DISCRETE MATHEMATICS I

Soumyashant Nayak, notes by Ramdas Singh

Third Semester

List of Symbols

Placeholder

Contents

Chapter 1

DISCRETE STRUCTURES

1.1 A Brief Introduction

July 22nd.

Discrete mathematics is primarily the study of tools for reasoning precisely the systematically about digital systems, logical problems, and combinatorial structures such as the integers, graphs, logical statements, and finite automata. Furthermore, combinatorics is the mathematics of counting and configuration; the counting, organizing, and analyzing discrete structures.

Bloch's principle, or Bloch's heuristics, states that every proposition on whose statement the actual infinity occurs can always be considered as a consequence of a proposition where it does not occur as a proposition on finite terms. The *Ramsey principle* states that complete disorder is impossible. In any sufficiently large structure, order or regularity must emerge. These two principles may be considered complimentary to each other.

1.2 Useful Methods

The method of *double counting* can be thought of a creative device or trick. Before strictly showing the statement, we utilise some examples.

Example 1.1. Suppose we wish to show that $\sum_{k=0}^{n} \binom{n}{k} = 2^n$. We first ask how many ways can a subset be chosen from $\{1, 2, \ldots, n\}$. The first method is to build a subset by deciding whether we want i to be a part of our subset for $i \in \{1, 2, ..., n\}$. The second method is find the number of subsets of caridnality i for $i \in \{1, 2, \ldots, n\}$ and add up all the results. This leads us to conclude $2^n = \sum_{k=0}^{n} \binom{n}{k}$ after equating the answers from both methods.

Theorem 1.2 (The *q-binomial theorem*)**.** *We use the following notation:*

$$\binom{n}{k}_q = \frac{(q^n - 1) \cdots (q^{n-k+1} - 1)}{(q^k - 1) \cdots (q - 1)}. \tag{1.1}$$

Simply stated, the q-binomial theorem is

$$\sum_{k=0}^{n} q^{\binom{k}{2}} \binom{n}{k}_q z^k = \prod_{i=0}^{n-1} (1 + q^i z). \tag{1.2}$$

The proof of the above theorem is performed by double counting; counting the number of pairs (U, B) where U is a k-dimensional subspace of $\mathbb{F}_q^n$ and B is the flag of nested subspaces of U.

Reccurrence Relations and Generating Functions

Perhaps, the most important example of a recurrence relation is the Fibonacci sequence, where the terms in the sequence are defined as $F_0 = 0$, $F_1 = 1$, $F_{n+2} = F_{n+1} + F_n$ for all $n \geq 0$. Let us find the *generating function* of this sequence; we start by creating

$$F(t) = F_0 + F_1 t + F_2 t^2 + F_3 t^3 + \cdots, \tag{1.3}$$

the generating function of $(F_n)_{n=0}^{\infty}$. We can then work as follows—

$$\begin{aligned} tF(t) &= F_0 t + F_1 t^2 + \cdots, \\ t^2 F(t) &= F_0 t^2 + \cdots, \\ \implies (1 - t - t^2)F(t) = t \implies F(t) &= \frac{-t}{t^2 + t - 1}. \end{aligned} \tag{1.4}$$

If we look at $F_{n+1} = 1 \cdot F_n + 1 \cdot F_{n-1}$ and $F_n = 1 \cdot F_n + 0 \cdot F_{n-1}$, we may notice a matrix as $\begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_n \\ F_{n-1} \end{bmatrix}$. Back substituting multiple times leads us to conclude $\begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^n \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. We can then diagonalize the centre matrix to decompose it as $P \begin{bmatrix} (1+\sqrt{5})^n/2^n & 0 \\ 0 & (1-\sqrt{5})^n/2^n \end{bmatrix} P^{-1}$; thus, the terms of the sequence are really linear combinations of the diagonal elements that appear.

Principle of Inclusion-Exclusion

July 24th.

Simply stated, for sets A and B, $\#(A \cup B) = \#A + \#B + \#(A \cap B)$. For three sets A, B and C, we have $\#(A \cup B \cup C) = \#A + \#B + \#C - \#(A \cap B) - \#(B \cap C) - \#(A \cap C) + \#(A \cap B \cap C)$. This can be extended to any finite number of finite sets.

Theorem 1.3 (The *principle of inclusion-exclusion*)**.** *Let S be an N-set ($\#S = N$), and let $E_1, E_2, \ldots, E_r$ be, not necessarily distinct, subsets of S. For any subset M of the indexing set $\{1, 2, \ldots, r\}$, let $N(M)$ denote the number of elements of S in $\bigcap_{i \in M} E_i$, and for $0 \leq j \leq r$, define*

$$N_j = \sum_{\#M = j} N(M). \tag{1.5}$$

Then the number of elements of S not in any of the E_i's is

$$\#(S \setminus \bigcup_{i=1}^{r} E_i) = N - N_1 + N_2 - N_3 + \cdots + (-1)^r N_r. \tag{1.6}$$

Proof. For $x \in S$, define $M : S \to \{0, 1\}$ as $M(x) = 1$ if $x \in \bigcap_{i \in M} E_i$ and 0 otherwise. Thus,

$$\sum_{x \in S} M(x) = \#(\bigcap_{i \in M} E_i) = N(M) \implies N_j = \sum_{\#M = j} \sum_{x \in S} M(x) = \sum_{x \in S} \sum_{\#M = j} M(x). \tag{1.7}$$

The alternating sum then becomes

$$\sum_{x \in S} 1 - \sum_{x \in S} \sum_{\#M = 1} M(x) + \cdots + (-1)^r \sum_{x \in S} \sum_{\#M = r} M(x) = \sum_{x \in S} \left(1 - \sum_{\#M = 1} M(x) + \cdots + (-1)^r \sum_{\#M = r} M(x) \right). \tag{1.8}$$

Call the term within the parentheses as $F(x)$. We deal with cases; if $x \notin \bigcup_{i=1}^{r} E_i$, then $F(x) = 1$. If x is in exactly $k \geq 1$ of the sets $E_1, E_2, \ldots, E_r$, then

$$F(x) = 1 - \binom{k}{1} + \binom{k}{2} - \binom{k}{3} + \cdots + (-1)^k \binom{k}{k} = (1 - 1)^k = 0. \tag{1.9}$$

This is independent of k; we conclude that the alternating sum reduces to the number of elements in S not in any of the E_i's. ∎

Corollary 1.4. *Retaining notation from the previous theorem, if* $S = \bigcup_{i=1}^{r} E_r$*, then*

$$N = N_1 - N_2 + \cdots + (-1)^{r-1} N_r. \tag{1.10}$$

We look at some examples of the principle in use.

Example 1.5. Let d_n be the number of permutations π of the set $\{1, 2, \ldots, n\}$ such that $\pi(i) \neq i$ for all $1 \leq i \leq n$. Such a permutation is called a *derangement*, where no point is fixed. We wish to count all such permutations. Let the set of all permutations of the set be S. Let E_i denote the set of all permutations that fix i, for $1 \leq i \leq n$. Thus, S without $\bigcup_{i=1}^{n} E_i$ would then denote the set of all derangements. Making use of the principle, we have

$$\#(S \setminus \bigcup_{i=1}^{n} E_i) = n! - \binom{n}{1}(n-1)! + \binom{n}{2}(n-2)! - \cdots = n!\left(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!}\right) \tag{1.11}$$

which is approximately $\frac{n!}{e}$ for larger n. Thus, the probability of choosing a derangemnet is e^{-1}.

Example 1.6. Suppose we have two sets X and Y with $\#X = n$ and $\#Y = k$. We ask how many surjective maps exists from X to Y. The set S, this time, is the set of all functions from X to Y, being Y^X. E_i denotes the set of functions from X to Y such that y_i is not in the image of X. The elments within S not in any of the E_i's are surjective maps. Clearly, $N_i = \binom{k}{i}(k-i)^n$, and the cardinality of $S \setminus \bigcup_{i=1}^{k} E_i$ is then

$$\sum_{i=0}^{k} (-1)^i \binom{k}{i} (k-i)^n. \tag{1.12}$$

Example 1.7. We wish to show that the expression $\sum_{i=0}^{n} (-1)^i \binom{n}{i}\binom{m+n-i}{k-i}$ evaluates to $\binom{m}{k}$ if $m \geq k$, and 0 otherwise. To this end, fix $X = \{x_1, \ldots, x_n\}$ and $Y = \{y_1, \ldots, y_m\}$ and set $Z = X \cup Y$. We now ask how many k-subsets of Z consist of only points form Y. Let S be the set of all k-subsets of Z, and denote E_i to be the set of k-subsets of Z containing x_i for $1 \leq i \leq n$. The left hand side of our inclusion-exclusion principle evaluates to $\binom{m+n}{k}$. Each N_i evaluates to $\binom{n}{i}\binom{m+n-i}{k-i}$, proving our expression above.

The next example relates to the Euler totient function.

Example 1.8. Recall that, from the *fundamental theorem of arithmetic*, each natural number may be expressed uniquely (upto order) as the product of distinct primes raised to values, that is, $n = p_1^{a_1} \cdots p_r^{a_r}$. The *Euler totient function* $\phi : \mathbb{N} \to \mathbb{C}$ acts on the naturals and returns $\phi(n)$, the number of positive integers $k \leq n$ such that $\gcd(k, n) = 1$. Certainly, $\phi(p) = p - 1$ for a prime p. Our task is to find a closed form formula for $\phi(n)$.
Set $S = \{1, 2, \ldots, n\}$, and set E_i to be the set of integers in S divisible by p_i for $1 \leq i \leq r$. Clearly, the value $\#(S \setminus \bigcup_{i=1}^{r} E_i)$ returns the set of all numbers in $\{1, 2, \ldots, n\}$ coprime to n. Note that $N = n$. The value of N_1 is $\sum_{i=1}^{r} \frac{n}{p_i}$, the value of N_2 is is $\sum_{1 \leq i \leq j \leq r} \frac{n}{p_i p_j}$. The closesd form formula then becomes

$$\phi(n) = \#(S \setminus \bigcup_{i=1}^{r} E_i) = n - n\sum_{i=1}^{r} \frac{1}{p_i} + n \sum_{1 \leq i \leq j \leq r} \frac{1}{p_i p_j} + \cdots = n\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right)\cdots\left(1 - \frac{1}{p_r}\right). \tag{1.13}$$

Number Theory

We continue with the Euler totient function.

Theorem 1.9. $\sum_{d|n} \phi(d) = n.$

Proof. For each integer $m \in \{1, 2, \ldots, n\}$, the value $\gcd(m, n)$ is a divides n. Fix a divisor d of n. The number of integers m such that $\gcd(m, n) = d$ is equal to the number of integers m such that $\gcd\left(\frac{m}{d}, \frac{n}{d}\right) = 1$, where $\frac{m}{d}$ runs over integers between 1 and $\frac{n}{d}$. Therefore, the number of such m is $\phi\left(\frac{n}{d}\right)$. Summing over all divisors d of n, we get:

$$n = \sum_{d|n} \phi\left(\frac{n}{d}\right)$$

which is the same as $\sum_{d|n} \phi(d)$. ■

The *Möbius function* is defined as

$$\mu(d) := \begin{cases} 1, & \text{if } d \text{ is a product of even number of distinct primes,} \\ -1, & \text{if } d \text{ is a product of odd number of distinct primes,} \\ 0, & \text{if otherwise; the number } d \text{ is not square-free.} \end{cases} \tag{1.14}$$

Theorem 1.10.

$$\sum_{d|n} \mu(d) = \begin{cases} 1, & \textit{if } n = 1, \\ 0, & \textit{if otherwise.} \end{cases} \tag{1.15}$$

Proof. For $n = 1$, it is clear. For $n > 1$, rewriting n as $p_1^{a_1} \cdots p_r^{a_r}$ helps us see that

$$\sum_{d|n} \mu(d) = 1 - \binom{r}{1} + \binom{r}{2} - \binom{r}{3} + \cdots = (1-1)^r = 0. \tag{1.16}$$

■

This property of the Möbius function proves to be useful.

July 29th.

Theorem 1.11 (The *Möbius inversion formula*)**.** *Suppose we have two function $f : \mathbb{N} \to \mathbb{R}$ and $g : \mathbb{N} \to \mathbb{R}$ which relate as*

$$f(n) = \sum_{d|n} g(d). \tag{1.17}$$

Then the function g satisfies

$$g(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) f(d). \tag{1.18}$$

Proof. We work as

$$\begin{aligned} \sum_{d|n} \mu\left(\frac{n}{d}\right) f(d) &= \sum_{d|n} \mu\left(\frac{n}{d}\right) \left(\sum_{d'|d} g(d')\right) = \sum_{d|n,\ d'|d} \mu\left(\frac{n}{d}\right) g(d') = \sum_{d'|n,\ m|\frac{n}{d'}} g(d')\mu(m) \\ &= \sum_{d'|n} \left(g(d') \left(\sum_{m|\frac{n}{d'}} \mu(m)\right)\right) = g(n). \end{aligned} \tag{1.19}$$

■

Example 1.12. Let N_n denote the number of distinct circular binary sequences of length n, up to rotation. That is, two sequences are considered the same if one is a rotation of the other. We aim to compute N_n explicitly.
Let $M(d)$ denote the number of aperiodic circular binary sequences of length d, meaning sequences that are not periodic with any smaller period. Note that each such aperiodic sequence of length d contributes to sequences of length n whenever $d \mid n$. Indeed, every binary circular sequence of length n can be viewed as made up of $\frac{n}{d}$ repetitions of a primitive block of length d.
Thus, we have:

$$N_n = \sum_{d|n} M(d).$$

Now consider the total number of binary strings of length n, which is 2^n. Each such string can be arranged in a circle in n different ways, one for each rotation. However, many of these circular sequences are identical under rotation, so we overcounted by a factor of the size of the symmetry group.
Let $f(n) = 2^n$ be the total number of binary strings of length n. Each such string is generated by repeating an aperiodic sequence of length d exactly $\frac{n}{d}$ times, for some $d \mid n$. Since each aperiodic circular sequence of length d has d rotations, we get:

$$f(n) = 2^n = \sum_{d|n} d \cdot M(d).$$

Applying Möbius inversion to this relation, we obtain:

$$n \cdot M(n) = \sum_{d|n} \mu(d) \cdot 2^{n/d},$$

and hence,

$$M(n) = \frac{1}{n} \sum_{d|n} \mu(d) \cdot 2^{n/d}.$$

Substituting back into the formula for N_n, we get:

$$N_n = \sum_{d|n} M(d) = \sum_{d|n} \frac{1}{d} \sum_{k|d} \mu(k) \cdot 2^{d/k}.$$

Interchanging the order of summation, we arrive at the classical formula:

$$N_n = \frac{1}{n} \sum_{d|n} \phi(d) \cdot 2^{n/d},$$

where ϕ is Euler's totient function.

Lemma 1.13 (*Burnside's lemma*)**.** *Let G be a permutation group acting on some finite set X. Let $\psi(g)$ denote the number of points of X fixed by $g \in G$. Then the number of orbits of G is $\frac{1}{|G|} \sum_{g \in G} \psi(g)$.*

In the above, by the set of points fixed by $g \in G$, we mean the set $\{x \mid g \cdot x = x\}$. By the *orbit* of x, we mean the set $\{g \cdot x \mid g \in G\}$.

Such inversion formulae are common in discrete math. The following are some examples.

Example 1.14.

- For an integer n, $f(n) = \sum_{i=1}^{n} g(i)$ if and only if $g(n) = f(n) - f(n-1)$. This is known as a *telescoping sum*.
- For an integer n, $f(n) = \sum_{d|n} g(d)$ if and only if $g(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) f(d)$. This is the Möbius inversion formula.

- For a set S, $f(S) = \sum_{T \subseteq S} g(T)$ if and only if $g(S) = \sum_{T \subseteq S} (-1)^{|S|-|T|} f(T)$.

Partially Ordered Sets

Definition 1.15. A *poset* S, or a *paritally ordered set*, is a (countable or finite) set of objects with a binary relation $\leq$ satisfying the following properties.

1. Reflexivity: $x \leq x$ for all $x \in S$.
2. Antisymmetry: if $x \leq y$ and $y \leq x$, for some $x, y \in S$, then $x = y$.
3. Transitivity: if $x \leq y$ and $y \leq z$, for some $x, y, z \in S$, then $x \leq z$.

Some examples are as follows.

Example 1.16.
1. $(\{1, 2, \ldots, n\}, \leq)$ is a poset, with $a \leq b$ if $b - a$ is a non-negative integer.
2. $(\{1, 2, \ldots, n\}, \leq_1)$ is also a poset, where $a \leq_1 b$ if $a \mid b$.
3. $(\mathcal{P}(\{1, 2, \ldots, n\}), \leq_2)$ is also a poset; here, $S \leq_2 T$ if $S \subseteq T$. The power set is also denoted as $2^{\{1,2,\ldots,n\}}$.
4. The set of partitions of $\{1, 2, \ldots, n\}$, equippied with the partial ordering of refinement. By a partition, we mean $\{S_1, S_2, \ldots, S_r\}$ such that $S_i \cap S_j = \emptyset$ for $i \neq j$, and $\bigcup_{i=1}^{r} S_i = \{1, 2, \ldots, n\}$. Similarly, let $\{T_1, T_2, \ldots, T_d\}$ be another partition. Then $\{S_1, \ldots, S_r\} \leq \{T_1, \ldots, T_d\}$ if for $1 \leq i \leq r$, $S_i \subseteq T_k$ for some $1 \leq k \leq d$. Here, we term $\{S_1, \ldots, S_r\}$ a *refinement* of $\{T_1, \ldots, T_d\}$.

The idea of the Möbius function really comes from posets, where its definition is more generalized. Here, $\mu(d, n)$ can be thought of as a place-in for $\mu\left(\frac{n}{d}\right)$.

Definition 1.17. The *Möbius function of a poset* is defined as

$$\mu(x, y) = \begin{cases} 0 & \text{if } x \not\leq y, \\ 1 & \text{if } x = y, \\ -\sum_{x \leq z < y} \mu(x, z) & \text{if } x \lneq y. \end{cases} \tag{1.20}$$

Note that we need only compute $\mu(x, y)$ on all intervals $[x, y]$ $(x \leq y)$. We call an element y a *successor* of x if there exists no z satisfying $x \lneq z \lneq y$. Note that the successor may not be unique. Let us denote any successor by $\text{succ}(x)$.

If g and f are two functions defined and related as

$$g(x) = \sum_{y \leq x} f(y). \tag{1.21}$$

We now introduce the *zeta function* ζ defined as $\zeta(x, y) = 1$ if $x \leq y$ and 0 otherwise. Thus, the above equation can be rewritten as

$$g(x) = \sum_{y \leq x} f(y) = \sum_{y \in S} \zeta(y, x) f(y). \tag{1.22}$$

August 5th.

Lemma 1.18. *A finite partial order can always be embedded in a total ordering; that is, there exists an indexing $S = \{x_1, \ldots, x_n\}$ such that $x_i \leq x_j$ in S implies $i \leq j$.*

As a proof outline, pick a maximal element x of S. Label it x_n. Repeat the process with $S \setminus \{x\}$, then proceed inductively. The embedding is then clear.

Thus, using the lemma, we can rewrite the relation between g and f in matrix form as

$$\begin{pmatrix} g(x_1) & \cdots & g(x_n) \end{pmatrix} = \begin{pmatrix} f(x_1) & \cdots & f(x_n) \end{pmatrix} \begin{pmatrix} \zeta(x_1,x_1) & \cdots & \zeta(x_1,x_n) \\ \vdots & \ddots & \vdots \\ \zeta(x_n,x_1) & \cdots & \zeta(x_n,x_n) \end{pmatrix}. \tag{1.23}$$

Since $\zeta(x_i, x_j) = 0$ when $i > j$, the matrix on the right is upper triangular. Also, all the diagonal entries are 1. Denote the above matrix on the right by Z. Note that $Z = I + N$ where I is the identity matrix and N is upper triangular with 0's on the diagonal. Z^{-1} can be computed by taking a power series, and noting that N^n is 0.

$$Z^{-1} = (I+N)^{-1} = I - N + N^2 - N^3 + \cdots + (-1)^{n-1}N^{n-1} \tag{1.24}$$

Let $M = [\mu(x_i, x_j)]$. We find the (x_i, x_j) entry of MZ as

$$\sum_{y\in P} M_{x_i,y} Z_{y,x_j} = \sum_{y\in P} \mu(x_i,y)\zeta(y,x_j) = \sum_{y\le x_j} \mu(x_i,y) = \sum_{x_i\le y\le x_j} \mu(x_i,y). \tag{1.25}$$

Noting that $\mu(x_i, x_j) = -\sum_{x_i \le z < x_j} \mu(x_i, z)$, we get the above expression to be 1 if $x_i = x_j$ and 0 otherwise. Thus, $MZ = I$. This is the used definition of the the Möbius functon. $ZM = I$ may be verified similarly.

Theorem 1.19. *Let $(P, \le)$ be a finite poset, with $f, g : P \to \mathbb{Z}$ functions. Then,*

1. *$f(x) = \sum_{y\le x} g(y)$ if and only if $g(x) = \sum_{y\le x} \mu(y,x) f(y)$, and*
2. *$f(x) = \sum_{x\le y} g(y)$ if and only if $g(x) = \sum_{x\le y} \mu(x,y) f(y)$.*

This is the Möbius inversion formula for a poset.

Proof. We have

$$\sum_{y\le x} \mu(y,x) f(y) = \sum_{y\le x} \left(\sum_{z\le y} \mu(y,x) g(z) \right) = \sum_{z\le x} \sum_{z\le y\le x} \mu(y,x) g(z) = \sum_{z\le x} g(z) \sum_{z\le y\le x} \mu(y,x) = g(x) \tag{1.26}$$

since $\mu(y, x)$ at the end will be zero if $z \neq x$. To show the converse, we have

$$\sum_{y\le x} g(y) = \sum_{y\le x} \sum_{z\le y} \mu(z,y) f(z) = \sum_{z\le x} f(z) \left(\sum_{z\le y\le x} \mu(z,y) \right) = f(x) \tag{1.27}$$

since $\mu(z, y)$ is zero if $z \neq x$. ■

Example 1.20. Verify that if the poset is the positive integers with the standard ordering $\le$, then

$$\mu(i,j) = \begin{cases} 1 & \text{if } i = j, \\ -1 & \text{if } i = j-1, \\ 0 & \text{if otherwise.} \end{cases} \tag{1.28}$$

Example 1.21. Let our poset be $(2^S, \subseteq)$ for a set S. For fixed subsets $U, T \in 2^S$, we have

$$\sum_{U\subseteq R\subseteq T} (-1)^{\#T - \#S} = \begin{cases} 1 & \text{if } U = T, \\ 0 & \text{if otherwise.} \end{cases} \tag{1.29}$$

To show this, without the loss of generality, we will assume that $U = \emptyset$. Denoting $\#T = n$, we have

$$\sum_{R\subseteq T} (-1)^{\#R} = \sum_{k=0}^{n} \binom{n}{k} (-1)^k = 0. \tag{1.30}$$

Here, $Z(R,T) = 1$ if $R \subseteq T$ and 0 otherwise. Also, $M(R,T) = (-1)^{\#T-\#R}$ if $R \subseteq T$ and 0 otherwise. Since $MZ = I$, $M(R,T)$ must be the Möbius function for $(2^S, \leq)$.

Chapter 2

RECURRENCE RELATIONS AND GENERATING FUNCTIONS

2.1 Generating Functions

August 7th.

We begin with ordinary ones.

Definition 2.1. For a sequence $(a_n)_{n\geq 0} \subseteq \mathbb{R}$, the *ordinary generating function* associated with this sequence is defined as

$$f(x) = \sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + \cdots . \tag{2.1}$$

Note that we are not concerned with convergence right now. We deconstruct our abstraction of ideas into levels, starting with the first level as regarding ordinary generating functions as algebraic objects. One can multiply and add them to create new generating functions. The second level is regarding them as analytic objects, only if the radius of convergence is positive.

The above is known as *Z-transform*, where a sequence is mapped onto a function. When using the word transform, we generally mean a 'change of basis'; in this case, we are changing from a sequence space to a function space.

Definition 2.2. For a sequence $(a_n)_{n\geq 0} \subseteq \mathbb{R}$, the *exponential generating function* associated with this sequence is defined as

$$f(x) = \sum_{n=0}^{\infty} \frac{a_n}{n!} x^n = a_0 + a_1 x + \frac{a_2}{2!} x^2 + \cdots . \tag{2.2}$$

Again, we have transformed from a sequence space to a function space. One can also transform from a random variable space to a function space.

Definition 2.3. For a random variable X taking values in $\mathbb{R}$, the *moment generating function* associated with this random variable is defined as

$$M_X(t) = E[e^{tX}] = \sum_{n=0}^{\infty} \frac{E[X^n]}{n!} t^n = 1 + E[X]t + \frac{E[X^2]}{2!} t^2 + \cdots . \tag{2.3}$$

2.1.1 Algberaic Operaions

We give a kind of correspondence between algebraic operations and combinatorial interpretations.

1. Multiplying by x^k maps $a_0 + a_1x + a_2x^2 + \cdots$ to $a_0x^k + a_1x^{k+1} + a_2x^{k+2} + \cdots$. This corresponds to shifting the sequence $(a_0, a_1, \ldots)$ right by k places. This is known as the *shift operator*.
2. Multiplication is also defined; for two functions $a_0 + a_1x + a_2x^2 + \cdots$ and $b_0 + b_1x + b_2x^2 + \cdots$, their product is given by $a_0 + (a_1b_0 + a_0b_1)x + (a_2b_0 + a_1b_1 + a_0b_2)x^2 + \cdots$. This corresponds to combining objects of size k and size $n - k$ chosen independently.
3. Differentiation maps $a_0 + a_1x + a_2x^2 + \cdots$ to $a_1 + 2a_2x + 3a_3x^2 + \cdots$. This corresponds to weighing the sequence values by their index, with a shift of one place to the right.

Example 2.4. Suppose we have k boxes labelled 1 through k, and box i contains r_i balls for $1 \le i \le k$. We wish to encode all possible configurations in a kind of book-keeping device. For a particular $(r_1, \ldots, r_k)$, we have

$$\sum_{r_i \ge 0} x_1^{r_1} \cdots x_k^{r_k} = (1 + x_1 + x_1^2 + \cdots)(1 + x^2 + x_2^2 + \cdots) \cdots (1 + x_k + x_k^2 + \cdots). \tag{2.4}$$

We find the number of partitions of n (balls) into k numbers (boxes), where each number if non-negative. Disregarding the order, we set all the x_i's equal to each other. Thus, we wish to find the coefficient of x^n where $r_1 + \cdots + r_k = n$. From the sum above, we have

$$(1 + x + x^2 + \cdots)^k = (1 - x)^{-k} = \sum_{j=0}^{\infty} \binom{k - 1 + j}{j} x^j x^j. \tag{2.5}$$

Therefore, the required coefficient is $\binom{k-1+n}{n}$.

We briefly introduce the idea of rings. A *ring* $(R, +, *)$ is a set R with two operations $+$ and $*$ such that $(R, +)$ is an abelian group, $(R, *)$ is a monoid, and the distributive law holds. Some examples of rings include $\mathbb{Z}$, $M_n(\mathbb{C})$, and $\mathbb{C}[x]$. Another example is the ring of formal power series $\mathbb{C}[[x]]$, which consists of all series of the form $a_0 + a_1x + a_2x^2 + \cdots$ where $a_i \in \mathbb{C}$. We ask which elements of this ring are invertible.

We claim that $a_0 + a_1x + a_2x^2 + \cdots$ is invertible if and only if $a_0 \neq 0$. We find $b_0 + b_1x + b_2x^2 + \cdots$ such that

$$(a_0 + a_1x + a_2x^2 + \cdots)(b_0 + b_1x + b_2x^2 + \cdots) = 1. \tag{2.6}$$

This first gives us $a_0b_0 = 1$, so $b_0 = \frac{1}{a_0}$. The next term gives us $a_0b_1 + a_1b_0 = 0$, so $b_1 = -\frac{a_1}{a_0^2}$. Continuing this process, we find that the coefficients of b can be expressed in terms of the coefficients of a as

$$b_n = -\frac{1}{a_0} \sum_{k=1}^{n} a_k b_{n-k}. \tag{2.7}$$

There is also the ring homomorphism $\mathrm{ev}_z : \mathbb{C}[[x]] \to \mathbb{C}$ where $x \mapsto z$, with $z \in \mathbb{C}$.

Example 2.5. Let d_n denote the number of derangements of $\{1, 2, \ldots, n\}$. We consider a derangement Π of $\{1, 2, \ldots, n + 1\}$ where

- Case I: $\Pi(n + 1) = i$ and $\Pi(i) = n + 1$ for some i. The number of such derangements is nd_{n-1}.
- Case II: $\Pi(n + 1) = i$ and $\Pi(j) = n + 1$ for some $i \neq j$. The number of such derangements is $d_{n+1} = nd_n$.

Thus, the total number of derangements is $d_{n+1} = n(d_n + d_{n-1})$. Here, $d_0 = 1$, $d_1 = 0$, and $d_2 = 1$. The exponential generating function, here, is

$$D(x) = \sum_{n=0}^{\infty} d_n \frac{x^n}{n!} \implies D'(x) = \sum_{n=1}^{\infty} nd_n \frac{x^{n-1}}{(n-1)!} = \sum_{n=0}^{\infty} d_{n+1} \frac{x^n}{n!} = \sum_{n=1}^{\infty} \frac{d_n}{(n-1)!} x^n + \sum_{n=1}^{\infty} \frac{d_{n-1}}{(n-1)!} x^n. \tag{2.8}$$

This gives us

$$D'(x) = xD'(x) + xD(x) \implies \frac{D'(x)}{D(x)} = \frac{x}{1-x} \implies D(x) = C\left(\frac{e^{-x}}{1-x}\right). \tag{2.9}$$

Plugging in $x = 0$ givens $C = 1$. Thus, the exponential generating function is

$$D(x) = \frac{e^{-x}}{1-x} = (1 + x + x^2 + x^3 + \cdots)(1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \cdots). \tag{2.10}$$

The coefficient of x^n in the above series is $\frac{d_n}{n!}$, giving us

$$d_n = n! \sum_{k=0}^{n} a_k b_{n-k} = n!\left(1 - \frac{1}{1!} + \frac{1}{2!} + \cdots + (-1)^n \frac{1}{n!}\right). \tag{2.11}$$

Example 2.6. Suppose we wish to find number of ways to make n change with the denominations 1, 2, and 5. We use generating functions. Thus,

$$(1 + x + x^2 + \cdots)(1 + x^2 + x^4 + \cdots)(1 + x^5 + x^{10} + \cdots) = \frac{1}{(1-x)(1-x^2)(1-x^5)}. \tag{2.12}$$

From above, taking the n^{th} derivative of the fraction,dividng it by $n!$, and evaluating at $x = 0$ provides the number of ways to make change for n.

2.1.2 Extended Binomial Theorem

August 8th.

We begin by extending the definition of a binomial coefficient.

Definition 2.7. For any $u \in \mathbb{R}$ and positive integer k, we define the *extended binomial coefficient* as

$$\binom{u}{k} = \frac{u(u-1)(u-2)\cdots(u-k+1)}{k!} \tag{2.13}$$

with $\binom{u}{0} = 1$.

Theorem 2.8. *For positive integers n and r, we have*

$$\binom{-n}{r} = (-1)^r \binom{n+r-1}{r}. \tag{2.14}$$

Proof. We simply have

$$\binom{-n}{r} = (-1)^r \frac{n(n+1)\cdots(n+r-1)}{r!} = (-1)^r \binom{n+r-1}{r}. \tag{2.15}$$

■

Theorem 2.9 (The *extended binomial theorem*)**.** *For any $u \in \mathbb{R}$ and positive integer k, we have*

$$(1+x)^u = \sum_{k=0}^{\infty} \binom{u}{k} x^k. \tag{2.16}$$

.

The extended binomial theorem helps to compute coefficients of generating functions such like $\frac{1}{(1+x)^5}$ or even $\sqrt{1+x}$.

Example 2.10. We compute the coefficient of x^{2026} in the generating function $G(x) = \frac{1}{(1-x)^2(1+x)^2}$. We break down as partial fractions to get

$$G(x) = \frac{1}{(1-x)^2(1+x)^2} = \frac{1}{4}\left(\frac{1}{1-x} + \frac{1}{(1-x)^2} + \frac{1}{1+x} + \frac{1}{(1+x)^2}\right) \tag{2.17}$$

$$= \frac{1}{4}\sum_{k=0}^{\infty}\left((-1)^k\binom{-1}{k} + (-1)^k\binom{-2}{k} + \binom{-1}{k} + \binom{-2}{k}\right)x^k. \tag{2.18}$$

The odd terms vanish, so we set k to be even to get the coefficient of x^k as

$$\frac{1}{4}\left(2\binom{-1}{k} + 2\binom{-2}{k}\right) = 1 + \frac{k}{2}. \tag{2.19}$$

Setting $k = 2026$ gives the desired solution of 1014.

2.1.3 Bernoulli Numbers

Definition 2.11. In power series of $\frac{t}{e^t-1} = \sum_{k=0}^{\infty} B_k \frac{t^k}{k!}$, the coefficients B_k are known as the *Bernoulli numbers*.

Here, B_0 is defined to be 1. One can also recursively define them as $B_0 = 1$ and B_j for $j \geq 1$ such that $\sum_{k=0}^{n}\binom{n+1}{k}B_k = 0$ for $n \geq 1$

One also has a useful formula for the Bernoulli numbers.

Theorem 2.12 (*Faulhaber's formula*)**.** *We have*

$$\sum_{m=1}^{n} m^k = \frac{1}{k+1}\sum_{j=0}^{k}\binom{k+1}{j}B_j n^{k+1-j}. \tag{2.20}$$

Setting $k = 1$, we get

$$1 + 2 + \cdots + n = \frac{1}{2}\left(\binom{2}{0}B_0(n+1)^2 + \binom{2}{1}B_1(n+1)\right) = \frac{1}{2}\left((n+1)^2 - (n+1)\right) = \frac{1}{2}(n+1)n. \tag{2.21}$$

Similarly, one may verify for $k = 2$ or $k = 3$.

Proof. To this end, we use the *Bernoulli polynomials* B_k which are interpreted as coefficients in

$$\frac{te^{xt}}{e^t - 1} = \sum_{k=0}^{\infty} B_k(x)\frac{t^k}{k!} \tag{2.22}$$

with $B_k(0) = B_k$. We claim that $B_n(x) = \sum_{k=0}^{n}\binom{n}{k}B_k x^{n-k}$. We have

$$\frac{t}{e^t-1}e^{xt} = \left(\sum_{k=0}^{\infty} B_k\frac{t^k}{k!}\right)e^{xt} = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{\infty}\binom{n}{k}B_k x^{n-k}\right)\frac{t^n}{n!}. \tag{2.23}$$

Summing gives us $\dfrac{t(e^{nt}-1)}{(e^t-1)^2}$

$$= \sum_{m=0}^{n-1}\frac{te^{mt}}{e^t-1} = \frac{t}{e^t-1}\sum_{m=0}^{n-1}(1 + mt + m^2\frac{t^2}{2!} + \cdots) = \frac{t}{e^t-1}\left(n + \sum_{m=0}^{n-1} m\frac{t}{1!} + \cdots + \sum_{m=0}^{n-1} m^k\frac{t^k}{k!} + \cdots\right)$$

$$\implies \frac{t(e^{nt}-1)}{e^t-1} = \frac{te^{nt}}{e^t-1} - \frac{t}{e^t-1} = \sum_{k=0}^{\infty}(B_n(x) - B_k(0))\frac{t^k}{k!} = t\left(n + (\cdots)\right). \tag{2.24}$$

Matching the terms gives us

$$1 + 2^k + \cdots + n^k = \frac{1}{k+1}\left(B_{k+1}(n+1) - B_{k+1}(0)\right) = \frac{1}{k+1}\left(\sum_{j=0}^{k+1} B_j (n+1)^{k-j}\right) - B_{k+1}(0). \quad (2.25)$$

■

2.2 The Pigeonhole Principle

August 12th.

The *pigeonhole principle* simply states that if there are N objects places into k boxes, then some box contains at least $\lceil \frac{N}{k} \rceil$ objects.

Proof. Assume, if possible, that every box has less than $\lceil \frac{N}{k} \rceil$ objects. Then the total number of objects is at most $k \cdot \lceil \frac{N}{k} \rceil$. We take cases.

- Case I, where $k \mid N$. Then $\frac{N}{k}$ is a positive integer, and the total number of objects is less than $k \cdot \frac{N}{k} = N$. This is a contradiction.
- Case II, where $k \nmid N$. Then every box has at most $\lfloor \frac{N}{k} \rfloor$ objects, which is less than $\frac{N}{k}$. Multiplying by k tell us that the total number of objects is less than $k \cdot \frac{N}{k}$ which is, again, a contradiction.

■

Theorem 2.13 (The *Erdös-Szekeres theorem*)**.** *Any sequence of $mn+1$ distinct real numbers either contains an increasing subsequence of length $n+1$ or a decreasing subsequence of length $m+1$.*

Proof. Suppose the sequence of numbers is $a_1, a_2, \ldots, a_{mn+1}$. Assign a pair of numbers (b_i, c_i) to each index i, where b_i is the length of the longest increasing subsequence starting at i and c_i is the length of the longest decreasing subsequence starting at i.

If $b_i \geq n+1$ for some i, or if $c_i \geq m+1$ for some i, we are done. Else, we have $b_i \leq n$ and $c_i \leq m$ for all i. Since both are less than or equal to their respective bounds, the most number of distinct pairs (b_i, c_i) is mn. Thus, by the pigeonhole principle, there exists some $i \neq j$ such that $(b_i, c_i) = (b_j, c_j)$. Without loss of generality, assume $i < j$.

- Case I, where $a_i < a_j$. Then a_i can be (pre)appended to any increasing sequence starting at a_j which is a contradiction since $b_i > b_j$.
- Case II, where $a_i > a_j$. Then a_i can be (pre)appended to any decreasing sequence starting at a_j which is a contradiction since $c_i > c_j$.

■

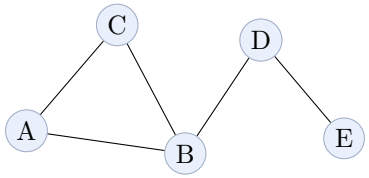
Dirichlet's principle states that in any set of $n+1$ integers, two of them must leave the same remainder modulo n. This is easy to see since there are only n possible remainders (namely $0, 1, \ldots, n-1$) and $n+1$ integers. By the pigeonhole principle, at least two of the integers must fall into the same remainder class.

Chapter 3

GRAPHS

3.1 Introduction

A *graph* is a pair $G = (V, E)$, where V is a set whose elements are called *vertices* and $E \subseteq V \times V$ is a set of unordered pairs $\{v_1, v_2\}$ of vertices, whose elements are called edges. Here, (v_1, v_2) and (v_2, v_1) are undistinguishable, and are simply denoted by $\{v_1, v_2\}$ or v_1v_2.



The above shows a simple undirected graph on five vertices $V = \{A, B, C, D, E\}$ with edges $E = \{AB, AC, BC, BD, DE\}$. Here, simple and undirected are also terms to be defined in the context of graph theory.

Definition 3.1. A graph is called a *simple graph* if it has no loops (edges connecting a vertex to itself) and no multiple edges (more than one edge connecting the same pair of vertices). Otherwise, it is termed a *multigraph*. A graph is called an *undirected graph* if its edges have no orientation; that is, the edge uv is identical to the edge vu. Otherwise, it is termed a *directed graph*.

In directed graphs, or *digraph*s, one deals with $G = (V, E, s, t)$, where $s : E \to V$ gives the *source node* of an edge and $t : E \to V$ gives the *target node* of an edge. In this edge set E, $uv \neq vu$, unlike the case of a simple graph.

Structure-preserving maps are useful in graph theory too.

Definition 3.2. Suppose we have two graphs $G = (V(G), E(G))$ and $H = (V(H), E(H))$. A function $f : V(G) \to V(H)$ is said to be a *graph homomorphism* if f preserves adjacency; that is, if $v_1v_2 \in E(G)$, then $f(v_1)f(v_2) \in E(H)$. If f is also bijective and f and f^{-1} are both graph homomorphisms, then f is termed a *graph isomorphism*.

We also term the group $\mathrm{Aut}(G)$ as the group of all graph isomorphisms of G, with the group operation of composition.

Definition 3.3. Suppose we have two digraphs $G_1 = (V_1, E_1, s_1, t_1)$ and $G_2 = (V_2, E_2, s_2, t_2)$. A *digraph homomorphism* is two maps $f_V : V_1 \to V_2$ and $f_E : E_1 \to E_2$ such that

$$s_2(f_E(e)) = f_V(s_1(e)) \quad \text{and} \quad t_2(f_E(e)) = f_V(t_1(e)). \tag{3.1}$$

That is, the source node of every image edge is the image node of every source node, and the target

node of every image edge is the image node of every target node.

One also discusses the neighbours of nodes.

Definition 3.4. The *degree of a node*, or the *valency of a node*, is simply defined as the number of edges incident with the vertex. If v is such a node in a graph (V, E), then $\deg(v) = \#\{u \in V \mid vu \in E\}$. In digraphs, one defines the *out-degree of a node* v as the number of edges with v as the source node, and the *in-degree of a node* v as the number of edges with v as the target node.

A *regular graph* is one where every vertex has the same degree. We now discuss the first ever theorem (historically) in graph theory.

Theorem 3.5. *A finite (simple) graph G has an even number of vertices of odd degree.*

Proof. Let $G = (V, E)$ be a graph. One can deduce that

$$2 \cdot \#E(G) = \sum_{v \in V(G)} \deg(v). \tag{3.2}$$

Thus, there must be an even number of vertices of odd degree to keep the term on the left even. ■

3.2 Walks, Paths, and Cycles

Definition 3.6. A *walk on a graph* G is an alternating sequence of vertices and edges

$$(v_0, e_1, v_1, e_2, v_2, \ldots, e_k, v_k) \tag{3.3}$$

such that for all i, e_i is an edge between v_{i-1} and v_i. The *length of a walk*, in this case, is termed k.

Definition 3.7. If the edges $e_1, e_2, \ldots, e_k$ are distinct, then the walk is called a *path on a graph*. A *simple path* is one where the vertices $v_0, v_1, \ldots, v_k$ are also all distinct. Finally, a *simple closed path*, or a *cycle on a graph*, is one where $v_0 = v_k$ and the rest are distinct.

A *metric on a graph* between two vertices $d(v_1, v_2)$ is defined as the length of the shortest walk between v_1 and v_2. This walk is always a path since if it's not, there is a repetition of edges, and appropriate middle edges and vertices can be deleted to form a path or a shorter walk. If no such path exists, then $d(v_1, v_2) = \infty$. Thus, a finite *connected graph* G is one where $d(v_1, v_2) < \infty$ for all $v_1, v_2 \in V(G)$.

August 21st.

3.2.1 The Königsberg Bridge Problem

The following figure illustrates the famous problem of the *Königsberg bridge problem*.

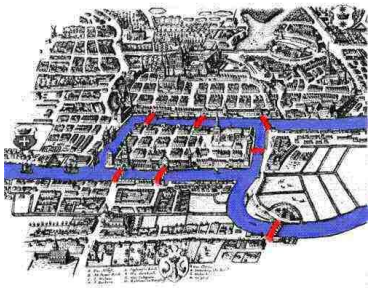


Figure 3.1: The Seven Bridges of Königsberg

Euler asked the question if one could cross each of the seven bridges exactly once and come back to the same side of the riverbank; this is formally considered as the first ever problem in graph theory.

Here, an *Eulerian circuit* is defined, which is a closed path using every edge in the graph exactly once. A graph with an Eulerian circuit is termed an *Eulerian graph*.

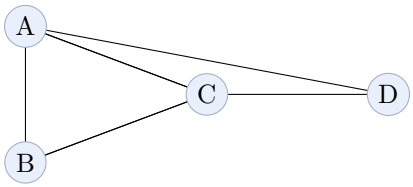


Figure 3.2: Graph Representation of the Seven Bridges of Königsberg

(Above graph to be fixed.)

Theorem 3.8. *A finite multigraph G is Eulerian if and only if G is connected and is a edge-disjoint union of cycle $G = C_1 \cup C_2 \cup \cdots \cup C_m$ where C_i's are cycles with no common edges.*

Proof. Suppose $G = C_1 \cup C_2 \cup \cdots \cup C_m$ where C_i's are edge-disjoint cycles. For $m = 2$, choose $v \in V(C_1) \cap V(C_2)$; such a v must exist or else the graph is disconnected. Starting at v, exhaust all edges in C_1 using the trivial Eulerian circuit and return to v. Do the same with C_2, and you have found the Eulerian circuit. Now apply the induction hypothesis; for an arbitrary m, choose $v \in V(C_1 \cup C_2 \cup \cdots \cup C_{m-1}) \cap V(C_m)$. Again, such a v must exist since G is connected. Starting at v, exhaust all edges in C_1 using the trivial Eulerian circuit and return to v, then use the Eulerian circuit in $C_1 \cup C_2 \cup \cdots \cup C_{m-1}$ formed via the induction hypothesis.

For the converse, an Eulerian circuit on the graph involves all edges and returns to the same vertex, so G must be connected. To show the disjoint union of cycles, find a cycle in the Eulerian circuit; there must exist at least one since, if not, the circuit itself is a cycle. Delete the edges from this cycle, and join the starting vertex and ending vertex in of this cycle in the circuit. Repeat the same until the resulting Eulerian circuit is a cycle. The disjoint union of this cycle and the cycles removed is the starting circuit. ■

One can show a better result.

Theorem 3.9. *A finite multigraph G is Eulerian if and only if G is connected and every vertex has an even degree.*

Proof. If G is Eulerian, then G is connected by the previous theorem, and $G = C_1 \cup C_2 \cup \cdots \cup C_m$, a union of disjoint cycles. Also, for $v \in V(G)$,

$$\deg_G(v) = \sum_{i=1}^{m} \deg_{C_i}(v) = 2\left(\sum_{i=1}^{m} \mathbf{1}_{\{v \in C_i\}}\right). \tag{3.4}$$

Thus, $\deg_G(v)$ is even for all $v \in V(G)$.

For the converse implication, assume G is connected and every vertex has an even degree. We will show that G is Eulerian. Start by choosing any cycle C_1 in G. Remove the edges of C_1 from G to form a subgraph G'. Since every vertex in G has even degree, removing the edges of C_1 leaves every vertex in G' with even degree. If G' is connected, repeat the process to find another cycle C_2 in G'. Continue this process until no edges remain. If G' is disconnected at any step, then each connected component of G' must also have all vertices of even degree. By the same argument, we can find cycles in each connected component and remove their edges. Eventually, all edges of G are partitioned into disjoint cycles. Since G is connected, these cycles can be combined into a single Eulerian circuit by appropriately traversing edges between cycles. Thus, G is Eulerian. ■

3.3 Adjacency

For a simple graph $G = (V, E)$, an *adjacency matrix* can be defined of dimension $\#V \times \#V$, with $a_{v,w} = 1$ if $vw \in E$, and $a_{v,w} = 0$ otherwise. For a multigraph, $a_{v,w}$ is the number of edges between vertices v and w. Similarly, for a directed graph, $a_{v,w} = 1$ if $vw \in E$ and $a_{w,v} = 0$ otherwise. For an undirected graph, the adjacency matrix A is symmetric and consists of only 1's and 0's.

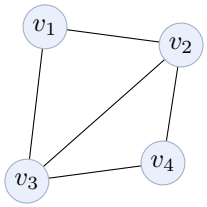


Figure 3.3: A simple graph with four vertices

The adjacency matrix for the graph in Figure 3.3 is given as $A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$. For two graphs G_1 and G_2 to be isomorphic, one can show that $A(G_1)$ must be similar to $A(G_2)$. Moreover, to count the number of walks from v to w of length k, one can use the k-th power of the adjacency matrix: the entry (i, j) of A^k gives the number of walks of length k from vertex v_i to vertex v_j.

August 28th.

Theorem 3.10. *Let $G = (V, E)$ be a simple graph with n vertices $V = \{1, 2, \ldots, n\}$ and adjacency matrix A. Then the $(i, j)^{th}$ entry of A^k gives the number of walks of length k from vertex i to vertex j.*

Proof. $k = 1$ is trivial, as a walk of length 1 is just showing there exists an edge between the two vertices. Let $k = 2$. Then the $(i, j)^{\text{th}}$ entry of A^2 is given as

$$(A^2)_{i,j} = \sum_{m=1}^{n} a_{im} a_{mj} \tag{3.5}$$

where a_{ij} is 1 if i and j are neighbours and zero otherwise. Thus, $a_{im}a_{mj}$ represents if vertex m is an immediate intermediate vertex between i and j. Thus, the number of walks of length 2 from i to j is equal to the number of such intermediate vertices m.

Assume the result holds for an arbitrary k; that is, $(A^k)_{i,j}$ gives the number of walks from i and j of length k. Then, for $k + 1$, we have

$$(A^{k+1})_{i,j} = \sum_{m=1}^{n} (A^k)_{i,m} a_{mj}. \tag{3.6}$$

Any walk from i to j must have a neighbour of j at the k^{th} step. By the induction hypothesis, $(A^k)_{i,m}$ represents the number of walks from i to m of length k. Thus, the total number of walks from i to j of length $k + 1$ is the sum over all possible intermediate vertices m. ■

Theorem 3.11. *Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of $A(G)$ where $G = (V, E)$ is a simple graph. Then the number of closed walks of length k is given by $\sum_{i=1}^{n} \lambda_i^k$.*

Proof. The number of such closed walks of length k is, clearly, $\operatorname{tr} A^k$. The eigenvalues of A^k are $\lambda_1^k, \ldots, \lambda_n^k$, and the trace is the sum of all eigenvalues; the result immediately follows. ■

Remark 3.12.

- For $k = 2$, $\frac{1}{2} \operatorname{tr} A^2$ provides the number of edges in the graph.
- For $k = 3$, $\frac{1}{6} \operatorname{tr} A^3$ provides the number of triangles in the graph.

Example 3.13. G is connected if and only if the largest eigenvalue of A has multiplicity 1. This is known as the *Perron-Frobenius theorem* will be taken as granted for now, without providing a proof.

Proposition 3.14. *For any eigenvalue λ of $G = (V, E)$, it holds that $|\lambda| \leq \max_{v \in G} \deg(v)$.*

Proof. Pick a corresponding eigenvector $x \neq 0$ with $Ax = \lambda x$, where A is the adjacency matrix. Pick $x_j = \|x\|_\infty = \max_i |x_i|$. Then we have

$$|\lambda|\,|x_j| = |\lambda x_j| = \left|\sum_{i=1}^{n} A_{ji} x_i\right| \leq \sum_{i=1}^{n} A_{ji}\,|x_i| \leq |x_j| \sum_{i=1}^{n} A_{ji} = |x_j| \deg(j) \implies |\lambda| \leq \deg(j) \leq \max_{v \in G} \deg(v). \tag{3.7}$$

■

Proposition 3.15. *Let spectrum of $G = \{\lambda_1, \ldots, \lambda_n\}$ be the list of eigenvalues where $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. Then*

$$\frac{1}{n} \sum_{v \in G} \deg(v) \leq \lambda_n \leq \max_{v \in G} \deg(v). \tag{3.8}$$

Proof. Let $e = (1\ 1\ \cdots\ 1)^t$, and let $d = Ae = (\deg(v_1)\ \deg(v_2)\ \cdots\ \deg(v_n))^t$. Then $e^t A e = ed = \sum_{v \in G} \deg(v)$. Then $\sup_{\|x\|=1} \langle x, Ax \rangle = \lambda_{\max}(A) \geq \frac{e^t}{\sqrt{n}} A \frac{e}{\sqrt{n}} = \frac{1}{n} \sum_{v \in G} \deg(v)$. The second inequality is just the one above. ■

Theorem 3.16. *A finite multigraph G is Eulerian if and only if G is connected and every vertex has an even degree.*

Proof. If G is Eulerian, then it is connected and every vertex has an even degree by definition. Conversely, suppose G is connected and every vertex has an even degree. Take 'a' longest trail, one where edges are not repeated, starting at $x \in V(G)$. We claim that this trail is actually closed. Let, if possible, the endpoint be $y \neq x$. There must have been an "entering" edge followed by an "exiting" edge. In the last step, one has entered y but never exited it. So, an odd number of edges incident with y appear in this trail. As $\deg(y)$ is even, there must be an unused edge incident with y. If we append that edge to the trail, we have a longer trail, contradicting our assumption.

We make another claim that this trail exhausts all edges. To show this, start by deleting all edges appearing in this trail. Assume that the graph so-obtained has at least one edge. In the trail, T, if a vertex y appears all its incident edges must also appear. Let z be a vertex not appearing on T. Then none of the vertices in T are neighbours of z, showing G not connected which is a contradiction. ■

Hamiltonian Graphs

August 29th.

Definition 3.17. A cycle in $G = (V, E)$ is said to be a *Hamiltonian cycle* if the vertices in the cycle are all distinct, except the starting and ending vertices, and all vertices are exhausted. A simple graph with a Hamiltonian cycle is termed a *Hamiltonian graph*.

Example 3.18.
- Trivially, all polygons with n vertices C_n are Hamiltonian graphs.
- The complete graph K_n is Hamiltonian for all $n \geq 3$.

If $G = (V, E)$ is a graph, then a *subgraph* $H = (V(H), E(H))$ is such that $V(H) \subseteq V(G)$ and $E(H) \subseteq V(H)$ where $E(H)$ are edges connecting vertices in $V(H)$. A *spanning subgraph* is such that $V(H) = V(G)$. We note that G is Hamiltonian if and only if C_n is a spanning subgraph of G.

Our goal now is to find a characterizing condition for Hamiltonian graphs (similar to the characterization of Eulerian graphs), and if a graph is Hamiltonian, then finding the (a) Hamiltonian cycle. The proof of providing a characterization is out of the scope of this course, while the latter problem is NP-hard and not always computationally tactible.

Definition 3.19. The *Hamiltonian closure of a graph* G, denoted by $\mathrm{cl}(G)$, is the graph obtained by repeatedly adding an edge between non-adjacent vertices u, v such that $\deg(u)+\deg(v) \geq n = \#V(G)$.

Proposition 3.20. *With* $\#V(G) = n \geq 3$*, let* G *be a simple graph. If* $\deg(v) \geq \frac{n}{2}$ *for all* $v \in V(G)$*, then* G *is Hamiltonian.*

Lemma 3.21. *A graph* G *is Hamiltonian if and only if* $\mathrm{cl}(G)$ *is Hamiltonian.*

Proof. If G is Hamiltonian, then clearly $\mathrm{cl}(G)$ is Hamiltonian since adding edges cannot remove Hamiltonian cycles. Conversely, suppose $\mathrm{cl}(G)$ is Hamiltonian. Assume the contrary that there exists G with $\#V(G) = n$ and u an v are *not* neighbours in G with $\deg(u) + \deg(v) \geq n$, but G is not Hamiltonian. $G + uv$ is Hamiltonian, however. Suppose the intermiedate graphs obtained to get the closure are given as

$$G = G_0 \subseteq G_1 \subseteq G_2 \subseteq \cdots \subseteq G_t = \mathrm{cl}(G) \tag{3.9}$$

where $\mathrm{cl}(G)$ is Hamiltonian. Then every Hamiltonian cycle in $G + uv$ must contain the edge uv. Let this Hamiltonian cycle be $(v, v_1, \ldots, v_{n-1}, v_n = u, v_{n+1} = v)$. Let $P = \{v_i \mid 2 \leq i \leq 2 \text{ and } v_1 v_i \in E(G)\}$ and $Q = \{v_i \mid 2 \leq i \leq n \text{ and } v_{i-1} v_n \in E(G)\}$ (P and Q are defined with respect to G and *not* $G + uv$). Then $\#P = \deg(v)$, $\#Q = \deg(v_n)$, and $\#P + \#Q = \deg(u) + \deg(v) \geq n$. Moreover, $P \cup Q \subseteq \{v_2, \ldots, v_n\}$. $P \cap Q$ is non-empty since

$$\#(P \cup Q) = \#P + \#Q - \#(P \cap Q) \geq n - \#(P \cap Q) \implies \#(P \cap Q) \geq n - \#(P \cup Q) \geq 1. \tag{3.10}$$

Thus, there exists some $v_i \in P \cap Q$ with $2 \leq i \leq n$, that is, $v_1 v_i \in E(G)$ and $v_{i-1} v_n \in E(G)$. Then the cycle $(v = v_1, v_i, v_{i+1}, \ldots, v_n = u, v_{i-1}, v_{i-2}, \ldots, v_2, v_1)$ is a Hamiltonian cycle not using the added edge uv—a contradiction. ∎

3.4 Bipartite Graphs

A simple graph $G = (V, E)$ is said to be bipartite if there exists a partition of $V(G)$ as $V = V_1 \sqcup V_2$ such that no pairs of vertices in V_1 are edges, and no pairs of vertices in V_2 are edges; that is, there exist $V_1, V_2 \subseteq V$ such that

$$V_1 \cap V_2 = \emptyset,\ V_1 \sqcup V_2 = V,\ E \subseteq V_1 \times V_2. \tag{3.11}$$

September 2nd.

Theorem 3.22 (*Kőnig's theorem*)**.** *A graph* G *is bipartite if and only if it has no odd cycles.*

Proof. Suppose G is bipartite with the required vertex sets $V_1 \sqcup V_2 = V(G)$. Take a cycle of length n in G and suppose it is an odd cycle, say $C = (v_1, v_2, \ldots, v_{2k}, v_{2k+1}, v_1)$ where $k \geq 1$. Without the loss of generality, assume $v_1 \in V_1$. Since a vertex's neighbours must be in a different vertex set, we have $v_2 \in V_2$. Proceeding inductively, we have $\{v_1, v_3, \ldots, v_{2k+1}\} \subseteq V_1$ and $\{v_2, v_4, \ldots, v_{2k}\} \subseteq V_2$. But $v_1 v_{2k+1} \in E(G)$, which is a contradiction since both v_1 and v_{2k+1} are in V_1. Thus, every cycle in G must be even.

For the converse, let $v_0 \in V$ and put it in a vertex set V_1. Put the neighbours of v_0 in V_2, that is, $N(v_0) \subseteq V_2$. We note that for $w_1, w_2 \in N(v_0)$, there cannot be the edge $w_1 w_2$, since having so would make (v_0, w_1, w_2, v_0) an odd cycle. Now put the neighbours of $N(v_0)$ in V_1, that is, $N(N(v_0)) \subseteq V_1$. Note that $v_0 \in N(N(v_0))$. Again, for $w_1, w_2 \in N(N(v_0))$, there cannot be the edge $w_1 w_2$ since having so would make $(v_0, *_1, w_1, w_2, *_2, v_0)$ an odd cycle, where $*_1, *_2$ are neighbours of w_1 and w_2 respectively, and both are in $N(v_0) \subseteq V_1$. Proceeding so, we have $V_1 = \{v \in V \mid d_G(v_0, v) \text{ is even }\}$ and $V_2 = \{v \in V \mid d_G(v_0, v) \text{ is odd }\}$. This shows bipartition for a connected G. Note that G is bipartite if and only if every

connected component of G is bipartite (a *connected component* of G is an equivalence relation on V, with $x \sim y$ if and only if there exists a path from x to y or $x = y$). If we find a bipartition of every connected component with $V_1^{(i)}, V_2^{(i)}$ where i signifies the i^{th} connected component, then defining

$$V_1 = \bigsqcup_{i=1}^{k} V_1^{(i)}, \quad V_2 = \bigsqcup_{i=1}^{k} V_2^{(i)} \tag{3.12}$$

results in the required bipartition. Coming back to the definition of V_1, all vertices even length away from v_0, and V_2, all vertices odd length away from v_0, we claim that no two vertices in V_1 are neighbours. Let us assume for contradiction that there exist $v, w \in V_1$ which are neighbours. Let P be the shortest path from v_0 to v and P' be the shortest path from v_0 to w. Then (v_0, P', w, v, P, v_0) is a walk of odd length which is a contradiction. Hence, any v and w in V_1 cannot be neighbours. Similarly, one can show fro V_2. ■

$K_{m,n}$ denotes the *complete bipartite graph* with m vertices in one vertex set and n vertices in the other vertex set. Here, uv is an edge in $K_{m,n}$ if and only if u and v are in different vertex sets. Thus the number of edges in $E(K_{m,n})$ is mn.

Corollary 3.23. *A bipartite graph has no triangles.*

To guarantee a triangle, how many edges must have a simple graph G have? We have the following theorem.

Theorem 3.24 (*Mantel's theorem*)**.** *If G, with $\#V(G) = n$, has more than $\left\lfloor \frac{n^2}{4} \right\rfloor$ edges, then G contains a triangle.*

Alternate proof. Consider G with no triangles, with m edges and n vertices. Let x and y be two vertices in G such that $xy \in E(G)$. Then $\deg(x) + \deg(y) \leq n$ since each vertex in G can only be connected to one of x or y; otherwise, if a vertex z were adjacent to both, the set $\{x, y, z\}$ would form a triangle. Now observe that

$$\sum_{xy \in E(G)} (\deg(x) + \deg(y)) \leq m \cdot n. \tag{3.13}$$

On the other hand, each degree appears in this sum once for each incident edge, so:

$$\sum_{xy \in E(G)} (\deg(x) + \deg(y)) = \sum_{v \in V(G)} \deg(v)^2 \leq mn. \tag{3.14}$$

Applying the Cauchy-Schwarz inequality results in

$$\sum_{v \in V(G)} \deg(v)^2 \geq \frac{1}{n} \left(\sum_{v \in V(G)} \deg(v) \right)^2 = \frac{(2m)^2}{n}. \tag{3.15}$$

Combining the two inequalities gives

$$\frac{4m^2}{n} \leq mn \implies m \leq \frac{n^2}{4}. \tag{3.16}$$

Hence, the number of edges is at most $\left\lfloor \frac{n^2}{4} \right\rfloor$, as desired. ■

3.4.1 Coloring

September 4th.

A *proper coloring* of a graph G is a mapping $f : V \to C$, where C is a finite set of a colors, such that for every edge $uv \in E(G)$, we have $f(u) \neq f(v)$. In other words, adjacent vertices must be assigned different colors. The *chromatic number* $\chi(G)$ of a graph G is the smallest number of colors needed for a

proper coloring of G, that is, the minimal cardinality of C required. For a planar graph G, $\chi(G)$ is at most 4.

One also has the *chromatic polynomial* $P_G(k)$, or $P(G,k)$, which specifies the number of proper k-colorings of G for a given k. From here, one can define

$$\chi(G) := \min\{k \mid P(G,k) > 0\}. \tag{3.17}$$

Proposition 3.25 (The *deletion-contraction principle*)**.** *For any edge $e \in E(G)$, G a graph,*

$$P(G,k) = P(G-e,k) - P(G \setminus e, k) \tag{3.18}$$

holds, where $G - e$ is the graph obtained by removing the edge e from G (termed deletion), and $G \setminus e$ is the graph obtained by joinin the vertices of edge e (termed contraction).

Proof. The proof is left as an exericse to the reader. ■

The following are some properties of $P(G,x)$.

1. $P(G,x)$ is a monic polynomial of degree $\#V(G)$.
2. $\chi(G) = \min\{k \in \mathbb{N} \mid P(G,k) > 0\}$.
3. The constant term of the polynomial, a_0, is zero.
4. One has either $\sum_{i=1}^{n} a_i = 0$ or $P(G,x) = x^n$.
5. $a_{n-i} = (-1)^i |a_{n-i}|$.
6. $a_{n-1} = -\#E(G)$.

Proof. We induct on $\#E$. Look at the base, D_n, a graph with n vertices and no edges. Then $P(D_n,x) = x^n$, a monic polynomial of degree 1. For any graph G, we have $P(G,k) = P(G-e,k) - P(G \setminus e,k)$. By the induction hypothesis, $P(G-e,k)$ is a monic polynomial of degree n, and $P(G \setminus e,k)$ is monic polynomial of degree $n-1$. Thus, $P(G,k)$ is one too. The second property is by definition. The third statement follows the same argument as the first statement.

For the fourth statement, note that $P(G,1) = 0$ if there is at least one edge. By induction, one can see that $\sum_{i=1}^{n} a_i = 0$, unless no edge is present, in which case $P(G,x) = P(D_n,x) = x^n$. Fifth and sixth statements also follow an argument by induction ■

Theorem 3.26. *The chromatic polynomial of a graph $G = (V,E)$ can be written in the form*

$$P(G,k) = \sum_{X \subseteq E} (-1)^{\#X} x^{\beta_0(X)} \tag{3.19}$$

where $\beta_0(X)$ denotes the number of connected components of the subgraph (V,X).

Thus, to find the coefficient of x^m, one must collect all subgraphs with m connected components. Let k_E denote the number of spanning subgraphs of G with m connected components and even number of edges, and let k_O denote the number of spanning subgroups of G with m connected components and odd number of edges. Then the coefficient of x^m comes out to be $k_E - k_O$. In the case where $m = n$, we have $k_E = 1$ and $k_O = 0$.

Proof. Here, $P(G,k) = k^n - \#(IC)$, where IC is the number of improper colorings. Denote, for an edge $e = uv$, $B_e = \{c \in IC \mid c(u) = c(v)\}$. Then

$$P(G,k) = k^n - \left|\bigcup_{e \in E} B_e\right| = k^n - \sum_{X \neq \emptyset,\ X \subseteq E} (-1)^{\#X} \left|\bigcap_{e \in X} B_e\right| \tag{3.20}$$

via the principle of inclusion-exclusion. Let $X = \{e_1, \ldots, e_k\} \subseteq E$ with $e_i = u_i v_i$. Then

$$\bigcap_{e \in X} B_e = \{c : V \to \{1,2,\ldots,k\} \mid c(u_i) = c(v_i) \text{ for all } i\} = k^{\beta_0(x)} \tag{3.21}$$

since the function c has to be constant on each connected component of (V, X). Hence,

$$P(G,k) = k^n - \sum_{X \neq \emptyset,\ X \subseteq E} (-1)^{\#X} k^{\beta_0(X)}. \tag{3.22}$$

■

3.5 Trees and Cayley's

September 16th.

Definition 3.27. A graph G with no simple cycles is called a *forest*. A connected forest is called a *tree*; essentially, a tree is a connected graph with no simple cycles.

Lemma 3.28. *A finite tree on n vertices, for $n \geq 2$, has at least two vertices of degree 1.*

The vertex in a tree with degree 1 is called a *leaf*.

Proof. Pick the longest path in the tree T, say $v_0 v_1 \cdots v_k$. We claim that both v_0 and v_k are leaves. If not, say $\deg v_0 \geq 2$, then there exists $u \neq v_1$ such that $uv_0 \in E(T)$. If u is not in the path, then $uv_0 v_1 \cdots v_k$ is a longer path, contradicting our assumption. If u is in the path, then there exists $uv_0 v_1 \cdots v_i u$ which is a simple cycle, contradicting the definition of a tree. Thus, v_0 must have degree 1. Similarly, one can show for v_k. ■

Lemma 3.29. *A graph G with n vertices is a tree if and only if it is connected and has $n-1$ edges.*

Proof. If G is a tree, then it is connected by definition. We induct on n. The base case $n = 1$ is trivial. Assume the result holds for all trees with up to $n-1$ vertices. Let G be a tree with n vertices. By the previous lemma, there exists a leaf v in G. Remove v and the edge incident with it to form a subgraph G'. Then G' is a tree with $n-1$ vertices, and by the induction hypothesis, it has $(n-1)-1 = n-2$ edges. Thus, G has $(n-2)+1 = n-1$ edges.

For the converse, start with an n-vertex graph G. If G contains a simple cycle, then removing an edge from that cycle leaves the graph connected; repeat until there are no such cycles. After (finitely) many steps, we end up with a tree, which we have shown to have $n-1$ edges. This meas that the original graph G, which had at least one cycle, must have had more than $n-1$ edges. ■

Similarly, one can show that G is a forest if and only if it has $n - \beta_0(G)$ edges, where $\beta_0(G)$ is the number of connected components of G.

Definition 3.30. A *labelled tree* on $[n] = \{1, 2, \ldots, n\}$ is a tree with vertex set $[n]$. Two labelled trees are considered distinct if they are not isomorphic via an isomorphism that preserves the labels.

September 18th.

Theorem 3.31 (*Cayley's theorem*)**.** *The number of labelled trees on $[n]$ is n^{n-2}.*

Proof. We will show a bijection between the set of labelled trees on $[n]$ and $[n]^{n-2}$, the set of all sequences of length $n-2$ with entries from $[n]$. For a labelled tree T on vertex set $[n]$, generate a sequence of labelled trees $T_1, T_2, \ldots, T_{n-1}$ inductively as follows: let $T_1 = T$, and obtain T_2 by removing the leaf with the smallest label from T_1 along with its incident edge. Continue this process: at each step, remove the leaf with the smallest label from the current tree and its incident edge. This process terminates with T_{n-1}, which is a tree on two vertices. Thus, each T_i is a labelled tree on $n - i + 1$ vertices.

Let x_i denote the leaf removed from T_i (i.e., the leaf of T_i with the smallest label), and let y_i denote its unique neighbor in T_i. Define a sequence $(y_1, y_2, \ldots, y_{n-2})$, which we call the *Prüfer code* of the tree T. Since at each step the removed leaf and its neighbor are well-defined, and since the process continues

for $n-2$ steps, this produces a sequence of length $n-2$ with entries from $[n]$. Hence, every labelled tree on $[n]$ gives rise to a unique Prüfer code in $[n]^{n-2}$.

To show that this map is a bijection, it remains to show that every sequence in $[n]^{n-2}$ corresponds to a unique labelled tree on $[n]$. Given a sequence $(y_1, y_2, \ldots, y_{n-2})$ in $[n]^{n-2}$, we reconstruct the tree as follows:

1. Initialize the degree of each vertex $v \in [n]$ by $\deg(v) = 1 + \#\{i \mid y_i = v\}$, which counts the number of times v appears in the sequence plus one (show this to be true).
2. For $i = 1$ to $n-2$, find the smallest vertex x_i such that $\deg(x_i) = 1$. Add the edge (x_i, y_i) to the tree, and decrease both $\deg(x_i)$ and $\deg(y_i)$ by 1.
3. After processing all entries in the sequence, two vertices remain with degree 1. Connect these two vertices to complete the tree.

This algorithm constructs a unique labelled tree from any given Prüfer code. Since both the encoding and decoding procedures are well-defined and inverse to one another, the correspondence is a bijection between the set of labelled trees on $[n]$ and $[n]^{n-2}$. ■

A few key observations may be inferred:

1. If $(y_1, y_2, \ldots, y_{n-2})$ is the Prüfer code of tree T, then $(y_2, \ldots, y_{n-2})$ is the Prüfer code of the tree T_2.
2. The degree of a vertex i is $d_i = \sum_{j=1}^{n-2} \mathbf{1}_{[y_j = i]} + 1$.
3. The leaf $x_k = \min_{i \in [n]}\{i \mid i \notin \{x_1, \ldots, x_{k-1}, y_k, y_{k+1}, \ldots, y_{n-2}\}\}$.

Note that 1. follows from the definition, and 2. can be shown as a result of induction.

Theorem 3.32 (The *tree counting theorem*)**.** *The number of labelled spanning trees on n vertices with degree sequence $(d_1, d_2, \ldots, d_n)$ is given by*

$$\binom{n-2}{d_1 - 1, \ldots, d_n - 1} = \frac{(n-2)!}{(d_1-1)!(d_2-1)! \cdots (d_n-1)!}. \tag{3.23}$$

Proof. In the Prüfer code of a labelled tree T, a vertex i appears exactly $d_i - 1$ times. Thus the counting problem is equivalent to the number of Prüfer codes which contain $d_i - 1$ copies of i. This is given by the multinomial coefficient above. ■

3.6 Ramsey Theory

In any graph G on 6 vertices, either $K \subseteq G$ or $K_3 \subseteq \overline{G}$, where $\overline{G}$ is the complement of G. G and $\overline{G}$ cannot both be triangle-free. Equivalently, any edge-colouring of K_6 with two colours must have a monochromatic triangle.

Definition 3.33. The *Ramsey number* $R(m, n)$ is defined as

$$R(m,n) = \inf\{t \mid \text{ for any } G \subseteq K_t, \text{ either } K_m \subseteq G \text{ or } K_n \subseteq \overline{G}\} \tag{3.24}$$

One can easily see that $R(m, 2) = m$ and $R(m, n) = R(n, m)$, and $R(m, n) < \infty$ for all $m, n \in \mathbb{N}$.

Lemma 3.34. *$R(m, n) \leq R(m-1, n) + R(m, n-1)$ for all $m, n \geq 2$.*

Proof. Let $t = R(m-1, n) + R(m, n-1)$ and consider $v \in V(K_t)$, joined to $t-1$ other vertices. Bicolour all edges red or blue. Let P be the set of all vertices connected to v by a red edge, and let Q be the set of all vertices connected to v by a blue edge. Then $\#P + \#Q = p + q = t - 1$. Thus either $p \geq R(m-1, n)$ or $q \geq R(m, n-1)$. If $p \geq R(m-1, n)$, then either there is a red $(m-1)$-clique with vertices in P or there is a blue n-clique with vertices in P. In the former case, adding v gives a red m-clique; in the latter

case, we are done. Similarly, if $q \geq R(m, n-1)$, then either there is a red m-clique with vertices in Q or there is a blue $(n-1)$-clique with vertices in Q. In the latter case, adding v gives a blue n-clique; in the former case, we are done. ■

Lemma 3.35. *$R(n,n) \leq 4^n$ for all natural n.*

Proof. Using the above inequality, a proof by induction on n suffices. ■

Lemma 3.36. *$R(n,n) > \lfloor 2^{n/2} \rfloor$ for all natural n.*

Proof. Let $t = \lfloor 2^{n/2} \rfloor$. There is an edge colouring of K_t such that there is no monochromatic K_n; we wish to show this. Set of all possible colourings on K_t is $\Omega_t = E(K_t)^{\{R,B\}}$. Let $f : E(K_t) \to \{R, B\}$ be such a colouring. THe number of such functions is clearly $2^{\binom{t}{2}}$. Let $A_t = \{f \in \Omega_t : f \text{ has no monochromatic } n-\text{clique}\}$. $A_t \neq \emptyset$ and $\Omega_t \setminus A_t \neq \Omega_t$. If $f \notin A_t$, there exists $S \subseteq [t]$ and $\#S = n$ such that f is constant on E_S. Thus,

$$\Omega_t \setminus A_t = \bigcup_{S \subseteq [t],\ \#S=n} \{f \in \Omega_t : f \text{ is constant on } E_S\}. \tag{3.25}$$

We now look at the probability of a random colouring X being in $\Omega_t \setminus A_t$.

$$P(X \in \Omega_t \setminus A_t) \leq \sum_{S \subseteq [t],\ \#S=n} P(X|_{E_S} = R) + P(X|_{E_S} = B) = \sum_{S \subseteq [t],\ \#S=n} \frac{2}{2^{\binom{n}{2}}} = \binom{t}{n} \frac{1}{2^{\binom{n}{2}-1}}. \tag{3.26}$$

■

3.6.1 de Bruijn Sequences and Graphs

September 23rd.

Suppose we have a circular rotating drum with a 'window' of length n. The rotating drum has a tape of l symbols on it coming from a finite alphabet Σ with $|\Sigma| = q$. The drum rotates one position at a time, and at each position, we can see the n symbols in the window. We want to design a circular tape (sequence) of length l such that as the drum rotates, every possible sequence of length n appears in the window at least once.

Clearly, the number of 'words' that the tape must cover is q^n. For our best case scenario, we can have that every window is distinct, giving us a lower bound of q^n for l. An upper bound is nq^n, which can be achieved by concatenating all possible words of length n together. Thus we obtain

$$q^n \leq l \leq nq^n. \tag{3.27}$$

Definition 3.37. A *de Bruijn sequence* of order n over an alphabet Σ of size q is a cycle sequence of length q^n in which every possible word of length n over Σ appears exactly once as a contiguous subsequence.

We show that the lower bound can be achieved.

Theorem 3.38. *For every alphabet Σ of size q and $n \geq 1$, there exists a de Bruijn sequence of order n over Σ. In particular, there are $\dfrac{(q!)^{q^{n-1}}}{q^n}$ distinct de Bruijn sequences of order n over Σ.*

Before we show this theorem, we introduce the *de Bruijn graph*, which is a directed graph defined as follows. The vertex set is Σ^{n-1}, the set of all words of length $n-1$ over Σ. There is a directed edge from vertex $v = a_1 a_2 \cdots a_{n-1}$ to vertex $w = b_1 b_2 \cdots b_{n-1}$ if and only if $a_2 = b_1, a_3 = b_2, \ldots, a_{n-1} = b_{n-2}$; that is, the last $n-2$ symbols of v are the same as the first $n-2$ symbols of w. In this case, we label the edge vw with the symbol b_{n-1}, the last symbol of w. Note that each vertex has exactly q outgoing edges and exactly q incoming edges. Thus the graph has q^{n-1} vertices and q^n edges. The in-degree and the out-degree of each vertex is q. Let $B(q,n)$ denote the de Bruijn graph with alphabet size q and vertices consisting of words of length $n-1$.

Theorem 3.39. *There exists a bijection between Eulerian cycles in $B(q,n)$ and de Bruijn sequences of order n over an alphabet of size q.*

One can show that if G is a connected digraph then G has an Eulerian cycle if and only if the in-degree of each vertex is equal to its out-degree.

Proof. Let $\mathcal{E}$ denote the set of Eulerian cycles in $B(q,n)$ and let $\mathcal{D}$ denote the set of de Bruijn sequences of order n over an alphabet of size q. Also let $m = q^n$ be the number of directed edges in the graph $B(q,n)$. Suppose $C = w_0 w_1 \cdots w_{m-1} w_0$ is an Eulerian cycle in $\mathcal{E}$ of length n. Let $s(v)$ denote the suffix of the word v, and let $p(v)$ denote the prefix of the word v. Then $s(w_i) = p(w_{i+1})$ for all $0 \le i \le m-1$, where indices are taken modulo m. Define a sequence $S = a_1 a_2 \cdots a_m$, where a_i is the label of the edge $w_i w_{i+1}$ in $B(q,n)$, or $a_i = s(w_{i+1})$. Since C is an Eulerian cycle, every edge in $B(q,n)$ is traversed exactly once, and thus every word of length n appears exactly once as a contiguous subsequence of S. Therefore, S is a de Bruijn sequence of order n over Σ.

Conversely, given a de Bruijn sequence $S = a_1 a_2 \cdots a_m$ of order n, construct a cycle $C = w_0 w_1 \cdots w_{m-1} w_0$ in $B(q,n)$, where $w_i = a_i a_{i+1} \cdots a_{i+n-2}$ (indices taken modulo m). Since S is a de Bruijn sequence, every word of length n appears exactly once, and thus every edge in $B(q,n)$ is traversed exactly once. Therefore, C is an Eulerian cycle in $B(q,n)$. ■

Corollary 3.40. *de Bruijn sequences exist.*

3.7 Linear Algebra Interlude

If the legs of a right triangle are a and b, and the hypotenuse is c, then $a^2 + b^2 = c^2$. This is the Pythagorean theorem. THe converse, that if a triangle's sides satisfy $a^2 + b^2 = c^2$, then the triangle is a right triangle, is known as Carpenter's theorem. We provide a generalization of the Pythagorean theorem.

Theorem 3.41. *Let P be an orthogonal projection in $M_n(\mathbb{R})$, that is, $P^2 = P = P^T$. If k is the rank of the projection, then $\operatorname{tr} P = k$.*

An analogous generalization of the Carpenter's theorem is as follows.

Theorem 3.42. *Let $a_1, \ldots, a_n$ be numbers such that $0 \le a_i \le 1$ for all i and $\sum_{i=1}^n a_i = k$, where k is an integer. Then there exists an orthogonal projection P in $M_n(\mathbb{R})$ such that $\operatorname{tr} P = k$ and the diagonal entries of P are $a_1, \ldots, a_n$.*

Let $(\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$. Then the *permutation polytope* of $(\lambda_1, \ldots, \lambda_n)$ is defined as the convex hull of $\{(\lambda_{\sigma(1)}, \ldots, \lambda_{\sigma(n)}) \mid \sigma \in S_n\}$. Majorization: we define a pre-order on $\mathbb{R}^n$ as follows. For $(\lambda_1, \ldots, \lambda_n), (\mu_1, \ldots, \mu_n) \in \mathbb{R}^n$, we say that $(\lambda_1, \ldots, \lambda_n)$ is majorized by $(\mu_1, \ldots, \mu_n)$, denoted $(\lambda_1, \ldots, \lambda_n) \preceq (\mu_1, \ldots, \mu_n)$, if

$$\sum_{i=1}^{k} \lambda_{(i)} \le \sum_{i=1}^{k} \mu_{(i)} \quad \text{for all } 1 \le k \le n. \tag{3.28}$$

For example, $(\frac{1}{n}, \ldots, \frac{1}{n}) \preceq (1, 0, \ldots, 0)$.

Theorem 3.43. *$(\lambda_1, \ldots, \lambda_n) \preceq (\mu_1, \ldots, \mu_n)$ if and only if the permutation polytope of $(\lambda_1, \ldots, \lambda_n)$ is contained in the permutation polytope of $(\mu_1, \ldots, \mu_n)$. Equivalently, there exists a sequence of transpositions such that $(\mu_1, \ldots, \mu_n) \ge \cdots \ge (\lambda_1, \ldots, \lambda_n)$ where consecutive vectors differ only in two vectors.*

September 25th.

Let $G = (V_1 \sqcup V_2, E)$ be a finite bipartite graph. A *perfect matching* from V_1 to V_2 is a one-to-one function $f : V_1 \to V_2$ such that $vf(v) \in E$ for all $v \in V_1$. A *matching* from V_1 to V_2 is a one-to-one

function $f : U \to V_2$, where $U \subseteq V_1$, such that $vf(v) \in E$ for all $v \in U$. A matching is *maximum* if there is no matching with a larger domain.

Theorem 3.44 (*Hall's marriage theorem*)**.** *Let $G = (V_1 \sqcup V_2, E)$ be a finite bipartite graph. A perfect matching from V_1 to V_2 exists if and only if for every subset $S \subseteq V_1$, we have $\#N(S) \geq \#S$, where $N(S)$ it the set of all neighbours of S in V_2.*

Proof. The forward implication is easy to see. For the converse implication, let us perform induction on $n = \#V_1$. The base case $n = 1$ is trivial. Assume the result holds for all bipartite graphs with $\#V_1 < n$. We do casework.

Case I: $\#N(S) \geq \#S + 1$ for all non-empty $S \subsetneq V_1$. Fix $v \in V_1$ and let $w \in N(\{v\})$. Remove v from V_1 and w from V_2 to form a new bipartite graph G'. If $w \in N_G(S')$, then for any non-empty $S' \subseteq V_1 \setminus \{v\}$, we have $\#N_{G'}(S') = \#N_G(S') - 1 \geq \#S' + 1 - 1 = \#S'$. Thus, by the induction hypothesis, there exists a perfect matching from $V_1 \setminus \{v\}$ to $V_2 \setminus \{w\}$ in G'. Adding the pair (v, w) gives a perfect matching from V_1 to V_2 in G. If $w \notin N_G(S')$, then $\#N_{G'}(S') = \#N_G(S') \geq \#S'$. Again, by the induction hypothesis, there exists a perfect matching from $V_1 \setminus \{v\}$ to $V_2 \setminus \{w\}$ in G'. Adding the pair (v, w) gives a perfect matching from V_1 to V_2 in G.

Case II: There exists a non-empty $S \subsetneq V_1$ such that $\#N(S) = \#S$. Look at the subgraph G' induced by the vertex subset $S \cup N(S)$. For any non-empty $S' \subseteq S$, we have $\#N_{G'}(S') = \#N_G(S') \geq \#S'$. Thus, by the induction hypothesis, there exists a perfect matching from S to $N(S)$ in G'. We now show that the subgraph G'' induced by the vertex set $(V_1 \setminus S) \cup (V_2 \setminus N(S))$ also satisfies Hall's condition. Assuming the contrary, there exists a subset $S' \subseteq V_1 \setminus S$ such that $\#N_{G''}(S') < \#S'$. Then $N_{G''}(S') = N_G(S') \cap (V_2 \setminus N_G(S))$. Since $N_G(S' \cup S) = N_{G''}(S') \cup N_G(S)$, we have

$$\#N_G(S' \cup S) \leq \#N_{G''}(S') + \#N_G(S) < \#S' + \#S = \#(S' \cup S). \tag{3.29}$$

Hall's condition for G suggests otherwise ($\#N_G(S' \cup S) \geq \#(S' \cup S)$), a contradiction. Thus, by the induction hypothesis, there exists a perfect matching from $V_1 \setminus S$ to $V_2 \setminus N(S)$ in G''. Combining this with the perfect matching from S to $N(S)$ in G' gives a perfect matching from V_1 to V_2 in G. ∎

Definition 3.45. A matrix $A \in M_n(\mathbb{R})$ is said to be doubly stochastic if $a_{ij} \geq 0$ for all i, j, and $\sum_{i=1}^n a_{ij} = 1$ for all j and $\sum_{j=1}^n a_{ij} = 1$ for all i.

Theorem 3.46 (The *Birkhoff-von Neumann theorem*)**.** *The set of doubly stochastic matrices $DS_n(\mathbb{R})$ is the convex hull of the set of $n \times n$ permutation matrices.*

Proof. Let A be a doubly stochastic matrix. Let $V_1 = \{r_1, \ldots, r_n\}$ be the row indices and $V_2 = \{c_1, \ldots, c_n\}$ be the column indices. Consider the bipartite graph $G = (V_1 \sqcup V_2, E)$, where $r_i c_j \in E$ if and only if $a_{ij} > 0$. We claim that G satisfies Hall's condition. For any non-empty $S \subseteq V_1$, we have

$$\sum_{i \in S} \sum_{j=1}^n a_{ij} = \#S. \tag{3.30}$$

Each (strictly) positive entry in the rows corresponding to S lies in one of the columns of $N(S)$. Thus,

$$\#S = \sum_{i \in S} \sum_{j=1}^n a_{ij} \leq \sum_{j \in N(S)} \sum_{i=1}^n a_{ij} = \#N(S), \tag{3.31}$$

By Hall's marriage theorem, there exists a perfect matching from V_1 to V_2 in G corresponding to a permutation σ of $\{1, 2, \ldots, n\}$ such that $a_{i,\sigma(i)} > 0$ for all i. Let P be the permutation matrix corresponding to σ. Let $\epsilon = \min_{1 \leq i \leq n} a_{i,\sigma(i)} > 0$. Then $(\epsilon P)_{ij} \leq a_{ij}$ for all i, j. If $\epsilon = 1$, then $A = P$ and we are done. So assume that $\epsilon < 1$. Then

$$A = (A - \epsilon P) + \epsilon P = (1 - \epsilon) \left(\frac{A - \epsilon P}{1 - \epsilon} \right) + \epsilon P. \tag{3.32}$$

Note that the number of zero entries in $\frac{A - \epsilon P}{1 - \epsilon}$ is strictly greater than the number of zero entries in A. We perform induction on the number of non-zero entries of A. If A has exactly n non-zero entries, then A is

a permutation matrix and we are done. Assume the result holds for all doubly stochastic matrices with up to k non-zero entries. Let A be a doubly stochastic matrix with $k+1$ non-zero entries. Then $\frac{A-\epsilon P}{1-\epsilon}$ is a doubly stochastic matrix with at most k non-zero entries, and by the induction hypothesis, it can be written as a convex combination of permutation matrices. Thus, A can also be written as a convex combination of permutation matrices. ■

Definition 3.47. A linear mapping $T : \mathbb{R}^n \to \mathbb{R}^n$ is called a *T-transform* if there exist $0 \leq t \leq 1$ and $1 \leq j < k \leq n$ such that

$$T\begin{pmatrix} y_1 \\ \vdots \\ y_{j-1} \\ y_j \\ \vdots \\ y_k \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_{j-1} \\ ty_j + (1-t)y_k \\ \vdots \\ (1-t)y_j + ty_k \\ \vdots \\ y_n \end{pmatrix}. \tag{3.33}$$

October 7th.

Theorem 3.48. *The following are equivalent for $x, y \in \mathbb{R}^n$.*

1. *$x \preceq y$.*
2. *x is obtained from y by a finite sequence of T-transforms.*
3. *x is in the permutation polytope of y.*
4. *$x = Ay$ for a doubly stochastic matrix A.*

Proof. 1. implies 2.: We perform induction on n. We have assumed $x_1 \geq \cdots \geq x_n$ and $y_1 \geq \cdots \geq y_n$. This shows that $y_n \leq x_1 \leq y_1$. Chosse k such that $y_k \leq x_1 \leq y_{k-1}$. Rewrite $x_1 = ty_1 + (1-t)y_k$ for some $0 \leq t \leq 1$. Define a T-transform T_1 by

$$T_1 z = (tz_1 + (1-t)z_k, z_2, \ldots, z_{k-1}, (1-t)z_1 + tz_k, z_{k+1}, \ldots, z_n). \tag{3.34}$$

We claim that $x \preceq T_1 y$. We have

$$\sum_{j=1}^{m} y'_j = \sum_{j=2}^{k-1} y_j + (1-t)y_1 + ty_k + \sum_{j=k+1}^{m} y_j = \left(\sum_{j=1}^{m} y_j\right) - ty_1 + (t-1)y_k \geq \sum_{j=1}^{m} x_j - x_1 = \sum_{j=2}^{m} x_j \tag{3.35}$$

We use induction hypothesis to give us $(y'_2, \ldots, y'_n) \preceq (x_2, \ldots, x_n)$. Thus, x is obtained from y by a finite sequence of T-transforms.

For 2. implies 3., we assume $T_r \cdots T_1 y = x$ for $y \in \mathbb{R}^n$ and each T a T-transform. Note that the permutation polytope of y contains the permutation polytope of Ty for a T-transform T. This is becuase for any $\sigma \in S_n$, one can show that $\sigma(Ty)$ is in the permutation polytope of y. Thus, by induction, x is in the permutation polytope of y.

For 3. implies 4., let x be in the permutation polytope of y. Then $x = \sum_{\sigma \in S_n} t_\sigma(\sigma y)$ for some $t_\sigma \geq 0$ with $\sum_\sigma t_\sigma = 1$. Letting $A = \sum_{\sigma \in S_n} t_\sigma P_\sigma$, where P_σ is the permutation matrix corresponding to σ, we have $x = Ay$. Note that A is doubly stochastic.

For 4. implies 3., since any doubly stochastic matrix is a convex combination of permutation matrices, the result follows.

For 4. implies 1., let $x = Ay$ for a doubly stochastic matrix A. We have

$$\sum_{j=1}^{k} x_j = \sum_{j=1}^{k} \sum_{i=1}^{n} a_{ij} y_i \tag{3.36}$$

where $t_i = \sum_{j=1}^{k} a_{ij}$. Thus

$$\sum_{j=1}^{k} x_j - \sum_{j=1}^{k} y_j = \sum_{i=1}^{n} t_i y_i - \sum_{j=1}^{k} y_j + (k - \sum_{i=1}^{n} t_i) y_k = \sum_{i=1}^{n} (t_i - 1)(y_i - y_k) + \sum_{i=k+1}^{n} t_i (y_i - y_k) \leq 0. \quad (3.37)$$

■

Theorem 3.49 (*Schur-Horn theorem*)**.** *Let $A \in H_n(\mathbb{R})$, the set of real symmetric matrices, with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$. Then $\operatorname{diag}(A) \preceq (\lambda_1, \ldots, \lambda_n)$. Conversely, if $x \preceq (\lambda_1, \ldots, \lambda_n)$, then there exists $A \in H_n(\mathbb{R})$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ and diagonal entries $x_1, \ldots, x_n$.*

Proof. For the forward implication, let $A = U \operatorname{diag}(\lambda_1, \ldots, \lambda_n) U^t$ be the spectral decomposition of A. Then simply take $D = (|U_{ij}|^2)$, which is doubly stochastic, and note that $\operatorname{diag}(A) = D(\lambda_1, \ldots, \lambda_n)$. For the converse implication, let us show for a 2×2 matrix first. For $A = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$, a 2×2 doubly stochastic matrix with $0 \leq a, b \leq 1$ and $a + b = 1$, note that $\begin{pmatrix} \sqrt{a} & -\sqrt{b} \\ \sqrt{b} & \sqrt{a} \end{pmatrix}$ is an orthogonal matrix. Let $(r_1, \ldots, r_k)$ be majorized by $(\lambda_1, \ldots, \lambda_k)$ via the sequence of T-transforms $T_k \cdots T_1$, where each T_i is an orthostochastic matrix, say O_i. Then $r = (O_k \cdots O_1)(\lambda_1, \ldots, \lambda_k)$. ■

The linear mapping $\phi : M_n(\mathbb{R}) \to D_n(\mathbb{R})$ defined by $\phi(A) = \operatorname{diag}(A_{11}, A_{22}, \ldots, A_{nn})$ is a positive linear map. Moreover, for a diagonal matrix D, $\phi(DA) = D\phi(A)$ and $\phi(AD) = \phi(A)D$. Note that $\phi(I) = I$.

Theorem 3.50 (*Hadamard's determinant inequality*)**.** *If A is a positive semidefinite real matrix, then $\det A \leq \det \phi(A)$, with equality if and only if A is diagonal, or $\phi(A) = 0$.*

Proof. Assume A is positive definite. If we look at $\operatorname{tr}(\phi(A)^{-1/2} A \phi(A)^{-1/2})$, we have

$$\operatorname{tr}(\phi(A)^{-1/2} A \phi(A)^{-1/2}) = \operatorname{tr}(\phi(\phi(A)^{-1/2} A \phi(A)^{-1/2})) = \operatorname{tr}(I) = n. \quad (3.38)$$

Thus, if we denote $X = \phi(A)^{-1/2} A \phi(A)^{-1/2}$, then $\frac{1}{n} \operatorname{tr} X = 1$. By the AM-GM inequality, $\det X \leq (\frac{1}{n} \operatorname{tr} X)^n = 1$. Therefore, $\det A \leq \det \phi(A)$, with equality if and only if $X = I$, or $A = \phi(A)$. ■

3.8 Extremal Set Theory

How large or small a family of sets can be if we require that it satisfies certain restrictions? This is the basic question in extremal set theory. Given a finite set X, how large can a family $\mathcal{F} \subseteq \mathcal{P}(X)$ be if it has to avoid certain configurations.

3.8.1 Sperner's Problem

Pick as many subsets of an n-element set X such that no one subset contains another. The problem asks what the maximum possible number of sets is. It's all sets of size $\lfloor \frac{n}{2} \rfloor$. For all sets of size k with $1 \leq k \leq n$ denote it by $\mathcal{F}_k$. Then $\#\mathcal{F}_k = \binom{n}{k}$. We have

$$\max_{1 \leq k \leq n} \#\mathcal{F}_k = \binom{n}{\lfloor \frac{n}{2} \rfloor}. \quad (3.39)$$

The assertion of Sperner's theorem is that $\binom{n}{\lfloor \frac{n}{2} \rfloor}$ is also the maximum size of a family of subsets of X such that no one subset contains another.

Theorem 3.51 (*Dilworth's theorem*)**.** *Let $(P, \leq)$ be a finite poset. Then the maximum size of an antichain in P is equal to the minimum number of chains into which P can be partitioned. That is,*

$$\max\{\#A \mid A \text{ is an antichain in } P\} = \min\{k \mid P = C_1 \cup C_2 \cup \cdots \cup C_k, C_i \text{ are chains}\}. \quad (3.40)$$

Note that $S \subseteq P$ is termed a chain if all elements in S are comparable, and an antichain if no two elements in S are comparable. Note that, by definition, singletons are both chains and antichains.

If P was totally ordered, then the maximum size of an antichain is 1, and the minimum number of chains into which P can be partitioned is 1 since P itself is a chain. If P was an antichain, then the maximum size of an antichain is $\#P$, and the minimum number of chains into which P can be partitioned is $\#P$ since each element of P is a singleton chain.

Proof. Let $M = \max\{\#A \mid A \text{ is an antichain in } P\}$ and $m = \min\{k \mid P = C_1 \cup C_2 \cup \cdots \cup C_k, C_i \text{ are chains}\}$. We want to show that $M = m$. We first show that $m \geq M$. Let $\{a_1, \ldots, a_M\}$ be an antichain in P of size M. Let $P = C_1 \cup C_2 \cup \cdots \cup C_m$ be a partition of P into m chains. Since no two elements of the antichain are comparable, each a_i must belong to a different chain. Thus, $m \geq M$.

Now we show $M \geq m$. We proceed with induction on $\#P$. We showed the base case $\#P = 1$ above. Assume the result holds for all posets of size less than $\#P$. Let C be a maximal chain in P, that is, a chain which is not properly contained in any other chain. Let A' be an antichain in $P \setminus C$ of maximum size $M(A')$. We consider cases.

- Case I; $M(A') \leq M - 1$. Then, by the induction hypothesis, there is a partition of $P \setminus C$ which has at most $M - 1$ chains $C_1, \ldots, C_{M-1}$. Then $P = C \cup C_1 \cup \cdots \cup C_{M-1}$ is a partition of P into at most M chains, giving us $m \leq M(A') + 1 \leq M$.
- Case II; There is an antichain $A = \{a_1, \ldots, a_M\}$ in $P \setminus C$ of size M. Construct two sets as follows:

$$S^- := \{x \in P \mid x \leq a_i \text{ for some } 1 \leq i \leq M\}, \quad S^+ := \{x \in P \mid x \geq a_i \text{ for some } 1 \leq i \leq M\}. \tag{3.41}$$

  Note that A is an antichain of size M in S^- and S^+, both of which are strictly contained in P. S^+ and S^- can both be decomposed as unions of M disjoint chains as $S^- = \bigcup_{i=1}^{M} S_i^-$ and $S^+ = \bigcup_{i=1}^{M} S_i^+$. Let b_i denote the maximum element in S_i^-. Then b_i can only be one of the a_j's, otherwise we can add b_i to A to get a larger antichain in P. By this logic, $\max(S_i^-) = a_i$ and $\min(S_i^+) = a_i$ for $1 \leq i \leq M$. Now construct chains as

$$C_i = S_i^- \cup S_i^+, \text{ where } S_i^- \cap S_i^+ = \{a_i\}, 1 \leq i \leq M. \tag{3.42}$$

  One can show that the C_i's are really chains, and $P = \bigcup_{i=1}^{M} C_i$, with $C_i \cap C_j = \emptyset$ for $i \neq j$. Thus, $m \leq M$.

■

Theorem 3.52 (*Sperner's theorem*)**.** *Let X be an n-element set and $\mathcal{F}$ be a family of subsets of X such that none of the sets in $\mathcal{F}$ contain each other. Then $\#\mathcal{F} \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$. In fact, this is the maximum size of such a family.*

Proof. We look at the poset $(\mathcal{P}(X), \subseteq)$; $\mathcal{F}$ is then an antichain in this poset. By Dilworth's theorem, the maximum size of an antichain is equal to the minimum number of chains into which $\mathcal{P}(X)$ can be partitioned. We will show that $\mathcal{P}(X)$ can be partitioned into at most $\binom{n}{\lfloor \frac{n}{2} \rfloor}$ chains. This will prove the theorem. We proceed by induction on n. A symmetric chain is a chain of the form $A_a \subseteq A_{a+1} \subseteq \cdots \subseteq A_{n-a}$ where $\#A_i = i$ for $a \leq i \leq n - a$. Note that the chain is symmetric about the middle layer of the poset, so a symmetric chain must contain a set of size $\lfloor \frac{n}{2} \rfloor$. We claim that there is a symmetric chain partition of $\mathcal{P}([n])$ into $\binom{n}{\lfloor \frac{n}{2} \rfloor}$ components. The based case $n = 1$ is easy to see since the only symmetric chain in $\mathcal{P}([1])$ is $\emptyset \subseteq \{1\}$. Assume that the resul is true for $\mathcal{P}([n-1])$, with symmetric chain partition $\bigcup \mathcal{C}$ where each chain $\mathcal{C}$ is of the form

$$A_a \subseteq A_{a+1} \subseteq \cdots \subseteq A_{n-1-a}, \quad \#A_i = i, a \leq i \leq n - 1 - a. \tag{3.43}$$

For a symmetric chain $\mathcal{C}$ of the above form, define

$$L(\mathcal{C}) = A_{a+1} \subseteq \cdots \subseteq A_{n-1-a}, \quad U(\mathcal{C}) = A_a \cup \{n\} \subseteq A_{a+1} \cup \{n\} \subseteq \cdots \subseteq A_{n-1-a} \cup \{n\}. \tag{3.44}$$

Note that for two different symmetric chains $\mathcal{C}$ and $\mathcal{C}'$ in $\mathcal{P}([n-1])$, $L(\mathcal{C}) \cap L(\mathcal{C}') = \emptyset$ and $U(\mathcal{C}) \cap U(\mathcal{C}') = \emptyset$, and $L(\mathcal{C}) \cap U(\mathcal{C}) = \emptyset$ and $L(\mathcal{C}) \cap U(\mathcal{C}') = \emptyset$. Note that $A_{\lfloor \frac{n}{2} \rfloor}$ is a set of size $\lfloor \frac{n}{2} \rfloor$. Each chain has a distinct middle element $A_{\lfloor \frac{n}{2} \rfloor}$ and all sets of size $\lfloor \frac{n}{2} \rfloor$ appear in some symmetric chain in the partition (as a middle element). Thus, the number of symmetric chains in the partition is $\binom{n}{\lfloor \frac{n}{2} \rfloor}$. ■

To show the above theorem, one may also make use of the YLM inequality which states that if $\mathcal{F}$ is a family of subsets of an n-element set X such that no one subset contains another, and a_k denotes the number of k-element sets in $\mathcal{F}$ for $0 \leq k \leq n$, then

$$\sum_{k=0}^{n} \frac{a_k}{\binom{n}{k}} \leq 1. \tag{3.45}$$

This can be used as

$$\frac{1}{\binom{n}{k}} \geq \frac{1}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} \implies \frac{1}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} \sum_{k=0}^{n} a_k \leq \sum_{k=0}^{n} \frac{a_k}{\binom{n}{k}} \leq 1 \implies \#\mathcal{F} = \sum_{k=0}^{n} a_k \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}. \tag{3.46}$$

To show the YLM inequality, we 'record' certain permutations of X. Note that for some $S \in \mathcal{F}$, the number of permutations of $[n]$ such that the first $\#S$ elements of the permutation are exactly the elements of S is $\#S!(n - \#S)!$. Since no two sets in $\mathcal{F}$ contain each other, each permutation of $[n]$ can be recorded at most once. Thus, if a_k denotes the number of k-element sets in $\mathcal{F}$ for $0 \leq k \leq n$, then

$$\sum_{k=0}^{n} a_k k!(n-k)! \leq n!. \tag{3.47}$$

The inequality follows.

October 10th.

There exists a third proof and it follows from *König-Egevary theorem*. If we have a bipartite graph $G = (V_1 \sqcup V_2, E)$, the *matching number* is defined as the maximum number of pairwise disjoint edges. A *vertex cover* is a subset of the vertices such that every edge has at least one vertex in the subset. The *vertex cover number* is the minimum size of a vertex cover. The König-Egevary theorem states that in a bipartite graph, the matching number is equal to the vertex cover number. For the proof, let us label $L = V_1$, $R = V_2$.

Proof. Take a maximum matching $M \subseteq E$. A *free vertex* is a vertex in G which does not appear in any edge in M. Let U be the set of free vertices in L; note that U may be empty. If U were empty, then M would be a perfect matching and the vertex cover number would be less than or equal to $\#L$. Thus, one could get both values equal to $\#L$. Assume $U \neq \emptyset$. We look at an *alternating path*, starting from a vertex in U and alternating between edges in the matching M and edges not in M. An *augmented path* is an alternating path starting and ending at a free vertex. In such an augmented path, the number of edges in the path not in M is one more than the number of edge in the path in M. If P is the path, then $M' = M \Delta P = (M \setminus P) \cup (P \setminus M)$ is a matching larger than M, contradicting the maximality of M. Thus, there are no augmented paths in G, and an alternating path cannot end in U.

Let Z be the set of vertices in G reachable from $U \subseteq L$ via alternating paths. If we define $Z_L = Z \cap L$ and $Z_R = Z \cap R$. Letting $C = (L \setminus Z_L) \cup Z_R$, we claim that C is a vertex cover. To see this, we need to show that if $xy \in E$, then at least one of x, y is in C. Without the loss of generality, assume $x \in L$ and $y \in R$. We do cases.

- Case I: $x \notin Z_L$. Then $x \in L \setminus Z_L$, and thus $x \in C$.
- Case II: $x \in Z_L$. Then either $xy \notin M$, in which case $y \in Z_R$ and thus $y \in C$, or $xy \in M$ in which case the alternating path must pass through y to reach x, and thus $y \in Z_R$ and $y \in C$.

Thus, C is a vertex cover. We now show that $\#C = \#M$. Since the size of any vertex cover is at least the size of any matching, this equality will prove the theorem. The number of matching edges with both endpoints in Z is $\#Z_R$ since each vertex in Z_R is matched to a vertex in Z_L. The number of matching edges where one (neither) endpoint(s) is reachable from U is $\#(L \setminus Z_L)$. Thus $\#M = \#(L \setminus Z_L) + \#Z_R = \#C$. ■

Corollary 3.53. *Dilworth's theorem.*

Proof. Let $(P, \leq)$ be a finite poset. We construct a bipartite graph $G = (L \sqcup R, E)$ where both L and R are copies of P, and $x \in L$ is connected to $y \in R$ if and only if $x \lneq y$ in P. We claim that the minimum number of chains into which P can be partitioned is $\#P-$ the maximum matching in G, and that the maximum cardinality of an antichain in P is $\#P-$ the minimum vertex cover in G. This claim is left as an exercise to the reader. By König-Egevary theorem, the cardinality of the maximum matching in G is equal to the cardinality of the minimum vertex cover in G, and thus Dilworth's theorem follows. ■

Chapter 4

LATIN SQUARES & DESIGNS

4.1 Latin Squares

Definition 4.1. A *latin square* of order n is an $n \times n$ array filled with n distinct symbols, each occuring exactly once in each row and exactly once in each column.

Euler's generals problem in latin squares is as follows. Given n^2 military personnel, each with a rank and regiment of possible n and n options respectively, the problem is to arrange them in an $n \times n$ array such that any choice of a row or a column gives a group of n personnel with n distinct ranks. If we also require that any choice of a row or a column gives a group of n personnel with n distinct regiments, then it is *not* possible to do so for $n = 6$.

Definition 4.2. Two latin squares L_1, L_2 of order n, on the same set of row and column indices, are said to be *orthogonal latin squares* if the n^2 ordered pairs $(l_{ij}^{(1)}, l_{ij}^{(2)})$ for $1 \leq i, j \leq n$ are all distinct.

Euler's generals problem then becomes to find two orthogonal latin squares of order 6.

Theorem 4.3. *For integer $n \geq 1$, the maximum number $N(n)$ of pairwise mutually orthogonal latin squares of order n satisfies $N(n) \leq n - 1$.*

Proof. The key is to convert this problem into a linear algebra one; Let $L_1, \ldots, L_t$ be mutually orthogonal latin squares of order n. Let S be their common symbol set. For $s \in S$ and $1 \leq i \leq t$, define $L_{i,s}$ to be the indicator matrix of the symbol s in the latin square L_i. Then $L_{i,s}$ is a permutation matrix. If J denotes the matrix will all entries 1, then

$$A_{i,s} = L_{i,s} - \frac{1}{n}J \tag{4.1}$$

has row and column sums 0. As a small lemma, if V denotes the set of matrices with row sum and column sum 0, then $\dim(V) = (n-1)^2$. This is because we can choose the first $n-1$ rows and $n-1$ columns freely, and the last row and column are determined by the previous ones. Consider the Frobenius inner product on $M_n(\mathbb{R})$ defined by $\langle A, B\rangle = \mathrm{tr}(A^T B)$. Note that $\langle A_{i,s}, J\rangle = 0$. Taking two matrices $A_{i,s}$ and $A_{j,s'}$, we have

$$\langle A_{i,s}, A_{j,s'}\rangle = \langle L_{i,s} - \frac{1}{n}J, L_{j,s'} - \frac{1}{n}J\rangle = \langle L_{i,s}, L_{j,s'}\rangle - \frac{1}{n}\langle J, L_{j,s'}\rangle = 1 - 1 = 0 \text{ for } i \leq j,\ s, s' \in S. \tag{4.2}$$

Moreover, $\sum_{s\in S} A_{i,s} = 0$ for each i. Thus $A_{i,s} \in V$. For a fixed i, we get

$$\dim(\mathrm{span}_{s\in S}\{A_{i,s}\}) = n - 1 \tag{4.3}$$

by comparing inner products for both $i = j$ and $i \neq j$. So for $1 \leq i \leq t$, the $\text{span}_{s\in S}\{A_{i,s}\}$ form mutually orthogonal subspaces of V. Therefore,

$$\sum_{i=1}^{t} \dim(\text{span}_{s\in S}\{A_{i,s}\}) \leq \dim V = (n-1)^2 \implies t(n-1) \leq (n-1)^2. \tag{4.4}$$

We then have $\langle \sum_{s\in S} \alpha_s A_{i,s}, \sum_{s\in S} \alpha_s A_{i,s'} \rangle = n\alpha_{s'} - \sum_{s\in S} \alpha_s = 0$ which forces $\alpha_s = \alpha_{s'}$ for all $s, s' \in S$. ■

Theorem 4.4. *If q is a prime power then there exists a projective plane of order q.*

4.1.1 Projective Planes and Fields

We move our discussion on to projective planes.

Definition 4.5. A *finite projective plane* of order n is an incidence structure of points and lines satisfying

1. any two distinct points lie on exactly one line,
2. any two distinct lines meet at exactly one points,
3. there exist four points with no three collinear.

Lemma 4.6. *Each line in a finite projective plane contains the same number of points $(n+1)$. Each point is on $(n+1)$ lines. The total number of points is equal to the total number of lines n^2+n+1.*

Proof. There is a one-to-one correspondence between lines passing through p_2 and points on l other than p_1. The number of lines passing through p_2 equals the number of lines passing through p_3. There is a one-to-one correspondence between points on l_1, lines passing through p, and points on l_2. The number of points on l_1 equals the number of points on l_2, or $n+1$. The total number of points is $(n+1)^2 - n = (n+1)n + 1$. ■

Theorem 4.7. *There exists a projective plane of order n if and only if there exists a set of $n-1$ mutually orthogonal latin squares of order n.*

Proof. We prove the converse implication. Choose a line ℓ_∞ and label the points on it as $p_0, \ldots, p_n$. The lines through p_0, except ℓ_∞, are $R_1, \ldots, R_n$. The lines through p_1, except ℓ_∞, are $C_1, \ldots, C_n$. For a point $p \in P \setminus \ell_\infty$, $R_i \cap C_j = p$ uniquely determines via (i, j). If we use p_2 to determine the symbol set $\{s_1, \ldots, s_n\}$, then define latin squares as $L(P_2)(R_i, C_j)$ to be the line passing through $R_i \cap C_j$ and p_2. Note that $L(p_2)(R_i, C_j) \neq L(p_2)(R_{i'}, C_{j'})$ for $(i, j) \neq (i', j')$ since $p_2, p_{ij}, p_{i'j'}$ are not collinear. So $L(p_2)$ is a latin square. Similarly, we can define $L(p_3), \ldots, L(p_n)$. To show that they are mutually orthogonal, we show that if $(i, j) \neq (i', j')$, then

$$(L(p_2)(R_i, C_j), L(p_3)(R_i, C_j)) \neq (L(p_2)(R_{i'}, C_{j'}), L(p_3)(R_{i'}, C_{j'})). \tag{4.5}$$

Case I, where $i \neq i'$ (the case $j \neq j'$ is similar). If $L(p_2)(R_i, C_j) = L(p_2)(R_{i'}, C_{j'})$, then $p_2, R_i \cap C_j, R_{i'} \cap C_{j'}$ are collinear. So $L(p_3)(R_i, C_j) \neq L(p_3)(R_{i'}, C_{j'})$. Case II, where $i = i'$ and $j = j'$. This is trivial.

For the forward implication, let $L_1, \ldots, L_{n-1}$ be mutually orthogonal latin squares of order n on the symbol set $\{s_1, \ldots, s_n\}$. Define $P' = \{(r, c) \mid r, c \in \{1, 2, \ldots, n\}\}$ (as our set of points), and $L_{k,s} = \{(r, c) \mid L_k(r, c) = s\}$ (as our set of lines). Note that $\bigcup_{i=1}^{n} L_{k,s_i} = [n] \times [n]$ for each k. The lines L_{k,s_i} are, thus, 'parallel' in the sense that they partition P'. For $1 \leq k \leq n-1$, we have a class of parallel lines. Thus we have $n^2 - n$ lines so far. Add n more lines $R_1, \ldots, R_n$ where $R_i = \{(i, j) \mid j \in [n]\}$, and n more lines $C_1, \ldots, C_n$ where $C_j = \{(i, j) \mid i \in [n]\}$. This gives $n^2 + n$ lines. Each class of parallel lines will be associated with a point at infinity. Let $\mathcal{F}_k$ denote the class of parallel lines corresponding to L_k. Then $\mathcal{F}_k \cap \mathcal{F}_{k'} = \emptyset$ for $k \neq k'$. Thus we have $(n+1)$ classes $\mathcal{F}_1, \ldots, \mathcal{F}_{n-1}, R, C$ and $(n+1)$ points $P = P' \cup \{n+1 \text{ points}\}$. We now verify the projective plane axioms.

Note that $L_{k,s} \cap L_{k',s'} = \{(r,c)\}$ for some unique (r,c) since L_k and $L_{k'}$ are orthogonal. Also, $R_i \cap C_j = \{(i,j)\}$ for all i,j. Moreover, $L_{k,s} \cap R_i = \{(i,c)\}$ for some unique c, and $L_{k,s} \cap C_j = \{(r,j)\}$ for some unique r. Finally, $R_i \cap R_{i'} = \emptyset$ for $i \neq i'$, and $C_j \cap C_{j'} = \emptyset$ for $j \neq j'$. Thus any two distinct lines meet at exactly one point. ■

Let $\mathbb{K}$ be a finite field, with $\mathbb{Z} \hookrightarrow \mathbb{K}$. Note that since this field is finite, for all $a \in \mathbb{K}$ there exists an integer k such that $a = ka$ or $(k-1)a = 0$. The *characteristic* of $\mathbb{K}$ is the smallest $n \in \mathbb{N}$ such that $na = 0$ for all $a \in \mathbb{K}$. The characteristic is always a prime.

If $na = 0$ for $a \neq 0$, then $naa^{-1} = (a + a + \cdots + a)a^{-1} = (1 + 1 + \cdots + 1) = 0$. If we let $n = pq$, then $p(qa) = 0$ for all $a \in \mathbb{K}$ implies either $qa = 0$ for all $a \in \mathbb{K}$ or $p1 = 0$. In either case, a factor of n is a 'smaller' characteristic. This process only ends if n is prime.

Proposition 4.8. *The order of a finite field is a prime power.*

Proof. Let p be the characteristic of $\mathbb{K}$ and q be another prime factor of $\#\mathbb{K}$. Then $qa = 0$ for some $a \neq 0$. By Cauchy's theorem, $pa = 0$ for some $a \neq 0$. So $p = q$. Thus the order of $\mathbb{K}$ is p^k for some $k \in \mathbb{N}$. ■

4.2 Designs

October 23rd.

Theorem 4.9 (*Fisher's inequality*)**.** *Let $A_1, \ldots, A_m \subseteq \mathcal{P}([n])$ be distinct subsets such that $\#(A_i \cap A_j) = k$ for all $i \neq j$. Then $m \leq n$.*

Proof. The idea is to build an incidence vector for A_i as follows: define $v_{A_i} \in \mathbb{R}^n$ such that $v_{A_i}(j) = 1$ if $j \in A_i$ and 0 otherwise. The inner product $\langle v_{A_i}, v_{A_j} \rangle = \#(A_i \cap A_j) = k$ for $i \neq j$. We claim that the set $\{v_{A_1}, \ldots, v_{A_m}\}$ is linearly independent. Suppose; then there exist $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ not all zero such that $\sum_{i=1}^m \alpha_i v_{A_i} = 0$. Taking inner product with itself gives

$$\langle \sum_{i=1}^m \alpha_i v_{A_i}, \sum_{j=1}^m \alpha_j v_{A_j} \rangle = \sum_{i=1}^m \alpha_i^2 \#A_i + \sum_{1 \leq i < j \leq n} 2\alpha_i \alpha_j k = \sum_{i=1}^m \alpha_i^2 (\#A_i - k) + k(\sum_{i=1}^m \alpha_i)^2 = 0 \tag{4.6}$$

If $\#A_i \gneq k$ for all i, then $\alpha_i = 0$ for all i, a contradiction. Now suppose $\#A_1 = k$. Then there cannot exist another A_i such that $\#A_i = k$ since $A_1 \cap A_i = k$ implies $A_1 = A_i$. Thus $\#A_i > k$ for all $i \neq 1$. Then from the above equation, we have $\alpha_i = 0$ for all $i \neq 1$. So $\alpha_1 v_{A_1} = 0$ implies $\alpha_1 = 0$, a contradiction. Thus the set $\{v_{A_1}, \ldots, v_{A_m}\}$ is linearly independent, and so $m \leq n$. ■

We discuss this problem: suppose $A \in M_n(\mathbb{R})$ with entries $|a_{ij}| \leq 1$ for all i,j. What is the maximum possible value of $\det(A)$? If all the columns of A are orthogonal, that would be good. Since eigenvalues have to satisfy $Ax = \lambda x$, we have $|\lambda| \leq n$ and we have a simple bound

$$|\det(A)| = \left| \prod_{i=1}^n \lambda_i \right| \leq n^n. \tag{4.7}$$

A better bound is

$$|\det(A)| \leq \left| \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)} \right| \leq n! < n^n. \tag{4.8}$$

The Hadamard inequality for general matrices states that if the columns of A are $c_1, c_2, \ldots, c_n \in \mathbb{R}^n$, then

$$|\det(A)| \leq \|c_1\|_2 \|c_2\|_2 \cdots \|c_n\|_2 \leq n^{n/2} < n! < n^n. \tag{4.9}$$

with equality if and only if there exists a matrix H with all entries ± 1 such that $HH^t = H^t H = nI_n$. This inequality can be shown via $A^t A$ being positive definite.

Definition 4.10. A matrix $H \in M_n(\mathbb{R})$ is said to be a *Hadamard matrix* if all its entries are ± 1 and $\frac{1}{\sqrt{n}}H$ is orthogonal.

For $n = 1$, $H = \begin{pmatrix}1\end{pmatrix}$ is a Hadamard matrix. For $n = 2$, $H = \begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}$ works. For n odd, we may assume the first row is all 1s. Then the inner product of the first row with second row $(x_1, \ldots, x_n)$ gives $x_1 + x_2 + \cdots + x_n = 0$, which is impossible since n is odd. Thus no Hadamard matrix exists for odd $n > 1$. Note that if H is a Hadamard matrix, then so is the block matrix $\begin{pmatrix}H & H\\ H & -H\end{pmatrix}$ (verify). Thus if a Hadamard matrix exists for n, then one exists for $2n$. This gives us Hadamard matrices for all powers of 2. This is known as *Sylvester's construction*.

For a general Hadamard matrix, let the first row be all 1s. Let $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$ be the second and third rows respectively. Then the inner product of these two rows gives $x_1y_1 + x_2y_2 + \cdots + x_ny_n = 0$. The inner product with the first row gives $x_1 + x_2 + \cdots + x_n = 0$ and $y_1 + y_2 + \cdots + y_n = 0$. Let $I_1 = \{i \mid x_i = 1, y_i = 1\}$, $I_2 = \{i \mid x_i = 1, y_i = -1\}$, $I_3 = \{i \mid x_i = -1, y_i = 1\}$, and $I_4 = \{i \mid x_i = -1, y_i = -1\}$. We have

$$\#I_1 + \#I_2 + \#I_3 + \#I_4 = n, \tag{4.10}$$
$$\#I_1 + \#I_2 - \#I_3 - \#I_4 = 0, \tag{4.11}$$
$$\#I_1 - \#I_2 + \#I_3 - \#I_4 = 0, \tag{4.12}$$
$$\#I_1 - \#I_2 - \#I_3 + \#I_4 = 0. \tag{4.13}$$

Solving, we get $\#I_1 = \#I_2 = \#I_3 = \#I_4 = n/4$. Thus n must be divisible by 4. This gives us the *Hadamard conjecture*: for all n divisible by 4, there exists a Hadamard matrix of order n.

Let q be a prime power congruent to 3 mod 4. Recall that in the field $\mathbb{F}_q$, half of the non-zero elements are quadratic residues (or squares), and half are quadratic non-residues (or non-squares); in particular, $+1$ is a square and -1 is a non-square.

The *quadratic character* of $\mathbb{F}_q$ is the function χ defined by

$$\chi(x) = \begin{cases} 0, & \text{if } x = 0,\\ +1, & \text{if } x \text{ is a quadratic residue,}\\ -1, & \text{if } x \text{ is a quadratic non-residue.}\end{cases} \tag{4.14}$$

Now let A be the matrix whose rows and columns are indexed by elements of $\mathbb{F}_q$, and whose (x, y) entry is given by $a_{xy} = \chi(y - x)$. The matrix A is skew-symmetric, with zero diagonal and ± 1 elsewhere, and satisfies the equation $A^2 = J - qI$, where J is the all-ones matrix and I is the identity matrix. Now, if we replace the diagonal zeros of A by -1s and border A with a row and column of $+1$s, we obtain a Hadamard matrix of order $q + 1$, called a *Paley matrix*.

We now consider complex Hadamard matrices of order n. Here, the entries come from the unit circle S^1. (The definition changes to $\frac{1}{\sqrt{n}}H$ being unitary). For $n = 3$, we already have a matrix as

$$\begin{pmatrix}1 & 1 & 1\\ 1 & \omega & \omega^2\\ 1 & \omega^2 & \omega\end{pmatrix}. \tag{4.15}$$

In fact, for any n we can construct a Hadamard matrix with entries $a_{jk} = \exp(2\pi i(j-1)(k-1)/n)$ for $1 \le j, k \le n$. Let us look at the inner product of two rows R_j and $R_{j'}$ in such a matrix:

$$\langle R_j, R_{j'}\rangle = \sum_{k=1}^{n} a_{jk}\overline{a}_{j'k} = \sum_{k=1}^{n} \exp(2\pi i(j-j')(k-1)/n). \tag{4.16}$$

This summation is n for $j = j'$. For $j \neq j'$, this summation is 0. Thus, $H = (a_{jk})_{n\times n}$ is a complex Hadamard matrix. $\frac{1}{\sqrt{n}}H_n$, for a $n \times n$ complex Hadamard matrix H_n, is a unitary matrix. Applying it to a vector $v \in \mathbb{C}^n$ gives a new vector $\hat{v} = \frac{1}{\sqrt{n}}H_nv$. This $\hat{v}$ is known as the *discrete Fourier transform* of v. In this sense, complex Hadamard matrices are more analytic, and real Hadamard matrices are more combinatorial.

Definition 4.11. Let $v, k, t, \lambda \in \mathbb{Z}_{\geq 0}$ be such that $v \geq k \geq t \geq 0$ and $\lambda \geq 1$. A t-(v, k, λ) *design* consists of

1. v points forming the set $\mathcal{P}$,
2. a collection of distinct k-subsets $\mathcal{B}$ of $\mathcal{P}$ (called *blocks*),
3. such that every t-subset of $\mathcal{P}$ is contained in exactly λ blocks.

In other words, $\#\mathcal{P} = v$, $\#B = k$ for all $B \in \mathcal{B}$, and for any $T \subseteq \mathcal{P}$ with $\#T = t$, we have $\#\{B \in \mathcal{B} \mid T \subseteq B\} = \lambda$.

If $t = 2$, then we simply call it a (v, k, λ) design. For the case $\lambda = 1$, a $t - (v, k, 1)$ design is called a *Steiner system*, denoted by $S(t, k, v)$. If $(\#\mathcal{B} =)b = v$, then the design is called a *symmetric design*.

Example 4.12. Suppose q is a prime power. Then $S(2, q+1, q^2+q+1)$ is a design with q^2+q+1 points and blocks of size $q+1$. Moreover, for any two points, there is a unique block containing them. This design arises from the finite projective plane of order q. Moreover, this is a symmetric design since the number of blocks is equal to the number of points.

One can show that $b = \#\mathcal{B}$ satisfies

$$b = \frac{\lambda\binom{v}{t}}{\binom{k}{t}} \tag{4.17}$$

in a $t - (v, k, \lambda)$ design, via double counting.

October 28th.

Proof. Consider $\{(T, B) \mid T \subseteq B, B \text{ is a block}, T \text{ is a } t\text{-subset of } B\}$. The number of t-subsets of $\mathcal{P}$ is $\binom{v}{t}$. Each t-subset is contained in exactly λ blocks. Thus the total number of such pairs is $\lambda\binom{v}{t}$. On the other hand, each block B has $\binom{k}{t}$ t-subsets. Thus the total number of such pairs is also $b\binom{k}{t}$. Equating gives us the desired result. ■

Lemma 4.13. *In a $2 - (v, k, \lambda)$ design, every point in $\mathcal{P}$ is contained in exactly r blocks, where r satisfies both $r(k-1) = \lambda(v-1)$ and $bk = vr$.*

Theorem 4.14 (*Fisher's inequality for designs*)**.** *For a $2 - (v, k, \lambda)$ design with b blocks and $v > k$, we have $b \geq v$.*

Proof. Let $B_1, \ldots, B_b$ be the blocks. Consider the set $S_p = \{i \mid p \in B_i\}$ for each point $p \in \mathcal{P}$. Thus $\#(S_p \cap S_q) = \lambda$ for $p \neq q$. Note that $S_p = S_q$ is not possible since this would imply $r(k-1) = \lambda(v-1)$ fails. Thus, by Fisher's inequality, the number of such sets S_p is at most b. Since there are v points, we have $v \leq b$. ■

4.2.1 Hadamard Designs

In some sense, every Hadamard matrix of dimension $4n \times 4n$ gives rise to a symmetric $2-(4n-1, 2n-1, n-1)$ design. Let H be a Hadamard matrix of order $4n$. We may assume that, without the loss of generality, the first row and first column of H are all 1s. Replace every -1 in H with 0. Delete the first row and first column to obtain a $(4n-1) \times (4n-1)$ matrix H' with entries 0 and 1. Let $\mathcal{P}$, the points, be the set of row indices of H' and let $\mathcal{B}$, the blocks, be the set of column indices of H'. We define a block $B_j \in \mathcal{B}$ to contain a point $p_i \in \mathcal{P}$ if and only if the (i, j)-entry of H' is 1. Note that $\mathcal{P} = 4n - 1$. Each column corresponds to the rows which contain 1 for that column. Each column of H' has $2n - 1 (= k)$ 1s since each column of H has $2n$ 1s. For any choice of 2 rows, there are exactly $n - 1$ columns for which both rows have 1s since the inner product of any two rows of H is 0. Thus, we have constructed a symmetric $2 - (4n-1, 2n-1, n-1)$ design, called a *Hadamard design*.

Index