = 0\}$
$Col(A)$ = span of columns of A
$Row(A)$ = span of rows of A
$\langle v, w\rangle$ = inner (i.e. dot) product of v and w
$V^\perp$ = orthogonal complement of $V = \{w \mid \langle w, v\rangle = 0 \;\forall\; v \in V\}$

10.2 Basic Notions

Given a graph X, we can associate with it a matrix A, called its *adjacency* matrix, by defining $A_{uv} = 1$ if vertices u and v are adjacent, and $A_{uv} = 0$ otherwise. We use the notation $u \sim v$ to say that u and v are adjacent. While A is a convenient way of writing down the edges in a graph, it is also much more than that. For instance, suppose that X have n vertices. Then A, as an $n \times n$ matrix, is a linear transformation on an n-dimensional vector space. What do elements of this space look like? Each row and column of A is associated with a vertex of X, so we can associate the coordinates of any n-dimensional vector with the vertices of X as well. In this sense, vectors act as functions on X, where $f(v)$ is the value of the vth coordinate of f (remember, v is a vertex, not an integer). The right way to think about vectors are as functions, so we will call vectors f instead of v to reflect this. (Also, it avoids the unfortunate clash of notation caused by the fact that "vector" and "vertex" both start with v.)

Once we begin thinking about vectors as functions, it is natural to ask what the adjacency matrix does to our space of functions. What should go in the equation below?

Theorem 10.1

$$(Af)(v) = \sum \tag{*}$$

10.3 USAMO 2008, Problem 6

Consider the following problem (re-worded from this year's USAMO):

Problem You are given a graph X. You want to partition $V(X)$ into two sets, S and T, such that

1. For all $s \in S$, s is adjacent to an even number of vertices in S.
2. For all $t \in T$, t is adjacent to an even number of vertices in T.

Show that this is always possible, and furthermore that the number of ways to do this is a power of 2.

Question How can we express this condition algebraically (e.g., in terms of some function)?

Question What field should be working in (hint: there is somewhere better than $\mathbb{C}$ to express the conditions of the problem)?

Question What constraints does the function need to satisfy in order to represent a valid partition?

Lemma 10.2 *f represents a valid partition iff*

$$\sum_{u \sim v} \qquad\qquad\qquad =$$

for all $v \in V(X)$.

We can express the above relation in the form $Lf = d$, where L is whatever matrix represents the system of equations. In fact, it turns out that L is a special matrix called the *Laplacian* of X.

Lemma 10.3 *$L = D - A = D + A$, where D is the diagonal matrix such that $D_{vv} = d_v$.*

Proof

$$(Lf)(v) = \sum_{u \sim v} f(v) - f(u) = \sum_{u \sim v} f(v) - \sum_{u \sim v} f(u) = d(v)f(v) - (Af)(v) = ((D - A)f)(v)$$

In this formulation, the second part of the problem becomes easy. If there is a solution, i.e. some f with $Lf = d$, then $L(f + h) = d$ iff h is in the null space $Null(L)$ of L (why?). Also, elements of $Null(L)$ have a unique representation in the form

$$\sum_i c_i b_i$$

where $c_i \in \mathbb{Z}/2\mathbb{Z}$ and $\{b_1, \ldots, b_k\}$ is a basis for $Null(L)$. Thus $|Null(L)| = 2^k$, so we have shown that if a solution exists, then the total number of solutions is a power of 2.

It turns out that the hardest part of this problem is showing that a solution exists. Recall that a solution exists iff there exists f with $Lf = d$, which is the same as saying that d is in the column space of L. Note that an elements is in the null space of L iff it is orthogonal to every row vector of L (this is just the definition of matrix multiplication). Taken backwards, this shows that the row space of L is equal to $Null(L)^{\perp}$. However, L is symmetric, so the row and column spaces of L are equal. Thus, the $Col(L)$ is equal to $Null(L)^{\perp}$. So, to show that $d \in Col(L)$, it suffices to show that, for all $k \in Null(L)$, k is orthogonal to d.

Lemma 10.4 *Let $h \in Null(L)$. Then $\langle d, h \rangle = 0$.*

Proof $h \in Null(L) \implies Lh = 0 \implies h^T Lh = 0 \implies h^T(D - A)h = 0 \implies h^T Dh = h^T Ah$. But $h^T Dh = \sum_v h(v)d(v)h(v) = \sum_v d(v)h(v)^2 = \sum_v d(v)h(v)$ since we are working in $\mathbb{Z}/2\mathbb{Z}$. This last expression is just $\langle d, h \rangle$, so if we can show that $h^T Ah = 0$ then we will be done.

Sublemma 10.5 $h^T Ah = 0$

Proof $h^T Ah = \sum_{u,v} h(u)A_{u,v}h(v)$. If $u = v$, then $A_{u,v} = 0$. If $u \neq v$, then $h(u)A_{u,v}h(v) + h(v)A_{v,u}h(u) = 0$ since $A_{u,v} = A_{v,u}$ and $x + x = 0$ in $\mathbb{Z}/2\mathbb{Z}$. We have thus paired off all non-zero terms in our sum to add up to zero, so the entire sum must be zero, as stated.

The above sublemma completes the proof of the lemma.

But now, we are done, since we have shown that $d \in Col(L)$, so there is some f with $Lf = d$, and this f represents a valid partition.

I think arguments like this are really cool. What started out as a messy combinatorial problem turned into a relatively nice algebra problem whose solution (modulo knowing the requisite linear algebra) is mostly straightforward. However, while we are looking at combinatorial problems algebraically, we might as well develop some machinery to deal with them. We will see some of this in the next problem.

10.4 USA Mock Olympiad 2, 2007-08, Problem 4

The next problem doesn't fall quite so easily to algebraic techniques, but an algebraic attack is still far easier than any combinatorial attack. Here it is:

Problem Start with an $m \times n$ grid of integers. Every second, simultaneously replace each integer with the sum of the integers in adjacent squares. For what ordered pairs (m, n) will all of the integers eventually be even, no matter what the original integers were?

Note that this is again a question in $\mathbb{Z}/2\mathbb{Z}$ (which I will start referring to as the field $\mathbb{F}_2$), rather than something like $\mathbb{R}$ or $\mathbb{C}$. We can create a graph X where adjacency in the graph is equivalent to adjacency in the grid. Then, as before, we have a natural function to associate with any assignment of integers to points in the grid (namely, the function that assigns the relevant integer to each vertex).

Question At the end of each second, if the current assignment of integers is f, then what will the next assignment of integers be?

So, we really want to determine if there is some integer k such that $A^k f = 0$ in $\mathbb{F}_2$ for all f, or equivalently (since there are only finitely many $f : X \to \mathbb{F}_2$) $A^k = 0$ for some k. This condition is called *nilpotence.*

It turns out that $A^k = 0$ iff A has no non-zero eigenvalues (as in $\mathbb{R}$, where eigenvalues can be complex, A could have eigenalues not in $\mathbb{F}_2$). We can show one direction (which?) by noting that λ^k is an eigenvalue of A^k whenever λ is an eigenalue of A. So, call a grid *bad* if there exists an initial assignment of integers such that the integers will never be even (equivalently, A is not nilpotent). In order to show that a grid is bad, all we have to do is find a single non-zero eigenvalue of A.

As an aside, we introduce the notion of field extensions of $\mathbb{F}_2$. I won't rigorously define them, but we can think of a field extension of $\mathbb{F}_2$ as a field that somehow "contains" $\mathbb{F}_2$ (in much the same way that $\mathbb{C}$ contains $\mathbb{R}$ and $\mathbb{R}$ contains $\mathbb{Q}$). The field $\mathbb{F}_{2^k}$ is the roots of the polynomial $x^{2^k} - x$. I haven't shown that it makes any sense to talk about these roots, or that they are all distinct, but assuming that this is true, we can check that the roots do indeed form a field. The hardest part is checking closure under addition, which can be done by using the binomial theorem and considering when $\binom{2^k}{i}$ can be odd.

Knowing about these additional fields gives us more tools with which to find non-zero eigenvalues of A. What are the eigenvalues of A? It turns out that this can be answered by looking at the eigenvalues of a much simpler matrix. To do this, we need to develop some machinery.

Definition If $f : X \to F$ and $g : Y \to F$, then $f \otimes g : X \times Y \to F$ is defined as $(f \otimes g)(u, v) = f(u)g(v)$.

Definition If A, B are matrices, then $A \otimes B$ is the matrix with $(A \otimes B)_{(u,v),(u',v')} = A_{u,u'} B_{v,v'}$.

Exercise Show that $(A \otimes B)(f \otimes g) = (Af) \otimes (Bg)$.

Definition Given two graphs X and Y, define the *Cartesian product* $X \Box Y$ to be the graph with vertex set $V(X)\ V(Y)$ such that $(v, w) \sim (v', w')$ iff either $v = v'$ and $w \sim w'$ or $v \sim v'$ and $w = w'$.

Exercise Show that $A(X \Box Y) = A(X) \otimes I_Y + I_X \otimes A(Y)$. Here I_X is the identity matrix on X.

Now, suppose that f is an eigenvector of $A(X)$ with eigenvalue λ, and g is an eigenvector of $A(Y)$ with eigenvalue μ. Then

$$(A(X) \otimes I + I \otimes A(Y))(f \otimes g) = (A(X)f) \otimes g + f \otimes (A(Y)g) = (\lambda f) \otimes g + f \otimes (\mu g) = (\lambda + \mu)(f \otimes g)$$

So, $f \otimes g$ is an eigenvector of $X \square Y$ with eigenvalue $\lambda + \mu$. In particular, $X \square Y$ cannot be nilpotent unless all of the eigenvalues of $A(X)$ and $A(Y)$ are equal and also equal to each other.

Exercise Let P_k be the path on k vertices, and let $G_{(m,n)}$ be an mxn grid. Then $G_{(m,n)} = P_m \square P_n$.

What this means is that we can get a lot of information about our original combinatorial problem by investigating the eigenvalues of P_k for each k. It turns out that the answer to the original problem is as follows:

Proposition 10.6 *(m, n) is bad unless $(m, n) = (2, 2)$ or $(2^i - 1, 2^j - 1)$.*

One can prove that $(2, 2)$ isn't bad with an exhaustive search. Showing that $(2^i - 1, 2^j - 1)$ isn't bad is on the homework. I am going to devote the rest of this class to proving Lemmas 10.7, 10.8, and 10.9 below, which you should verify imply that all other grids are bad.

Lemma 10.7 *If $2 \mid k - 1$, then P_k has an eigenvalue of 0.*

Lemma 10.8 *If $3 \mid k - 1$, then P_k has an eigenvalue of 1.*

Lemma 10.9 *If $m \geq 5$, m odd, and $m \mid k - 1$, then P_k has at least two distinct eigenvalues.*

Finding eigenvalues in $\mathbb{F}_2$ is rather annoying, so first let's get some intuition for the eigenvalues of P_k over $\mathbb{C}$. The best way to think about an eigenvector of the adjacency matrix is really as an *eigenfunction* f on the graph such that, at any vertex v, the sum of the values of f on neighbouring vertices is a fixed multiple of $f(v)$. For paths P_k, a solution that "almost" works is to guess $f(v_n) = \zeta^n$ for some ζ, where v_n is the nth vertex along the path. Note that this would have eigenvalue $\zeta + \zeta^{-1}$. The problem is that the eigenfunction relation doesn't hold on the two ends of the graph. If f were a function that went to zero just past the edge of the graph, then we would be fine. We can do this by instead taking $f(v_n) = \zeta^n - \zeta^{-n}$, where $\zeta^{k+1} = 1$. (A more familiar form of this would be $f(v_n) = 2i \sin(\frac{n\pi}{k+1})$, or just $f(v_n) = \sin(\frac{n\pi}{k+1})$, but we actually don't want to use this because it's hard to generalize trigonometry to fields other than $\mathbb{R}$ and perhaps $\mathbb{C}$, but it's easy to talk about roots of unity like ζ in any field.)

So, how can we find a function like f in $\mathbb{F}_2$? We can just do the exact same thing! If $\zeta^{k+1} = 1$, then $f(v_n) = \zeta^n - \zeta^{-n}$ works as an eigenfunction with eigenvalue $\zeta + \zeta^{-1}$. We just need to be careful of two things: (1) that $\zeta^n - \zeta^{-n}$ isn't always 0 (or else our eigenfunction is trivial), and also that $\zeta + \zeta^{-1}$ takes on at least two different values.

Lemmas 10.7 and 10.8 are left as exercises. We will focus on proving Lemma 10.9. It is a result from field theory that the non-zero elements of $\mathbb{F}_{2^k}$ form a cyclic group under multiplication. Thus for any d such that $d \mid 2^k - 1$, there is an element of $\mathbb{F}_{2^k}$ of multiplicative order d, or in other words a primitive dth root of unity. To prove Lemma 10.9, we want to pick k such that $m \mid 2^k - 1$. By Euler's theorem, for any odd m we can pick $k = \phi(m)$

and get the ζ we want. To see that $f(v_n) = \zeta^n - \zeta^{-n}$ is non-degenerate, take $n = 1$. If $\zeta - \zeta^{-1} = 0$, then $\zeta^2 = 1$, so in particular ζ does not have multiplicative order m. Similarly, $f(v_n) = (\zeta^2)^n - (\zeta^2)^{-n}$ is non-degenerate since then we would have $\zeta^4 = 1$, so ζ again does not have have multiplicative order m (here we see why $m \geq 5$ is necessary).

Now, suppose that $\zeta + \zeta^{-1} = \zeta^2 + \zeta^{-2}$. Then $\zeta^4 + \zeta^3 + \zeta + 1 = 0 \iff (\zeta^3 + 1)(\zeta + 1) = 0$. But we have already barred both $\zeta = 1$ or $\zeta^3 = 1$ (both imply that ζ has multiplicative order less than m), so this is impossible, and both of the eigenvalues we found were actually distinct. This is precisely what we wanted to show, so we are done with Lemma 10.9, and, modulo some details that you should work out on your own, with the entirety of the problem.

In this problem, we used significantly more machinery from abstract algebra and algebraic graph theory. However, once we had the general ideas down, the proof is again not too bad. After developing the general machinery, all that was left for us was essentially computations.

To wrap up, I'd like to note that, in both of these cases, we were able to start using linear algebra tools on a graph theoretic problem because we could write down nice functions related to the problem in question. In practice, you may not start off by picking the "right" function – for example, it may not satisfy any linear relations. Hopefully, with some modification it will be possible to make it linear, or perhaps you can apply more powerful tools in algebraic combinatorics like Combinatorial Nullstellensatz. Of course, there's also the possibility that you really can't pick an appropriate function – then you'll be stuck with plain old pure combinatorial methods. Oh well. With this tidbit in mind, I'll leave you with some exercises. Some will patch up things I skipped over in class, and others will be opportunities to practice the methods presented here.

10.5 USA Mock Olympiad 8, 2007-08, Problem 1

Problem: Let p be a prime greater than 2. Find the number of pairs of 6-tuples (a, b, c, d, e, f) with $a^2 + b^2 + c^2 \equiv d^2 + e^2 + f^2 \pmod{p}$.

Answer: $p^5 + p^3 - p^2$.

We begin with a lemma:

Lemma: There exists an ordered tuple (a, b, c) with $a^2 + b^2 + c^2 \equiv 0$ and $c \not\equiv 0$.

Proof: Let S be the set of quadratic residues mod p. Note that $|S| = \frac{p+1}{2}$. Take any r and note that $S \cap (r - S)$ must be non-empty by Pigeonhole (both sets contain more than half of the residues mod p – here $r - S = \{r - s \mid s \in S\}$). Thus every residue is the sum of two quadratic residues, so that in particular there exists a pair (a, b) where $a^2 + b^2 = -1$. Thus there exists a triple $(a, b, 1)$ with $a^2 + b^2 + 1 = 0$. This proves the lemma.

Now take such a pair a, b, c. We treat all the triples (x, y, z) as a vector space over $\mathbb{F}_p$ and consider two cases.

Case one: (x, y, z) is not in the orthogonal space to (a, b, c). Then take any such triple (x, y, z). We can group it with all triples of the form $(x + ka, y + kb, z + kc)$. Such a triple has sum of squares equal to $x^2 + y^2 + z^2 + 2k(xa + yb + zc) + k^2(a^2 + b^2 + c^2) = (x^2 + y^2 + z^2) + 2k(xa + yb + zc)$. Since $xa + yb + zc \neq 0$ by assumption, this group contains every non-zero residue exactly once. Thus all triples outside the orthogonal space of (a, b, c) have evenly distributed norms across the residues mod p. Since there are $p^3 - p^2$ such vectors (a vector space of dimension d has p^d elements), every residue r gets $p^2 - p$ triples (u, v, w) with $u^2 + v^2 + w^2 = r$ in this case.

Case two: (x, y, z) is in the orthogonal space to (a, b, c). Let (x, y, z) be a vector in this space that is linearly independent from (a, b, c). Then all vectors in the space are of the form $m(x, y, z) + n(a, b, c)$, and have norm equal to $m^2(x^2 + y^2 + z^2)$. (This is because (a, b, c) is itself in the orthogonal space, since $a^2 + b^2 + c^2 = 0$.) The claim, however, is that $x^2 + y^2 + z^2 \not\equiv 0$. This is because, were it equal to zero, then it would be orthogonal to both (a, b, c) and (x, y, z). However, the dimension of the space of vectors orthogonal to both (a, b, c) and (x, y, z) is 1. But (a, b, c) is in this space. This contradicts the assumption that (a, b, c) and (x, y, z) are linearly independent, so it must be that $x^2 + y^2 + z^2 \not\equiv 0$. Since, as m ranges from 0 to $p - 1$, $m^2(x^2 + y^2 + z^2)$ takes on the value of zero once and then either all quadratic residues or non-residues twice, this space contributes p triples such that $u^2 + v^2 + w^2 = 0$, and $2p$ triples to $\frac{p-1}{2}$ non-zero values of r such that $u^2 + v^2 + w^2 = r$.

This analysis shows that there are p^2 triples (u, v, w) with $u^2 + v^2 + w^2 = 0$. For the remaining residues, for $\frac{p-1}{2}$ of them there are $p^2 + p$ such triples, and for the other $\frac{p-1}{2}$ there are $p^2 - p$ triples. Now note that the number of 6-tuples with $a^2 + b^2 + c^2 = d^2 + e^2 + f^2$ is equal to the sum of the squares of the number of triples for each residue. Our answer is thus the sum of the squares of all these numbers, so it is $(p^2)^2 + \frac{p-1}{2}\left((p^2 - p)^2 + (p^2 + p)^2\right) = p^4 + (p - 1)(p^4 + p^2) = p^5 + p^3 - p^2$, as stated.

10.6 Chromatic Number of the Odd-Distance Graph

The *odd-distance graph* is the graph $\mathcal{O}$ with $V(\mathcal{O}) = \mathbb{R}^2$ and where two vertices are connected if their Euclidean distance is an odd integer. A *coloring* of a graph with k colors is a partition of the vertices into k sets (color classes) so that no two vertices within the same set are connected. The purpose of this section is to show that there exists no finite coloring of $\mathcal{O}$ where the color classes are measurable (measurable means that they have a well-defined size in the sense that a function that is 1 on the color class and 0 outside of it can be integrated). The proof is as follows:

Consider the operator $B_\alpha : L^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2)$ defined by

$$(B_\alpha f)(x, y) = \int_{-\pi}^{\pi} \sum_{k=0}^{\infty} \alpha^{-k} f(x + (2k + 1)\cos(\theta), y + (2k + 1)\sin(\theta))d\theta \tag{1}$$

Clearly, B_α is a linear operator. We also make the following observation:

Lemma 10.10 *Let I be an independent set in $\mathcal{O}$, and let g be any function that is zero outside of I. Then $\langle f, B_\alpha f\rangle = 0$.*

Proof

$$\begin{aligned}
\langle f, B_\alpha f\rangle &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)(B_\alpha f)(x,y)dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\int_{-\pi}^{\pi}\sum_{k=0}^{\infty}\alpha^{-k}f(x+(2k+1)\cos(\theta), y+(2k+1)\sin(\theta))d\theta dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\pi}^{\pi}\sum_{k=0}^{\infty}\alpha^{-k}f(x,y)f(x+(2k+1)\cos(\theta), y+(2k+1)\sin(\theta))d\theta dxdy \\
&= 0
\end{aligned}$$

In the last equality we used the fact that $f(x,y)f(x+(2k+1)\cos(\theta), y+(2k+1)\sin(\theta)) = 0$ since not both (x,y) and $(x+(2k+1)\cos(\theta), y+(2k+1)\sin(\theta))$ can be in I (they are at odd distance), so at least one of the two must be zero.

We can use this to bound the chromatic number χ of $\mathcal{O}$. Let $C_\alpha = I - \frac{\alpha-1}{2\pi}B_\alpha$, where I is the identity. Then C_α is equivalent to convolution by some function, and so is diagonalized by the Fourier transform on $\mathbb{R}^2$. Therefore, its operator norm is equal to its largest eigenvalue. We thus have the following:

Lemma 10.11

$$\chi \geq \frac{\rho(C_\alpha)}{\rho(C_\alpha)-1} \tag{2}$$

Proof By the preceeding comments, it suffices to show that $\chi \geq \frac{||A||}{||A||-1}$. Suppose that there exists a χ-coloring of $\mathcal{O}$ with color classes $I_1, \ldots, I_\chi$. Let S_r be a circle with radius r centered at the origin. Let f_i be defined as

$$f_i(x) = \left\{ \begin{array}{ll} 1 & x \in I_i \cap S_r \\ 0 & x \notin I_i \cap S_r \end{array} \right\} \tag{3}$$

Let $f = f_1 + \ldots + f_\chi$. We note that each f_i satisfies the conditions of Lemma 10.10. Therefore, $\langle f_i, C_\alpha f_i\rangle = \langle f_i, f_i\rangle$. We then have:

$$\begin{aligned}
2(\chi-1)||A||||f||^2 &= \sum_{i,j=1}^{\chi} ||A||||f_i - f_j||^2 \\
&\geq \sum_{i,j=1}^{\chi} \langle f_i - f_j, C_\alpha(f_i - f_j)\rangle \\
&= \sum_{i,j=1}^{\chi} \langle f_i, C_\alpha f_i\rangle + \langle f_j, C_\alpha f_j\rangle - \langle f_i, C_\alpha f_j\rangle - \langle f_j, C_\alpha f_i\rangle \\
&= \left(\sum_{i,j=1}^{\chi} ||f_i||^2 + ||f_j||^2\right) - 2\sum_{i,j=1}^{\chi} \langle f_i, C_\alpha f_j\rangle \\
&= 2\chi||f||^2 - 2\langle \sum_{i=1}^{\chi} f_i, C_\alpha(\sum_{j=1}^{\chi} f_j\rangle \\
&= 2\chi||f||^2 - 2\langle f, C_\alpha f\rangle
\end{aligned}$$

So $2(\chi-1)||A||||f||^2 \geq 2\chi||f||^2 - 2\langle f, C_\alpha f\rangle$. This re-arranges to $\chi(||A||||f||^2 - ||f||^2) \geq ||A||||f||^2 - \langle f, C_\alpha f\rangle$, or $\chi \geq \frac{||A||}{||A||-1}\left(1 - \frac{\langle f, C_\alpha f\rangle}{||f||^2}\right)$. We will bound $\langle f, C_\alpha f\rangle$ in terms of α and r. Let $r = 2k+1$, where r is an integer. Let $D_\alpha = I - C_\alpha$. Then it suffices to show show that $\frac{\langle f, D_\alpha f\rangle}{||f||^2}$ approaches 1 as $k \to \infty$. For a point at distance between $2j$ and $2j+2$ from the origin, $(D_\alpha f)(x,y) \geq (1-\alpha)\left(1+\alpha+\ldots+\alpha^{k-j}\right)$ for $j = 0, \ldots, k-1$. Therefore, $\langle f, D_\alpha f\rangle$ is bounded below by the sum

$$\sum_{j=0}^{k-1}(1-\alpha^{k+1-j})\pi((2j+2)^2-(2j)^2) \tag{4}$$

This simplifies to $\pi\left((2k)^2 - 8\frac{\alpha^{k+2}-(k+1)\alpha^2+k\alpha}{(\alpha-1)^2} + 4\frac{\alpha^{k+1}-\alpha}{\alpha-1}\right)$. On the other hand, $||f||^2 = \pi(2k+1)^2$, so we want to look at the quantity

$$\frac{(2k)^2 - 8\frac{\alpha^{k+2}-(k+1)\alpha^2+k\alpha}{(\alpha-1)^2} + 4\frac{\alpha^{k+1}-\alpha}{\alpha-1}}{(2k+1)^2} \tag{5}$$

We break this up into the two quantities

$$\frac{(2k)^2}{(2k+1)^2} - 4\frac{\alpha^{k+2}+\alpha^{k+1}-(2k+1)\alpha^2+2k\alpha-\alpha}{(2k+1)^2(\alpha-1)^2} \tag{6}$$

Clearly $\frac{(2k)^2}{(2k+1)^2}$ tends to 1 as $k \to \infty$. If the numerator of the other quantity does not tend to ∞, then we are done since the denominator does tend to ∞. Otherwise, we can use L'hopital's rule, from which we get that the second quantity tends to

$$4\frac{\alpha^{k+2}\ln(\alpha) + \alpha^{k+1}\ln\alpha - 2\alpha^2 + 2\alpha}{(8k+4)(\alpha-1)^2} \tag{7}$$

The top is clearly bounded as $k \to \infty$ (remember $\alpha < 1$), and the bottom is clearly unbounded, so this expression goes to 0 as $k \to \infty$, so that $\frac{\langle f, D_\alpha f}{||f||^2}$ does indeed tend to 1. Therefore, we can let $r \to \infty$, so that $\frac{\langle f, C_\alpha f\rangle}{||f||^2} \to 0$, and we get the desired bound.

We next compute the eigenvalues of B_α (if λ is an eigenvalue of B_α, then $1 - \frac{\alpha-1}{2\pi}\lambda$ is an eigenvalue of C_α). Since B_α is diagonalized by the Fourier transform, $f_{(r,s)}(x,y) = e^{i(rx+sy)}$ with $r, s \in \mathbb{R}$ are the eigenfunction of B_α. We see that the eigenvalue of the eigenfunction $f_{(r,s)}$ is given by

$$\lambda_{(r,s)} = \int_{-\pi}^{\pi} \sum_{k=0}^{\infty} \alpha^{-k} e^{i(2k+1)(r\cos(\theta)+s\sin(\theta))} d\theta = \int_{-\pi}^{\pi} \sum_{k=0}^{\infty} \alpha^{-k} e^{i(2k+1)\sqrt{r^2+s^2}\cos(\theta+\phi)} d\theta \tag{8}$$

for an appropriately chosen ϕ. Thus we need only actually consider $\lambda_{(r,0)}$, which we from now on denote $\lambda(r)$. Then we have

$$\lambda(r) = \int_{-\pi}^{\pi} \sum_{k=0}^{\infty} \alpha^{-k} \left(e^{ir\cos(\theta)}\right)^{2k+1} = \int_{-\pi}^{\pi} \frac{e^{ir\cos(\theta)}}{1-\alpha^{-1}e^{2ir\cos(\theta)}} d\theta \tag{9}$$

Here we have simply summed the geometric series. Since B_α is symmetric, $\lambda(r)$ must be real, so we can take the real part of the integral:

$$\begin{aligned}
\lambda(r) &= Re\left[\int_{-\pi}^{\pi} \frac{(\cos(r\cos(\theta)) + i\sin(r\cos(\theta)))(1-\alpha^{-1}\cos(2r\cos(\theta)) + i\alpha^{-1}\sin(2r\cos(\theta)))}{(1-\alpha^{-1}\cos(2r\cos(\theta)))^2 + \alpha^{-2}\sin(2r\cos(\theta))^2} d\theta\right] \\
&= \int_{-\pi}^{\pi} \frac{\cos(r\cos(\theta))(1-\alpha^{-1}\cos(2r\cos(\theta))) - \alpha^{-1}\sin(r\cos(\theta))\sin(2r\cos(\theta))}{1+\alpha^{-2}-2\alpha^{-1}\cos(2r\cos(\theta))} d\theta \\
&= \int_{-\pi}^{\pi} \alpha \frac{\alpha\cos(r\cos(\theta)) - \cos(r\cos(\theta))\cos(2r\cos(\theta)) - \sin(r\cos(\theta))\sin(2r\cos(\theta))}{\alpha^2+1-2\alpha\cos(2r\cos(\theta))} d\theta \\
&= \int_{-\pi}^{\pi} \alpha \frac{\alpha\cos(r\cos(\theta)) - \cos(r\cos(\theta))}{\alpha^2+1-2\alpha\cos(2r\cos(\theta))} d\theta \\
&= \int_{-\pi}^{\pi} \frac{\alpha(\alpha-1)\cos(r\cos(\theta))}{(\alpha-1)^2+4\alpha\sin^2(r\cos(\theta))} d\theta
\end{aligned}$$

In the second-to-last step, we used the identity $\cos(a-b) = \cos(a)\cos(b) + \sin(a)\sin(b)$. We will show that the magnitude of λ_{min} is at most $O((\alpha-1)^{-\frac{3}{4}})$, which shoes that $\rho(C_\alpha) = 1 + O((\alpha-1)^{\frac{1}{4}}$. This will show that as α approaches 1, $\frac{\rho(C_\alpha)}{\rho(C_\alpha)-1}$ grows without bound, so that there cannot exist any finite coloring of $\mathcal{O}$.

Note that for $r \leq \frac{\pi}{2}$, $\lambda(r)$ is necessarily positive since the integrand is always positive ($\cos(r\cos(\theta))$ being the only thing that can go negative in the expression). We thus assume that $r > \frac{\pi}{2}$. It suffices to show that

$$\int_0^{\frac{\pi}{2}} \frac{(\alpha-1)\cos(r\cos(\theta))}{(\alpha-1)^2+4\alpha\sin^2(r\cos(\theta))} d\theta \geq -c(\alpha-1)^{-\frac{3}{4}} - d \tag{10}$$

for all r for some constants c, d (as this, neglecting a factor of 4α, is clearly an upper bound for the integral above). Let h be the function we are integrating. Let $\mathcal{R}_k$ denote the region for which $|h(\theta)| \geq 1$ and that contains the value of θ where $\cos(\theta) = \frac{k\pi}{r}$. Then we note that $|\int_{\mathcal{R}_k} h(x)dx| > |\int_{\mathcal{R}_{k-1}} h(x)dx|$ since $\cos(\theta)$ decreases faster as θ increases from 0 to $\frac{\pi}{2}$. Also, the signs of these integrals alternate, so we can either throw out all of them or all but the first one, depending on whether the integral of h across $\mathcal{R}_{\lfloor \frac{r}{\pi} \rfloor}$ is positive or negative. If it is positive, then we have thrown out all of the integral, except for a part where $|h(x) < 1|$, so that the remaining part of the integral is obviously bounded. Thus we will assume that the integral of h across $\mathcal{R}_{\lfloor \frac{r}{\pi} \rfloor}$ is negative. We will bound the area of $\mathcal{R}_{\lfloor \frac{r}{\pi} \rfloor}$. First, we determine when

$$\frac{\alpha - 1}{(\alpha - 1)^2 + 4\alpha \sin^2(r\cos(\theta))} \geq 1 \tag{11}$$

as this is clearly a superset of the area where $h(\theta) \geq 1$. But this happens when $\alpha - 1 \geq (\alpha-1)^2 + 4\alpha\sin^2(r\cos(\theta))$, or $\sin^2(r\cos(\theta)) \leq \frac{(\alpha-1)-(\alpha-1)^2}{4\alpha} = (\alpha-1)\frac{2-\alpha}{4\alpha} < \frac{\alpha-1}{4}$. So the area for which (11) holds is contained in the area for which $\sin(r\cos(\alpha)) \in [-\frac{\sqrt{\alpha-1}}{2}, \frac{\sqrt{\alpha-1}}{2}]$. On the other hand, this is contained in the area in which $r\cos(\theta)$ is within $\sqrt{\frac{\alpha-1}{2}}$ of a multiple of π, as $\sin(\sqrt{\frac{\alpha-1}{2}}) > \sqrt{\frac{\alpha-1}{2}} - \frac{(\alpha-1)^{1.5}}{12\sqrt{2}} > \frac{\sqrt{\alpha-1}}{2}$ for $\alpha - 1$ small enough. So we want to find when

$$-\frac{1}{r}\sqrt{\frac{\alpha-1}{2}} \leq \frac{k\pi}{r} - \cos(\theta) \leq \frac{1}{r}\sqrt{\frac{\alpha-1}{2}} \tag{12}$$

We claim that, if $\cos(\theta_0) = \frac{k\pi}{r}$, then it suffices to take $\theta \in [\theta_0 - \frac{2\sqrt[4]{\alpha-1}}{\sqrt{r}}, \theta_0 + \frac{2\sqrt[4]{\alpha-1}}{\sqrt{r}}]$. First of all, if $\theta_0 - \frac{\sqrt{\alpha-1}}{r} < 0$ or $\theta_0 + \frac{\sqrt{\alpha-1}r}{>}\frac{\pi}{2}$, then θ is outside of our range of integration and so we are definitely covering at least the area we need on that end of the interval. Thus we may assume otherwise, and we have the following lemma:

Lemma 10.12 *If $d > 0$ and $\theta, \theta + d \in [0, \frac{\pi}{2}]$, then $\cos(\theta) - \cos(\theta + d) \geq 1 - \cos(d)$.*

Proof Take $\frac{d}{d\theta}[\cos(\theta) - \cos(\theta + d)] = \sin(\theta + d) - \sin(\theta)$. This is clearly increasing for $\theta \in [0, \frac{\pi}{2} - d]$, so we might as well take $\theta = 0$, as this gives a smaller value for $\cos(\theta) - \cos(\theta + d)$ than any legal value of θ. Then we get $1 - \cos(d)$ as our answer, as claimed.

With Lemma 10.12 in hand, we need only show that $1 - \cos(\frac{2\sqrt[4]{\alpha-1}}{\sqrt{r}}) > \frac{1}{r}\sqrt{\frac{\alpha-1}{2}}$. This is evident once again from the Taylor approximation as, for $\alpha - 1$ small enough, $1 - \cos(\frac{2\sqrt[4]{\alpha-1}}{\sqrt{r}}) > \frac{2\sqrt{\alpha-1}}{r} - \frac{2(\alpha-1)}{3r^2} > \frac{1}{r}\sqrt{\frac{\alpha-1}{2}}$. Thus for any given value of k, the area for which (11) holds is at most $\frac{4\sqrt[4]{\alpha-1}}{\sqrt{r}}$. We only care about $\mathcal{R}_{\lfloor \frac{r}{\pi} \rfloor}$, so in particular we can take $k = \lfloor \frac{r}{\pi} \rfloor$ and the preceding argument holds. On the other hand, $\frac{\alpha-1}{(\alpha-1)^2 + 4\alpha\sin^2(r\cos(\theta))} < \frac{1}{\alpha-1}$, so integrating across this entire region gives us a value whose magnitude is at most $\frac{4}{\sqrt{r}(\alpha-1)^{\frac{3}{4}}}$. Integrating across the rest of the interval $[0, \frac{\pi}{2}]$ gives us a value of magnitude at most $\frac{\pi}{2}$, since we have shown that the integral across all of the remaining $\mathcal{R}_k$, $k < \lfloor \frac{r}{\pi} \rfloor$, must yield a positive number, and for all other portions of the interval $|h(\theta)| < 1$ by design. Also, recall that we established that $r > \frac{\pi}{2}$, so in particular $r > 1$. Thus we have that

$$\int_0^{\frac{\pi}{2}} \frac{\alpha - 1}{(\alpha-1)^2 + 4\alpha \sin^2(r\cos(\theta))} d\theta \geq -4(\alpha-1)^{-\frac{3}{4}} - \frac{\pi}{2} \tag{13}$$

as desired. This establishes that the measurable chromatic number of the odd-distance graph is infinite.

11 Linear Algebra Exercises

1. Do all of the exercises posed earlier in the linear algebra section.

2. Prove any of the other statements in this packet that you feel need more detail, and make sure you can answer all the questions posed in the text.

3. Prove Lemma 10.7.

4. Prove Lemma 10.8.

5. Show that Proposition 10.6 actually does show that $P_m \square P_n$ cannot be nilpotent unless $m = n = 2$ or $m = 2^i - 1, n = 2^j - 1$.

6. Prove that $P_2 \square P_2$ is nilpotent.

7. Prove that $P_{2^i-1} \square P_{2^j-1}$ is nilpotent. (Hint: given a function on this graph that is not eventually sent to zero by A, you can "fold the graph in half" to produce a function on $P_{2^i-1} \square P_{2^{j-1}-1}$ that is also not sent to zero. So, we can use induction (the base case being that $P_1 \square P_1$ is nilpotent) to show that $P_{2^i-1} \square P_{2^j-1}$ is nilpotent for all i, j.)

   The following set of exercises explores extensions of material covered above, as well as a few results from spectral graph theory over $\mathbb{C}$ (which is where people usually work).

8. Generalize the grid problem to $\mathbb{F}_p$ (that is, $\mathbb{Z}/p\mathbb{Z}$, where p is prime) instead of just $\mathbb{F}_2$.

9. Generalize the grid problem to $\mathbb{Z}/n\mathbb{Z}$ for (not necessarily prime) n.

10. Generalize the grid problem to grids in arbitrarily many dimensions.

11. Let X be a graph, and let A be its adjacency matrix over $\mathbb{C}$. Since A is symmetric, all of its eigenvalues are real. Let λ be the largest eigenvalue of A. Show that the eigenspace of λ has dimension 1 iff X is connected, and that $-\lambda$ is an eigenvalue of A iff X is bipartite. (Note: showing that the space must have dimension 1 when X is connected is quite hard. Come see me if you want hints.)

12. Let X, Y be graphs. Define their *weak product* $X \times Y$ such that $V(X \times Y) = V(X) \times V(Y)$ and $(u, v) \sim (u', v')$ iff $u \sim u'$ and $v \sim v'$. If A is the adjacency matrix of X, and B is the adjacency matrix of Y, then show that the adjacency matrix of $X \times Y$ is $A \otimes B$, and that $A \otimes B$ has precisely the eigenvalues $\lambda \cdot \mu$, where λ ranges over the eigenvalues of A and μ ranges over the eigenvalues of B. (Hint: for the last part, you will need the spectral theorem.)

13. Use the last two exercises to show that $A \times B$ is connected iff not both A and B are bipartite.

14. Show that if X is a bipartite graph on an odd number of vertices, then its adjacency matrix is not invertible (over $\mathbb{C}$).