
# Enumerative Combinatorics Graph Theory

**A mathematician is a device for turning coffee into theorems.**

**– Alfréd Rényi**


**Author:** Haixiao Wang

**Institute:** Department of Mathematics, University of California San Diego


**Date:** October 22, 2024

**Info**: Based on materials of Math 154 and Math 184.

ElegantLaTeX Program

*Template derived from ElegantBook 4.5*

https://github.com/ElegantLaTeX/ElegantBook

# Contents

# Chapter 1 Combinatorics

## 1.1 Induction

### 1.1.1 Weak Induction

1. The Initial Step. Prove that the statement is true for the smallest value of m for which it is defined, usually 0 or 1.
2. The Induction Step. Prove that from the fact that the statement is true for $n$ ("the induction hypothesis"), it follows that the statement is also true for $n+1$.

### 1.1.2 Strong Induction

1. The Initial Step. Prove that the statement is true for the smallest value of m for which it is defined, usually 0 or 1.
2. The Induction Step. Prove that from the fact that the statement is true for all integers less than $n+1$ ("the induction hypothesis"), it follows that the statement is also true for $n+1$.

## 1.2 Pigeonhole Principle

**Theorem 1.1**

*Given $n$ pigeons each assigned to one of $m$ holes for some $m < n$, there must be some hole with at least two pigeons.* ♡

**Proof** Assume for sake of contradiction that each hole has at most 1 pigeon, then $n =$ number of pigeons $= \sum_{\text{hole } h}[\text{number of pigeons in } h] \leq m$, contradiction.

**Theorem 1.2 (Dirichlet's Theorem)**

*Let $\alpha$ be any real number and $q$ a positive integer. There exist integers $n$ and $m$ with $0 < m \leq q$ so that* ♡

**Proof** For every $k = 0, 1, ..., q$ we can write $k\alpha = m_k + x_k$ such that $m_k$ is an integer and $0 \leq x_k < 1$. One can divide the interval $[0, 1)$ into $q$ smaller intervals of measure $\frac{1}{q}$. Now, we have $q + 1$ numbers $x_0, x_1, ..., x_q$ and $q$ intervals. Therefore, by the pigeonhole principle, at least two of them are in the same interval. We can call those $x_i, x_j$ such that $i < j$. Now:

$$|x_j - x_i| = |(j\alpha - m_j) - (i\alpha - m_i)| = |(j-i)\alpha - (m_j - m_i)| < \frac{1}{q}$$

Let $m = j - i$ and $n = m_j - m_i$. Dividing both sides by $m$ will result in:

$$\left|\alpha - \frac{n}{m}\right| < \frac{1}{mq} \leq \frac{1}{m^2}.$$

**Theorem 1.3 (Generalized Pigeonhole)**

*Given $n$ pigeons each assigned to one of $m$ holes for some $m$ with $(k-1)m < n$, then there must be some hole with at least $k$ pigeons.* ♡

### 1.2.1 Exercises

**Exercise 1.1(Symmetric Polynomials)** Call a polynomial $P$ in the variables $x_1, x_2, \ldots, x_n$ *symmetric* if switching any of the variables leaves $P$ unchanged. So for example $x_1^2 + x_2^2 + x_3^2 - x_1x_2x_3$ is a symmetric polynomial in $x_1, x_2, x_3$ but $x_1 + 2x_2 + 3x_3$ is not. A particular example of this are the *power-sum symmetric polynomials* defined as $p_k = \sum_{i=1}^n x_i^k$. Show that any symmetric polynomial can be written as a polynomial in the power-sum symmetric polynomials. For example, if $P(x, y, z) = x_1^2 + x_2^2 + x_3^2 - x_1x_2x_3$, then $P = p_2 - p_1^3/6 + p_1p_2/2 - p_3/3$. (Hint: You will want to use induction, but not on the number of variables. Start with a polynomial $P$ and find a way to add or subtract products of the power-sum polynomials to simplify it. Repeat this until there is nothing left.)

**Proof** Let $P(x_1, \cdots, x_n)$ be the symmetric polynomial. Since whenever $P$ contains a monomial, it must also have the symmetric monomials with the same coefficient, we can see that $P$ must be a linear combination of terms of the form:

$$m_l = \sum_{\substack{i_1,\cdots,i_l \in [n] \\ i_1,\cdots,i_l \text{ are distinct}}} x_{i_1}^{\lambda_1} x_{i_2}^{\lambda_2} \cdots x_{i_l}^{\lambda_l}. \tag{1.1}$$

We claim that each such $m_l$ can be expressed as a polynomial in terms of *power-sum symmetric polynomials* $p_k$ for each $l \in \{1, 2, \cdots, n\}$. We will prove the desired argument by induction on $l$, i.e., the number of variables appearing in the relevant monomials.

1. $l = 1$, then $m_1 = \sum_{i_1 \in [n]} x_{i_1}^{\lambda_1} = p_{\lambda_1}$.
2. $l = 2$, then

$$\begin{aligned} m_2 = \sum_{\substack{i_1,i_2 \in [n] \\ i_1 \neq i_2}} x_{i_1}^{\lambda_1} x_{i_2}^{\lambda_2} &= \sum_{i_1 \in [n]} \sum_{i_2 \in [n]} x_{i_1}^{\lambda_1} x_{i_2}^{\lambda_2} - \sum_{\substack{i_1,i_2 \in [n] \\ i_1 = i_2}} x_{i_1}^{\lambda_1} x_{i_2}^{\lambda_2} && \text{(1.2a)} \\ &= (\sum_{i_1 \in [n]} x_{i_1}^{\lambda_1})(\sum_{i_2 \in [n]} x_{i_2}^{\lambda_2}) - \sum_{i_1 \in [n]} x_{i_1}^{\lambda_1+\lambda_2} && \text{(1.2b)} \\ &= p_{\lambda_1} p_{\lambda_2} - p_{\lambda_1+\lambda_2} && \text{(1.2c)} \end{aligned}$$

   Note that second term in Equation (1.2b) is a monomial with only 1 distinct variables.
3. Suppose that for $l = t$, we can express

$$m_t = \sum_{i_1,\cdots,i_t \in [n]} x_{i_1}^{\lambda_1} x_{i_2}^{\lambda_2} \cdots x_{i_t}^{\lambda_t}, \quad \text{distinct } i_1, \cdots, i_t, \tag{1.3a}$$

   by a polynomial in terms of $p_{\lambda_1}, \cdots, p_{\lambda_t}$. As for case $l = t + 1$, we expand $p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_{t+1}}$, then

$$\begin{aligned} p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_{t+1}} &= \left(\sum_{i_1 \in [n]} x_{i_1}^{\lambda_1}\right) \left(\sum_{i_2 \in [n]} x_{i_2}^{\lambda_2}\right) \cdots \left(\sum_{i_t \in [n]} x_{i_t}^{\lambda_t}\right) \left(\sum_{i_{t+1} \in [n]} x_{i_{t+1}}^{\lambda_{t+1}}\right) && \text{(1.4a)} \\ &= \sum_{i_1,\cdots,i_{t+1} \in [n]} x_{i_1}^{\lambda_1} x_{i_2}^{\lambda_2} \cdots x_{i_{t+1}}^{\lambda_{t+1}} \quad \text{distinct } i_1, \cdots, i_{t+1} && \text{(1.4b)} \\ &\quad + (\text{monomials with number of distinct variables } \leq t) && \text{(1.4c)} \end{aligned}$$

   As we can see, Equation (1.4c) should only contain terms like $m_l$ with $l \leq t$ with various constants in front of each $m_l$ to match coefficients, since some of the indices $i_1, \cdots, i_l$ may coincide with each other, thus the number of distinct variables decreases. Consequently, $m_{t+1}$, a.k.a. the term in Equation (1.4b), can be written as

$$m_{t+1} = p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_{t+1}} - [C_t m_t + C_{t-1} m_{t-1} + \cdots + C_1 m_1], \tag{1.5}$$

   where $C_t, \cdots, C_1$ are some constants to match the coefficients. By inductive assumption, second term on the right hand side of Equation (1.5) can be represented by a polynomial with at most $t$ terms from $\{p_{\lambda_1}, p_{\lambda_2}, \cdots, p_{\lambda_{t+1}}\}$, which finishes the proof.

**Proof** [Alternative proof] Let $P(x_1, \cdots, x_n)$ be the symmetric polynomial and $k$ denote the order of $P$. We claim that $P$ can be represented by combinations of $p_1, \cdots, p_k$ for all $k \in \mathbb{N}$. We prove this argument by induction on $k$.

1. If $k = 1$, then trivially, $P = \sum_{i=1}^n x_i = p_1$.
2. If $k = 2$, the terms to consider, which makes $P$ still being symmetric, are $\sum_{i=1}^n x_i^2 = p_2$ and $\sum_{i \neq j} x_i x_j = (p_1^2 - p_2)/2$.
3. If $k = 3$, by assumption, lower order terms can be written as combinations of $p_1$ and $p_2$. Then the order 3 terms to consider, which makes $P$ still being symmetric, are $\sum_{i=1}^n x_i^3 = p_3$, $\sum_{i \neq j} x_i^2 x_j = p_1 p_2$ and $\sum_{i \neq j, j \neq t, i \neq t} x_i x_j x_t = p_1^3/6 - p_1 p_2/2 + p_3/3$.
4. Suppose that the argument holds true for $k = l$. For the case $k = l+1$, we only need to consider the terms of order $l+1$ being added to $P$, e.g., $\sum_{i=1}^n x_i^{l+1} = p_{l+1}$, $\sum_{i \neq j} x_i^l x_j = p_l p_1 - p_{l+1}$, $\sum_{i \neq j} x_i^{l-1} x_j^2 = p_{l-1} p_2 - p_{l+1}$. Generally, let $\lambda^{(l)} = (\lambda_1^{(l)}, \lambda_2^{(l)}, \cdots, \lambda_n^{(l)})$ be a partition of $l$, i.e., $\lambda_1^{(l)} + \cdots + \lambda_n^{(l)} = l$ with $\lambda_i^{(l)} \geq 0$, then the monomial $m_{\lambda^{(l)}}(x)$ can be written as
$$m_{\lambda^{(l)}}(x) = \sum_{\sigma \in P([n])} x_{\sigma(i_1)}^{\lambda_1^{(l)}} x_{\sigma(i_2)}^{\lambda_2^{(l)}} \cdots x_{\sigma(i_n)}^{\lambda_n^{(l)}}, \tag{1.6}$$
where $P([n])$ denote the set of all the permutations on $[n]$. As we can see here, the monomial contains all the symmetric terms in $P$. Let $\Lambda^l = \{m_{\lambda^{(l)}} | \lambda^{(l)} \text{ is a partition of } l\}$ denote the space spanned by $m_{\lambda^{(l)}}$. Denote
$$p_{\lambda^{(l)}} = \prod_{\substack{i=1 \\ \lambda_i^{(l)} \neq 0}}^{n} p_{\lambda_i^{(l)}}. \tag{1.7}$$
We can express $p_{\lambda^{(l)}}$ in terms of $m_{\lambda^{(l)}}$ [1]
$$p_{\lambda^{(l)}} = c_{\lambda^{(l)}\lambda^{(l)}} m_{\lambda^{(l)}} + \sum_{\mu^{(l)} \triangleright \lambda^{(l)}} c_{\lambda^{(l)}\mu^{(l)}} m_{\mu^{(l)}}. \tag{1.8}$$
As we know from assumption, any $m_{\lambda^{(l)}}(x) \in \Lambda^l$ can be represented by combinations of $p_1, \cdots, p_l$, i.e., $\{p_1, \cdots, p_l\}$ is also a basis of $\Lambda^l$, then $c_{\lambda^{(l)}\lambda^{(l)}} \neq 0$. There exists a triangular matrix $C^{(l)} = [c_{\lambda^{(l)}\mu^{(l)}}]$ with non-zero diagonal entries used to express $p_{\lambda^{(l)}}$ in terms of $m_{\lambda^{(l)}}$, thus $(C^{(l)})^{-1}$ exists.
The analysis above can be generalized to case $k = l+1$, i.e., we can write
$$p_{\lambda^{(l+1)}} = c_{\lambda^{(l+1)}\lambda^{(l+1)}} m_{\lambda^{(l+1)}} + \sum_{\mu^{(l+1)} \triangleright \lambda^{(l+1)}} c_{\lambda^{(l+1)}\mu^{(l+1)}} m_{\mu^{(l+1)}}, \tag{1.9}$$
and the invertible matrix $C^{(l+1)}$ exists. It is easy to see that $C^{(l+1)}$ is triangular due to dominance order $\triangleright$, and $c_{\lambda^{(l+1)}\lambda^{(l+1)}}$ is non-zero by expanding both sides of Equation (1.9). As a result, we know that $\{p_1, \cdots, p_l, p_{l+1}\}$ is also a basis of $\Lambda^{l+1} = \{m_{\lambda^{(l+1)}} | \lambda^{(l+1)} \text{ is a partition of } l+1\}$. Therefore proved.

✍ **Exercise 1.2(Simultaneous Rational Approximation)** Dirichlet's Theorem is useful when you want to approximate one number by rationals, but what if you have two? Suppose that you have two real numbers $x$ and $y$ and want to find integers $n, k, m$ so that $|x - n/m|$ and $|y - k/m|$ are both small. Prove that for any integer $q$, one can always find $n, k, m$ with $|m| \leq q^2$ so that $|x - n/m|$ and $|y - k/m|$ are each at most $1/(mq)$.

**Proof** We consider the real numbers $lx$, $ly$, and the decomposition $lx = \lfloor lx \rfloor + a_l$, $ly = \lfloor ly \rfloor + b_l$ for $l = 0, 1, 2, \cdots, q^2$, where $a_l$, $b_l$ denote the decimal parts of $lx$, $ly$ respectively. Consider the pairs

[1] Here, $\mu^{(l)} \triangleright \lambda^{(l)}$ means $\sum_{j=1}^i \mu_i^{(l)} \geq \sum_{j=1}^i \lambda_i^{(l)}$ for all $1 \leq i \leq n$.

$(a_0, b_0), \cdots, (a_{q^2}, b_{q^2})$, which are in total $q^2+1$ pairs. One can divide the unit square $[0,1) \times [0,1)$ into $q^2$ smaller pieces of area $\frac{1}{q^2}$. Now, we have $q^2+1$ pairs with $q^2$ squares. Therefore, by the pigeonhole principle, at least two of them are in the same square. Without loss of generality, we call those pairs $(a_i, b_i)$ and $(a_j, b_j)$ with $i < j$. Now

$$|a_j - a_i| = |(jx - \lfloor jx \rfloor) - (ix - \lfloor ix \rfloor)| = |(j-i)x - (\lfloor jx \rfloor - \lfloor ix \rfloor)| < \frac{1}{q} \tag{1.10a}$$

$$|b_j - b_i| = |(jy - \lfloor jy \rfloor) - (iy - \lfloor iy \rfloor)| = |(j-i)y - (\lfloor jy \rfloor - \lfloor iy \rfloor)| < \frac{1}{q} \tag{1.10b}$$

Let $m = j - i$, $n = \lfloor jx \rfloor - \lfloor ix \rfloor$ and $k = \lfloor jy \rfloor - \lfloor iy \rfloor$. Dividing both sides by $m$ will result in

$$\left| x - \frac{n}{m} \right| < \frac{1}{mq}\,, \tag{1.11a}$$

$$\left| y - \frac{k}{m} \right| < \frac{1}{mq}\,. \tag{1.11b}$$

## 1.3 Counting

### 1.3.1 Binomial Theorem

**Definition 1.1 (Binomial Coefficient)**

*The number of ways to pick $k$ things from a set of size $n$ is defined as the binomial coefficient, given by*

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{1.12}$$

♣

**Remark** $\binom{n}{k} = \binom{n}{n-k}$, $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$

**Theorem 1.4 (Binomial Theorem)**

*For all nonnegative integers $n$,*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}\,. \tag{1.13}$$

♡

**Definition 1.2**

*Let $a$ be any real number, and let $k$ be a nonnegative integer. We define $(a)_0 = 1 = \binom{a}{0}$, and for $k \geq 0$,*

$$(a)_k = \binom{a}{k} = \frac{a(a-1)(a-2)\cdots(a-k+1)}{k!}\,. \tag{1.14}$$

♣

**Theorem 1.5 (Binomial Theorem, General Version)**

*For all nonnegative integers $n$,*

$$(1+x)^a = \sum_{k=0}^{\infty} (a)_k x^k\,. \tag{1.15}$$

♡

**Example 1.1**

$$\sqrt{1-4x} = (1-4x)^{1/2} = \sum_{k=0}^{\infty} \binom{\frac{1}{2}}{k} x^k = 1 - 2x - \sum_{n\geq 2} \frac{2^n(2n-3)!!}{n!} x^n$$
$$= 1 - 2x - \frac{2}{n} \sum_{n\geq 2} \binom{2n-2}{n-1} x^n .$$

### 1.3.2 Multinomial Theorem

Suppose we want to put things in order:

- $a_1$ things of type $1$,
- $a_2$ things of type $2$,
- $\cdots$
- $a_m$ things of type $m$,

with $a_1 + \cdots + a_m = n$. There are 2 ways of counting.

- If we give different labels to things in type $i$: $i_1, i_2, \cdots, i_{a_m}$, then all the things are different. The total number of ordering is $n!$.
- Given the number of orderings of those $m$ different types things, there are $a_i!$ ways to add labels to the type $i$ objects, thus the total number of ordering is $\#\text{ordering} \cdot a_1! \cdots a_m!$

**Definition 1.3 (Multinomial Coefficient)**

*The number of orderings of $m$ different types things with each size $a_1, a_2, \cdots, a_m$ is*

$$\binom{n}{a_1, a_2, \cdots, a_m} = \frac{n!}{a_1! a_2! \cdots a_m!} \tag{1.16}$$

♣

**Theorem 1.6 (Multinomial Theorem)**

*For all nonnegative integers $n$,*

$$(x_1 + x_2 + \cdots + x_m)^n = \sum_{a_1, a_2, \cdots, a_m} \binom{n}{a_1, a_2, \cdots, a_m} x_1^{a_1} x_2^{a_2} \cdots x_m^{a_m}, \tag{1.17}$$

*where the sum is taken over all $k$-tuples of non-negative integers $a_1, a_2, \cdots, a_m$ such that $n = \sum_{i=1}^m a_i$.* ♡

**Theorem 1.7**

*For all nonnegative integers $n$ and $a_1, a_2, \cdots, a_m$ such that $n = \sum_{i=1}^m a_i$, the equality*

$$\binom{n}{a_1, a_2, \cdots, a_m} = \binom{n}{a_1} \binom{n-a_1}{a_2} \cdots \binom{n - a_1 - \cdots - a_i}{a_{i+1}} \cdots \binom{n - a_1 - \cdots - a_{k-1}}{a_k}$$

♡

**Proof** The left-hand side counts all linear orderings of a multiset that consists of $a_i$ copies of the symbol $x_i$, for all $i \in [k]$. We show that the right-hand side counts the same objects. Indeed, let us first choose the $a_1$ positions we place all our symbols $x_1$. This can be done in $\binom{n}{a_1}$ ways. Let us now choose the $a_2$ positions where we place our symbols $x_2$. As a1 positions are already taken, this can be done in $\binom{n-a_1}{a_2}$ ways. Then we can choose the $a_3$ positions where we place our symbols $x_3$. As $a_1 + a_2$ positions are already taken, this can be done in $\binom{n-a_1-a_2}{a_3}$ ways. Iterating this procedure, we will choose the positions of all symbols, and we see that the total number of possible outcomes is indeed the right-hand side.

### 1.3.3 Exercises

**Exercise 1.3(Counting Matchings)** Let $[12]$ denote the set $\{1, 2, 3, \ldots, 12\}$. A *matching* of $[12]$ is a way of partitioning the elements into pairs so that each element is in exactly one pair. For example, one matching is $\{1, 3\}, \{2, 7\}, \{4, 10\}, \{5, 6\}, \{8, 11\}, \{9, 12\}$. For each of the following count the number of matchings with the given property both as a formula and by giving the exact number. Remember to justify your answer.

(a) The number of all matchings of $[12]$. [5 points]

(b) The number of matchings of $[12]$ where each even number is paired with another even number. [5 points]

(c) The number of matchings of $[12]$ where each even number is paired with an odd number. [5 points]

(d) The number of matchings of $[12]$ where each number is paired with another number at most 2 away from it (for this you will want to relate the number of such pairings of $[2n]$ to the number of such pairings of $[2(n-1)]$ and of $[2(n-2)]$ and produce a recurrence). [5 points]

(e) The number of matchings of $[12]$ where each of $1, 2, 3$ is paired to one of $1, 2, 3, 4, 5, 6, 7, 8, 9$. [5 points]

(f) The number of matchings of $[12]$ where there are exactly 2 pairs of even numbers that are matched together. [5 points]

**Proof**

(a) We start pairing from 1, and there are 11 options, then the first pair is removed and 10 numbers left. We then find the matching for smallest remaining element, and there are 9 options, then the second pair is removed and 8 elements left. By repeating this process, the number of choices for next iterations should be $7, 5, 3, 1$. Therefore, the total number of matchings is $11 * 9 * 7 * 5 * 3 * 1 = 10395$.

(b) We choose 3 pairs from odd and even separately. Following the same strategy in part (a), the number of matchings from 6 distinct elements should be $5 * 3 * 1 = 15$. The pairings for even numbers and odd numbers are independent from each other, then the number of such type matchings of $[12]$ is $15*15 = 225$.

(c) We consider the matching of each even number from 2 to 12 sequentially. There are 6 options for 2, then the first pair is removed. For the second round, there are 5 options for 4. By repeating this process, the number of choices for next iterations should be $4, 3, 2, 1$. Therefore, the total number of matchings is $6! = 720$.

(d) Let $k_n$ denote the number of such pairings in $[2n]$. We want to find the relationship between $k_n$, $k_{n-1}$ and $k_{n-2}$ and then establish the recurrence equation. We consider the matching of the last number $2n$. It can be only matched to $2n-1$ or $2n-2$, since it can only be matched at most 2 away from itself. That's why $k_{n-3}$ is not necessary to our consideration. In first case where $2n$ is matched to $2n-1$, $[2n]$ can be treated as $\{2n-1, 2n\}$ being added to $[2(n-1)]$ and there are $k_{n-1}$ options. In the second case where $2n$ is matched to $2n-2$ and $\{2n-3, 2n-1\}$ is a pair automatically, $[2n]$ can be viewed as $\{2n-3, 2n-2, 2n-1, 2n\}$ being added to $[2(n-2)]$. As a result, $k_{n-2}$ options are available there. Consequently, we have $k_n = k_{n-1} + k_{n-2}$ as shown in Figure 1.1, which is a Fibonacci Sequence. Obviously, we have $k_1 = 1$ and $k_2 = 2$ for the base cases by simple counting. According to the recurrence relation above, we have $k_3 = 3$, $k_4 = 5$, $k_5 = 8$, and finally $k_6 = 13$ which is of our interest.

For readers who are interested in the result for any $n \in \mathbb{N}_+$, we claim that the formula for $k_n$ is

$$k_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1}\right). \tag{1.18}$$

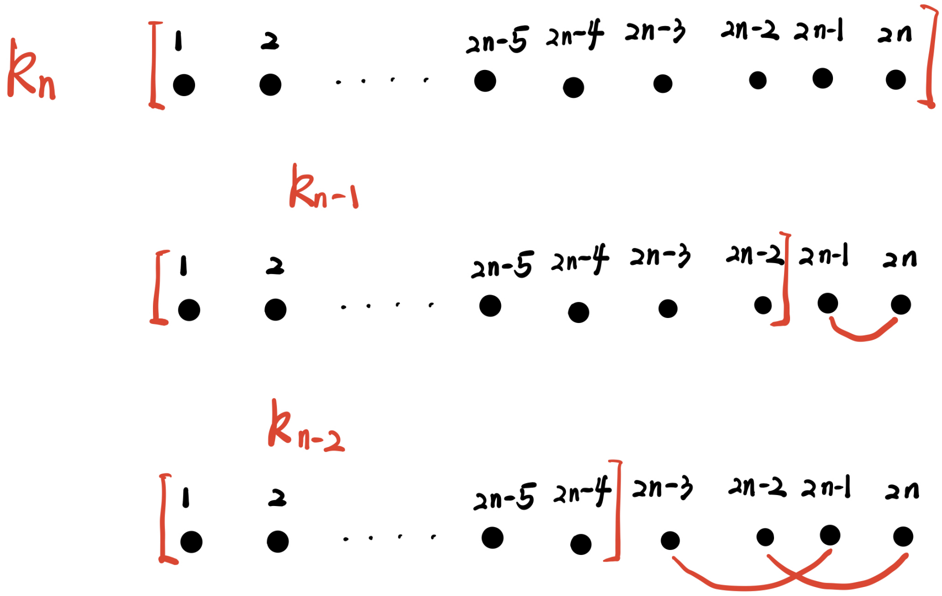


**Figure 1.1:** Illustration of $k_n = k_{n-1} + k_{n-2}$.

Let $a = \frac{1+\sqrt{5}}{2}$ and $b = \frac{1-\sqrt{5}}{2}$, then $k_n = \frac{1}{\sqrt{5}}(a^{n+1} - b^{n+1})$. Note that

$$a^2 = \left(\frac{1+\sqrt{5}}{2}\right)^2 = \frac{1+2\sqrt{5}+5}{4} = \frac{3+\sqrt{5}}{2} = a+1\,, \tag{1.19a}$$

$$b^2 = \left(\frac{1-\sqrt{5}}{2}\right)^2 = \frac{1-2\sqrt{5}+5}{4} = \frac{3-\sqrt{5}}{2} = b+1\,. \tag{1.19b}$$

We now start our induction.

(i) $n = 1$, obviously, $k_1 = \frac{1}{\sqrt{5}}(a^2 - b^2) = \frac{1}{\sqrt{5}}[(a+1)-(b+1)] = 1$.

(ii) $n = 2$,

$$k_2 = \frac{1}{\sqrt{5}}(a^3 - b^3) = \frac{1}{\sqrt{5}}[(a+1)a - (b+1)b] = \frac{1}{\sqrt{5}}[(a+b+1)(a-b)] = 2\,. \tag{1.20}$$

(iii) For $n = l$, we write

$$k_l = \frac{1}{\sqrt{5}}(a^{l+1} - b^{l+1}) = \frac{1}{\sqrt{5}}(a^2a^{l-1} - b^2b^{l-1}) \tag{1.21a}$$

$$= \frac{1}{\sqrt{5}}[(a+1)a^{l-1} - (b+1)b^{l-1}] = \frac{1}{\sqrt{5}}[(a^l - b^l) + (a^{l-1} - b^{l-1})] \tag{1.21b}$$

$$= k_{l-1} + k_{l-2}\,. \tag{1.21c}$$

The proof is finished here since the recurrence formula is finally obtained.

(e) There are 2 cases. First, 2 numbers from $\{1,2,3\}$ are matched together ($\binom{3}{2}$ choices), then the remaining one is matched to one of $\{4,5,6,7,8,9\}$(6 choices), and then we choose 4 pairs from the remaining 8 elements(similar to part (b), $7*5*3*1 = 105$ choices). Thus the number of this case is $\binom{3}{2}\cdot 6\cdot 105 = 1890$. On the contrary, we first match $1,2,3$ sequentially to numbers from $\{4,5,6,7,8,9\}(6,5,4$ choices each round), then choose 3 pairs from the remaining 6 numbers(by part (b), $5*3*1 = 15$ choices), thus the number of this case is $6\cdot 5\cdot 4\cdot 15 = 1800$. Therefore, the total number of such pairing is $1890 + 1800 = 3690$.

(f) We first divide even numbers to 3 pairs, by part (b), there are $5*3*1 = 15$ choices. We then choose 1 from the 3 pairs, and match each even number in this pair to odd numbers, then there are $6,5$ choices each round. Finally, we divide the remaining $4$ odd numbers to $2$ groups, and there are $3 * 1$ choices. In total, the number of such pairings should be $15 * 3 * 6 * 5 * 3 = 4050$.

## 1.4 Divide and Conquer

We will distribute $n$ objects into $k$ boxes, and ask in how many ways this can be done.

### 1.4.1 Compositions

Identical objects to distinct boxes.

**Definition 1.4 (Weak composition and composition)**

*A sequence $(a_1, a_2, \cdots, a_k)$ of integers, fulfilling $a_1 + a_2 + \cdots + a_k = n$ and $a_i \geq 0$ for all $i$, is called a **weak composition** of $n$. If in addition, $a_i > 0$ for all $i \in [k]$, then $(a_1, a_2, \cdots, a_k)$ is called a **composition** of $n$.* ♣

If we put $n$ balls into $k$ boxes, letting $a_i$ be the number of balls in the $i$th box, we get a (weak) composition of $n$ into $k$ parts. Weak composition allows empty boxes, however composition does not.

**Theorem 1.8**

*For all positive integers $n$ and $k$, the number of weak compositions of $n$ into $k$ parts (distribute $n$ objects to $k$ boxes while allowing empty boxes) is*

$$\binom{n+k-1}{k-1} = \binom{n+k-1}{n}, \tag{1.22}$$

*which is the same as the number of ways distributing $k$ objects to $n$ boxes while allowing empty boxes.* ♡

**Proof** We draw $n$ stars $*$ in a line, and we want to insert $k-1$ vertical lines $|$ to separate things. Weak composition allows empty boxes, thus the number of available positions increases by 1 every time we finished inserting $|$.

**Corollary 1.1**

*For all positive integers $n$ and $k$, the number of compositions of $n$ into $k$ parts is $\binom{n-1}{k-1}$.* ♡

**Proof** Composition doesn't allow empty boxes, thus the number of available positions decreases by 1 every time we finished inserting $|$.

**Corollary 1.2**

*For all positive integers $n$, the number of all compositions of $n$ is $2^{n-1}$.* ♡

### 1.4.2 Set Partitions

Distinct objects to identical boxes.

**Definition 1.5**

*A partition of the set $[n] := \{1, \cdots, n\}$ is a collection of non-empty blocks so that each element of $[n]$ belongs to exactly one of these blocks.* ♣

**Remark** The ways to put $n$ labeled balls into $k$ nonempty unlabeled bins correspond exactly to the partitions of $[n]$ into $k$ subsets.

**Definition 1.6 (Stirling numbers of the second)**

*The number of partitions of $[n]$ into $k$ non-empty blocks is denoted by $S(n,k)$.* ♣

**Theorem 1.9**

*For all positive integers $k \le n$,*

$$S(n,k) = S(n-1,k-1) + k \cdot S(n-1,k)\,. \tag{1.23}$$

♡

**Proof** Consider the assignment of last element $n$. It is either a single block, where $S(n-1,k-1)$ ways available to arrange $[n-1]$ into $k-1$ boxes, or contained in one of the $k$ block with other elements, where $k \cdot S(n-1,k)$ ways available to distribute $n$.

**Definition 1.7 (Bell number)**

*The number of all set partitions of $[n]$ into non-empty parts is denoted by $B(n)$, called $n$th **Bell number**, which by definition should satisfy*

$$B(n) = \sum_{i=0}^{n} S(n,i)\,. \tag{1.24}$$

*We set $B(0) = 1$.* ♣

### 1.4.3 Integer Partition

**Definition 1.8**

*Let $a_1 \ge a_2 \ge \cdots \ge a)k \ge 1$ be integers so that $a_1 + a_2 + \cdots + a_k = n$. Then the sequence $(a_1, a_2, \cdots, a_k)$ is called a **partition** of the integer $n$.*

1. *The number of partitions of $n$ into exactly $k$ parts is denoted by $p_k(n)$.*
2. *The number of all partitions of $n$ is denoted by $p(n) = \sum_{k=1}^{n} p_k(n)$.*

♣

**Example 1.2** The positive integer 5 has 7 partitions. Indeed, they are $(5)$; $(4,1)$; $(3,2)$; $(3,1,1)$; $(2,2,1)$; $(2,1,1,1)$; $(1,1,1,1,1)$. Therefore, $p(5) = 7$.

**Definition 1.9 (Ferrers Diagram)**

*A **Ferrers shape** of a partition $p = (a_1, a_2, \cdots, a_k)$ is a set of $n$ square boxes with horizontal and vertical sides so that in the $i$th row we have $a_i$ boxes and all rows start at the same vertical line. It is named after the British mathematician Norman Macleod Ferrers.* ♣

**Definition 1.10 (Conjugate partition of $p$.)**

***Conjugate** partition of $p = (a_1, a_2, \cdots, a_k)$ is obtained by reflecting a Ferrers shape of $p$ with respect to its main diagonal*

*A partition of $n$ is called **self-conjugate** if it is equal to its conjugate.* ♣

**Theorem 1.10 (Theorem 5.17, [Miklos2017AWalkTC])**

*The number of partitions of $n$ into at most $k$ parts is equal to that of partitions of $n$ into parts not larger than $k$.* ♡

**Proof** The first number is equal to that of Ferrers shapes of size $n$ with at most $k$ rows. The second number is equal to that of Ferrers shapes with at most $k$ columns. Finally, these two sets of Ferrers shapes are equinumerous as one can see by taking conjugates.

**Theorem 1.11**

*The number of partitions of $n$ into distinct odd parts is equal to that of all self-conjugate partitions of $n$.* ♡

**Theorem 1.12 (Theorem 5.18 in [Miklos2017AWalkTC])**

*The number of partitions of $n$ into distinct odd parts is equal to that of all self-conjugate partitions of $n$.* ♡

**Theorem 1.13**

*Let $q(n)$ be the number of partitions of $n$ in which each part is at least two. Then $q(n) = p(n) - p(n-1)$, for all positive integers $n \geq 2$.* ♡

**Theorem 1.14**

*Let $a = (a_1, a_2, \cdots, a_k)$ be a partition of the integer $n$, and let $m_i$ be the multiplicity of $i$ as a part of $a$, that is, $m_i$ denotes the number of blocks of length $i$. Then the number of set partitions of $[n]$ that are of type $a$ is equal to*

$$P_a = \frac{\binom{n}{a_1, a_2, \cdots, a_k}}{(m_1!) \cdot (m_2!) \cdots (m_n!)} \tag{1.25}$$

♡

### 1.4.4 Summary

#### 1.4.4.1 Allowing empty boxes

| | parameters | formula |
|---|---|---|
| Functions | $n$ distinct objects<br>$k$ distinct boxes | $k^n$ |
| Weak Compositions | $n$ identical objects<br>$k$ distinct boxes | $\binom{n+k-1}{k-1}$ |
| Set Partitions | $n$ distinct objects<br>$k$ identical boxes | $\sum_{i=1}^{k} S(n, i)$ |
| Integer Partitions | $n$ identical objects<br>$k$ identical boxes | $\sum_{i=1}^{k} p_i(n)$ |

#### 1.4.4.2 No empty boxes

| | parameters | formula |
|---|---|---|
| Surjections | $n$ distinct objects<br>$k$ distinct boxes | $S(n,k)k!$ |
| Surjections | $n$ distinct objects<br>any number of distinct boxes | $\sum_{i=1}^{n} S(n,i)i!$ |
| Compositions | $n$ identical objects<br>$k$ distinct boxes | $\binom{n-1}{k-1}$ |
| Compositions | $n$ identical objects<br>any number of distinct boxes | $\sum_{i=1}^{n} \binom{n-1}{i-1} =$ $2^{n-1}$ |
| Set Partitions | $n$ distinct objects<br>$k$ identical boxes | $S(n,k)$ |
| Set Partitions | $n$ distinct objects<br>any number of identical boxes | $\sum_{i=0}^{n} S(n,i) =$ $B(n)$ |
| Integer Partitions | $n$ identical objects<br>$k$ identical boxes | $p_k(n)$ |
| Integer Partitions | $n$ identical objects<br>any number of identical boxes | $\sum_{i=1}^{n} p_i(n) = p(n)$ |

### 1.4.5 Exercises

✍ **Exercise 1.4(Composition Bijections)**

(a) Give a bijection between the set of compositions of $n$ into parts of size 1 and 2 and the set of compositions of $n+2$ into parts of size at least 2. [15 points]

(b) Give a bijection between the set of weak compositions of $n$ into $k+1$ parts and the set of weak compositions of $k$ into $n+1$ parts. Hint: use stars and bars. [15 points]

**Proof** (a): Let $X$ denote the set of composition of $n$ into parts of size 1 or 2, and $Y$ denote the set of composition of $n+2$ into parts of size at least 2. We first define a function $f : X \to Y$ as follows: Given an element $x \in X$, we can write it as

$$(a_1, a_2, \ldots, a_k) \text{ with } \sum_{i=1}^{k} a_i = n \text{ and } a_i \in \{1,2\}, \forall i.$$

We append an auxiliary entry $a_0 := 2$ to the front of $x$, so we obtain $x' = (a_0, a_1, \ldots, a_k)$. The basic idea of our construction is to replace the sequence $2+1+1+1+1+\ldots+1$ in the composition into parts of size 1 and 2 with $k+2$ (where there were $k$ 1's in the first sum).

More specifically, we consider the indices $0 = i_1 < i_2 < \cdots < i_l \leq k$ that satisfies

$$a_{i_1} = a_{i_2} = \cdots = a_{i_l} = 2 \text{ and } a_i = 1 \text{ for any } i \notin \{i_1, \ldots, i_l\}.$$

In words, $a_{i_j}$ are those entries of $x'$ that equals 2. For the sake of presentation, we define $i_{l+1} = k+1$. We now define, for all $j = 1, 2, \ldots, l$,

$$b_j = \sum_{i=i_j}^{i_{j+1}-1} a_i.$$

Clearly, we have

$$\sum_{j=1}^{l} b_j = \sum_{j=1}^{l} \sum_{i=i_j}^{i_{j+1}-1} a_i = \sum_{i=0}^{k} a_i = a_0 + \sum_{i=1}^{k} a_i = n+2.$$

This means $y = (b_1, b_2, \ldots, b_l)$ is a composition of $n+2$. Also, we know that $b_j \geq a_{i_j} = 2$, so $y \in Y$. We define our function $f$ by taking $f(x) = y$.

Now we show that $f$ is a bijection, by constructing its inverse $g : Y \to X$ as follows: Given an element $y \in Y$, we can write it as

$$(b_1, b_2, \ldots, b_l) \text{ with } \sum_{j=1}^{l} b_j = n+2 \text{ and } b_j \geq 2, \forall j.$$

For each $j = 1, \ldots, l$ we define a small composition $x_j$ for $b_j$ by setting

$$x_j = (2, 1, \ldots, 1) \text{ where } 1 \text{ repeats } b_j - 2 \text{ times.}$$

Now we take $x' = (a_0, a_1, \ldots, a_k)$ to be the sequence obtained by connecting $x_1, x_2, \ldots, x_l$ altogether in a row. Clearly, $a_0 = 2$ as it's the first entry of $x_1$ and each entry of $x'$ is either 1 or 2. We also have

$$\sum_{i=1}^{k} a_i = \sum_{j=1}^{l} b_j = n+2.$$

So if we take $x$ to be $(a_1, \ldots, a_k)$, $x$ is a composition of $n+2-2 = n$. This means $x \in X$ and we define $g$ by taking $g(y) = x$.

Finally, based on how we constructed $f$ and $g$, if $f(x) = y$, we must have $g(y) = x$ and vice versa. We conclude that $f$ is a bijection.

(b): Let $X$ be the set of weak compositions of $n$ into $k+1$ parts and $X'$ be the set of sequences of $n$ stars and $k$ bars. Similarly, let $Y$ be the set of composition of $k$ into $n+1$ parts and $Y'$ be the set of sequences of $k$ stars and $n$ bars. We learned in class (or see the proof of Theorem 5.2 in textbook, where the author used balls versus walls instead of stars versus bars) that there's a bijection between $X$ and $X'$ and there's a bijection between $Y$ and $Y'$. To show there's a bijection between $X$ and $Y$, it suffices for us to find a bijection $f : X' \to Y'$.

We can define the $f$ as follows: Given a sequences, denoted by $x$, of $n$ stars and $k$ bars, we consider a new sequence $y$ of $k$ stars and $n$ bars, where each star of $x$ is changed to a bar and each bar of $x$ is changed to a star. We can define $f$ by letting $f(x) = y$. This function $f$ is a bijection since we can construct its inverse $g : Y' \to X'$ similarly: Given a sequence $y$ of $k$ stars and $n$ bars, we switch the role of stars and bars and define the new sequence $x$ as the image of $y$ under $g$.

✍ **Exercise 1.5(Sterling Number Inequalities)** Prove that for all $n \geq k > 0$ that

$$k^{n-k} \leq S(n,k) \leq k^n / k!$$

Hint: Relate $S(n,k)$ to the number of functions from $\{1, 2, \ldots, n\}$ to $\{1, 2, \ldots, k\}$.

**Proof** Let $N$ be the number of surjective functions from $\{1, 2, \ldots, n\}$ to $\{1, 2, \ldots, k\}$. We have the identity

$$N = k! \cdot S_{n,k}.$$

We give a justification to this identity. (It's also proved in textbook as Corollary 5.9.) Consider the process to construct an arbitrary surjective function from $\{1, 2, \ldots, n\}$ to $\{1, 2, \ldots, k\}$. First we partition $\{1, 2, \ldots, n\}$ into $k$ parts $P_1, P_2, \ldots, P_k$. There are $S(n,k)$ ways to do this; Then we take a permutation $\sigma$ of $\{1, 2, \ldots, k\}$. There are $k!$ ways to do this; Finally, we define our surjective function to map every element in each $P_i$ to $\sigma(i)$. There's only one way to do this. The claimed identity follows from the principle of multiplication.

There are only $k^n$ functions from $\{1,2,\ldots,n\}$ to $\{1,2,\ldots,k\}$ in total, because for each element in $\{1,2,\ldots,n\}$, there are $k$ options for its image. So $N \leq k^n$ and by the previous identity, we have $S_{n,k} \leq k^n/k!$ as wanted.

There are at least $k!k^{n-k}$ surjective functions from $\{1,2,\ldots,n\}$ to $\{1,2,\ldots,k\}$, because we can explicitly construct this many according to the following procedure: First, we choose a permutation $\sigma$ of $\{1,2,\ldots,k\}$. There are $k!$ ways to do this; Then, for each $i \in \{k+1,k+2,\ldots,n\}$, we choose a number $a_i \in \{1,2,\ldots,k\}$. There are $k^{n-k}$ ways to do this; Finally, we define our surjective function by mapping each $i \in \{1,2,\ldots,k\}$ to $\sigma(i)$ and mapping each $i \in \{k+1,k+2,\ldots,n\}$ to $a_i$. Note that such a function is surjective as each element in $\{1,2,\ldots,k\}$ is an image of $\sigma$. Therefore, we have $N \geq k!k^{n-k}$ and by the identity in first paragraph, we have $S(n,k) \geq k^{n-k}$ as wanted.

**Exercise 1.6(Partition Identity)** Prove that:

$$p(n) = \sum_{k=1}^{\lfloor\sqrt{n}\rfloor} \sum_{m=0}^{n-k^2} p_{\leq k}(m)p_{\leq k}(n-k^2-m).$$

Where $p_{\leq k}(n)$ denotes the number of partitions of $n$ with at most $k$ parts. Hint: Count the number of partitions with a $k \times k$ box in the upper left of the Ferrers diagram.

**Proof** Let $F(n)$ be the set of Ferrers diagrams of size $n$ and $F_{\leq k}(n)$ be the set of Ferrers diagrams of size $n$ and at most $k$ rows. From the lecture we know that $p(n) = |F(n)|$ and $p_{\leq k}(n) = |F_{\leq k}(n)|$.

We define $F(n,k)$ to be the set of Ferrers diagrams of size $n$ such that the largest square grid of boxes in the upper left corner is of size $k \times k$. See Figure 1.2 for an example. Since every Ferrers diagram has its upper left corner in some square grid and such a square grid is unique given it's the largest, we have $F(n) = \bigcup_{i=1}^{\infty} F(n,k)$ and $F(n,k_1) \cap F(n,k_2) = \emptyset$ whenever $k_1 \neq k_2$.

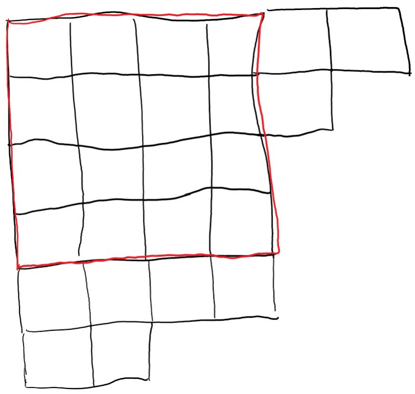


**Figure 1.2:** The diagram belongs to $F(25,4)$, the largest square grid of boxes in the upper left corner is marked red.

If $k > \lfloor\sqrt{n}\rfloor$, a $k \times k$ grid of boxes contains at least $(\lfloor\sqrt{n}\rfloor + 1)^2 > n$ small boxes, so a Ferrers diagram of size $n$ cannot contain a $k \times k$ grid, which implies $F(n,k) = \emptyset$ as well. So we have $F(n) = \bigcup_{k=1}^{\lfloor\sqrt{n}\rfloor} F(n,k)$ and this implies

$$|F(n)| = \sum_{k=1}^{\lfloor\sqrt{n}\rfloor} |F(n,k)|.$$

For each $1 \leq k \leq \lfloor\sqrt{n}\rfloor$, we define $F(n,k,m)$ to be the set of Ferrers diagrams that's in $F(n,k)$ and have $m$ small squares below the largest square grid in the upper left corner. See Figure 1.3 for examples. We have

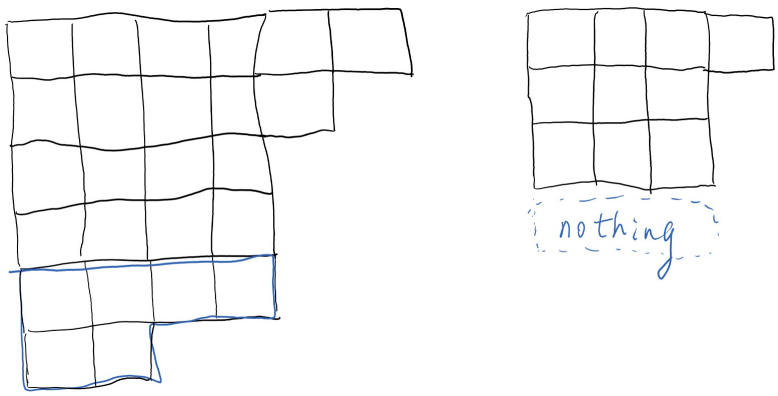



**Figure 1.3:** The left diagram belongs to $F(25, 4, 6)$. The right diagram belongs to $F(10, 3, 0)$. The part below the largest square grid in the upper left corner is marked blue.

$F(n, k) = \bigcup_{m=0}^{\infty} F(n, k, m)$ and $F(n, k, m_1) \cap F(n, k, m_2) = \emptyset$ whenever $m_1 \neq m_2$. Notice that any Ferrers diagram in $F(n, k, m)$ contains at least $k^2 + m$ small squares, so $F(n, k, m) = \emptyset$ whenever $m > n - k^2$, and this implies

$$|F(n, k)| = \sum_{m=0}^{n-k^2} |F(n, k, m)|.$$

we now argue that $|F(n, k, m)| = |F_{\leq k}(m)||F_{\leq k}(n - k^2 - m)|$. Consider the following procedure to construct an arbitrary element in $F(n, k, m)$: First, choose a Ferrers diagram $F_1$ of size $m$ that has at most $k$ columns, by conjugation, it's the same as choosing a Ferrers diagram of size $m$ with at most $k$ rows. So there are $|F_{\leq k}(m)|$ options for this; Then choose a Ferrers diagram $F_2$ of size $n - k^2 - m$ with at most $k$ rows. There are $|F_{\leq k}(n - k^2 - m)|$ options for this; Finally, take a $k \times k$ grid of small square boxes, put $F_1$ below the grid and align their leftmost columns, and put $F_2$ at the right of the grid and align their upmost rows. There's only one way to do this. In this way, we can construct every element in $F(n, k, m)$ exactly once. So, by principle of multiplication, we have the claim identity.

Finally, we combine all the identities we have so far.

$$\begin{aligned}
p(n) = |F(n)| &= \sum_{k=1}^{\lfloor\sqrt{n}\rfloor} |F(n, k)| \\
&= \sum_{k=1}^{\lfloor\sqrt{n}\rfloor} \sum_{m=0}^{n-k^2} |F(n, k, m)| \\
&= \sum_{k=1}^{\lfloor\sqrt{n}\rfloor} \sum_{m=0}^{n-k^2} |F_{\leq k}(m)||F_{\leq k}(n - k^2 - m)| \\
&= \sum_{k=1}^{\lfloor\sqrt{n}\rfloor} \sum_{m=0}^{n-k^2} |p_{\leq k}(m)||p_{\leq k}(n - k^2 - m)|,
\end{aligned}$$

which is exactly what we wish to prove.

## 1.5 Cycles and Permutations

### 1.5.1 Cycles in Permutations

**Lemma 1.1**

*Let $p : [n] \mapsto [n]$ be a permutation, and let $x \in [n]$. Then there exists a positive integer $1 \leq i \leq n$ so that $p^i(x) = x$.* ♡

**Proof** Consider the entries $p(x), p^n(x), \cdots, p^n(x)$. If none of them is equal to $x$, then the Pigeon-hole Principle implies that there are two of them that are equal, say $p^j(x) = p^k(x)$, with $j < k$. Then, applying $p^{-1}$ to both sides of this equation, we get $p^{j-1}(x) = p^{k-1}(x)$. Repeating this step, we get $p^{j-2}(x) = p^{k-2}(x)$, and repeating this step $j-3$ more times, we get $p(x) = (p(x))^{k-j+1}$, and the desired thing shows up.

**Definition 1.11 (Permutation)**

*Let $p : [n] \mapsto [n]$ be a permutation. Let $x \in [n]$, and let $i$ be the smallest positive integer so that $p^i(x) = x$. Then we say that the entries $x, p(x), p^2(x), \cdots, p^{i-1}(x)$ form an $i$-cycle in $p$.* ♣

**Corollary 1.3**

*All permutations can be decomposed into the disjoint unions of their cycles.* ♡

**Remark** **Canonical cycle form**: That is, each cycle will be written with its largest element first, and the cycles will be written in increasing order of their first elements.

**Theorem 1.15**

*Let $a_1, a_2, \cdots, a_n$ be nonnegative integers so that the equality $\sum_{i=1}^{n} i \cdot a_i = n$ holds. Then the number of $n$-permutations with $a_i$ cycles of length $i$ where $i \in [n]$, is*

$$\frac{n!}{a_1!a_2!\cdots a_n! \cdot 1^{a_1}2^{a_2}\cdots n^{a_n}} \tag{1.26}$$

♡

**Definition 1.12 (Stirling number of the first kind)**

*The number of $n$-permutations with $k$ cycles is called a* ***signless Stirling number of the first kind****, and is denoted by $c(n,k)$. The number $s(n,k) = (-1)^{n-k}c(n,k)$ is called a* ***Stirling number of the first kind****.* ♣

**Theorem 1.16**

*Let $n$ and $k$ be positive integers satisfying $n \geq k$. Then*

$$c(n,k) = c(n-1,k-1) + (n-1)c(n-1,k). \tag{1.27}$$

♡

**Lemma 1.2**

*Let $n$ be a fixed positive integer. Then*

$$\sum_{k=0}^{n} c(n,k)x^k = x(x+1)\cdots(x+n-1)\,. \tag{1.28}$$

*and replacing $x$ by $-x$,*

$$\sum_{k=0}^{n} s(n,k)x^k = (x)_n = x(x-1)\cdots(x-n+1)\,. \tag{1.29}$$

♡

**Remark** For the Stirling number of the second kind, we have

$$x^n = \sum_{k=0}^{n} S(n,k)(x)_k \,. \tag{1.30}$$

It is well-known that the set of all polynomials with real coefficients is a vector space $V$ over the field of real numbers. The most obvious basis of $V$ is $B = \{1, x, x^2, x^3, \cdots\}$, but it is not the only interesting basis. It is easy to show that $B' = \{1, (x)_1, (x)_2, (x)_3, \cdots\}$ is also a basis of $V$.

Now let $S$ (resp. $s$) be the infinite matrix whose entry in position $(n,k)$ is $S(n,k)$ (resp. $s(n,k)$). Then formulas above shows that $s$ is the transition matrix from $B$ to $B'$, while $S$ is the transition matrix from $B'$ to $B$. This proves the promised connection between the two different kinds of Stirling numbers.

### 1.5.2 Permutations with restricted Cycle Structure

**Lemma 1.3 (Transition Lemma)**

*Let $p : [n] \mapsto [n]$ be a permutation written in canonical cycle notation. Let $g(p)$ be the permutation obtained from $p$ by removing the parentheses and reading the entries as a permutation in the one-line notation. Then $g$ is a bijection from the set $S_n$ of all permutations on $[n]$ onto $S_n$.* ♡

**Example 1.3** Let $p = (412)(53)$, then $g(p) = g((412)(53)) = 41253$.

**Proof** It suffices to show that for each permutation $q = q_1 q_2 \cdots q_n$ written in the one-line notation, there exists exactly one permutation $p \in S_n$ so that $q = g(p)$.

**Remark** The entries of $q$ that are larger than all entries on their left are called **left-to-right maxima**. Note that if $q$ has $t$ left-to-right maxima, then $g^{-1}(q) = p$ has $t$ cycles. Also note that the leftmost left-to-right maximum of $q$ is always $q_1$, and the rightmost left-to-right maximum of $q$ is always the entry $n$. A surprising application is the following.

**Lemma 1.4**

*Let $i$ and $j$ be two elements of $[n]$. Then $i$ and $j$ are in the same cycle in exactly half of all $n$-permutations.* ♡

**Lemma 1.5**

*Let $i \in [n]$. Then for all $k \in [n]$, there are exactly $(n-1)!$ permutations of length $n$ in which the cycle containing $i$ is of length $k$.* ♡

**Lemma 1.6 (Lemma 6.20 in [Miklos2017AWalkTC])**

*Let $|\mathrm{ODD}(m)|$, resp. $|\mathrm{EVEN}(m)|$ be the set of $m$-permutations with all cycle lengths odd, resp. even. For all positive integers $m$, the equality $|\mathrm{ODD}(2m)| = |\mathrm{EVEN}(2m)|$ holds.* ♡

**Theorem 1.17**

*For all positive integers $m$,*

$$|\mathrm{ODD}(2m)| = |\mathrm{EVEN}(2m)| = 1^2 \cdot 3^2 \cdot 5^2 \cdots (2m-1)^2 \,. \tag{1.31}$$

♡

**Proof** We construct a bijection $\Phi$ from $\mathrm{ODD}(2m)$ onto $\mathrm{EVEN}(2m)$. Let $\pi \in \mathrm{ODD}(2m)$. Then $\pi$ consists of an even number $2k$ of odd cycles. Denote by $C_1, C_2, C_{2k}$ the cycles in canonical order. For all $i$ with $1 \le i \le k$, take the last elementof $C_{2i-1}$, and put it to the end of $C_{2i}$ to get $\Phi(\pi)$, the image of $\pi$.

**Example 1.4** If $p = (4)(513)(726)(8)$, then $\Phi(p) = (5134)(72)(86)$.

Note that if $C_{2i-1}$ is a singleton, it disappears. Also note that the canonical form is maintained.

We claim that $\Phi$ is a bijection from $\mathrm{ODD}(2m)$ onto $\mathrm{EVEN}(2m)$. Let $\sigma \in \mathrm{EVEN}(2m)$, with cycles $c_1, c_2, c_h$. To prove that $\Phi$ is a bijection, it suffices to show that we can recover the only permutation $\pi \in \mathrm{ODD}(2m)$ for which $\Phi(\pi) = \sigma$. While recovering $\pi$, we must keep in mind that it might have more than $h$ cycles, because some of its singletons might have been absorbed by the cycles immediately after them. If the last entry in $c_h$ is larger than the first entry in $c_{h-1}$, then create a singleton cycle with the last entry in $c_h$, placing it in front of $c_h$, and repeat the whole procedure using $c_{h-2}$ and $c_{h-1}$. Otherwise, move the last entry in $c_h$ from $c_h$ to the end of $c_{h-1}$, and repeat the whole procedure using $c_{h-3}$ and $c_{h-2}$. If at any point only one cycle remains, create a singleton cycle with the last entry in that cycle. It is then straightforward to check that the permutation $\pi$ obtained this way fulfills $\Phi(\pi) = \sigma$. It also follows from the simple structure of $\Phi$ that at no point of the recovering procedure could we have done anything else.

**Example 1.5** The preimage of $(41)(62)(75)(83)$ under $\Phi$ is $(412)(6)(753)(8)$. The preimage of $(21)(53)(64)(87)$ under $\Phi$ is $(1)(2)(534)(6)(7)(8)$.

**Theorem 1.18 (Theorem 6.25 in [Miklos2017AWalkTC])**

*For all positive integers $m$,*

$$|\mathrm{ODD}(2m+1)| = (2m+1)|\mathrm{ODD}(2m)| = 1^2 \cdot 3^2 \cdot 5^2 \cdots (2m-1)^2 \cdot (2m+1)\,. \tag{1.32}$$

♡

### 1.5.3 Exercises

✍ **Exercise 1.7(Matchings and Permutations)**

(a) How many permutations of $[2n]$ consist of $n$ cycles of length 2?

**Proof** This is a standard cycle type counting calculation. We want to count the number of permutations of $[2n]$ of cycle type $2\ldots2$. Hence, the number will be

$$\frac{(2n!)}{2^n n!} \tag{1.33}$$

(b) Show a bijection between such permutations and matchings of $[2n]$.

**Proof** Given such a permutation as above, it can be written as

$$(a_1\ a_2)(a_3\ a_4)\ldots(a_{2n-1}\ a_{2n}) \tag{1.34}$$

Therefore, we could simply bijectively map this permutation to the matching of $a_1$ to $a_2$, $a_3$ to $a_4$, etc. This is clearly a bijection (with inverse making each pair of matching into a cycle).

(c) How many permutations of $[3n]$ consist of $n$ cycles of length 3?

**Proof** Using the counting formula for permutations of the cycle type $3\ldots3$, the number will be

$$\frac{(3n)!}{3^n n!} \tag{1.35}$$

(d) Does the answer in part (c) equal to the number of partitions of $[3n]$ into sets of size 3? Why or why not?

**Proof** No it does not. Each such permutation *does* partition $[3n]$ into $n$ cycles of length 3 which gives a partition into sets of size 3. However, the permutation also arranges these triples of elements into a cycle. Since 3 elements can be arranged into a cycle in 2 different ways (for example $1, 2, 3$ can be arranged as $(123)$ or $(132)$), this means that there are more such permutations than set partitions.

✍ **Exercise 1.8(Square Permutations)** Let $\pi$ be a permutation of $[n]$. Show that there exists a permutation $\sigma$ with $\pi = \sigma^2$ if and only if $\pi$ has a even number of cycles of length $k$ for every even number $k$.

**Proof** Suppose $\pi = \sigma^2$ for some permutation. Since distinct cycles commute with each other, it suffices to investigate the effect of squaring cycles of $\sigma$. Let $(a_1\ a_2\ \ldots\ a_k)$ be a cycle in the cycle decomposition of $\sigma$.

Observe that if $k$ is odd, we have

$$(a_1 \ \dots \ a_k)^2 = (a_1 \ a_3 \ \dots \ a_k \ a_2 \ a_4 \ \dots \ a_{k-1}) \tag{1.36}$$

Therefore, squaring a cycle of odd length gives you a cycle with same length. If $k$ is even, we have

$$(a_1 \ \dots \ a_k)^2 = (a_1 \ a_3 \ \dots \ a_{k-1})(a_2 \ a_4 \ \dots \ a_k) \tag{1.37}$$

Summarizing the above result, it is clear that the only way to obtain an even cycle is using case 2. In that case, it is clear that every even cycle comes with pairs. Hence, there can only be even numbers of cycles of even length.

Suppose $\pi$ has a even number of cycles of length $k$ for every even number $k$. We first need to show that every cycle $\tau$ of odd length can be written as a square. Suppose $\tau$ has length $k$. Then, it is clear that $\tau^k = 1$. Therefore, $\tau = \tau^{k+1}$. Since $k$ is odd, $k+1$ is even. Hence, we have $\tau = (\tau^{\frac{k+1}{2}})^2$, which is a square. Suppose we are given two cycles $\tau_1 = (a_1 \ \dots \ a_k)$ and $\tau_2 = (b_1 \ \dots \ b_k)$ of even length $k$ in the cycle decomposition, they must be distinct by assumption of cycle decomposition. Then, we have

$$\tau_1\tau_2 = (a_1 \ b_1 \ a_2 \ b_2 \ \dots \ a_k \ b_k)^2 \tag{1.38}$$

by direct computation.

**Exercise 1.9(Largest in its Cycle)** For integers $1 \le k \le n$, how many permutations of $[n]$ have $k$ as the largest element in its cycle? Hint: Consider the canonical cycle representation of such permutations.

**Proof** [Proof by Haixiao] Let $\pi \in S_n$, where $S_n$ is the set of all permutations on $[n]$, be a permutation of its canonical cycle notation(CCN), satisfying the requirements that $k$ is the largest element in its cycle, which can be written as

$$(\cdots)\cdots(k\cdots)\cdots(\cdots) \tag{1.39}$$

By the requirements of CCN, the cycles are arranged in the increasing order of their largest elements. Consequently, if $j > k$, then $j$ should only appear in cycles of $\pi$ that are right to $k$. Let $g(\pi)$ be the permutation obtained from $\pi$ by removing the parentheses and reading the entries as a permutation in the one-line notation. For example, $\pi = (412)(53)$, then $g(\pi) = g((412)(53)) = 41253$. By **Transition Lemma** [**Miklos2017AWalkTC**], $g$ is a bijection from the set $S_n$ onto $S_n$. Therefore, it is enough to count the number of one-line permutations where $j$ is on the right of $k$ if $j \in \{k+1, k+2, \cdots, n\}$. We first choose $n-k+1$ positions from $n$, where $\binom{n}{n-k+1}$ choices available. Then we put $\{k, k+1, \cdots, n-1, n\}$ to those $n-k+1$ positions, with the only constraint that $k$ takes the first position, and there are in total $(n-k)!$ ways to arrange them in the desired order. Finally, we put the numbers $\{1, 2, \cdots, k-1\}$ into the $k-1$ positions left, where we have $(k-1)!$ ways to arrange them. Therefore, the total number of the desired permutations should be

$$\binom{n}{n-k+1}(n-k)!(k-1)! = \frac{n!}{(n-k+1)!(k-1)!}(n-k)!(k-1)! = \frac{n!}{n-k+1}. \tag{1.40}$$

**Proof** [Alternative proof] Let $\pi$ be a permutation of $[n]$. From the definition of the canonical cycle presentation, $k$ is the largest element in the permutation if and only if the numbers before $k$ in the canonical cycle presentation are less than $k$. Therefore, we count the number of permutation in the alternative description.

Define the following bijection $S_n \to [n] \times [n-1] \times \cdots \times [1]$, where $\pi$ is sent to $(a_1, \dots, a_n)$ with $a_i$ is the number of $j$ such that $i \le j$ and $j$ is not to the left of $i$ (in the canonical cycle presentation). For example, $(1)(423)(65)$ is sent to $(6, 4, 3, 3, 1, 1)$. Then, based on our criterion, the permutations we want are precisely those whose image under the bijection has $k$th entry $n-k+1$ (all element that are larger than $k$ is to the right). Hence, since there are in total $n-k+1$ possibility of image of the $k$th entry, it is clear that the number of such permutation is

$$\frac{n!}{n-k+1} \tag{1.41}$$

✍ **Exercise 1.10(Sterling Number Bounds)** Show the following size bounds on Sterling numbers of the first kind for $1 \le k \le n$:

(a) $c(n,k) \ge (n-1)!/(k-1)!$ [10 points]

**Proof** We note that $c(n,k)$ is at least the number of permutations of $[n]$ with one cycle of length $n-k+1$ and $k-1$ cycles of length $1$. If $k=n$, there is only one such cycle. Otherwise, using our formula for counting the number of permutations with a given cycle structure we have that

$$c(n,k) \ge \frac{n!}{(n-k+1)1^{k-1}(k-1)!1!} \ge \frac{n!}{n(k-1)!} = \frac{(n-1)!}{(k-1)!}.$$

**Proof** [Alternative proof] Here we use the formula in the hint. Translation: $c(n,k)$ is the coefficient of $x^k$ of the polynomial $x(x+1)\dots(x+n-1)$. To get the degree $k$ coefficients, we need to sum all choice of $n-k$ numbers in the product. But to get a bound, we can sum a subset of all possible choice. We choose to combine $x$'s from the term $x, x+1, \dots, x+k-1$ and multiply the numbers from the term $x+k, x+k+1, \dots, x+n-1$. By multiplying these out, we get a contribution

$$(n-1)(n-2)\dots kx^k = \frac{(n-1)!}{(k-1)!}x^k \tag{1.42}$$

Since this is only one of the combination, we conclude that

$$c(n,k) \ge \frac{(n-1)!}{(k-1)!} \tag{1.43}$$

(b) For any positive integer $a$,

$$c(n,k) \le n!\binom{n+a}{a}/(a+1)^k.$$

Hint: you may want to use the relation that

$$\sum_{k=1}^{n} x^k c(n,k) = x(x+1)\cdots(x+n-1).$$

[20 points]

**Proof** Using the hint,

$$(a+1)^k c(n,k) \le \sum_{k=1}^{n}(a+1)^k c(n,k) = (a+1)(a+2)\dots(a+n) = n!\binom{n+a}{a} \tag{1.44}$$

Therefore, we get the desired inequality by dividing $(a+1)^k$ on both side.

## 1.6 The Sieve

**Theorem 1.19 (Sieve Formula, or Principle of Inclusion-Exclusion)**

*Let $A_1, A_2, \cdots, A_n$ be finite sets. Then*

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = \sum_{j=1}^{n} (-1)^{j-1} \sum_{i_1, i_2, \cdots, i_j} |A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_j}|, \tag{1.45}$$

*where $\{i_1, i_2, \cdots, i_j\}$ ranges over all $j$-element subsets of $[n]$.* ♡

**Definition 1.13 (Derangement)**

*A **derangement** is a permutation without fixed points.* ♣

**Theorem 1.20**

*The number of derangement is $\sum_{j=0}^{n} (-1)^j \frac{n!}{j!}$.* ♡

**Proof** Let $A_i$ be the set of all permutations of $[n]$ in which the element $i$ is in the $i$th position.

1. It is clear that $|A_i| = (n-1)!$, then the total contribution of the first term of the right-hand side of Sieve's formula is $(n-1)! \cdot n = n!$.
2. Consider $A_i \cap A_j$, then $|A_i \cap A_j| = (n-2)!$ since we have 2 fixed and the remaining $n-2$ free. The total contribution of the second term is
$$(-1)^{2-1} \binom{n}{2} (n-2)! = -\frac{n!}{2!}$$
3. In general, considering the contributions of $j$ terms $|A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_j}|$, a similar argument shows that the contribution of the $j$th term is
$$(-1)^{j-1} \binom{n}{j} (n-j)! = (-1)^{j-1} \frac{n!}{j!}$$

The Sieve's formula yields

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = \sum_{j=1}^{n} (-1)^{j-1} \sum_{i_1, i_2, \cdots, i_j} |A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_j}| = \sum_{j=1}^{n} (-1)^{j-1} \frac{n!}{j!}, \tag{1.46}$$

which is the number of permutations of $[n]$ with at least one fixed point. Consequently, the number $D(n)$ of permutations of $[n]$ with no fixed points, or the number of derangements, is

$$D(n) = n! - |A_1 \cup A_2 \cup \cdots \cup A_n| = n! - \sum_{j=1}^{n} (-1)^{j-1} \frac{n!}{j!} = \sum_{j=0}^{n} (-1)^j \frac{n!}{j!} \tag{1.47}$$

**Theorem 1.21**

*For all positive integers $n$ and $k$, the following equality holds*

$$S(n,k) = \frac{1}{k!} \sum_{j=0}^{k} (-1)^j \binom{k}{j} (k-j)^n = \sum_{j=0}^{k} (-1)^j \frac{(k-j)^n}{j!(k-j)!}. \tag{1.48}$$

♡

**Proof** We will find a formula for $k! \cdot S(n,k)$, which count the number of surjections from $[n]$ to $[k]$. It is clear that functions from $[n]$ to $[k]$ is $k^n$, however not all those functions are surjections.

Let $A_i$ denote the set of all functions from $[n]$ to $[k]$ whose image does not contain $i$. Obviously, $|A_i| = (k-1)^n$ since as such functions can map any element of $[n]$ into any one of $k-1$ elements other than

$i$. Similarly,

$$|A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_j}| = (k-j)^n\,, \quad \forall j \in [k]\,. \tag{1.49}$$

Therefore, the Sieve formula yields

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = \sum_{j=1}^{n}(-1)^{j-1} \sum_{i_1,i_2,\cdots,i_j} |A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_j}| = \sum_{j=1}^{n}(-1)^{j-1}\binom{k}{j}(k-j)^n\,,$$

which gives the number of functions from $[n]$ to $[k]$ whose range is not the entire set $[k]$. Then the number of surjections is

$$k!S(n,k) = k^n - |A_1 \cup A_2 \cup \cdots \cup A_n| = \sum_{j=0}^{k}(-1)^j\binom{k}{j}(k-j)^n\,. \tag{1.50}$$

**Theorem 1.22**

*Let $f$ and $g$ be functions that are defined on the subsets of $[n]$, and whose range is the set of real numbers. Let us assume that $f$ and $g$ are connected by*

$$g(S) = \sum_{T \subseteq S} f(T)\,, \tag{1.51}$$

*where $S, T \subset [n]$. Then*

$$f(S) = \sum_{T \subseteq S} g(T)(-1)^{|S-T|} \tag{1.52}$$

♡

**Proof** We consider the right hand side and write

$$\sum_{T \subseteq S} g(T)(-1)^{|S-T|} = \sum_{T \subseteq S}\sum_{U \subseteq T} f(U)(-1)^{|S-T|}\,. \tag{1.53}$$

Knowing that set $S$ is given. For each $f(U)$ with $U \subseteq T \subseteq S$, it is equipped with sign $(-1)^{|S-T|}$ depending on the difference between $S$ and $T$. Let $|S-T| = i$, then the total contribution of $f(U)$ is $\sum_{i=0}^{|S-U|}(-1)^i\binom{|S-U|}{i}f(U) = (1-1)^{|S-U|}f(U)$, which is 0 when $U \subsetneq S$ and $f(U)$ if $U = S$.

### 1.6.1 Exercises

**Exercise 1.11(Chromatic Polynomials)** A *graph* $G$ is a pair of a set $V$ of *vertices*, and a set $E$ of *edges* each connecting two vertices. We call a graph *simple* if no edge connects two of the same vertex and no two edges connect the same pair of vertices. An *$n$-coloring* of a graph is a way of assigning each vertex a number from $1, 2, \ldots, n$ so that no two vertices connected by an edge are assigned the same number.

(a) Show that for any finite, simple graph $G$ there is a polynomial $P_G(x)$ so that for any positive integer $n$, the number of $n$-colorings of $G$ equals $P_G(n)$. Hint: Use Inclusion-Exclusion.

(b) What are the three highest degree terms of $P_G(x)$ in terms of properties of the graph $G$?

**Proof** For given graph $G = (V, E)$, let $|V(G)|, |E(G)|$ denote the number of vertices and edges in graph $G$. Two vertices $u, v \in V(G)$ are called **adjacent** in $G$ if they are connected by an edge $(u, v) = e \in E(G)$. An $n$-coloring of $G$ is called **proper** if no two adjacent vertices are assigned with the same number. The goal of this question is to find a polynomial $P_G(x)$ such that $P_G(n)$ is exactly the number of **proper** proper colorings given $n$ colors.

(a) Given $n$ colors, the total number of possible coloring, either proper or non-proper, is $n^{|V(G)|}$. For each edge $e = (u, v) \in E(G)$, let $A_e$ denote the set of coloring where vertices $u, v \in e$ have the same color, then $\bigcup_{e \in E} A_e$ denotes the set of coloring where there exists 2 adjacent vertices in $G$ with the same color.

Consequently,

$$n^{|V(G)|} - \left| \bigcup_{e \in E(G)} A_e \right| \tag{1.54}$$

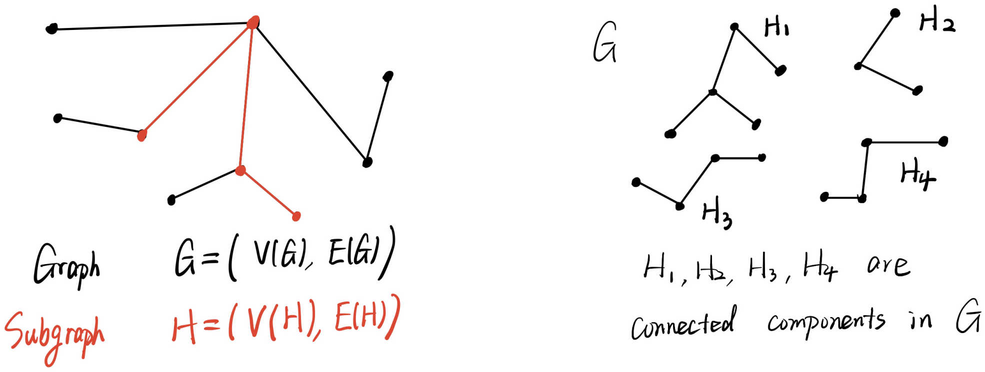
denotes the desired number of proper coloring. Now, we transform our problem to calculate $|\cup_{e\in E(G)} A_e|$. By Inclusion-Exclusion principle,

$$\left| \bigcup_{e \in E(G)} A_e \right| = \sum_{j=1}^{|E(G)|} (-1)^{j-1} \sum_{\{e_1,e_2,\cdots,e_j\} \subset E(G)} |A_{e_1} \cap A_{e_2} \cap \cdots \cap A_{e_j}| . \tag{1.55}$$

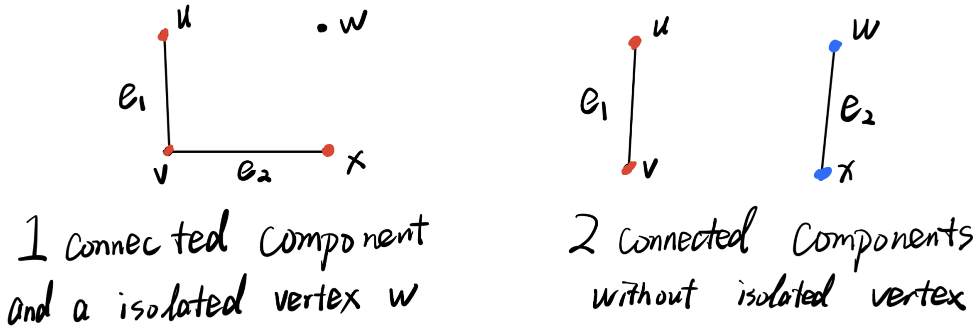
The cases where $|V(G)|$ equals 1 and 2 are trivial. We start our discussion for the case $|V(G)| \geq 3$ and $|E(G)| \geq 1$.

- $j = 1$, the 2 vertices on edge $e_1$ should have the same color($n$ choices), while the other $|V(G)| - 2$ vertices are free, thus $|A_{e_1}| = n \cdot n^{|V(G)|-2} = n^{|V(G)|-1}$ for all $e_1 \in E(G)$.
- $j = 2$, as shown in Figure 1.4b, if $e_1$ and $e_2$ are connected, then all vertices on $e_1$ and $e_2$($u$, $v$ and $x$) should have the same color while the other vertices outside $e_1$, $e_2$ (like $w$) are free. There are in total 3 vertices inside $e_1$ and $e_2$, thus $|A_{e_1} \cap A_{e_2}| = n^1 \cdot n^{|V(G)|-3} = n^{|V(G)|-2}$. Otherwise, if $e_1$ and $e_2$ are disjoint, then colors on them could be different and the total number of vertices on $e_1$ and $e_2$ are 4, thus $|A_{e_1} \cap A_{e_2}| = n^2 \cdot n^{|V(G)|-4} = n^{|V(G)|-2}$. Note that the second case is only meaningful when $|V(G)| \geq 4$.
- To generalize the above argument, we should introduce the concept of **subgraph** and **connected components**. A graph $H$ is called a **subgraph** of $G$ if $V(H) \subset V(G)$ and $E(H) \subset E(G)$. If we partition vertices in $V(G)$ into disjoint sets and obtain some subgraphs $H_1, \cdots, H_k$ such that every pair of vertices in $H_i$ is connected by some path in $H_i(1 \leq i \leq k)$ and no connection between different $H_i$'s, then each subgraph $H_i$ is called a connected component of $G$, as shown in Figure 1.4a. We revisited the argument for case $j = 2$. Let $H$ denote the subgraph of $G$ formed by edges $e_1$ and $e_2$. Let $V(H)$ and $c(H)$ denote the number of vertices and connected components in graph $H$, respectively. If $e_1$ and $e_2$ are connected, then there is 1 connected component occupying 3 vertices in $H$, thus $|A_{e_1} \cap A_{e_2}| = n^1 \cdot n^{|V(G)|-3} = n^{c(H)} \cdot n^{|V(G)|-|V(H)|} = n^{|V(G)|-2}$. If $e_1$ and $e_2$ are not connected, then there are 2 connected components occupying 4 vertices in $H$, hence $|A_{e_1} \cap A_{e_2}| = n^2 \cdot n^{|V(G)|-4} = n^{c(H)} \cdot n^{|V(G)|-|V(H)|} = n^{|V(G)|-2}$.
- In general, let $H$ be a subgraph of $G$ having no isolated vertices with $|E(H)| = j$. By construction, $H$ only contains $j$ edges and associated vertices. Without loss of generality, we denote $E(H) = \{e_1, e_2, \cdots, e_j\} \subset E(G)$. By construction of $|A_{e_1} \cap A_{e_2} \cap \cdots \cap A_{e_j}|$, vertices within the same connected component should have the same color, while different connected components could have different colors, and vertices outside $V(H)$ are free, thus

$$|A_{e_1} \cap A_{e_2} \cap \cdots \cap A_{e_j}| = n^{|c(H)|} \cdot n^{|V(G)|-|V(H)|} = n^{|V(G)|-|V(H)|+|c(H)|} . \tag{1.56}$$




**(a)** Subgraph and connected conponents.




**(b)** Example of one or two connected components when $j = 2$.

Therefore,

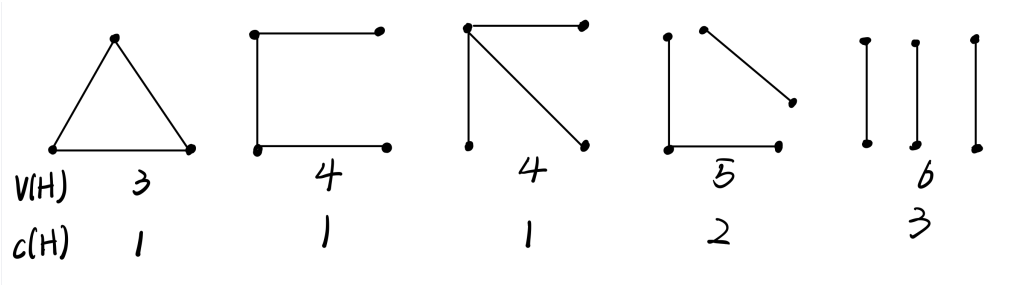
$$P_G(n) = n^{|V(G)|} - \left| \bigcup_{e \in E(G)} A_e \right| \tag{1.57a}$$

$$= n^{|V(G)|} + \sum_{j=1}^{|E(G)|} (-1)^j \sum_{\substack{E(H) \subset E(G) \\ |E(H)|=j}} |A_{e_1} \cap A_{e_2} \cap \cdots \cap A_{e_j}| \tag{1.57b}$$

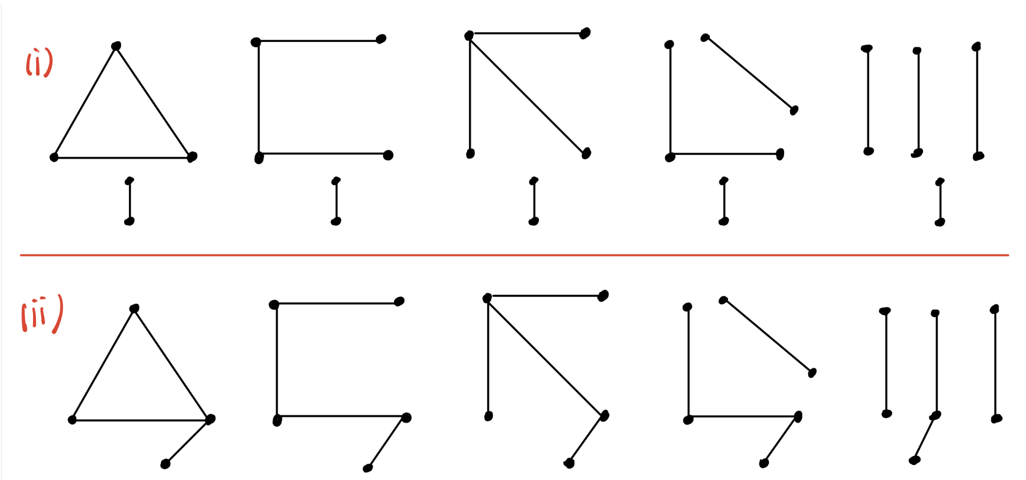
$$= n^{|V(G)|} + \sum_{j=1}^{|E(G)|} (-1)^j \sum_{\substack{E(H) \subset E(G) \\ |E(H)|=j}} n^{|V(G)|-|V(H)|+|c(H)|}. \tag{1.57c}$$

Note that as each term here is a power of $n$ that $P_G(n)$ is in fact a polynomial.
The **chromatic number** of graph $G$ is defined to be the smallest positive integer $n$ such that $P_G(n) > 0$, a.k.a., $\chi(G) = \min\{n \in \mathbb{N} | P_G(n) > 0\}$.

(b) From the formula above, it is clear that the highest degree term is $n^{|V(G)|}$, while the exponent is the number of vertices in graph $G$. The second highest degree term is $-\sum_{e \in E(G)} n^{|V(G)|-1} = -|E(G)| \cdot n^{|V(G)|-1}$, where the absolute value of the coefficient represents the number of edges in $G$. The third highest degree term $n^{|V(G)|-2}$, which corresponds to the subgraph $H$ in $G$ such that $|V(H)| - c(H) = 2$. We claim that this equality holds only if $H$ has exactly 2 edges or $H$ is a triangle, as shown in Figure 1.4b and Figure 1.5a. The coefficient of this term is therefore $\binom{|E(G)|}{2} - |\#$ of triangles in $G|$, by going through all possible choices of selecting 2 edges from edge set $E(G)$, and counting triangles in $G$, while the minus sign is due to the sign difference for the cases $j = 2$ and $j = 3$ in our formula.
Now we prove the claim. If $H$ only has 2 edges $e_1$, $e_2$ and they are connected, then $V(H) = 3$ and




**(a)** Case of 3 edges




**(b)** Case of 4 edges

$c(H) = 1$, otherwise $V(H) = 4$ and $c(F) = 2$. In both cases, $V(H) - c(H) = 2$. If $H$ is a triangle, then clearly, $V(H) = 3$, $c(H) = 1$ and $V(H) - c(H) = 2$. For all the other $H$, we claim that $V(H) - c(H) \geq 3$, which can be proved by an induction argument.

- Base case. We can easily check that $V(H) - c(H) \geq 3$ for cases $E(H) = 3$ except $H$ being a triangle, as shown in Figure 1.5a.
- If we add one more edge to graphs in Figure 1.5a, there are two possible cases shown in Figure 1.5b.
  - (i) The newly added edge forms a connected component itself, which means that $V(H)$ increases by 2 and $c(F)$ increases by 1, hence $V(H) - c(H) \geq 3$.
  - (ii) The newly added edge connected to some existing component, which means that $V(H)$ increases by 1 and $c(F)$ stays the same, hence $V(H) - c(H) \geq 3$.

**Exercise 1.12(Reverse Inclusion-Exclusion)** For finite sets $A, B$, and $C$ give a formula for $|A \cap B \cap C|$ in terms of $|A|, |B|, |C|, |A \cup B|, |B \cup C|, |C \cup A|, |A \cup B \cup C|$.

**Proof** Note that

$$|A \cup B| = |A| + |B| - |A \cap B|, \quad |B \cup C| = |B| + |C| - |B \cap C|, \quad |A \cup C| = |A| + |C| - |A \cap C|, \tag{1.58}$$

and Inclusion-Exclusion formula gives us

$$|A \cup B \cup C| = |A| + |B| + |C| - (|A \cap B| + |B \cap C| + |A \cap C|) + |A \cap B \cap C|. \tag{1.59}$$

Then

$$
\begin{aligned}
&|A \cap B \cap C| && \text{(1.60a)}\\
&= |A \cup B \cup C| - (|A| + |B| + |C|) + (|A \cap B| + |B \cap C| + |A \cap C|) && \text{(1.60b)}\\
&= |A \cup B \cup C| - (|A| + |B| + |C|) + (|A| + |B| - |A \cup B| + |B| + |C| - |B \cup C| + |A| + |C| - |A \cup C|) && \text{(1.60c)}\\
&= |A \cup B \cup C| + |A| + |B| + |C| - |A \cup B| - |B \cup C| - |A \cup C|\,. && \text{(1.60d)}
\end{aligned}
$$

## 1.7 Generating Functions

### 1.7.1 Ordinary Generating Functions

**Definition 1.14**

*Let $\{a_n\}_{n\geq 0}$ be a sequence of real numbers. Then the formal power series $A(x) = \sum_{n\geq 0} a_n x^n$ is called the **ordinary generating function** of the sequence $\{a_n\}_{n\geq 0}$.* ♣

#### 1.7.1.1 Basic Tools

Properties of generating function $A(x)$.

1. Identity: A generating function is determined by its coefficients. In particular, $\sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} b_n x^n$ if and only if $a_n = b_n$ for all $n \geq 0$.
2. Geometric series: $\frac{1}{1-cx} = \sum_{n=0}^{\infty} (cx)^n$ if $|cx| < 1$.
3. Shifts: $A(x) = \sum_{n\geq 0} a_n x^n$, then $xA(x) = \sum_{n\geq 1} a_{n-1} x^n$.
4. Sum of generating functions: $\sum_{n=0}^{\infty} a_n x^n + \sum_{n=0}^{\infty} b_n x^n = \sum_{n=0}^{\infty} (a_n + b_n) x^n$.

#### 1.7.1.2 Partial Fractions

Given a generating function $F(x) = p(x)/q(x)$ with $p(x)$ and $q(x)$ polynomials, we find a formula for the coefficients by **Partial Fractions**.

1. Reduce $p(x)$. Write $F(x) = r(x) + p_0(x)/q(x)$ with $\text{degree}(p_0(x)) < \text{degree}(q(x))$.
2. Factor $q(x)$. Write $q(x) = (x - r_1)(x - r_2)\cdots(x - r_k)$.
3. Rewrite
$$p_0(x)/q(x) = \frac{A_1}{x - r_1} + \frac{A_2}{x - r_2} + \cdots + \frac{A_k}{x - r_k}. \tag{1.61}$$
This is always possible if $q(x) = 0$ has distinct roots, otherwise a bit more complicated.
4. Get Formula. Note that
$$\frac{1}{x - r_i} = -\frac{1}{r_i} \sum_{n=0}^{\infty} \left(\frac{x}{r_i}\right)^n \tag{1.62}$$
Adding terms together, we would get
$$\frac{A_1}{x - r_1} + \frac{A_2}{x - r_2} + \cdots + \frac{A_k}{x - r_k} = \sum_{n} -\left(\frac{A_1}{r_1^{n+1}} + \cdots + \frac{A_k}{r_k^{n+1}}\right) x^n \tag{1.63}$$

**Techniques to Solve Simple Recurrence Relations**

1. Define generating function $F(x)$
2. Use recurrence relation to relate $F(x)$ to shifted generating functions.
3. Solve for $F(x)$ as a rational function.
4. Compute a partial fraction decomposition.
5. Use to obtain a formula for coefficients.

#### 1.7.1.3 Products of Generating Functions

Let $A(x) = \sum_{n=0}^{\infty} a_n x^n$, $B(x) = \sum_{n=0}^{\infty} b_n x^n$, and $C(x) = A(x)B(x)$, then

$$c_n = \sum_{m+k=n}^{n} a_m b_k = \sum_{k=0}^{n} a_k b_{n-k} . \tag{1.64}$$

**Combinatorial Interpretation.** Suppose that you have objects of type-A and objects of type-B. Each has a size which is a non-negative integer, and there are $a_m$ objects of type-A of size $m$, and $b_k$ objects of type-B of size $m$. Then $c_n$ is the number of ways to find a pair of an object of type-A and an object of type-B where the sum of the sizes is $n$.

**Example 1.6(Fibonacci Numbers)** Let $F(x) = \sum_{n=0}^{\infty} f_n x^n$ be the generating function. Recurrence relationship tells us $f_{n+2} - f_{n+1} - f_n = 0$, then

$$\sum_{n=0}^{\infty} (f_{n+2} - f_{n+1} - f_n) x^{n+2} = 0 \tag{1.65}$$

Note

1. $\sum_{n=0}^{\infty} f_n x^{n+2} = x^2 F(x)$,
2. $\sum_{n=0}^{\infty} f_{n+1} x^{n+2} = x[F(x) - f_0] = xF(x) - x$,
3. $\sum_{n=0}^{\infty} f_{n+2} x^{n+2} = F(x) - f_1 x - f_0 = F(x) - x - 1$

Adding together,

$$F(x) - x - 1 - xF(x) + x - x^2 F(x) = 0 \implies F(x) = \frac{1}{1 - x - x^2} . \tag{1.66}$$

By expanding $F(x)$, we have

$$f_n = \frac{1}{\sqrt{5}} \left[ \left( \frac{1+\sqrt{5}}{2} \right)^{n+1} - \left( \frac{1-\sqrt{5}}{2} \right)^{n+1} \right] . \tag{1.67}$$

**Example 1.7**

$$\frac{1}{1-(x+y)} = \sum_{n=0}^{\infty} (x+y)^n = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \binom{n+m}{m} x^n y^m \tag{1.68a}$$

$$= \frac{1/(1-x)}{1 - y/(1-x)} = \sum_{m=0}^{\infty} \left( \frac{1}{1-x} \right)^{m+1} y^m . \tag{1.68b}$$

Comparing $y^m$ coefficients, we find

$$\left( \frac{1}{1-x} \right)^{m+1} = \sum_{n=0}^{\infty} \binom{n+m}{m} x^n \tag{1.69}$$

If we substitute $x = z$ and $y = z^2$, then

$$\frac{1}{1-(z+z^2)} = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \binom{n+m}{m} z^{n+2m} = \sum_{k=0}^{\infty} \left( \sum_{m=0}^{k/2} \binom{k-m}{m} \right) z^k , \tag{1.70}$$

where $\frac{1}{1-(z+z^2)}$ is generating function of Fibonacci Numbers.

- As an **actual function** of a (real or complex) variable $x$, that should converge for $x$ in some range.
- A **formal power series**- namely a set of symbols that can be manipulated in the same way a real function could, but that can't necessarily be evaluated anywhere.
- If we treat $F(x)$ as a function, we need to worry about issues of convergence. For which (if any) values of $x$ does $F(x)$ converge? When we do manipulations on infinite sums, are we allowed to?
- Because of this, for most combinatorial applications, it is better to treat $F(x)$ as a formal power series.
- But sometimes actually having a function is useful.

#### 1.7.1.4 Integer partition

Consider the generating function

$$\frac{x^3}{(1-x)^3} = (x + x^2 + x^3 + \cdots)(x + x^2 + x^3 + \cdots)(x + x^2 + x^3 + \cdots) \tag{1.71}$$

$$= x^{1+1+1} + x^{2+1+1} + x^{1+2+1} + x^{1+1+2} + x^{1+1+3} + x^{1+2+2} + x^{1+3+1} \tag{1.72}$$

$$+ x^{2+1+2} + x^{2+2+1} + x^{3+1+1} + \cdots \tag{1.73}$$

Expanding out, we get the sum of $x^{a+b+c}$ over all triples of positive integers $(a, b, c)$. The coefficient of $x^n$ is the number of triples $(a, b, c)$ which is the solution to $a+b+c = n$. In other words the number of compositions of $n$ into three parts.

An integer partition of $n$ is a way of writing $n$ as the sum of some number of 1's plus some number of 2's and so on, that is $n = \sum_{k=1}^{\infty} k \cdot a_k$ for $a_k \geq 0$. Note that

$$\sum_{a=0}^{\infty} (x^k)^a = \frac{1}{1-x^k}, \quad \forall k \geq 1 \tag{1.74}$$

Then

$$F(x) = \sum_{n=0}^{\infty} p(n)x^n = \prod_{k=1}^{\infty} \frac{1}{1-x^k}. \tag{1.75}$$

It is not hard to show that $F(x) \leq e^{\frac{\pi^2}{6(1-x)}}$ for $0 < x < 1$. On the other hand, $F(x) > p(n)x^n$. Setting $x = 1 - \pi/\sqrt{6n}$, we find that $p(n) \leq e^{\pi\sqrt{2n/3}}$.

#### 1.7.1.5 Catlan Number

**Definition 1.15**

*The nth* ***Catalan Number*** *$C_n$ is the number of up-left (steps $(0, 1)$ or $(1, 0)$ ) lattice paths from $(0, 0)$ to $(n, n)$ that stay on or above the line $x = y$.* ♣

Let $(k, k)$ denote the first time this move hits $x = y$. Note that the first step must be $(0, 1)$ and the step to hit $(k, k)$ must be $(1, 0)$. After first step, this move is at $(0, 1)$, while before the step to $(k, k)$, this move stays at $(k-1, k)$. Since $(k, k)$ is the first time to hit $x = y$, this move should stay above the line $y = x + 1$ between $(0, 1)$ and $(k-1, k)$, which gives $C_{k-1}$ choices. After first hit at $(k, k)$, the remaining steps give $C_{n-k}$ choices. Thus the recurrence relationship is

$$C_n = \sum_{k=1}^{n} C_{k-1}C_{n-k}, \quad C_0 = 1. \tag{1.76}$$

Define the generating function $H(x) = \sum_{n=0}^{\infty} C_n x^n$, and using product formula

$$[H(x)]^2 = \sum_{n=0}^{\infty} \left( \sum_{k=0}^{n} C_k C_{n-k} \right) x^n = \sum_{n=0}^{\infty} C_{n+1} x^n. \tag{1.77}$$

Thus

$$x[H(x)]^2 = \sum_{n=0}^{\infty} C_{n+1} x^{n+1} = H(x) - 1. \tag{1.78}$$

Hence

$$H(x) = \frac{1 \pm \sqrt{1-4x}}{2x} \tag{1.79}$$

Only the $-$ term makes sense, since $H(0)$ needs to be finite. Recall that

$$\sqrt{1-4x} = 1 - 2x - 2x^2\binom{2}{1}/2 - 2x^3\binom{4}{2}/3 - \cdots \tag{1.80a}$$

$$\frac{1-\sqrt{1-4x}}{2x} = 1 + x\binom{2}{1}/2 + x^2\binom{4}{2}/3 + x^3\binom{6}{3}/4 \tag{1.80b}$$

Therefore

$$C_n = \frac{1}{n+1}\binom{2n}{n}. \tag{1.81}$$

#### 1.7.1.6 Composition of Generating Functions

How many compositions of $n$ into (any number of) odd parts are there?

**Example 1.8** If $n = 5$ we have: $5, 3+1+1, 1+3+1, 1+1+3, 1+1+1+1+1$, so 5 in total.

For any $k \in \mathbb{N}$, what is the number of compositions into exactly $k$ parts, while each part is odd? This is the number $k$-tuples $(a_1, \cdots, a_k)$ of solutions to $a_1 + \cdots + a_k = n$ with each $a_i$ being odd. The generating function is

$$(x + x^3 + x^5 + \cdots)(x + x^3 + x^5 + \cdots)\cdots(x + x^3 + x^5 + \cdots) = \left(\frac{x}{1-x^2}\right)^k. \tag{1.82}$$

Let $O_n$ be the number of compositions of $n$ into any number of odd parts and let $O_{n,k}$ be the number of compositions into exactly $k$ odd parts. We have

$$\sum_{n=0}^{\infty} O_{n,k}x^n = \left(\frac{x}{1-x^2}\right)^k, \quad O_n = \sum_{k=0}^{n} O_{n,k}. \tag{1.83}$$

$$\sum_{n=0}^{\infty} O_n x^n = \sum_{n=0}^{\infty}\sum_{k=0}^{n} O_{n,k}x^n = \sum_{k=0}^{\infty}\sum_{n=0}^{\infty} O_{n,k}x^n = \sum_{k=0}^{\infty}\left(\frac{x}{1-x^2}\right)^k = \frac{1-x^2}{1-x-x^2}. \tag{1.84a}$$

**Strategy to Compute Compositions**

- Define an $A$-structure on a set to be a thing where there are $a_k$ ways to build an A-structure on a set of size $k$, and define the generating function to be $A(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots$. (**Macro**)
- Define a $B$-structure as a thing where there are $b_m$ $B$-structures of size $m$, and generating function is $B(x) = b_0 + b_1x + b_2x^2 + b_3x^3 + \cdots$. (**Micro**) Then

$$A(B(x)) = a_0 + a_1B(x) + a_2[B(x)]^2 + a_3[B(x)]^3 + \cdots \tag{1.85a}$$

$$= c_0 + c_1x + c_2x^2 + c_3x^3 + \cdots \tag{1.85b}$$

  Note that we almost always want to take $b_0 = 0$ for this. Otherwise, $A(B(x))$ isn't necessarily defined even at $x = 0$.
- Let $[B(x)]^k$ have the coefficients given by $[B(x)]^k = \sum_{n=0}^{\infty} c_{n,k}x^n$, we then have

$$c_n = \sum_{k=0}^{\infty} a_k c_{n,k} \tag{1.86}$$

  Then $c_{n,k}$ is the number of ways to find an ordered list of $k$ $B$-structures whose total size is $n$.
- Define an $A$-structure on a set to be a thing where there are $a_k$ ways to build an $A$-structure on a set of size $k$. Then we have that $c_n$ is the number of ways to find an ordered list of $B$-structures of total size $n$ and then build an $A$-structure on top of them.

### 1.7.2 Exponential generating functions

Sometimes, generating function doesn't converge.

**Example 1.9** Let $a_n = c(n,2)$ (Sterling Number of the first kind).

$$a_n = c(n,2) = (n-1)c(n-1,2) + c(n-1,1) = (n-1)a_{n-1} + (n-2)!\,. \tag{1.87}$$

The generating function is

$$A(x) = \sum_{n=0}^{\infty} a_n x^n = x\sum_{n=0}^{\infty} n a_n x^{n+1} + \sum_{n=1}^{\infty} (n-1)! \cdot x^{n+1}\,, \tag{1.88}$$

where the second term doesn't converge. The ordinary generating function converges only if $a_n$ grows at most exponentially.

We want to introduce **exponential generating functions**, which in general converges.

**Definition 1.16**

*Let $\{a_n\}_{n\geq 0}$ be a sequence of real numbers. Then the formal power series $A(x) = \sum_{n\geq 0} a_n x^n/n!$ is called the **exponential generating function** of the sequence $\{a_n\}_{n\geq 0}$.* ♣

#### 1.7.2.1 Basic tools

| | Ordinary | Exponential |
|---|---|---|
| | $A(x) = \sum_{n\geq 0} a_n x^n$ | $A(x) = \sum_{n\geq 0} a_n x^n/n!$ |
| $a_n = 1$ | $A(x) = 1/(1-x)$ | $A(x) = e^x$ |
| $a_n = n$ | $A(x) = x/(1-x)^2$ | $A(x) = xe^x$ |
| Shift | $xA(x) = \sum_{n\geq 1} a_{n-1} x^n$ | $A'(x) = \sum_{n\geq 0} a_{n+1} x^n/n!$ |
| | Converges only if angrows at most exponentially. | Converges more generally. |

**Example 1.10** $a_0 = 2$, $a_{n+1} = 2a_n - 1$. The generating function is $A(x) = \sum_{n\geq 0} a_n x^n/n!$, then

$$A'(x) = \frac{d}{dx}A(x) = \sum_{n\geq 0} a_{n+1} x^n/n! = \sum_{n\geq 0} (2a_n - 1)x^n/n! \tag{1.89a}$$

$$= 2\sum_{n\geq 0} a_n x^n/n! - \sum_{n\geq 0} x^n/n! = 2A(x) - e^x \tag{1.89b}$$

By solving this ODE, we have $A(x) = e^x + e^{2x}$, hence

$$A(x) = e^x + e^{2x} = \sum_{n\geq 0} x^n/n! + \sum_{n\geq 0} (2x)^n/n! = \sum_{n\geq 0} (2^n+1)x^n/n! \tag{1.90}$$

Thus $A_n = 2^n + 1$ for $n \geq 0$.

**Example 1.11(Derangements)** A permutation without fixed point is called **derangement**. Let $D_n$ denote the number of derangements in $S_n$. We first prove a recurrence relationship.

**Lemma 1.7**

$$D_{n+1} = nD_n + nD_{n-1}\,. \tag{1.91}$$

♡

**Proof** [Proof of claim] We consider the position of last element $n+1$.

- $n+1$ is in a cycle of length 2. Then there are $n$ choices for the second element and $D_{n-1}$ choices for the remaining $n-1$ elements to form derangement, which are independent, and in total $nD_{n-1}$.
- $n+1$ is in a cycle of length more than 2. Removing $n+1$ from it's cycle gives a derangement of $[n]$. There are $D_n$ choices for derangements of $[n]$, and $n$ ways to insert $n+1$ in the existing permutation ($n$ numbers to map), in total $nD_n$.

Let $F(x) = \sum_{n\geq 0} D_n x^n/n!$ be the exponential generating function, then

$$F'(x) = \sum_{n\geq 0} D_{n+1}x^n/n! = \sum_{n\geq 0} nD_n x^n/n! + \sum_{n\geq 0} nD_{n-1}x^n/n! \tag{1.92a}$$

$$= x\sum_{n\geq 1} D_n x^{n-1}/(n-1)! + x\sum_{n\geq 1} D_{n-1}x^{n-1}/(n-1)! \tag{1.92b}$$

$$= xF'(x) + xF(x)\,. \tag{1.92c}$$

Hence $(1-x)F'(x) = xF(x)$, and using the fact $F(0) = 1$

$$\frac{F'(x)}{F(x)} = \frac{x}{1-x} = \frac{1}{1-x} - 1 \tag{1.93a}$$

$$\log(F(x)) = -\log(1-x) - x \tag{1.93b}$$

Therefore $F(x) = e^{-x}/(1-x)$. By expanding $F(x)$ in terms of power series, we would have $D_n = \sum_{j=0}^{n}(-1)^j n!/j!$.

#### 1.7.2.2 Multiplication of Exponential Generating Functions

$A(x) = \sum_{n\geq 0} a_n x^n/n!$, $B(x) = \sum_{n\geq 0} b_n x^n/n!$, and $C(x) = A(x)B(x) = \sum_{n\geq 0} c_n x^n/n!$, what is $c_n$?

$$C(x) = A(x)B(x) = \left(\sum_{m\geq 0} a_m x^m/m!\right)\left(\sum_{k\geq 0} b_k x^k/k!\right) \tag{1.94a}$$

$$= \sum_{n=0}^{\infty}\left(\sum_{m+k=n} a_m b_k/(m!\cdot k!)\right)x^n \tag{1.94b}$$

$$= \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}\binom{n}{k}a_k b_{n-k}\right)x^n/n! \tag{1.94c}$$

Thus $c_n = \sum_{k=0}^{n}\binom{n}{k}a_k b_{n-k}$.

**Combinatorial Interpretation**

- Define $A$-structure a thing so that there are $a_n$ $A$-structures on a set of size $n$.
- Define $B$-structure a thing so that there are $b_n$ $A$-structures on a set of size $n$.
- Ordinary generating function multiplication talks about the number of ways to find an $A$-structure and a $B$-structure of total size $n$.
- Exponential generating function multiplication has $c_n$ = number of ways to partition $[n]$ into two sets and put an $A$-structure on one and a $B$-structure on the other.
- If $A$-structure of size $k$, then $\binom{n}{k}$ ways to partition $[n]$, $a_k$ $A$-structures and $b_{n-k}$ $B$-structures.

**Example 1.12(Bell Number)** In order to get a set partition of $[n+1]$, we need to partition $[n]$ into:

- The set of elements that go with $n+1$. This is $A$-structure, and it has only one option, thus $a_m = 1$.
- The rest elements and a set partition of the rest. This $B$-structure, and the number of choices is set partition $b_k = B(k)$.

The recurrence relationship gives

$$B(n+1) = \sum_{k=0}^{n}\binom{n}{k}B(k)\,. \tag{1.95}$$

We calculate the generating function $F(x) = \sum_{n=0}^{\infty} B(n)x^n/n!$, and using multiplication of exponential generating functions, we have

$$F'(x)=\sum_{n=0}^{\infty}B(n+1)\frac{x^n}{n!}=\sum_{n=0}^{\infty}\frac{x^n}{n!}\sum_{k=0}^{n}\binom{n}{k}B(k) \tag{1.96}$$

$$=\sum_{m=0}^{\infty}\frac{x^m}{m!}\sum_{k=0}^{\infty}B(k)\frac{x^k}{k!}=e^xF(x) \tag{1.97}$$

Note that $F(0)=1$, then $\log(F(x))=e^x-1$, hence

$$F(x)=\sum_{n=0}^{\infty}B(n)\frac{x^n}{n!}=e^{e^x-1}\,. \tag{1.98}$$

#### 1.7.2.3 Generalized Inclusion-Exclusion

Recall principle of Inclusion-Exclusion. Let $A_1,A_2,\cdots,A_n$ be finite sets. Then

$$|A_1\cup A_2\cup\cdots\cup A_n|=\sum_{k=1}^{n}(-1)^{k-1}\sum_{1\le i_1<i_2<\cdots<i_k\le n}|A_{i_1}\cap A_{i_2}\cap\cdots\cap A_{i_k}|\,, \tag{1.99}$$

where $\{i_1,i_2,\cdots,i_k\}$ ranges over all $k$-element subsets of $[n]$.

Suppose we expected that there was a formula of this form, but didn't know the coefficients $(-1)^{k-1}$. How would we find them?

We make the argument slightly more general, and we replace with $(-1)^{k-1}$ with $a_k$'s. For some numbers $a_1,a_2,\cdots,a_k,\cdots,a_n$, we want to consider expressions of the form:

$$\sum_{k=1}^{n}a_k\sum_{1\le i_1<i_2<\cdots<i_k\le n}|A_{i_1}\cap A_{i_2}\cap\cdots\cap A_{i_k}|\,. \tag{1.100}$$

If element $x$ is in exactly $m$ of the $A_i$'s, how much does it contribute to this sum ? For $k$th term ($k$-wise $A_i$'s intersects ), $x$ is in $\binom{m}{k}$ of the $k$-wise intersections, and the contribution is $b_m=\sum_{k=1}^{m}a_k\binom{m}{k}$.

Now we refer to the generating functions. Let $A(x)=\sum_{k=0}^{\infty}a_kx^k/k!$ where $a_0=0$ and $B(x)=\sum_{m=0}^{\infty}b_mx^m/m!$, then by multiplication of exponential generating functions

$$B(x)=\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}a_k\binom{n}{k}\right)\frac{x^n}{n!}=\left(\sum_{k=0}^{\infty}a_kx^k/k!\right)\left(\sum_{m=0}^{\infty}x^m/m!\right)=e^xA(x)\,. \tag{1.101}$$

- For Inclusion-Exclusion we want to count each $x$ once. and we restrict the contribution of each $x$ to be 1, that is $b_m=1$ for each $m\ge 1$, then,

$$B(x)=\sum_{k=1}^{\infty}\frac{x^k}{k!}=e^x-1\,, \tag{1.102}$$

  then

$$A(x)=e^{-x}B(x)=1-e^{-x}=\sum_{k=1}^{\infty}(-1)^{k-1}\frac{x^k}{k!} \tag{1.103}$$

  Therefore $a_k=(-1)^{k-1}$.
- Suppose instead we wanted to count the number of elements that are in exactly two of the $A_i$'s. We want $b_2=1$ and all other $b_m=0$, hence $B(x)=x^2/2$ and

$$A(x)=e^{-x}B(x)=x^2e^{-x}/2=\frac{x^2}{2\cdot 0!}-\frac{x^3}{2\cdot 1!}+\frac{x^4}{2\cdot 2!}-\frac{x^5}{2\cdot 3!}+\cdots \tag{1.104}$$

The coefficient of $x^k$ is $\frac{(-1)^k}{2\cdot(k-2)!}$, thus $a_k = \frac{(-1)^k k!}{2\cdot(k-2)!} = \frac{(-1)^k k(k-1)}{2}$. Therefore

$$\#\text{ of elements in exactly two of } A_i = \sum_{k=2}^{n}(-1)^k\binom{k}{2}\sum_{1\le i_1<i_2<\cdots<i_k\le n}|A_{i_1}\cap A_{i_2}\cap\cdots\cap A_{i_k}|$$

### 1.7.2.4 Compositions of Exponential Generating Functions

Recall that for Bell number

$$F(x)=\sum_{n=0}^{\infty}B(n)\frac{x^n}{n!}=e^{e^x-1}=1+(e^x-1)+\frac{((e^x-1))^2}{2!}+\frac{(e^x-1)^3}{3!}+\cdots \tag{1.105}$$

We take a closer look of each power term.

1. Note that $e^x-1=\sum_{m=1}^{\infty}x^m/m!$.
2. For the second order term,

$$\begin{aligned}\frac{(e^x-1)^2}{2!}&=\frac{1}{2!}(x+\frac{x^2}{2!}+\frac{x^3}{3!}+\cdots)(x+\frac{x^2}{2!}+\frac{x^3}{3!}+\cdots)\\&=\sum_{n=0}^{\infty}\left[\frac{1}{2!}\sum_{\substack{m,k\ge1\\m+k=n}}\binom{n}{k}\right]\frac{x^n}{n!},\end{aligned}$$

where $\binom{n}{k}$ means the number of ways to partition $[n]$ into sets of size $m$ and $k$, $\frac{1}{2!}$ accounts for overcounting of which set is first, and $\left[\frac{1}{2!}\sum_{m,k\ge1,m+k=n}\binom{n}{k}\right]$ denotes the total number of partitions of $[n]$ into 2 sets. (We define $\binom{n}{k}=0$ if $k>n$.)
3. Higher powers,

$$\frac{(e^x-1)^k}{k!}=\sum_{n=0}^{\infty}\left[\frac{1}{k!}\sum_{\substack{m_1,m_2,\cdots,m_k\ge1\\m_1+m_2+\cdots+m_k=n}}\binom{n}{m_1,m_2,\cdots,m_k}\right]\frac{x^n}{n!},$$

where

- $\binom{n}{m_1,m_2,\cdots,m_k}$ represents the number of partitions of $[n]$ into $A_1,A_2,\cdots,A_k$ with $|A_i|=m_i$,
- $\sum_{\substack{m_1,m_2,\cdots,m_k\ge1\\m_1+m_2+\cdots+m_k=n}}\binom{n}{m_1,m_2,\cdots,m_k}$ represents the number of partitions of $[n]$ into **non-empty** $A_1,A_2,\cdots,A_k$.
- $\left[\frac{1}{k!}\sum_{\substack{m_1,m_2,\cdots,m_k\ge1\\m_1+m_2+\cdots+m_k=n}}\binom{n}{m_1,m_2,\cdots,m_k}\right]$ denotes the number of partitions of $[n]$ into $k$ **non-empty** sets (without labelling on boxes) = $S(n,k)$

Therefore,

$$\frac{(e^x-1)^k}{k!}=\sum_{n=0}^{\infty}S(n,k)\frac{x^n}{n!}.$$

**Theorem 1.23**

*Summing over $k$ we have that,*

$$\sum_{k=0}^{\infty}\sum_{n=0}^{\infty}S(n,k)\frac{x^n}{n!}y^k=\sum_{k=0}^{\infty}\frac{(e^x-1)^k}{k!}=e^{y(e^x-1)}.$$

*Setting $y=1$ and change the order of summation, we have $\sum_{n=0}^{\infty}B(n)\frac{x^n}{n!}$.* ♡

We now consider a more general argument. Define $A$-structure so that there are $a_m$ A-structures on a set of size $m$. Let $A(x) = \sum_{m=1}^{\infty} a_m/m!$, then

$$\frac{[A(x)]^k}{k!} = \sum_{n=0}^{\infty} \left[ \frac{1}{k!} \sum_{\substack{m_1, m_2, \cdots, m_k \geq 1 \\ m_1+m_2+\cdots+m_k=n}} \binom{n}{m_1, m_2, \cdots, m_k} a_{m_1} a_{m_2} \cdots a_{m_k} \right] \frac{x^n}{n!} \tag{1.107}$$

where

- $\binom{n}{m_1, m_2, \cdots, m_k} a_{m_1} a_{m_2} \cdots a_{m_k}$ represents the number of partitions of $[n]$ into $A_1, A_2, \cdots, A_k$ with $|A_i| = m_i$ and put $A$-structure on each.
- $\sum_{\substack{m_1, m_2, \cdots, m_k \geq 1 \\ m_1+m_2+\cdots+m_k=n}} \binom{n}{m_1, m_2, \cdots, m_k} a_{m_1} a_{m_2} \cdots a_{m_k}$ represents the number of partitions of $[n]$ into **non-empty** $A_1, A_2, \cdots, A_k$ and put $A$-structure on each.
- $\left[ \frac{1}{k!} \sum_{\substack{m_1, m_2, \cdots, m_k \geq 1 \\ m_1+m_2+\cdots+m_k=n}} \binom{n}{m_1, m_2, \cdots, m_k} a_{m_1} a_{m_2} \cdots a_{m_k} \right]$ denotes the number of partitions of $[n]$ into $k$ **non-empty** sets (without labelling on boxes) with an $A$-structure on each.

**Example 1.13** How many ways can we partition $[n]$ into two subsets and select one element from each?

**Proof** In this problem, $A$-structure on $S$ of size $m$ is to select one element of from $S$, thus $a_m = m$. Exponential generating function is given by

$$A(x) = \frac{1}{2} \left( \sum_{k=1}^{\infty} k \frac{x^k}{k!} \right)^2 = \frac{(xe^x)^2}{2} = \frac{1}{2} x^2 e^{2x} \tag{1.108a}$$

$$= \frac{1}{2} x^2 + \frac{1}{2} x^2 \frac{(2x)}{1!} + \frac{1}{2} x^2 \frac{(2x)^2}{2!} + \frac{1}{2} x^2 \frac{(2x)^3}{3!} + \cdots + \frac{1}{2} x^2 \frac{(2x)^k}{k!} + \cdots \tag{1.108b}$$

Thus coefficients of $x^k/k!$ is $2^{k-3} k(k-1)$.

So if $A(x) = \sum_{k=1}^{\infty} a_k x^k/k!$, and $B(x) = \sum_{k=1}^{\infty} b_k x^k/k!$. Then

$$B(A(x)) = \sum_{n=1}^{\infty} b_n [A(x)]^n/n! \,. \tag{1.109}$$

The $x^n/n!$ coefficient is

- $b_1$ times the number of partitions of $[n]$ into 1 part with an $A$-structure plus
- $b_2$ times the number of partitions of $[n]$ into 2 parts with an $A$-structure plus
- $\cdots$
- $b_k$ times the number of partitions of $[n]$ into $k$ parts with an $A$-structure plus $\cdots$

So the $x^n/n!$ coefficient of $B(A(x))$ counts the number of ways to partition $[n]$ into subsets, with an $A$-structure on each subset (**micro**), a $B$-structure on the collection of subsets (**macro**).

**Example 1.14** How many ways can $[n]$ be partitioned into an even number of subsets of size at least 2?

- $A$-structure: subset of size at least 2, thus $a_m = 1$ if $m \geq 2$, otherwise 0, then $A(x) = e^x - 1 - x$.
- $B$-structure: an even number of subsets, thus $b_k = 1$ if $k$ is even, otherwise 0. Then

$$B(y) = \sum_{k=0}^{\infty} \frac{1 + (-1)^k}{2} \frac{y^k}{k!} = \frac{1}{2}(e^y + e^{-y}) = \cosh(y) \tag{1.110}$$

- Let $y = e^x - 1 - x$, thus the generating function of this problem is $\cosh(e^x - 1 - x)$.

**Example 1.15** A permutation of $[n]$ is equivalent to partitioning $[n]$ into subsets and then arranging each subset into a cycle.

- $A$-structure: cycle. There are $(m-1)!$ cycles on a $m$-element subset so $a_m = (m-1)!$, then

$$A(x) = \sum_{m=1}^{\infty} \frac{(m-1)!x^m}{m!} = \sum_{m=1}^{\infty} \frac{x^m}{m} = \sum_{m=1}^{\infty} \int_0^x t^{m-1} dt = \int_0^x \sum_{m=1}^{\infty} t^{m-1} dt \tag{1.111a}$$

$$= \int_0^x \frac{1}{1-x} dx = -\log(1-x) = \log\left(\frac{1}{1-x}\right). \tag{1.111b}$$

- $B$-structure: just a set (just partition into any number of cycles without further actions). $B(y) = \sum_{k=1}^{\infty} y^k/k!$.
- Generating function for number of permutations is

$$B(A(x)) = e^{-\log(1-x)} = \frac{1}{1-x} = \sum_{n=0}^{\infty} n! \frac{x^n}{n!}. \tag{1.112}$$

By matching coefficients, there are $n!$ permutations of $[n]$.

**Example 1.16** What about permutations with an even number of cycles?

- $A$-structure: cycle, $A(x) = -\log(1-x)$.
- $B$-structure: $b_k = 1$ if $k$ is even, otherwise $0$, $B(y) = \cosh(y)$.
- Generating function for number of permutations is

$$B(A(x)) = \cosh(-\log(1-x)) = \frac{1}{2}\left(e^{-\log(1-x)} + e^{\log(1-x)}\right) = \frac{1}{2}\left[(1-x) + \frac{1}{(1-x)}\right] \tag{1.113a}$$

$$= 1 + \sum_{n=2}^{\infty} \frac{n!}{2} \frac{x^n}{n!} \tag{1.113b}$$

By matching coefficients, the numbers are $1$ for $n=0$, $0$ for $n=1$, and $n!/2$ otherwise.

#### 1.7.2.5 Sterling Numbers

What if we count permutations of $[n]$ weighted by $y^{\#\text{ of cycles}}$.

- $A$-structure: cycle, $A(x) = -\log(1-x)$.
- $B$-structure: $b_k = y^k$, $B(x) = 1 + \sum_{k=1} x^k y^k / k! = e^{xy}$.
- Generating function for number of permutations is

$$B(A(x)) = e^{-y\log(1-x)} = \left(\frac{1}{1-x}\right)^y \tag{1.114}$$

$y^k x^n/n!$-coefficient is $\#$ permutations of $[n]$ with k cycles, which is $c(n,k)$- (unsigned) Sterling Number of the first kind.

**Theorem 1.24**

*Summing over $k$ we have that,*

$$\sum_{k=0}^{\infty}\sum_{n=0}^{\infty} c(n,k)\frac{x^n}{n!}y^k = \left(\frac{1}{1-x}\right)^y.$$

*Setting $y=1$ and change the order of summation, we have $\sum_{n=0}^{\infty} B(n)\frac{x^n}{n!}$.*

**Example 1.17(Derangements)**

- $A$-structure: cycle of length $\geq 2$, $a_m = (m-1)!$ if $m \geq 2$ and $a_1 = 0$. $A(x) = -\log(1-x) - x$.
- $B$-structure: any partition, $B(y) = e^y$.
- Generating function for number of permutations is

$$B(A(x)) = e^{-\log(1-x)-x} = \frac{e^{-x}}{1-x}. \tag{1.115}$$

$y^k x^n/n!$-coefficient is $\#$ permutations of $[n]$ with k cycles, which is $c(n,k)$- (unsigned) Sterling Number of the first kind. Note that if we want permutations without 2-cycles we get $B(A(x)) = e^{-\log(1-x)-x^2/2} = \frac{e^{-x^2/2}}{1-x}$.

**Example 1.18(Permutations with Even/Odd Length Cycles)** What about permutations with only cycles of even length?

- $A$-structure: cycle of even length $a_m = (m-1)!$ if $m$ is even otherwise $0$. We define $a_0 = 1$, then $A(x) = \sum_{m \text{ even}} x^m/m$, and

$$A(x) = \sum_{m \text{ even}} (m-1)!\frac{x^m}{m!} = \sum_{m=0}^{\infty} \frac{1+(-1)^m}{2}(m-1)!\frac{x^m}{m!} = \frac{1}{2}A(x) + \frac{1}{2}A(-x) \tag{1.116a}$$

$$= \frac{1}{2}\left[\log\left(\frac{1}{1-x}\right) + \log\left(\frac{1}{1+x}\right)\right] = \log\left(\frac{1}{\sqrt{1-x^2}}\right). \tag{1.116b}$$

- $B$-structure: any partition, $B(y) = e^y$.
- Thus, the generating function for permutations with even length cycles is

$$B(A(x)) = \sum_{n=0}^{\infty} \mathrm{EVEN}(n)\frac{x^n}{n!} = \frac{1}{\sqrt{1-x^2}}. \tag{1.117}$$

What about permutations with only cycles of odd length?

- $A$-structure: cycle of odd length $a_m = (m-1)!$ if $m$ is odd otherwise $0$. We define $a_0 = 1$, then $A(x) = \sum_{m \text{ odd}} x^m/m$, and

$$A(x) = \sum_{m \text{ odd}} (m-1)!\frac{x^m}{m!} = \sum_{m=0}^{\infty} \frac{1-(-1)^m}{2}(m-1)!\frac{x^m}{m!} = \frac{1}{2}A(x) - \frac{1}{2}A(-x) \tag{1.118a}$$

$$= \frac{1}{2}\left[\log\left(\frac{1}{1-x}\right) + \log\left(\frac{1}{1+x}\right)\right] = \log\left(\sqrt{\frac{1-x}{1+x}}\right). \tag{1.118b}$$

- $B$-structure: any partition, $B(y) = e^y$.
- Thus, the generating function for permutations with even length cycles is

$$B(A(x)) = \sum_{n=0}^{\infty} \mathrm{ODD}(n)\frac{x^n}{n!} = \frac{1+x}{\sqrt{1-x^2}}. \tag{1.119}$$

Then we know

$$\sum_{n=0}^{\infty} \mathrm{ODD}(n)\frac{x^n}{n!} = (1+x)\sum_{n=0}^{\infty} \mathrm{EVEN}(n)\frac{x^n}{n!} \tag{1.120a}$$

$$= \mathrm{EVEN}(0) + \mathrm{EVEN}(0)x + \mathrm{EVEN}(2)x^2/2! + \mathrm{EVEN}(2)x^3/2! \tag{1.120b}$$

$$+ \mathrm{EVEN}(4)x^4/4! + \mathrm{EVEN}(4)x^5/4! + \cdots \tag{1.120c}$$

Thus

- $\mathrm{ODD}(2m) = \mathrm{EVEN}(2m)$.
- $\mathrm{ODD}(2m+1)/(2m+1)! = \mathrm{EVEN}(2m)/(2m)!$.
- $\mathrm{ODD}(2m+1) = (2m+1)\mathrm{EVEN}(2m)$

And the coefficients

$$(1-x^2)^{-1/2} = \sum_{m=0}^{\infty}\left(-\frac{1}{2}\right)_m \frac{(-x^2)^m}{m!} \tag{1.121a}$$

$$= \sum_{m=0}^{\infty}\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\cdots\left(-\frac{(2m-1)}{2}\right)\frac{(-1)^m x^{2m}}{m!} \tag{1.121b}$$

$$= \sum_{m=0}^{\infty}\left(\frac{1}{2}\right)\left(\frac{3}{2}\right)\cdots\left(\frac{(2m-1)}{2}\right)\frac{x^{2m}}{m!} \tag{1.121c}$$

$$= \sum_{m=0}^{\infty}(2m-1)(2m-3)\cdots(1)\frac{(2m)!}{(2m)(2m-2)(2m-4)\cdots(4)(2)}\frac{x^{2m}}{(2m)!} \tag{1.121d}$$

$$= \sum_{m=0}^{\infty}(2m-1)^2(2m-3)^2\cdots(1)^2\frac{x^{2m}}{(2m)!} \tag{1.121e}$$

Thus $\mathrm{EVEN}(2m) = (2m-1)^2(2m-3)^2\cdots(1)^2$.

### 1.7.3 Exercises

**Exercise 1.13(Finite Differences of $1/x$)** For $n$ a positive integer, give a formula for

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

Hint: Integrate the binomial theorem.

**Proof** Note that by binomial theorem,

$$(1+x)^n = \sum_{k=0}^{n}\binom{n}{k}x^k\,, \tag{1.122}$$

then by integrating both sides, we have

$$\sum_{k=0}^{n}\frac{(-1)^k}{k+1}\binom{n}{k} = -\sum_{k=0}^{n}\frac{(-1)^{k+1}}{k+1}\binom{n}{k} = \sum_{k=0}^{n}\binom{n}{k}\left(\frac{x^{k+1}}{k+1}\right)\Bigg|_{-1}^{0} = \sum_{k=0}^{n}\binom{n}{k}\int_{-1}^{0}x^k\,dx \tag{1.123a}$$

$$= \int_{-1}^{0}\sum_{k=0}^{n}\binom{n}{k}x^k\,dx = \int_{-1}^{0}(1+x)^n\,dx = \frac{1}{n+1}(1+x)^{n+1}\Big|_{-1}^{0} = \frac{1}{n+1}\,. \tag{1.123b}$$

**Proof** [Alternative Proof] Let $a_k = \frac{(-1)^k}{k+1}\binom{n}{k}$, then the generating function of $\{a_k\}_{k\ge 0}$ is defined by

$$F(x) = \sum_{k=0}^{n}a_k x^k = \sum_{k=0}^{n}\frac{(-1)^k}{k+1}\binom{n}{k}x^k\,, \tag{1.124}$$

where $F(1)$ is the desired result, then

$$xF(x) = \sum_{k=0}^{n}\frac{(-1)^k}{k+1}\binom{n}{k}x^{k+1}\,. \tag{1.125}$$

Differentiate both side with respect to $x$, then

$$\frac{d}{dx}[xF(x)] = \sum_{k=0}^{n}(-x)^k\binom{n}{k} = (1-x)^n\,, \tag{1.126}$$

where the last equality holds according to Binomial theorem. We then integrate both side of the equation above with respect to $x$ from 0 to 1, then

$$F(1) = \int_0^1(1-x)^n\,dx = -\frac{(1-x)^{n+1}}{n+1}\Bigg|_0^1 = \frac{1}{n+1}\,. \tag{1.127}$$

✍ **Exercise 1.14(Linear Homogeneous Recurrence Relations)** We say that a sequence of numbers $a_0, a_1, a_2, \ldots$ satisfies a **linear homogeneous recurrence relation with constant coefficients** if there exists a positive integer $k$ and real numbers $c_1, c_2, \ldots, c_k$ so that for all sufficiently large integers $n$

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \ldots + c_k a_{n-k}.$$

Show that a sequence of real numbers $a_0, a_1, a_2, \ldots$ satisfies a linear homogeneous recurrence relation with constant coefficients if and only if the corresponding generating function

$$A(x) := \sum_{n=0}^{\infty} a_n x^n$$

is a rational function (i.e. is the ratio of two polynomials in $x$).

**Proof** ($\Longrightarrow$) Suppose that $a_0, a_1, a_2, \ldots$ satisfies a linear homogeneous recurrence relation with constant coefficients $c_1, \cdots, c_k$, that is, there exists some sufficiently large $N \in \mathbb{N}$ such that

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \ldots + c_k a_{n-k}\,, \quad \forall n \geq N\,. \tag{1.128}$$

Denote $P_N(x) := \sum_{n=0}^{N} a_n x^n$. We substitute recurrent relationship for $a_n$'s in $A(x)$, then we obtain

$$A(x) := \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{N-1} a_n x^n + \sum_{n=N}^{\infty} a_n x^n \tag{1.129a}$$

$$= \sum_{n=0}^{N-1} a_n x^n + \sum_{n=N}^{\infty} (c_1 a_{n-1} + c_2 a_{n-2} + \ldots + c_k a_{n-k}) x^n \tag{1.129b}$$

$$= \sum_{n=0}^{N-1} a_n x^n + c_1 x \sum_{n=N}^{\infty} a_{n-1} x^{n-1} + c_2 x^2 \sum_{n=N}^{\infty} a_{n-2} x^{n-2} + \ldots + c_k x^k \sum_{n=N}^{\infty} a_{n-k} x^{n-k} \tag{1.129c}$$

$$= \sum_{n=0}^{N-1} a_n x^n + c_1 x \sum_{n=N-1}^{\infty} a_n x^n + c_2 x^2 \sum_{n=N-2}^{\infty} a_n x^n + \ldots + c_k x^k \sum_{n=N-k}^{\infty} a_n x^n \tag{1.129d}$$

$$= P_{N-1}(x) + c_1 x[A(x) - P_{N-2}(x)] + \cdots + c_{k-1} x^{k-1}[A(x) - P_{N-k}(x)] + c_k x^k [A(x) - P_{N-k-1}(x)] \tag{1.129e}$$

$$= A(x)(c_1 x + c_2 x^2 + \cdots + c_{k-1} x^{k-1} + c_k x^k) + P_{N-1}(x) - [c_1 x P_{N-2}(x) + \cdots + c_k x^k P_{N-k-1}(x)] \tag{1.129f}$$

Therefore,

$$A(x) = \frac{P_{N-1}(x) - [c_1 x P_{N-2}(x) + \cdots + c_k x^k P_{N-k-1}(x)]}{1 - (c_1 x + c_2 x^2 + \cdots + c_{k-1} x^{k-1} + c_k x^k)}\,, \tag{1.130}$$

which is rational function.

($\Longleftarrow$) Suppose that $A(x)$ is rational function, i.e.,

$$A(x) = \frac{L(x)}{B(x)}\,, \quad \text{where } L(x) = \sum_{n=0}^{d} l_n x^n\,,\; B(x) = \sum_{n=0}^{k} b_n x^n\,, \tag{1.131}$$

then

$$A(x)B(x) = L(x) \implies \left(\sum_{n=0}^{\infty} a_n x^n\right)\left(\sum_{n=0}^{m} b_n x^n\right) = \sum_{n=0}^{\infty}\left(\sum_{i=0}^{n} a_{n-i} \cdot b_i\right) x^n = \sum_{n=0}^{d} l_n x^n\,, \tag{1.132}$$

meaning that $l_n = \sum_{i=0}^{n} a_{n-i} \cdot b_i$ for all $n \in \mathbb{N}$. However, $L(x)$ only has finitely many terms and $l_n = 0$ when $n > d$, hence

$$\sum_{i=0}^{n} a_{n-i} \cdot b_i = 0\,, \quad \forall n > d\,. \tag{1.133}$$

With out loss of generality, we assume $b_0 \neq 0$ (otherwise we consider $B'(x) = B(x)/x$ instead, which will give us non-zero first term). For all sufficiently large integers $i$, with $c_i := -b_i/b_0$, we have $c_i = 0$ when $i > k$ since $B(x)$ only have finitely many terms, then for sufficiently large $n$ we have

$$a_n = \sum_{i=1}^{n} a_{n-i}\Big(-\frac{b_i}{b_0}\Big) = \sum_{i=1}^{k} a_{n-i}c_i\,. \tag{1.134}$$

**Exercise 1.15(Semi-Increasing Sequences)** Define a sequence $a_1, a_2, \dots, a_n$ to be *semi-increasing* if $a_j \geq a_i - 1$ for all $j > i$. Define $sem(n,k)$ to be the number of semi-increasing sequences of length $n$ consisting of integers from 1 to $k$. Determine (as a function of $k$) the generating function

$$F_k(x) = \sum_{n=0}^{\infty} sem(n,k)x^n.$$

Hint: Proceed by induction on $k$. Consider the first occurrence (if one exists) of $k$ in the semi-increasing sequence counted by $sem(n,k)$.

**Proof** Following the hint, we will prove by induction on $k$ that the generating function $F_k(x) = \frac{1}{1-x}\left(\frac{1-x}{1-2x}\right)^{k-1}$. If $k = 1$, since our semi-increasing sequences of length $n$ will only consist of integer 1, there is only one choice. Hence, the generating function in this case will be

$$F_1(x) = \sum_{n=1}^{\infty} x^n = \frac{1}{1-x} \tag{1.135}$$

Hence, our base case is verified. For inductive step, assume the generating function $F_{k-1}(x)$ is $\frac{1}{1-x}\left(\frac{1-x}{1-2x}\right)^{k-2}$. We want to calculate the generating function $F_k(x)$. We observe that we have the following recursive relation. Suppose we are given any length $n$ semi-increasing sequence consisting of integers from 1 to $k$, we record the first occurrence of $k$ in the sequence. Then, the numbers in front of $k$ form an semi-increasing sequence consisting of integers from 1 to $k-1$. Moreover, numbers after that $k$ can only be $k$ or $k-1$. If $k$ does not appear, this would simply be a semi-increasing sequence of length $n$ consisting of integers $1, \dots, k-1$ (counted by $sem(n, k-1)$). Hence, we conclude that

$$sem(n,k) = sem(n,k-1) + \sum_{i=1}^{n} sem(i-1,k-1)2^{n-i} \tag{1.136}$$

Therefore,

$$\begin{aligned}
F_k(x) &= \sum_{n=0}^{\infty} sem(n,k)x^n && (1.137)\\
&= \sum_{n=0}^{\infty}\left(sem(n,k-1)x^n + \sum_{i=1}^{n} sem(i-1,k-1)2^{n-i}x^n\right) && (1.138)\\
&= \sum_{n=0}^{\infty}\left(sem(n,k-1)x^n + \sum_{j=0}^{n-1} sem(j,k-1)2^{n-j-1}x^n\right) && (1.139)\\
&= F_{k-1}(x) + \sum_{n=0}^{\infty}\sum_{j=0}^{n-1} sem(j,k-1)2^{n-j-1}x^n && (1.140)
\end{aligned}$$

Using the product rule for infinite series, the double summation above can be rewritten as

$$F_{k-1}(x)+\sum_{n=0}^{\infty}\sum_{j=0}^{n-1} sem(j,k-1)2^{n-j-1}x^n = F_{k-1}(x)+\sum_{n=0}^{\infty} x\sum_{j=0}^{n-1} sem(j,k-1)2^{n-1-j}x^{n-1} \tag{1.141}$$

$$= F_{k-1}(x)+x\left(\sum_{n=0}^{\infty} sem(n,k-1)x^n\right)\left(\sum_{n=1}^{\infty}2^n x^n\right) \tag{1.142}$$

$$= F_{k-1}(x)\times\left(\frac{x}{1-2x}+1\right) \tag{1.143}$$

$$= F_{k-1}(x)\times\left(\frac{1-x}{1-2x}\right) \tag{1.144}$$

$$= \frac{1}{1-x}\left(\frac{1-x}{1-2x}\right)^{k-1} \qquad \text{(inductive hypothesis)}$$

We can also arrive at this conclusion more directly as follows. As observed above a semi-increasing sequence of length $n$ using the numbers $1,2,\ldots,k$ is either a semi-increasing sequence using only $1,2,\ldots,k-1$ or it is a semi-increasing sequence using the numbers $1,2,\ldots,k-1$ followed by an arbitrary sequence of $k$'s and $k-1$'s that starts with a $k$. The generating function for the former term is $F_{k-1}(x)$. For the latter, we need to count the number of ways to find (A) a semi-increasing sequence using numbers $1,2,\ldots,k-1$ and (B) a sequence using $k-1$'s and $k$'s starting with a $k$ so that the sum of their lengths is equal to $n$. This is asking for the number of ways to find an object of type A and an object of type B the sum of whose sizes (lengths) equals $n$. Using the combinatorial interpretation of a product of generating functions, the generating function for this is given by the product of the generating function for objects of type A and the generating function for objects of type B. This is

$$F_{k-1}(x)\left(\sum_{n=1}^{\infty}2^{n-1}x^n\right) = F_{k-1}(x)(x/(1-2x)).$$

**Exercise 1.16(Partition Identity)** Use generating functions to prove that $p(n)-p(n-1)$ is the number of partitions of $n$ into parts of size bigger than 1.

**Proof** Recall that the generating function of integer partition is given by

$$\sum_{n=0}^{\infty}p(n)x^n=\prod_{k=1}^{\infty}\frac{1}{1-x^k} \tag{1.145}$$

This is explained by the product interpretation of generating function and notice that

$$\prod_{k=1}^{\infty}\frac{1}{1-x^k}=(1+x+x^2+\ldots)(1+x^2+x^4+\ldots)\ldots \tag{1.146}$$

Hence, the using the same argument, the generating function for the integer partitions of $n$ into parts of size bigger than 1 is given by

$$\prod_{k=2}^{\infty}\frac{1}{1-x^k} \tag{1.147}$$

But notice that

$$\prod_{k=2}^{\infty}\frac{1}{1-x^k} = (1-x)\prod_{k=1}^{\infty}\frac{1}{1-x^k} \tag{1.148}$$

$$= (1-x)\sum_{n=0}^{\infty}p(n)x^n \tag{1.149}$$

$$= \sum_{n=0}^{\infty}p(n)x^n - \sum_{n=0}^{\infty}p(n)x^{n+1} \tag{1.150}$$

$$= \sum_{n=0}^{\infty}p(n)x^n - \sum_{n=1}^{\infty}p(n-1)x^n \tag{relabel}$$

$$= \sum_{n=0}^{\infty}[p(n)-p(n-1)]x^n \tag{$p(-1)=0$}$$

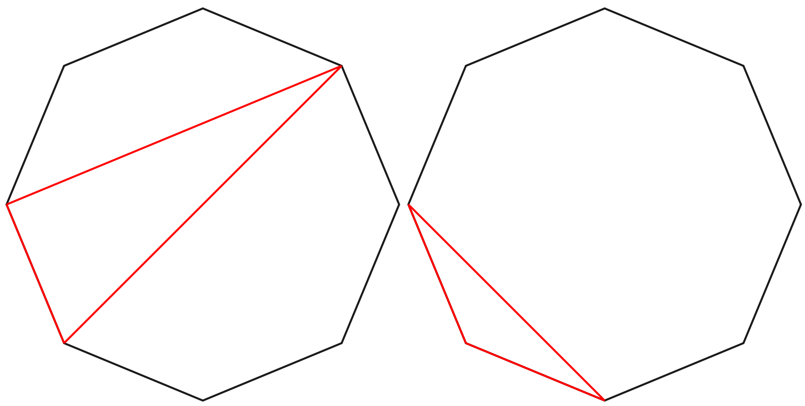
Hence, it follows from the definition of generating function that the number of partitions of $n$ into parts of size bigger than 1 is $p(n)-p(n-1)$.

✍ **Exercise 1.17(Triangulations)** Define a *triangulation* of a polygon to be a way of drawing line segments between some pairs of its non-adjacent vertices so that

1. No two of these line segments cross (except possibly at endpoints).
2. These line segments divide the interior of the polygon into triangles.

Prove that the number of triangulations of a convex $n$-gon is the Catalan number $C_{n-2}$. Hint: Show that it satisfies the recurrence relation for the Catalan numbers.

**Proof** Any polygon has to have at least three edges. Denote the number of triangulation of a $n$-gon by $t(n)$. For a triangle (3-gon) it is clear that there is only one possible triangulation (you don't connect any vertices as all of them are adjacent to each other). Hence, $t(3)=1$. Suppose we have an $(k+1)$-gon. Let we fixed one of the edge. Then, for any vertices not included in that specific edge, we can construct a unique triangle. The construction is illustrated below with the chosen edge and vertex marked by red:



We have two cases, illustrated in above diagram. If we are in the first case, we have split the $(k+1)$-gon three area: a triangle, a $(i+1)$-gon (to the right of the triangle), and a $(k-i)$-gon, where $i$ is the location of the chosen vertex. Hence, the total number of triangulation in this configuration can be calculated inductively using the product rule:

$$\# = t(i+1)t(k-i) \tag{1.151}$$

If we are in the second case, the $(k+1)$-gon is splited into 2 area: a triangle and a $(k-1)$-gon. Hence, the number of triangulation in this configuration is

$$\# = t(k-1) \tag{1.152}$$

With our above discussions, we now prove by induction that $t(n) = C_{n-2}$. The base case is clear as $t(3) = 1 = C_1$. Suppose $t(i)$ is the Catalan number $C_{i-2}$ for all $i < k$. Then, we observe that by our above discussion

$$t(k) = t(k-1) + t(k-1) + \sum_{i=3}^{k-2} t(i)t(k+1-i) \tag{1.153}$$

$$= C_0 C_{k-3} + C_{k-3} C_0 + \sum_{i=3}^{k-2} C_{i-2} C_{k-1-i} \tag{1.154}$$

$$= C_0 C_{k-3} + C_{k-3} C_0 + \sum_{j=1}^{k-4} C_j C_{k-3-j} \tag{relabel}$$

$$= \sum_{i=0}^{k-3} C_i C_{k-3-i} \tag{1.155}$$

$$= C_{k-2} \tag{1.156}$$

by the recursive definition of Catalan number. This completes our inductive step.

✍ **Exercise 1.18(Colored Compositions)**

(a) Let $a_n$ be the number of compositions of $n$ (into any number of parts) in which each part in the composition is colored either red or blue. Give a formula for the generating function

$$\sum_{n=0}^{\infty} a_n x^n.$$

**Proof** Here we are going to use the composition rule for generating functions (Theorem 8.17). Let $b_n$ be the structure of coloring the elements of an $n$-element set by red or blue (to put on the set of intervals) and $G(x)$ be the generating function:

$$G(x) = \sum_{n=0}^{\infty} b_n x^n = \sum_{n=0}^{\infty} 2^n x^n = \frac{1}{1-2x} \tag{1.157}$$

Let $F(x)$ be the generating function for the trivial structure (count every $n$-element set by 1 as we are not putting any structure within any interval), given by

$$F(x) = x + x^2 + \cdots = \frac{x}{1-x} \tag{1.158}$$

(to use the theorem we need $a_0 = 0$). Hence, the desired generating function is the composition

$$G(F(x)) = \frac{1}{1 - \frac{2x}{1-x}} = \frac{1-x}{1-3x} \tag{1.159}$$

(This can also be calculated using theorem 8.13, where the structure $a_n$ is choose either red or blue, i.e. $a_n = 2$ for all $n$.)

(b) Using the above generating function, obtain a formula for $a_n$.

**Proof** Then, we observe that

$$\frac{1-x}{1-3x} = (1-x)\sum_{n=0}^{\infty}(3x)^n \tag{1.160}$$

$$= \sum_{n=0}^{\infty} 3^n x^n - 3^n x^{n+1} \tag{1.161}$$

$$= 1 + \sum_{n=1}^{\infty} 2\cdot 3^{n-1} x^n \tag{1.162}$$

Hence, we have

$$\begin{cases} a_0 = 1 \\ a_n = 2\cdot 3^{n-1} \quad n \geq 1 \end{cases} \tag{1.163}$$

(c) Let $b_n$ be the number of compositions of $n$ (into any number of parts) in which exactly 2 parts are colored blue. Give a formula for the generating function

$$\sum_{n=0}^{\infty} b_n x^n.$$

**Proof** Let $b_n$ be the structure we desired (to put on the set of intervals) and $G(x)$ be the generating function:

$$G(x) = \sum_{n=0}^{\infty}\binom{n}{2}x^n = \sum_{n=0}^{\infty}\frac{x^2}{2}\frac{d^2}{dx^2}(x^n) = \frac{x^2}{2}\frac{d^2}{dx^2}\left(\frac{1}{1-x}\right) = \frac{x^2}{(1-x)^3} \tag{1.164}$$

Let $F(x)$ be the generating function for the trivial structure (count every $n$-element set by 1 as we are not putting any structure within any interval), given by

$$F(x) = x + x^2 + \cdots = \frac{x}{1-x} \tag{1.165}$$

(to use the theorem we need $a_0 = 0$). Hence, the desired generating function is the composition

$$G(F(x)) = \frac{x^2/(1-x)^2}{(1-2x)^3/(1-x)^3} = \frac{x^2(1-x)}{(1-2x)^3} \tag{1.166}$$

(d) Using the above generating function, obtain a formula for $b_n$. [10 points]

**Proof** We know that

$$\frac{8}{(1-2x)^3} = \frac{d^2}{dx^2}\frac{1}{1-2x} \implies \frac{1}{(1-2x)^3} = \frac{1}{8}\sum_{n=0}^{\infty}\frac{d^2}{dx^2}(2x)^n = \frac{1}{8}\sum_{n=2}^{\infty}n(n-1)2^n x^{n-2} \tag{1.167}$$

Therefore,

$$\frac{x^2(1-x)}{(1-2x)^3} = (x^2 - x^3)\sum_{n=2}^{\infty}\frac{n(n-1)2^n}{8}x^{n-2} \tag{1.168}$$

$$= \sum_{n=2}^{\infty}\frac{n(n-1)2^n}{8}x^n - \sum_{n=2}^{\infty}\frac{n(n-1)2^n}{8}x^{n+1} \tag{1.169}$$

$$= x^2 + \sum_{n=3}^{\infty}\frac{n(n-1)2^n - (n-1)(n-2)2^{n-1}}{8}x^n \tag{1.170}$$

Hence, we have

$$\begin{cases} b_0 = b_1 = 0 \\ b_2 = 1 \\ b_n = \frac{(n-1)(n\cdot 2^{n-1}+2^n)}{8} \quad n \geq 3 \end{cases} \tag{1.171}$$

✍ **Exercise 1.19(Matchings again)** Let $a_n$ be the number of matchings on a set of size $n$. Give a way of writing the exponential generating function for $a_n$ as a composition. Use this to compute an explicit formula for $a_n$.

**Proof** Recall that a matching is defined to be a set partition with each part having size 2. In the composition of exponential generating functions, we define the inner structure by setting

$$b_n = \begin{cases} 1 & \text{if n=2} \\ 0 & \text{otherwise} \end{cases}.$$

We define the outer structure to be trivial, i.e. $c_n = 1$ for all $n$. Then the exponential generating function for the number of matchings of $[n]$ should be their composition

$$A(x) = C(B(x)),$$

where $B(x) = \sum_{n=0}^{\infty} b_n x^n/n! = x^2/2$ and $C(x) = \sum_{n=0}^{\infty} c_n x^n/n! = e^x$. In particular,

$$A(x) = e^{\frac{x^2}{2}} = \sum_{n=0}^{\infty} \frac{(x^2/2)^n}{n!} = \sum_{n=0}^{\infty} \frac{(2n)!}{2^n n!} \frac{x^{2n}}{(2n)!} = \sum_{n \text{ even}} \frac{n!}{2^{n/2}(n/2)!} \frac{x^n}{n!}.$$

Therefore,

$$a_n = \begin{cases} \frac{n!}{2^{n/2}(n/2)!} & \text{if } n \text{ is even} \\ 0 & \text{otherwise} \end{cases}.$$

Another way to write the answer is by notice that, for any integer $n$, we have

$$\frac{(2n)!}{2^n} = \frac{(1\cdot 2)(3\cdot 4)\ldots((2i-1)\cdot(2i))\ldots((2n-1)\cdot 2n)}{2^n} = n!(1\cdot 3\cdot 5\ldots(2n-1)) = n!\prod_{i=1}^{n}(2i-1).$$

So we can also say $a_{2n} = \prod_{i=1}^{n}(2i-1)$ and $a_{2n+1} = 0$ for any $n$.

**Exercise 1.20(Binary Trees)** A *binary tree* is either empty (has no nodes) or has a root node and two more binary trees known as the left and right subtrees. Letting $b_n$ be the number of binary trees with nodes labelled $1, 2, \ldots, n$ and

$$B(x) = \sum_{n=0}^{\infty} b_n x^n/n!,$$

show that

$$B(x) = 1 + x(B(x))^2.$$

Conclude that $b_n = n!C_n$.

**Proof** In order for the identity to hold, we take $b_0 = 1$ by convention. For any $n > 0$, write the coefficient of $x^n$ in $x(B(x))^2$ as $c_n/n!$. Then according to the product formula of exponential generating functions, $c_n$ is the number of ways to partition $[n]$ into three parts $P_1, P_2, P_2$, and build a structure on each of them: For $P_1$, we build a trivial structure when $|P_1| = 1$ and we build 0 structure otherwise (this is equivalent to the requirement $|P_1| = 1$); For $P_2$, there are $b_{|P_2|}$ number of ways to build its structure, this is equivalent to build a binary tree using the nodes labelled by numbers in $P_2$; For $P_3$, we also build a binary tree using the nodes labelled by numbers in $P_3$.

Given any such a partition $\{P_1, P_2, P_3\}$ and their build structures, we can construct a binary tree with nodes labelled $1, 2, \ldots, n$ by taking $P_1$ as the root, $P_2$ as the left subtree, and $P_3$ as the right subtree; Conversely, given an binary tree with nodes labelled $1, 2, \ldots, n$, we let $P_1$ be the set containing its root only, $P_2$ be the set containing the nodes in its left subtree, and $P_3$ be the set containing the nodes in its right subtree, hence $\{P_1, P_2, P_3\}$ (along with the subtrees on $B$ and $C$) is exactly a set partition counted by $c_n$. So we conclude $c_n = b_n$ for $n > 0$. Therefore we have

$$B(x) = 1 + x(B(x))^2,$$

as wanted.

We proved during the class that from the functional equation $B(x) = 1 + x(B(x))^2$, we can solve for

$$B(x) = \sum_{n=0}^{\infty} C_n x^n,$$

where $C_n$ is the $n$-th Catalan number. Comparing this expansion with the definition of $B(x)$ we conclude $b_n = n!C_n$ as wanted.

## 1.8 Pattern Avoidance in Permutations

### 1.8.1 Stack Sortable Permutations

**Definition 1.17**
*A stack is an object that stores numbers. You can push more numbers onto it, or you can pop the most recently pushed number off.* ♣

**Definition 1.18**
*A permutation $\pi$ of $[n]$ is stack-sortable if there is a series of operations that involve $1, 2, \cdots, n$ being pushed into a stack in $\pi$ order, and popped off of the stack in sorted order, i.e., the popped result is $123 \cdots n$.* ♣

**Example 1.19** For example, 312 is stack-sortable using 'push, push, pop, push, pop, pop'.

**Theorem 1.25**
*A permutation is stack sortable unless there are entries $x < y < z$ with $\pi(z) < \pi(x) < \pi(y)$, i.e., three entries where the second smallest comes before the biggest comes before the smallest.* ♡

**Proof** (Impossibility) In a permutation $\pi$, if there is some $k < i < j$ such that $i, j, k$ appear subsequently, then $\pi$ is not stack sortable.

- If $i$ is popped before $j$ pushed, $i$ is before $k$ in final ordering.
- If $i$ is popped after $j$ is pushed, $i$ must be after $j$ in the final ordering.

(Achievability) If no such $i, j, k$ exist, we can use the following algorithm: Pop the stack when the next element to be added is bigger than the one on top or if there are no more elements to be added.

- Elements in stack always sorted since we only add elements smaller than current top.
- When we pop $i$, all smaller elements have already been popped, since we pop $i$ when $j > i$ is about to be pushed, no $k < i$ remains to be pushed, $k < i$ is below $i$ on the stack.

### 1.8.2 Pattern Avoidance

**Definition 1.19**
*Given two permutations $\pi$ of $[n]$ and $\rho$ of $[m]$, we say there is a copy of $\rho$ in $\pi$ if there are $1 \leq x_1 < x_2 < \cdots < x_m \leq n$ so that $\pi(x_1), \pi(x_2), \cdots, \pi(x_m)$ have the same relative orders as $\rho(1), \rho(2), \cdots, \rho(m)$.* ♣

**Example 1.20** A permutation $\pi$ is stack sortable if and only if it doesn't have a copy of 231.

**Definition 1.20 (Pattern Avoidance)**
*For a permutation $\rho$, let $S_n(\rho)$ be the set of permutations of $[n]$ that do not have a copy of $\rho$.* ♣

**Example 1.21** For example, $|S_n(231)|$ is the number of stack sortable permutations of $[n]$.

**Theorem 1.26**
$|S_n(231)| = C_n$. ♡

**Proof** Let $A_n = |S_n(231)|$. We will show that $A_n$ satisfies the same recurrence as $C_n$. Namely, $A_0 = 1$ and $A_n = \sum_{k=1}^{n} A_{k-1}A_{n-k}$. Consider $n$ in $k$th location. Entries on left must be smaller than entries on right,

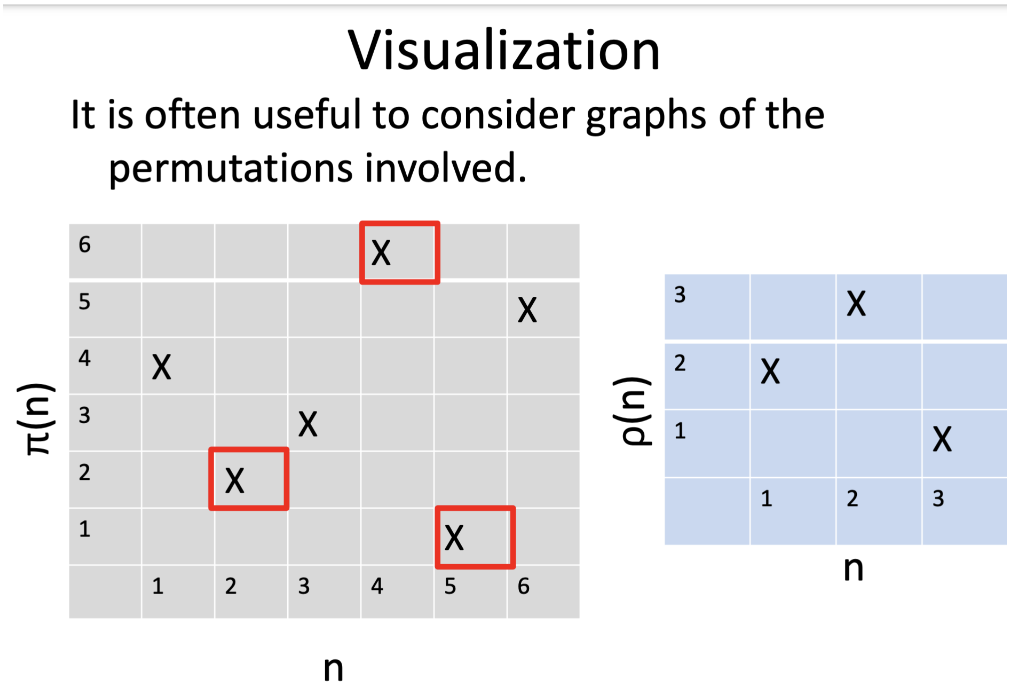



**Figure 1.6:** Permutation Graph and Copy

otherwise violating 231. Thus entries on left are $1, 2, \cdots, k-1$ and entries on right are $k+1, k+2, \cdots, n$. Meanwhile, entries on left/right must be 231-avoiding, hence the number of possibilities is $A_{k-1}A_{n-k}$. Summing over $k$ gives the full value of $A_n$.

**Definition 1.21**

*Complement of an $n$-permutation $p = p(1)p(2)\cdots p(n)$ to be the $n$-permutation $\bar{p}$ whose first entry is $n+1-p(1)$, whose second entry is $n+1-p(2)$, and in general, whose $i$th entry is $n+1-p(i)$. So for example, the complement of* 34152 *is* 32514. ♣

**Lemma 1.8**

$|S_n(231)| = |S_n(132)| = |S_n(312)| = |S_n(213)|$. ♡

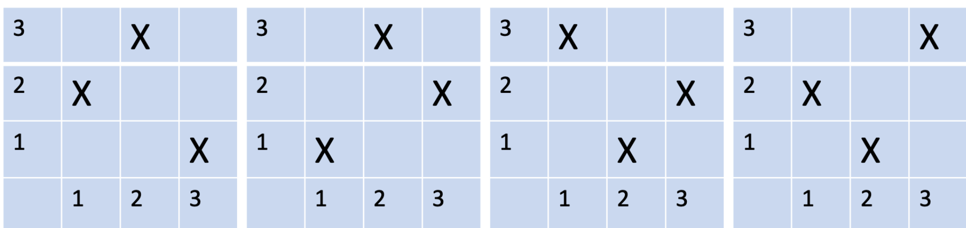
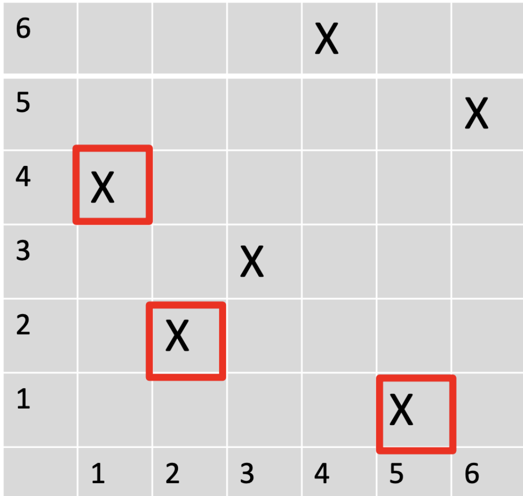


**Figure 1.7:** Rotations of Permutation Graph

**Proof** Note that the diagrams for these permutations are rotations of each other. For a permutation $\pi$, let $R\pi$ be its 90 degree rotation. $\pi$ contains a copy of $\rho$ iff $R\pi$ contains $R\rho$. Therefore, for any $\rho$, $S_n(\rho) = S_n(R\rho)$ since $R\pi \in S_n$.

**Definition 1.22**

***Left-to-Right minima*** *of a permutation $\pi$ are all of the indices $i$ so that $\pi(i) < \pi(j)$ for all $j < i$.* ♣



**Definition 1.23 (Valid Pairs)**

*A set of pairs $(i, j)$ with $1 \leq i, j \leq n$ is valid if:*

1. *They can be ordered $(i_1, j_1), (i_2, j_2), \cdots, (i_k, j_k)$ so that $i_1 < i_2 < \cdots < i_k$ and $j_1 > j_2 > \cdots > j_k$.*
2. *For each $1 \leq a \leq n$, if $i_{last}$ is the largest of the $i$'s less than or equal to $a$, $j_{last} \leq n + 1 - a$.* ♣

**Lemma 1.9**

*A set of pairs can be the set of left-to-right minima of a permutation of $[n]$ only if it is valid.* ♡

**Proof**

1. Condition (1) must hold since left-to-right minima must be decreasing from left-to-right because they are minima.
2. Condition (2) must hold since $\pi(1), \pi(2), \cdots, \pi(a)$ must all be at least $j_{last}$.

**Lemma 1.10**

*Let $\pi \in S_n$ be a 123-avoiding permutation. The pairs $(i, j)$ in $\pi$, which are not left-to-right minima, are in decreasing order.* ♡

**Proof** Proof by contradiction. Suppose this isn't the case.

- We have $(i_1, j_1)$ and $(i_2, j_2)$ which are not left-to-right minima with $i_1 < i_2$ and $j_1 < j_2$.
- Since $(i_1, j_1)$ is not a minima, there is another $(i_0, j_0)$ with $i_0 < i_1$ and $j_0 < j_1$.
- Thus $(i_0, j_0)$, $(i_1, j_1)$, $(i_2, j_2)$ is a copy of 123.

Therefore, for every valid set, there is exactly one 123-avoiding permutation with those left-to-right minimums.

- Only one since the other terms must come in decreasing order.
- Condition (2) implies none of these will be a new minimum.

- Left-to-right minimums decreasing, others decreasing, therefore cannot have three term increase.

**Proposition 1.1 ([Miklos2017AWalkTC] )**

*For every valid set $S$ of pairs, there is exactly one 123-avoiding permutation and exactly one 132-avoiding permutation with $S$ as its set of left-to-right minima, that is $|S_n(123)| = |S_n(132)|$ for all $n \in \mathbb{N}$.*

**Remark** It is clear by taking reverses, or by taking complements, that $|S_n(123)| = |S_n(321)|$. Therefore, it's enough to consider $S_n(123)$.

**Proof** We will construct a bijection f from the set of all 123-avoiding $n$-permutations onto the set of all 132-avoiding $n$- permutations which leaves all left-to-right minima fixed. (This last property is not needed for the proof of our Lemma, but it will be useful later.)

The bijection $f$ is defined as follows. We take any 123-avoiding $n$- permutation $p$, fix all its left-to-right minima, and remove all the elements that are not left-to-right minima, leaving their places empty. Then going from the left to the right, we put the elements which are not left-to-right minima into the empty slots between the left-to-right minima so that in each step we place the smallest element we have not placed yet which is larger than the previous left-to-right minimum. In other words, in each step, we place the smallest entry that is both available (that is, it is not a left-to-right minimum) and eligible (that is, it is not smaller than the previous left-to-right minimum). The reader is invited to verify that there is always at least one such entry, so the process will never get stuck.

Note that f($\rho$) is 132-avoiding, because if there were a 132-pattern in $f(\rho)$, then there would be one which starts with a left-to-right minimum, but that is impossible as elements larger than any given left-to-right minimum and to the right of it are written in increasing order.

The inverse of $f$ is even easier to describe: keep the left-to-right minima of $\rho$ fixed and put all the other elements into the empty slots between them in decreasing order. Note that this procedure will not change the set of left-to-right minima of p (why?). We obtain a permutation which is the union of two decreasing subsequences and is therefore 123-avoiding. If we apply this operation to $f(\rho)$, then we must get $\rho$ back, as the left-to-right minima have not changed, and the other elements must have been in decreasing order in $\rho$, otherwise $\rho$ would not have been 123-avoiding. This completes the proof of the lemma.

**Theorem 1.27**

*Let $\rho$ be any permutation pattern of length three. Then for all positive integers $n$, $|S_n(\rho)| = C_n = \binom{2n}{n}/(n+1)$.*

### 1.8.3 Pattern Comparison

**Theorem 1.28 ([Miklos2017AWalkTC] )**

*For all $n \geq 7$, the inequality $|S_n(1234)| < |S_n(1324)|$ holds.*

### 1.8.4 Exercises

**Exercise 1.21(Stack-Sortable Permutations)** It is not hard to show directly that the number of stack-sortable permutations of $[n]$ is given by the $n^{th}$ Catalan number. In particular, if a permutation of $[n]$ is stack sortable, there is some pattern of push and pop operations needed to sort it (a push adds the next element in line onto the stack and a pop removes the element on the top of the stack). Find a bijection between such sequences of pushes and pops and lattice paths from $(0,0)$ to $(n,n)$ that stay above the line $x=y$. Show that each such pattern of pushes and pops corresponds to exactly 1 unique stack-sortable permutation.

**Proof** Let $A$ denote the set of stack-sortable permutations of $[n]$, $B$ denote the set of all "meaningful" push-pop patterns with length $n$ (that is, a sequence of $n$ push and $n$ pop operations such that the number of appeared pushes is not less than the number of appeared pops at any time), and $C$ denote the set of lattice paths from $(0,0)$ to $(n,n)$ that stay above the line $x=y$.

First we find a bijection between $B$ and $C$. We define a function $f: B \to C$ as follows: given a push-pop pattern $x$, we construct a lattice path $f(x)$, by reading through each operation of $x$. We let the path start at $(0,0)$, go up when we encounter a push, and go down when we encounter a pop. Within $x$, the number of appeared pushes is not less than the number of appeared pops at any time. This guarantees that $f(x)$ always stays above the line $x=y$, which means $f$ is well-defined.

To see $f$ is a bijection, we can define its inverse $h: C \to B$ as follows: given a lattice path $y$, we construct a push-pop pattern $g(y)$, by following the path from its start. We add a push to our pattern when the path goes up and add a pop to our pattern when the path goes down. Similarly, the condition that $y$ stays above the line $x=y$ guarantees that $g(y)$ has the number of appeared pushes not less than the number of appeared pops at any time. So $g(y)$ is indeed a meaningful push-pop pattern and $g$ is well-defined.

Next we find a bijection between $B$ and $A$. We define a function $h: B \to A$ as follows: given a push-pop pattern $x$, we apply this push-pop pattern to the identity permutation, i.e. the sequence $1,2,3,\ldots,n$, and we call the resulting sequence $\pi_x(1),\pi_x(2),\ldots,\pi_x(n)$. Notice that $\pi_x$ is a permutation of $[n]$, and given any permutation $p$, written as a sequence $p(1),p(2),\ldots,p(n)$, if we apply the pattern $x$ to $p$, the resulting sequence is

$$p(\pi_x(1)),p(\pi_x(2)),\ldots,p(\pi_x(n)).$$

In particular, we set $h(x)=\pi_x^{-1}$, and the push-pop pattern $x$ takes $h(x)$ to be the identity permutation. So $h(x)$ is stack-sortable and $h$ is well-defined.

We show that $h$ is injective. Given $x_1 \neq x_2 \in B$, suppose the first position they differ is the $k$-th position, i.e. the $i$-th operation of $x_1$ and $x_2$ are the same for all $i<k$, and the $k$-th operations of $x_1$ and $x_2$ are different. Without loss of generality, let's assume the $k$-th position of $x_1$ is a pop but the that of $x_2$ is a push. Apply both patterns to the identity permutation, suppose after the $(k-1)$-th operation, the number at the top of the stack is $\alpha$, and the number waiting to be pushed is $\beta$. Then in $\pi_{x_1}$, $\alpha$ comes before $\beta$, since $\alpha$ will be pop-ed before $\beta$ gets into the stack. But in $\pi_{x_2}$, $\beta$ comes before $\alpha$, since $\beta$ will be on top of $\alpha$ after the $k$-th operation. Hence $\pi_{x_1} \neq \pi_{x_2}$, and $h(x_1)=\pi_{x_1}^{-1} \neq \pi_{x_2}^{-1}=h(x_2)$.

We show that $h$ is surjective. Given $y \in A$, by definition it's stack-sortable, so there exits a pattern $x$ such that applying this pattern to $y$ we obtain the identity permutation. However, by our analysis in the paragraph where we defined $h(x)$, we know that when we apply $x$ to any permutation $p$, we get the sequence

$$p(\pi_x(1)),p(\pi_x(2)),\ldots,p(\pi_x(n)).$$

There's a **only one** permutation $p$ making this sequence the identity permutation, which is $\pi_x^{-1}$. On the other hand, we know if we apply $x$ to the permutation $y$, we get the identity permutation. This means $y=\pi_x^{-1}=h(x)$, i.e. $y$ is indeed in the range of $h$.

**Exercise 1.22(Packing Patterns)** Show that for every positive integer $n$ there is a permutation $\pi$ of $[n^2]$ so that for *every* permutation $\rho$ of $[n]$, there is a copy of $\rho$ inside $\pi$.

**Proof** We define $\pi$ be describing its standard notation. The standard notation of $\pi$ the the concatenation of $n$

sequences, where the $i$-th sequence is

$$n-i+1, 2n-i+1, 3n-i+1, \ldots, n^2-i+1.$$

For example, when $n$ is 4, the standard notation of $\pi$ is

$$4, 8, 12, 16, 3, 7, 11, 15, 2, 6, 10, 14, 1, 5, 9, 13.$$

If we plot $(i, \pi(i))$ into the $xy$-coordinated plane, we get the left of Fig 1.9

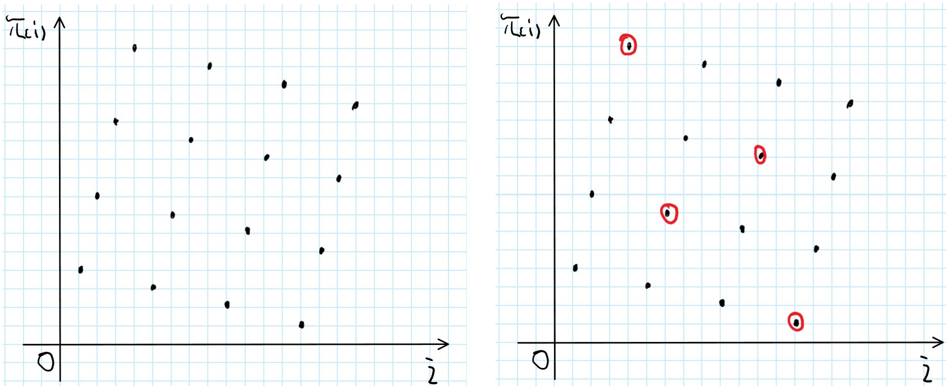



**Figure 1.8:** The sequence $\pi$ when $n=4$.(Left) Circled points represent a subsequence of $\pi$ that is a 4231-pattern.(Right)

For each $\rho$, consider the subsequence of $\pi$ defined by

$$\pi(\rho(1)), \pi(n+\rho(2)), \ldots, \pi((i-1)n+\rho(i)), \ldots, \pi((n-1)n+\rho(n)).$$

For example, when $\rho$ is the sequence $4, 2, 3, 1$, the subsequence of $\pi$ that is a $\rho$-pattern is $16, 7, 10, 1$. See the right of Fig 1.9 for example.

We wish to show this subsequence is a $\rho$-pattern. Indeed, for arbitrary pair $i \neq j$, we want to show $\rho(i) < \rho(j)$ is and only if $\pi((i-1)n+\rho(i)) < \pi((j-1)n+\rho(j))$. According to our definition of $\pi$, we know

$$\pi((i-1)n+\rho(i)) = \rho(i)n - i + 1 \quad \text{and} \quad \pi((j-1)n+\rho(j)) = \rho(j)n - j + 1.$$

Because $|(i-1)-(j-1)| < n$, then $\rho(i)n - i + 1 < \rho(j)n - j + 1$ if and only $\rho(i) < \rho(j)$, proving what we want.

## 1.9 Partial Orders

### 1.9.1 Basic Concepts

**Definition 1.24**

*Let $P$ be a set, a **partial order** is a relation $\leq$ on a $P$ satisfying three properties.*

- *$\leq$ is reflexive, that is, $x \leq x$ for all $x \in P$.*
- *$\leq$ is transitive, that is, if $x \leq y$ and $y \leq z$, then $x \leq z$.*
- *$\leq$ is is antisymmetric, that is, if $x \leq y$ and $y \leq x$, then $x = z$.*

*Then we say that $P_{\leq} = (P, \leq)$ is a partially ordered set, or **poset**.* ♣

**Example 1.22** Several examples.

- Let $P$ be the set of all subsets of $[n]$, and let $A \leq B$ if $A \subset B$. Then $P_{\leq}$ is a partially ordered set. This partially ordered set is denoted by $B_n$ and is often called a **Boolean algebra** of degree $n$.
- The set of all subspaces of a vector space, ordered by containment, is a partially ordered set.
- Let $P$ be the set of all positive integers, and let $x \leq y$ if $x$ is a divisor of $y$. Then $P_{\leq}$ is a partially ordered set.
- Let $P = \mathbb{R}$, the set of real numbers, and let $\leq$ be the traditional ordering. Then $P_{\leq}$ is a partial order, in which there are no two incomparable elements. Therefore, we also call $\mathbb{R}$ a **total order**, or **chain**.
- Let $P = \mathbb{R}^n$, and $\leq$ be domination, that is, $(x_1, \cdots, x_n) \leq (y_1, \cdots, y_n)$ if and only if $x_1 \leq y_1, \cdots, x_n \leq y_n$.

**Definition 1.25**

*Let $P$ be the set of all partitions of $[n]$. Let $\alpha$ and $\beta$ be two elements of $P$. Define $\alpha \leq \beta$ if each block of $\beta$ can be obtained as the union of some blocks of $\beta$. For instance, if $n = 6$, then $\{1,4\}\{2,3\}\{5\}\{6\} \leq \{1,4,6\}\{2,3,5\}$. Then $P_{\leq}$ is a partial order, which is often called the **refinement order**, and is denoted by $\Pi_n$.* ♣

**Definition 1.26**

*An element $x$ in a partial order $P_{\leq}$ is*

- *a **maximal** element if there is no $y$ such that $x < y$.*
- *a **maximum** element if there is $y \leq x$ for all $y \in P$.*
- *a **minimal** element if there is no $y$ such that $y < x$.*
- *a **maximum** element if there is $x \leq y$ for all $y \in P$.* ♣

**Remark** All finite posets have minimal and maximal elements. Not all finite posets have minimum or maximum elements, however.

**Definition 1.27 (Cover)**

*If $x < y$ in a poset $P$,but there is no element $z \in P$ so that $x < z < y$, then we say that $y$ covers $x$.* ♣

**Lemma 1.11**

*In any finite partial order $x \leq y$ if and only if there is some chain of elements $x = x_0, x_1, \cdots, x_n = y$ with $x_{i+1}$ covering $x_i$ sequentially.* ♡

**Proof** If $x_{i+1}$ covers $x_i$ sequentially, then by transitivity $x = x_0 < x_n = y$.

For the other direction, we proceed by induction on the number n of elements $z$ with $x < z < y$. The base case $n = 0$, which means $y$ covers $x$, then we take $x_0 = x$ and $x_1 = y$. For inductive step, we can find $x < z < y$. Fewer than $n$ elements between $x$ and $z$ or between $z$ and $y$. Inductive hypothesis gives $x = x_0 < x_1 < x_2 < \cdots < x_k = z < x_{k+1} < \cdots < x_m = y$.

**Definition 1.28 (Hasse diagram)**

*The Hasse diagram of a finite poset $P$ is a graph whose vertices represent the elements of the poset. If $x < y$ in $P$, then the vertex corresponding to $y$ is above that corresponding to $x$. If, in addition, $y$ covers $x$, then there is an edge between $x$ and $y$. Alternatively, if we want to avoid the imprecise (but intuitively obvious) notion of "above", we can say that the Hasse diagram of $P$ is the directed graph whose vertices are the elements of $P$, and in which there is an edge from $x$ to $y$ if $x$ is covered by $y$.* ♣

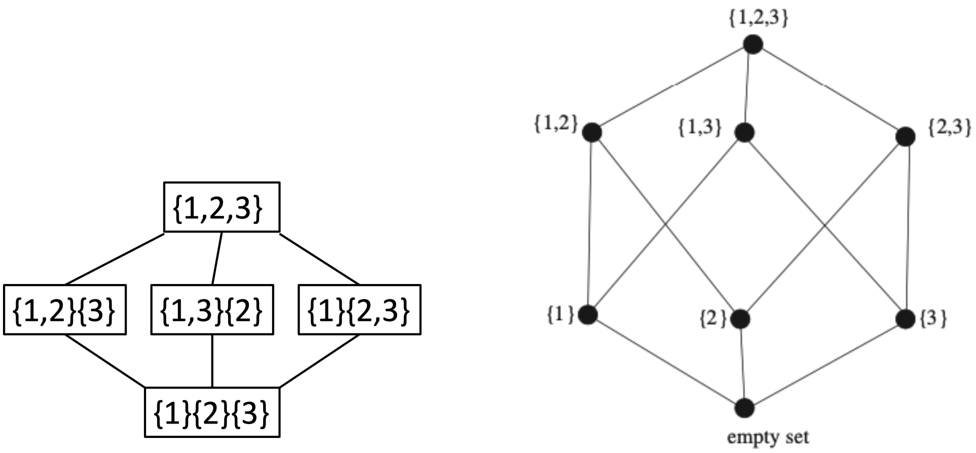



**Figure 1.9:** Hasse Diagram of $\Pi_3$(left) and $B_3$(right).

**Definition 1.29**

*Chains and Anti-Chains*

- *A **chain** in a partial order is a sequence of elements $x_1 < x_2 < \cdots < x_k$.*
- *A **chain cover** of a poset is a collection of disjoint chains whose union is the poset itself.*
- *An **anti-chain** in a partial order is a collection of elements $x_i$ so that no two of the $x_i$ are comparable.*
- *A chain, (resp. antichain) $X$ of $P$ is called **maximum** if $P$ has no larger chain (resp. antichain) than $X$, and $X$ is called **maximal** if it cannot be extended. That is, no element can be added to $X$ without destroying the chain (resp. antichain) property of $X$.*
- *Size of the largest antichain of a poset $P$ is often called the **width** of $P$.*
- *If $P$ is an $n$-element poset, then a **linear extension** of $P$ is just an order-preserving bijection $f$ from $P$ onto $[n]$. That is, if $x \leq y$ in $P$, then $f(x) \leq f(y)$ in $[n]$.* ♣

**Theorem 1.29 (Dilworth's Theorem, [Miklos2017AWalkTC])**

*In a finite partially ordered set $P$, the size of any maximum antichain is equal to the number of chains in any smallest chain cover.* ♡

### 1.9.2 Incidence algebras

If $x \leq y$ are elements of $P$, then the set of all elements $z$ satisfying $x \leq z \leq y$ is called the **closed interval** between $x$ and $y$, and is denoted by $[x, y]$. If all intervals of $P$ are finite, then $P$ is called locally finite. Note that this does not necessarily mean that $P$ itself is finite. The set of all positive integers with the usual ordering provides a counterexample.

A set of elements $I \subset P$ is called an ideal if $x \in I$ and $y \leq x$ imply $y \in I$. If an ideal is generated by one element, that is, $I = y : y \leq x$, then $I$ is called a **principal ideal**. For example, if $P = B_n$, then the ideal of all subsets of $[k]$ is a principal ideal, while the ideal of all subsets that have at most four elements is not. In some of our theorems, we will have to restrict ourselves to posets in which each principal ideal is finite, where each element is larger than a finite number of elements only. Note that this is a stronger requirement than being locally finite. The poset of all integers is locally finite, but has no finite principal ideals. Finally, we note that dual ideals, and principal dual ideals are defined accordingly.

**Definition 1.30**

*Let $P$ be a locally finite poset and $Int(P)$ be the set of all intervals of $P$. Then the **incidence algebra** $I(P)$ of $P$ is the set of all functions $f : Int(P) \mapsto \mathbb{R}$. Multiplication, which is also referred as **convolution**, in this algebra is defined by*

$$(f * g)(x, y) = \sum_{x \leq z \leq y} f(x, z)g(z, y)\,. \tag{1.172}$$

♣

**Remark** The incidence algebra can be thought of in terms of matrix multiplication, which behaves like upper triangular matrices. The function f corresponds to a matrix $F$ where $F_{x,y} = f(x, y)$ if $x \leq y$, otherwise $0$. Then composition corresponds to multiplication of the corresponding matrices.

**Corollary 1.4**

*Convolution $*$ is associative, that is $(f * g) * h = f * (g * h)$.*

♡

**Definition 1.31**

*Let $P$ be a locally finite poset. Let $\zeta \in I(P)$ be defined by $\zeta(x, y) = 1$ if $x \leq y$. Then $\zeta$ is called the **zeta** function of $P$.*

*The function $\delta$ satisfies $\delta(x, y) = 1$, if $x = y$, and $\delta(x, y) = 0$ if $x < y$. It is straightforward to verify that indeed, this function satisfies $\delta f = f\delta = f$ for all $f \in I(P)$, so it is indeed the **unit element** of $I(P)$.*

*A **multichain** in a poset is a multiset of elements $a_1, a_2, \cdots, a_m$ satisfying $a_1 \leq a_2 \leq \cdots \leq a_m$. Note that the inequalities are not strict, unlike in the definition of chains.*

♣

**Remark**

$$(f * \delta)(x, y) = \sum_{x \leq w \leq y} f(x, w)\delta(w, y) = f(x, y)\,, \tag{1.173}$$

where the only non-zero term in the sum is when $w = y$.

**Proposition 1.2 ([Miklos2017AWalkTC])**

*Let $x \leq y$ be elements of the locally finite poset $P$. Then the number of multichains $x = x_0 \leq x_1 \leq x_2 \leq \cdots, x_k = y$ is equal to $\zeta^k(x, y)$.*

♠

**Lemma 1.12**

*Let $P$ be a locally finite poset. Let $[x, y] \in Int(P)$. Then the number of chains of length $k$ that start at $x$ and end in $y$ is $(\zeta - \delta)^k(x, y)$.* ♡

**Lemma 1.13**

*For any function $f$ with $f(x, x)$ non-zero for all $x$ in $P$, there is a function $g$ so that $(f * g) = \delta$. Therefore, $g$ is called the* ***inverse*** *of $f$.* ♡

**Definition 1.32**

*The inverse of the $\zeta$ function of P is called the Möbius function of $P$, and is denoted by $\mu = \mu_P$.* ♣

**Theorem 1.30 ([Miklos2017AWalkTC])**

*Let $P$ be a locally finite poset. Let $[x, y] \in Int(P)$. Then $\mu(x, x) = 1$, and*

$$\mu(x, y) = -\sum_{x \leq z < y} \mu(x, z) \tag{1.174}$$

*if $x < y$. In other words, $\mu$ is the only function in $I(P)$ satisfying $\mu(x, x) = 1$ and $\sum_{x \leq z \leq y} \mu(x, z) = 0$.* ♡

**Proof** First, we have $1 = \delta(x, x) = (\mu\zeta)(x, x) = \mu(x, x)\zeta(x, x) = \mu(x, x)$. Second, we have

$$0 = \delta(x, y) = (\mu\zeta)(x, y) = \sum_{x \leq z \leq y} \mu(x, z)\zeta(z, y) = \sum_{x \leq z \leq y} \mu(x, z)\,, \tag{1.175}$$

if $x < y$. So the sum of $\mu(x, z)$, taken over all $z$ in a nontrivial interval $[x, y]$ is indeed 0 as we claimed.

**Corollary 1.5**

*Let $P$ be a locally finite poset. Let $[x, y] \in Int(P)$, and let us assume that $x \neq y$. Then*

$$\mu(x, y) = -\sum_{x < z \leq y} \mu(z, y)\,. \tag{1.176}$$

♡

**Proof** First, we have $1 = \delta(x, x) = (\zeta\mu)(x, x) = \zeta(x, x)\mu(x, x) = \mu(x, x)$. Second, if $x < y$, we have

$$0 = \delta(x, y) = (\zeta\mu)(x, y) = \sum_{x \leq z \leq y} \zeta(x, z)\mu(z, y) = \sum_{x \leq z \leq y} \mu(z, y)\,. \tag{1.177}$$

**Proposition 1.3**

*Let $P = B_n$, and let $S$ and $T$ be two elements of $P$, that is, two subsets of $[n]$ so that $S \subset T$. Then*

$$\mu(S, T) = (-1)^{|T| - |S|}\,. \tag{1.178}$$

♠

**Proof** Proof by induction on $k = |T| - |S|$. If $k = 0$, then $S = T$, so $\mu(S, T) = 1$ by definition, and the statement is true. Now let us assume that the statement is true for all nonnegative integers less than $k$, and let $|T| - |S| = k$. Then for all natural numbers $i$ satisfying $0 \leq i \leq k - 1$, the interval $[S, T]$ contains $\binom{k}{i}$ elements of $P$ that are $|S| + i$ element subsets $i$ of $[n]$. If $Z$ is such a subset, then it follows from the induction hypothesis that $\mu(S, Z) = (-1)^i$. Therefore, we have

$$\mu(S, T) = -\sum_{S \leq Z < T} \mu(S, Z) = -\sum_{i=0}^{k-1} \binom{k}{i} (-1)^i = (-1)^k\,, \tag{1.179}$$

where the last equality is a direct consequence of Binomial theorrem.

**Theorem 1.31 (M öbius Inversion Formula, [Miklos2017AWalkTC])**

*Let $P$ be a poset in which each principal ideal is finite, and let $f : P \mapsto \mathbb{R}$ be a function. Let the function $g : P \mapsto \mathbb{R}$ be defined by*

$$g(y) = \sum_{x \leq y} f(x)\,. \tag{1.180}$$

*Then*

$$f(y) = \sum_{x \leq y} g(x)\mu(x, y)\,. \tag{1.181}$$

♡

**Corollary 1.6 (Dual Version)**

*Let $P$ be a poset in which each principal ideal is finite, and let $f : P \mapsto \mathbb{R}$ be a function. Let the function $g : P \mapsto \mathbb{R}$ be defined by*

$$g(y) = \sum_{x \geq y} f(x)\,. \tag{1.182}$$

*Then*

$$f(y) = \sum_{x \geq y} g(x)\mu(y, x)\,. \tag{1.183}$$

♡

**Definition 1.33**

*Let $P$ and $Q$ be two posets. Then the direct product $P \times Q$ of these two posets is the poset whose elements are all the ordered pairs $(p, q)$, where $p \in P$, and $q \in Q$, and in which $(p, q) \leq (p', q')$ if $p \leq p'$ and $q \leq q'$.*

♣

# Chapter 2 Graph Theory

## 2.1 Basic Terminology

**Definition 2.1 (Graph)**

*A **graph** $G = (V, E)$ is a pair $(V, E)$, where $V$ is a set of vertices and $E$ is a collection of 2-subsets of $V$, i.e., $E = \{e = \{u, v\} | u, v \in V\}$. Let $V(G)$ denote the **vertex set** and $E(G)$ denote the **edge set**.* ♣

**Definition 2.2**

*Types of graphs*

1. *A **multigraph** is a pair $(V, E)$ where $V$ is a set and $E$ is a multiset of unordered pairs from $V$. In other words, repeated elements in our edge set are allowed.*
2. *A **pseudograph** is a pair $(V, E)$ where $V$ is a set and $E$ is a multiset of unordered multisets of size two from $V$. A pseudograph allows loops, namely edges of the form $a, a \in E$ for $a \in V$.*
3. *A **digraph** is a pair $(V, E)$ where $V$ is a set and $E$ is a multiset of ordered pairs from $V$. In other words, the edges now have a direction: the edge $(a, b)$ and edge $(b, a)$ are different, and denoted in a digraph by putting an arrow from $a$ to $b$ or from $b$ to $a$, respectively.*
   *An **orientation** of a graph $G$ is a digraph $\vec{G}$ obtained by replacing each edge $\{a, b\} \in E(G)$ with either the $arc(a, b)$ or the $arc(b, a)$. The graph $G$ is called the underlying graph of $\vec{G}$.*
4. *A **hypergaph** is a pair $(V, E)$ where edge set $E$ is collection of arbitrary subsets of vertices (rather than just 2-subsets). A hypergraph is $k$-uniform if each edge is incident to exactly $k$ vertices.*
5. ***Infinite** graphs are obtained by allowing $V$ or $E$ to be an infinite set.*
6. ***Simple graphs**: graphs without self-loops or multiple edges.* ♣

**Definition 2.3 (Adjacency, neighborhood and degree)**

*If $u$ and $v$ are two vertices of a graph $G = (V, E)$, then we say*

- *$u$ and $v$ are **adjacent** if $e = \{u, v\} \in E$, i.e., $\{u, v\}$ is an edge of $G$. Usually, we denote $(u, v) =: e = \{u, v\}$.*
- *$v$ is **incident** with edge $e$ if $v \in e$.*
- *The **neighborhood** of a vertex $v$ (denoted $N(v)$) is the set of vertices adjacent to $v$.*
- *The **degree** of $v$ (denoted $d(v)$) is the number of vertices adjacent to $v$.*
- *A graph is $d$-**regular** if all vertices have degree $d$. It is regular if it is $d$-regular for some $d$.* ♣

**Remark** The **indegree**(**outdegree**) of a vertex in a directed graph is the number of edges leading into(out) the vertex.

**Definition 2.4**

*Examples of simple graphs.*

1. *A **complete** graph on $n$ vertices (denoted $K_n$) is a graph with $n$ vertices and an edge between each pair $u, v \in V$.*
2. *A **cycle** on $n$ vertices (denoted $C_n$) is a graph with $n$ vertices connected in a loop.*
3. *A **path** on $n$ vertices (denoted $P_n$) is a graph with $n$ vertices connected in a chain.*

**Figure 2.1:** Subgraph(left) and induced subgraph(right).

4. *A **bipartite** graph is a graph whose vertices can be split into two disjoint parts $U$ and $W$, i.e., $U \cup W = V$ and $U \cap W = \emptyset$, where all edges connect one part to the other, i.e., each edge $e$ can be written as $e = (u, w)$ with $u \in U$ and $w \in W$.*
5. *A complete bipartite graph (denoted $K_{n,m}$) has an edge connecting every element of one part (of size $n$) to every element of the other (of size $m$).*

♣

**Lemma 2.1 (The Handshake Lemma)**

*For any finite simple graph $G = (V, E)$, we have,*

$$\sum_{v \in V} d(v) = 2|E| \,. \tag{2.1}$$

♡

**Proof** By counting the number of incident pairs. Each edge contains 2 vertices.

**Remark** The number of vertices with odd degrees in graph $G$ must be even.

The Handshake Lemma for $k$-uniform hypergraph: $\sum_{v \in V} d(v) = k|E|$.

**Definition 2.5 (Subgraphs)**

*A graph $H$ is a **subgraph** of $G$ if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$.*

*A subgraph $H$ is an **induced subgraph** if it contains all the edges of $G$ connecting two vertices in $V(H)$, i.e., for any two vertices $u, v \in V(H)$, if $(u, v) \in E(G)$, then $(u, v) \in E(H)$. See Figure 2.1.*

♣

**Definition 2.6 (Types of Walks)**

*A **walk** in a graph $G$ is a sequence of vertices $v_1, v_2, \cdots, v_n$ where $v_i$ and $v_{i+1}$ is connected by an edge for each $v \in [n-1]$. If the first and last vertices of a walk are $u$ and $v$, then we say the walk is a $uv$-walk, where vertices $u$ and $v$ are called the ends of the walk. The **length** of a walk is the number of edges in that walk.*

1. *A **trail** is a walk whose edges are distinct.*
2. *A **path** is a walk whose vertices are distinct. (We refer similarly to a $uv$-path.)*
3. *A **circuit** is a trail that starts and ends at the same vertex.*
4. *A **cycle** is a path with an additional edge connecting the ends.*

♣

**Lemma 2.2**

*In a graph $G$, every walk from vertex $u$ to vertex $v$ (a $u - v$ walk) contains a $u - v$ path (by removing some of the edges).*

♡

**Remark** Similarly, every circuit contains a cycle.

**Definition 2.7 (Connected graph)**

*A graph $G$ is* ***connected*** *if for any two distinct vertices $u$ and $v$, there is a $u - v$ path in $G$.* ♣

**Theorem 2.1**

*Any graph $G$ can be uniquely partitioned into connected components, where each component is a connected subgraph and no two components have any edges between them.* ♡

**Theorem 2.2**

*A graph $G$ is bipartite if and only if it has no cycles of odd length.* ♡

### 2.1.1 Exercises

✍ **Exercise 2.1(2-Regular Graphs)** Show that any finite, 2-regular graph $G$ is a disjoint union of cycles. In particular, show that $G$ has a number of induced subgraphs that are cycles and so that:

1. Each vertex is in exactly one of these induced subgraphs.
2. No edges connect these subgraphs to each other.

**Proof** Let $G = (V, E)$ be a finite 2-regular graph with $n$ vertices. For any vertex $v_1 \in V$, there are 2 edges incident to $v_1$. Let $e_1$ be one of them. Following $e_1$, we would move to another vertex $v_2 \in V$, which has 2 edges as well since $G$ is 2-regular. Starting at vertex $v_2$, one edge $e_1$ connects it back to $v_1$. The other edge $e_2$ will connect $v_2$ to some other vertex $v_3$. By repeating the process above, we would obtain a path $\{v_1, \cdots, v_{n_1}\}$, which must eventually reach a vertex we have seen before. This must be $v_1$ (since any other $v_k$ is adjacent only to $v_{k-1}$ and $v_{k+1}$), which gives one cycle $C_{n_1}$, as an induced subgraph of $G$. For any vertex $v_k \in C_{n_1}$, if it is connected to an outside vertex $v \in G \setminus C_{n_1}$, the degree of $v_k$ would be 3, since $v_k$ has already been adjacent to $v_{k-1}$ and $v_{k+1}$, which violates the 2-regularity assumption. Hence $C_{n_1}$ is a connected component of $G$. Repeating this procedure for each connected component of $G$, we find that $G$ is a disjoint union of cycles.

✍ **Exercise 2.2(Properties Inherited by Subgraphs)** For which of the following graph properties $P$ does the following hold: If $G$ is a graph satisfying $P$ and $H$ is an induced subgraph of $G$, then $H$ must also satisfy $P$. For each property, either give a proof or a counter-example.

(a) $G$ is a complete graph.

(b) $G$ is a bipartite graph.

(c) $G$ is a cycle.

(d) $G$ is a simple graph.

(e) $G$ is a path.

**Proof** Remember that a subgraph $H$ is an induced subgraph if it contains all the edges of $G$ connecting two vertices in $V(H)$.

(a) If $G = (V(G), E(G))$ is complete, then for any $v \in V(G)$, $v$ is connected to all $u \in V(G) \setminus \{v\}$. As a result, for any $v \in V(H) \subset V(G)$, $v$ is connected to all $u \in V(H) \setminus \{v\}$. Hence $H = (V(H), E(H))$ is complete.

(b) If $G = (V(G), E(G))$ is bipartite, then the vertex set $V(G)$ can be partitioned into two disjoint subsets $U(G)$ and $W(G)$, i.e., $V(G) = U(G) \cup W(G)$ and $U(G) \cap W(G) = \emptyset$, such that for any $(u, w) = e \in E(G)$, we have $u \in U(G)$ and $w \in W(G)$. Let $H = (V(H), E(H))$ be an induced subgraph of $G$. Let $U(H) = U(G) \cap V(H)$ and $W(H) = W(G) \cap V(H)$. For any edge $(u, w)$ of $H$ must also be an edge of $G$ and therefore, one of the vertices (say $u$) is in $U(G)$ and the other ($w$) in $W(G)$. However, since both are in $V(H)$, this means that $u \in U(H)$ and $w \in W(H)$. Thus, $H$ is bipartite.

(c) The induced subgraph $H$ may not be cycle. The counter-example can be seen in Figure 2.2.

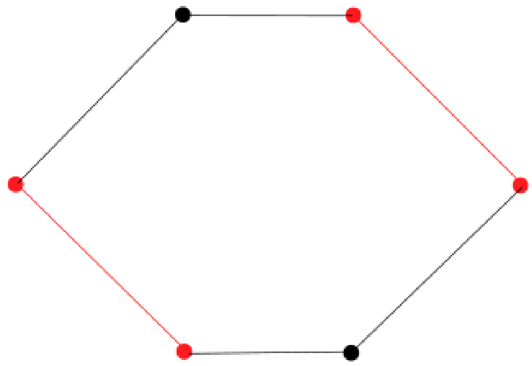


**Figure 2.2:** The induced subgraph $H$ of $C_6$ is marked red, which is not a cycle.

(d) A simple graph $G$ has neither multiedges, nor self loops. The induced subgraph $H$ doesn't contains any multiedges or self loops. Hence $H$ is simple.

(e) The counter-example can be seen in Figure 2.3, where $H$ has 2 disjoint components.

**Figure 2.3:** The induced subgraph $H$ of this path is marked red, which is not a path.

✍ **Exercise 2.3(Hypergraph Handshake Lemma)** Suppose that you have a hypergraph where each edge is incident on exactly $k$ vertices. Formulate and prove a version of the Handshake Lemma for this type of graph.

**Proof** Remember that in graph $G = (V, E)$ the degree of vertex $v \in V$, denoted by $d(v)$, is the number of vertices adjacent to $v$, i.e., the number of edges containing $v$. A hypergraph $H = (V, E)$ is called $k$-*uniform* if each edge $e \in E$ is incident on exactly $k$ vertices. The degree of $v$, denoted by $d(v)$, can be similarly defined as the number of edges containing $v$. The analogous Handshake Lemma is

$$\sum_{v \in V} d(v) = k|E| \,. \tag{2.2}$$

To prove this, we are going to count the number of vertex-edge incidence pairs in two different ways. On left hand side, each vertex $v$ is incident on $d(v)$ hyper-edges, thus the total number of pairs in this hypergraph is $\sum_{v \in V} d(v)$. On right hand side, each hyper-edge $e = (v_1, v_2, \cdots, v_k)$ contains $k$ incident vertices $v_1, v_2, \cdots, v_k$, thus the total number of pairs is $k|E|$.

✍ **Exercise 2.4(3-Regular Graphs)** Show that for every even integer $n \geq 4$ that there is a 3-regular graph with exactly $n$ vertices. What happens if $n$ is odd?

**Proof**

1. Let $n = 2k$ for some $k \geq 2$. Consider the cycle $C_{2k}$ on vertices $\{1, 2, \cdots, 2k-1, 2k\}$. The desired graph is obtained by connecting vertex $i$ and $i + k$ for all $1 \leq i \leq k$. The degree of vertex $i$ $(1 \leq i \leq k)$

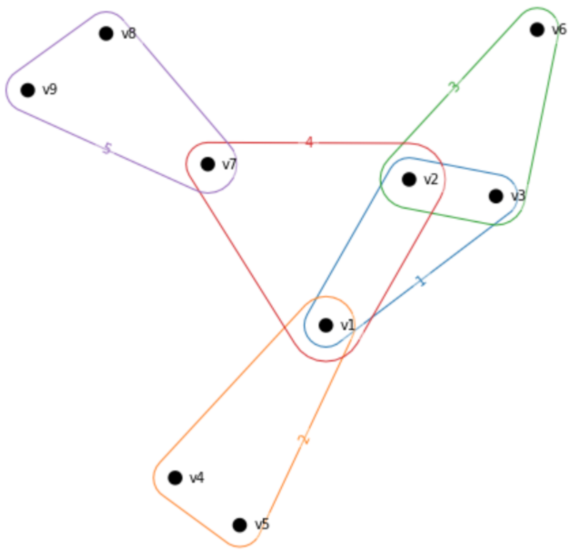



**Figure 2.4:** An example of 3-uniform hypergraph $H$ with 9 vertices and 5 hyperedges.

is 3 since $i$ is connected to $i-1$, $i+1$ and $i+k$. The degree of vertex $i$ $(k+1 \le i \le 2k)$ is also 3 since $i$ is connected to $i-1$, $i+1$ and $i-k$. Here vertex 1 is connected to $2k$.

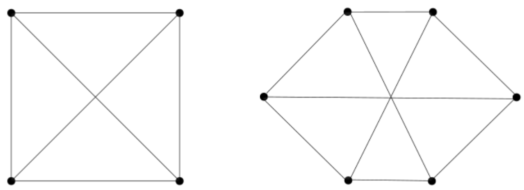


**Figure 2.5:** Examples for $n = 4$ and $n = 6$.

2. There is no such graph. The Handshake Lemma, $\sum_{v \in V} d(v) = 2|E|$, indicates that the sum of degrees should be even, however $3n$ is odd when $n$ is odd.

**Exercise 2.5** Let $G$ be a graph with at least two vertices. Prove or disprove:
(a) Deleting a vertex of maximim degree $\Delta(G)$ cannot increase the average degree.
(b) Deleting a vertex of mimimum degree $\delta(G)$ cannot reduce the average degree.

**Proof** *This problem tests your understanding of basic graph theory definitions (degrees).*

(a) True. Let $v$ be the vertex with maximum degree in $G$, i.e., $d(v) = \Delta(G)$. Define $G' = G \setminus \{v\}$. Using that $\Delta(G) \ge \bar{d}(G) = 2|E(G)|/|V(G)|$ by the Handshake-Lemma, it follows that

$$\bar{d}(G') = \frac{2|E(G')|}{|V(G')|} = \frac{2\big[|E(G)| - \Delta(G)\big]}{|V(G)|-1} \le \frac{2\big[|E(G)| - 2|E(G)|/|V(G)|\big]}{|V(G)|-1} = \frac{2|E(G)|}{|V(G)|} = \bar{d}(G).$$

(b) False. One possible counterexample is a complete graph $G = K_7$ on 7 vertices, which has average degree 6: deleting one vertex clearly yields a complete graph $K_6$ on 6 vertices, which has average degree 5.

**Exercise 2.6** A graph $H$ is called a subgraph of a graph $G$ if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$; in this case we write $H \subseteq G$. Prove that every graph $G$ with $|E(G)| \ge 1$ edges contains a subgraph $H$ with minimum degree $\delta(H) > \overline{d}(G)/2$.

(Hint: Construct a sequence of subgraphs $G = G_0 \supseteq G_1 \supseteq \cdots$ as follows: if $G_i$ has a vertex $v_i$ of degree $d_{G_i}(v_i) \leq \overline{d}(G_i)/2$, then obtain $G_{i+1}$ from $G_i$ by deleting $v_i$ (which includes deleting all edges containing $v_i$); if not, then terminate with $H = G_i$. Show that the resulting graph $H$ has the desired properties.)

**Proof** *Building upon the instructive Problem 3 concerning degree definitions, this problem trains applying induction (a classical proof technique) in the context of graph theory.*

Given the hint, this is a 'just-do-it' proof. Note that, by construction, the graph sequence terminates with a graph $H = G_i$ satisfying $\delta(H) > \bar{d}(H)/2$ if $\bar{d}(H) > 0$. Since $|E(G)| \geq 1$ implies $\bar{d}(G) > 0$, it thus suffices to prove that any graph $G_i$ in the sequence satisfies $\bar{d}(G_i) \geq \bar{d}(G)$. We prove this by induction:

- Base case $i = 0$: $G = G_0$ implies $\bar{d}(G_0) \geq \bar{d}(G)$.
- Induction step $i + 1 \geq 1$: Note that if the procedure does not terminate with $H = G_i$, then the next graph $G_{i+1} = G_i - \{v_i\}$ satisfies by similar reasoning as for Problem 3(a), using $d_{G_i}(v_i) \leq \overline{d}(G_i)/2 = |E(G_i)|/|V(G_i)|$ (by construction and the Handshake-Lemma) and the induction hypothesis for $G_i$,
$$\begin{aligned}\bar{d}(G_{i+1}) = \frac{2|E(G_{i+1})|}{|V_{G_{i+1}}|} &= \frac{2\big[|E(G_i)| - d_{G_i}(v_i)\big]}{|V_{G_{i+1}}|} \\ &\geq \frac{2\big[|E(G_i)| - |E(G_i)|/|V_{G_i}|\big]}{|V_{G_i}| - 1} = \frac{2|E(G_i)|}{|V_{G_i}|} = \bar{d}(G_i) \geq \bar{d}(G).\end{aligned}$$
(For the reader worried about the possibility of $|V(G_i)| - 1 = 0$, we remark that this can not happen since $\bar{d}(G_i) \geq \bar{d}(G_0) > 0$ implies that $G_i$ contains at least one edge and thus $|V(G_i)| \geq 2$ vertices.)

**Exercise 2.7** A graph $H$ is called a subgraph of a graph $G$ if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$; in this case we write $H \subseteq G$. Prove that every graph $G$ with $|V(G)| \geq 7$ vertices and at least $|E(G)| \geq 5|V(G)| - 14$ edges contains a subgraph $H \subseteq G$ with minimum degree $\delta(H) \geq 6$.

**Proof** *This problem trains applying induction (a classical proof technique) in the context of graph theory.*

This is a 'just-do-it' proof by induction on $|V(G)|$:

- Base case $|V(G)| = 7$: As discussed in class, any graph on 7 vertices can have at most $\binom{7}{2} = 21$ vertices, which is only attained by the complete graph $K_7$ on 7 vertices. Hence $G = K_7$, in which every vertex has degree 6, so that the 'trivial' subgraph $H = G$ satisfies $\delta(H) = 6$.
- Inductive step $|V(G)| \geq 8$: Let $G$ satisfy $|E(G)| \geq 5|V(G)| - 14$. If $\delta(G) \geq 6$, then $H = G$ satisfies the lemma. Otherwise there exists a vertex $v \in V(G)$ with degree $d_G(v) \leq 5$ in $G$. By removing that vertex from $G$ we then obtain a subgraph $G' \subseteq G$ with $|V(G')| = |V(G)| - 1 \geq 7$ vertices and
$$|E(G')| \geq |E(G)| - 5 \geq 5(|V(G)| - 1) - 14 = 5|V(G')| - 14$$
edges. Applying the induction hypothesis to $G'$, there is a subgraph $H \subseteq G'$ with minimum degree $\delta(H) \geq 6$. Since $H \subseteq G' \subseteq G$, we see that $H$ is also a subgraph of $G$, completing the induction step.

## 2.2 Trees

### 2.2.1 Basics

**Definition 2.8**

*Trees, leaves, and forests.*

1. *A **tree** is a connected graph without cycles.*
2. *A **leaf** in a tree is a vertex of degree 1.*
3. *A **forest** is a graph. where each connected component is a tree.*

♣

**Lemma 2.3**

*Let $T$ be a tree with vertices $u$ and $v$.*

1. *There exists a unique $u - v$ path in $T$.*
2. *Any tree $T$ on $n > 1$ vertices has at least two leaves.(Directly from Handshake Lemma.)*

♡

**Theorem 2.3 (Edge count of tree)**

*Any tree with $n$ vertices has exactly $n - 1$ edges.*

♡

**Proposition 2.1 (Generalization of previous result)**

*Any graph $G = (V, E)$ without cycles has $|V| - |E|$ connected components, namely $\#CCs = |V| - |E|$.*

♠

**Remark** If a new edge forms a cycle, it increases $|E|$ without decreasing $\#$CCs. In other words, if graph $G$ has at least one cycle, then $\#\text{CCs} > |V| - |E|$.

### 2.2.2 Spanning trees

**Definition 2.9**

*A subgraph $T$ is called a **spanning tree** of $G = (V, E)$ if*

1. *$T$ is a tree;*
2. *$V(T) = V(G)$.*

♣

**Remark** Every connected graph $G$ has a subgraph $T$ that is a tree connecting all of its vertices.

**Definition 2.10**

***Breadth First Search Tree** is constructed by following procedure:*

1. *Start at a base vertex $v$.*
2. *Connect $v$ to all its neighbors.*
3. *Connect them to all their neighbors without creating cycles.*
4. *Repeat until every vertex has been explored.*

***Depth First Search Tree** is constructed by following procedure:*

1. *Start at a base vertex $v$.*
2. *Follow path from $v$ until cannot extend anymore.*
3. *Backtrack until new branch.*
4. *Repeat until every vertex has been explored.*

♣

**Remark** BFS properties:

- BFS finds shortest paths from $v$ to other vertices. All vertices reachable from root $v$ with paths of length $d$ were found in $d^{th}$ round.
- No edges in $G$ provide shortcuts from a vertex $u$ to its descendants further down. If such an edge existed, it would have been used when exploring the neighbors of $u$.

DFS properties:

- $G$ has no extra edges that cross between different branches of the tree. If such an edge existed, it would have been used when exploring the first branch.

**Definition 2.11**

*Given a weighted graph $G = (V, E, W)$, where $W$ is the set of edge weights, such that there exists $w(e) \in W$ for every $e \in E$. The* ***minimum spanning tree*** *is defined as*

$$T^* = \arg \min_{\substack{T \text{ is a spanning} \\ \text{tree of } G}} \sum_{e \in E} w(e)\,.$$

♣

1. Repeatedly add lightest edge that does not create a cycle
2. Prim's Algorithm: Start at base vertex; Repeatedly add cheapest new edge connected base vertex to something new

**Theorem 2.4 (Cayley's Theorem)**

*There are $n^{n-2}$ labeled trees of order $n$.* ♡

**Theorem 2.5 (Generalization: Matrix Tree Theorem)**

*If $G$ is a connected graph with adjacency matrix $A$ and diagonal degree matrix $D$, then the number of unique spanning trees of $G$ is equal to the value of any cofactor of the matrix $D - A$.* ♡

### 2.2.3 Exercises

✍ **Exercise 2.8 (Non-Multiple of 3 Cycles and Circuits)** Prove that if a graph $G$ has a circuit whose number of edges is not a multiple of 3 then by removing edges from this walk, one can find a cycle whose number of edges is not a multiple of 3.

**Proof**

Let $C$ denote the circuit whose number of edges is not a multiple of 3. From Theorem 1.2 proven in class, a circuit contains a cycle, with possible repeated vertices. If $C$ is a cycle (no repeated vertices), then we are done. If $C$ has repeated vertices, for each vertex we can split it up (See Figure 1) and make it into two smaller circuits, and in each smaller circuit this vertex is not repeated. By continue splitting we can make $C$ into a collection of circuits where each circuit has no repeated vertices, hence $C$ is made into a collection of cycles. Note that this operation doesn't change the number of edges, so the number of edges of $C$ is the sum of the edges of the cycles. So there is at least one cycle whose number of edges is not a multiple of 3. By keeping this cycle and removing the other cycles, we are done.

✍ **Exercise 2.9(Distances and BFS)** In a graph $G$ define the *distance* between two vertices $u$ and $v$ to be the smallest number of edges in any path from $u$ to $v$ (or $\infty$ if no such path exists). Let $G$ be a connected graph with a vertex $v$ and let $T$ be the breadth first search tree rooted at $u$. Prove that for every other vertex $v$ in $G$ that the distance from $u$ to $v$ equals the length of the unique path from $u$ to $v$ in $T$. (In other words the path from $u$

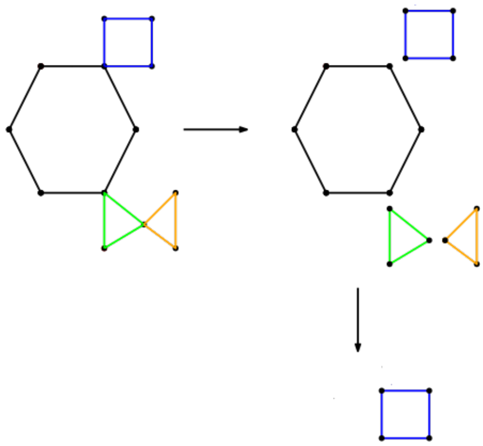


**Figure 2.6:** Splitting the repeated vertices and removing the cycles whose number of edges is a multiple of 3.

to $v$ in $T$ is a shortest path in $G$.)

**Proof** Since $G$ is connected and $T$ is the spanning tree, $v$ is in $T$, and hence there is a unique path $p$ that connects $u$ and $v$. Let's say it has length $l(p)$. By the definition of distance (the shortest length of a $uv$ walk), $d(u, v) \leqslant l(p)$. If $d(u, v) < l(p)$, there is a shorter $uv$ path outside $T$, since $p$ is the unique $uv$ path in $T$. Then this violates the property of Breadth First Search tree, which says that it finds the shortest path from $u$ to $v$ (we prove this in the lemma below). Hence $d(u, v) = l(p)$.

**Lemma** Let $G$ be a connected graph and $T$ be the Breadth First Search spanning tree rooted at $u$, then $T$ finds shortest path from $u$ to other vertices. In other words, if there is a $uv$ path of length $k$, then the BFS path has length at most $k$.

**Proof** We prove this by induction. When $k = 1$, $v$ is a neighbor of $u$, which is connected to $u$ in the first step of BFS, hence the length of the BFS path is 1.

Assume the lemma is true for $k - 1$. When the length of the $uv$ path is $k$, let $w$ be the vertex next to $v$ on this path (see Figure 2), so the length of the $uw$ path is $k - 1$. By inductive hypothesis, the BFS path from $u$ to $w$ has length at most $k - 1$. By the $(k - 1)$th step of BFS, $w$ will be connected to the tree, and $v$ is a neighbor, hence by $k$th step $v$ will also be connected to the tree, making the BFS path from $u$ to $v$ has length at most $k$.

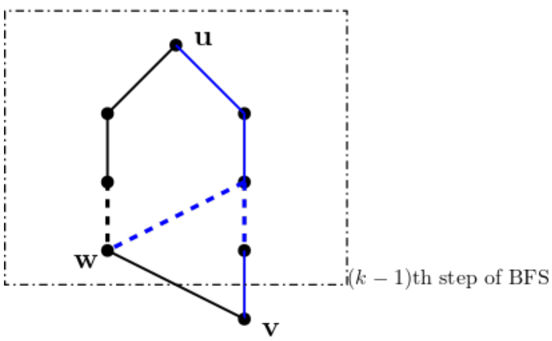



**Figure 2.7:** The blue paths are BFS paths. In $k$th step $v$ must be connected to the tree hence the length of $uv$ path is at most $k$.

✍ **Exercise 2.10(Bridges and Trees)** Let $G$ be a connected graph. We call an edge $e$ of $G$ a *bridge* if removing $e$ causes $G$ to become disconnected. Prove that an edge $e$ of $G$ is a bridge if and only if $e$ is part of every spanning tree of $G$.

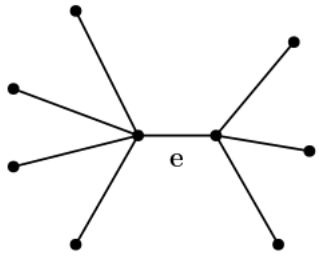



**Figure 2.8:** An example of a bridge $e$.

**Proof**

" $\Longrightarrow$ "

If $e$ is a bridge of $G$, and assume that there is a spanning tree $T$ that doesn't include $e$. Let $G'$ be the graph $G$ with edge $e$ removed. Then $G'$ is not connected, and so is $T$ since $T$ is the subgraph of $G'$. This is a contradiction, since a tree is by definition connected.

" $\Longleftarrow$ "

If $e$ is in every spanning tree of $G$, and assume that $e$ is not a bridge, which means that $G'$ is still connected. As we have shown in class, the connected graph $G'$ will have a spanning tree $T$, and by definition $T$ is also a spanning tree of $G$. But $T$ doesn't include $e$, which is a contradiction.

✍ **Exercise 2.11(Number of Bipartite Colorings)** Let $G$ be a finite, bipartite graph with $C$ connected components. How many ways can the vertices of $G$ be colored black and white so that each edge of $G$ connects a black vertex to a white vertex?

Let's first assume that $G$ is connected, i.e., it has one connected component. Pick an arbitrary vertex $u$ as the starting point. Note that if the color of $u$ is fixed, then the entire coloring of $G$ is fixed, since for any vertex $v$ in $G$ there is a $uv$ path with length $l$, and $v$ has the same color if $l$ is even and opposite color if $l$ is odd . Hence we have two ways of coloring ($u$ in black or $u$ in white). See Figure 4.

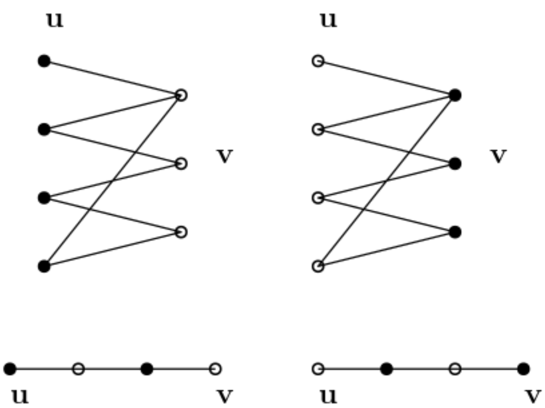



**Figure 2.9:** Two ways of coloring a connected bipartite graph.

Now if $G$ has $C$ connected components, note that the connected components are not connected to each other by any edges, any combination of the colorings of the connected components will satisfy the condition. Since each connected component has two ways of coloring, by the multiplication principle, in total $G$ has $2^C$

ways of coloring.

**Exercise 2.12(Short Cycles)** If $G$ is a connected graph on $n$ vertices with $n$ edges, we know it must contain a cycle. Prove that if $G$ has $n+1$ edges, it must contain a cycle of length at most $(2n/3)+2$. Hint: Consider $G$ as a tree plus two extra edges. Consider how the tree connects those 4 endpoints and how to connect them up to make a cycle.

**Proof** Let $G$ be a connected graph on $n$ vertices with $n+1$ edges. Assume for contradiction that all cycles in $G$ has length larger than $(2n/3)+2$. Let $T$ be a spanning tree of $G$. Since $T$ has $n-1$ edges, there are two different edges $e_1, e_2 \in E(G) \setminus E(T)$. Since $T \cup \{e_1\}$ and $T \cup \{e_2\}$ are both graphs on $n$ vertices with $n$ edges, there are cycles $C_1$ and $C_2$ such that $e_1 \in E(C_1) \setminus E(C_2)$ and $e_2 \in E(C_2) \setminus E(C_1)$. Then,

$$|E(C_1) \cap E(C_2)| = |E(C_1)| + |E(C_2)| - |E(C_1) \cup E(C_2)| > 2(2n/3+2) - n = n/3 + 4,$$

and $E(C_1) \cap E(C_2) \subseteq E(T)$. Let $P_1$ be $C_1$ with $e_1$ removed and $P_2$ be $C_2$ with $e_2$ removed. Then, both $P_1$ and $P_2$ are subgraphs of $T$, and $E(P_1) \cap E(P_2) = E(C_1) \cap E(C_2) \neq \emptyset$. We will show that their intersection is a path in $T$.

**Lemma 2.4**

*Let $T$ be a tree and $P$, $Q$ be paths in $T$. If $P$ and $Q$ share an edge, then their intersection $P \cap Q$ is a path in $T$.*

**Proof** Suppose that $P = (u_1, u_2, \ldots, u_j)$. Let $a = \min\{i : \{u_i, u_{i+1}\} \in E(P) \cap E(Q)\}$ and $b = \max\{i : \{u_{i-1}, u_i\} \in E(P) \cap E(Q)\}$. Since $E(P) \cap E(Q) \neq \emptyset$, $a$ and $b$ are well defined, and $a < b$. Let $R$ be the path $(u_a, u_{a+1}, \ldots, u_b)$. Then, $R$ contains $P \cap Q$. Since both $P$ and $Q$ contains a path from $u_a$ to $u_b$ and such path is unique in $T$, $R$ is a subgraph of $P \cap Q$. Thus, $P \cap Q = R$ is a path in $T$.

By the lemma above, the intersection of $P_1$ and $P_2$ is a path. Denote this path by $P_3$, and its endpoints by $s$ and $t$. Let $P_4$ be $C_1$ with edges $E(P_3)$ and vertices $V(P_3) \setminus \{s,t\}$ removed, and $P_5$ be $C_1$ with edges $E(P_3)$ and vertices $V(P_3) \setminus \{s,t\}$ removed. Then, both $P_4$ and $P_5$ are paths from $s$ to $t$ in $G$, and by construction, $E(P_4) \cap E(P_5) = \emptyset$. Let $C_3$ be the union of $P_4$ and $P_5$ in $G$. Then, $C_3$ is a cycle, and

$$\begin{aligned}
|E(C_3)| &= |E(P_4)| + |E(P_5)| \\
&= |E(C_1)| + |E(C_2)| - 2|E(P_3)| \\
&= |E(C_1) \cup E(C_2)| + |E(C_1) \cap E(C_2)| - 2|E(P_3)| \\
&= |E(C_1) \cup E(C_2)| + |E(P_3)| - 2|E(P_3)| \\
&= |E(C_1) \cup E(C_2)| - |E(P_3)| \\
&< n - (n/3 + 4) \\
&= 2n/3 - 4,
\end{aligned}$$

contradiction.

**Exercise 2.13(MSTs Have the Same Edge Weights)** Let $G$ be a weighted graph and let $T$ and $T'$ be two different Minimum Spanning Trees of $G$. Show that the set of weights of edges in $T$ is the same as the set of the weights of the edges in $T'$. Conclude that if the edges of $G$ all have distinct weights that there is a unique minimum spanning tree.

Hint: If $e \in T$ has a weight not in $T'$ try adding it to $T'$. This will allow you to improve $T'$ unless there is a cycle containing only $e$ and edges of lighter weight. In the latter case, find a way to improve $T$.

**Proof** Let $G$ be a weighted graph and $T$, $T'$ be two minimum spanning trees of $G$. Assume for contradiction that there exists $e \in T$ such that its weight $w(e)$ is not in the set of weight of edges in $T'$. Add $e$ to $T'$ and call

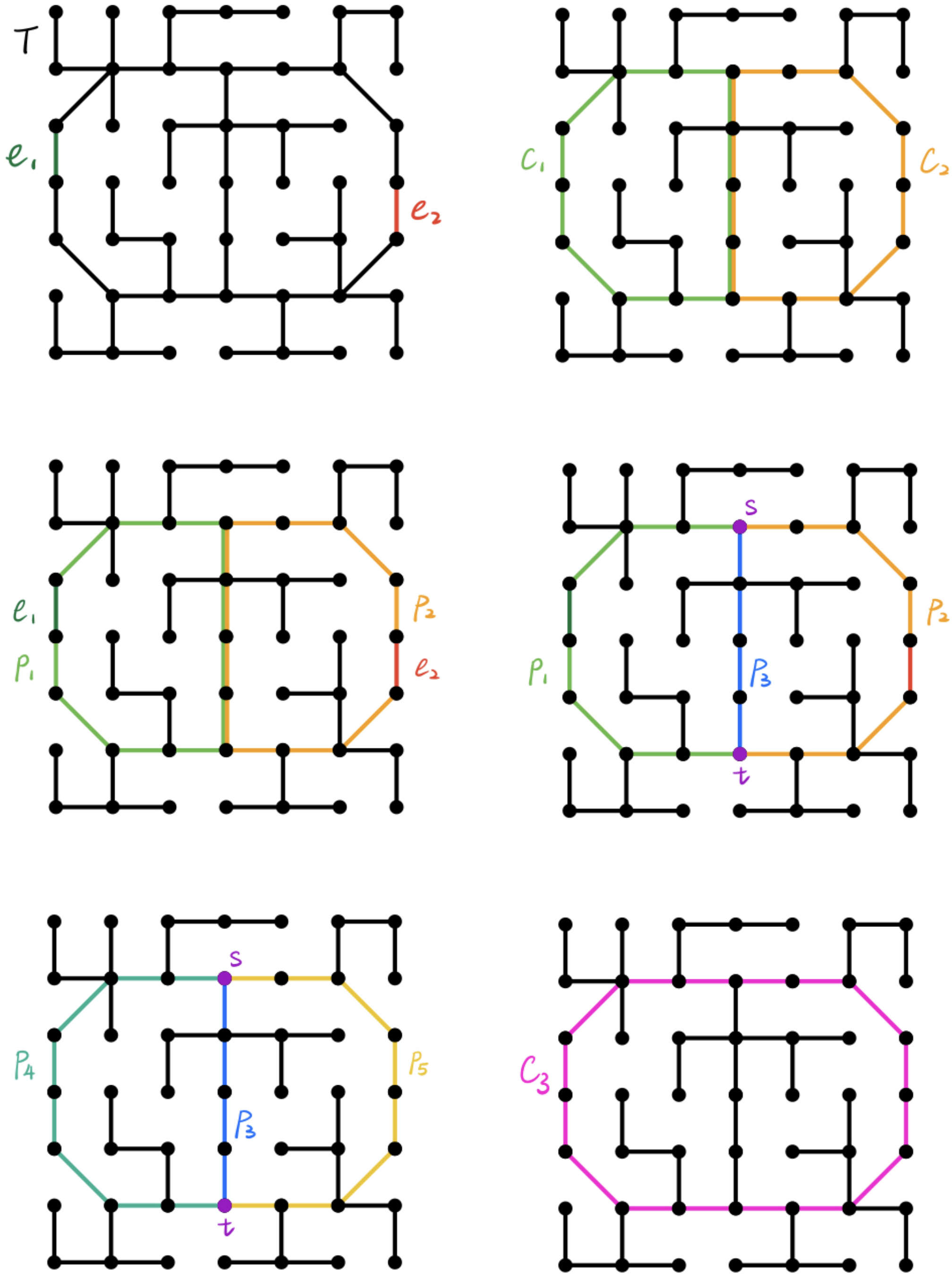



**Figure 2.10:** An example of the constructions in the proof

the new graph $S$. Then, $S$ contains a cycle $C$ which contains $e$. Suppose the weight of $e$ is not the largest among all edges in $C$. Let $e'$ be an edge of larger weight in $C$, and remove it from $S$, denote the new graph by $\hat{T}'$. Since $e'$ is in a cycle in $S$, it is not a bridge, and thus $\hat{T}'$ is connected. By the construction, $\hat{T}'$ has $n-1$ edges, so it is a spanning tree. Also, $w(\hat{T}') - w(T') = w(e) - w(e') < 0$, which is a contradiction to the fact that $T'$ has the minimal weight among all spanning trees. Therefore, $e$ has the largest weight in $C$. Since every edge $f$ in $T'$ has $w(f) \neq w(e)$, all other edges in $C$ have weight strictly smaller than $w(e)$. Suppose $e = \{u, v\}$. By deleting $e$ from $T$ we get two disjoint trees $T_1$ and $T_2$, one containing $u$ and the other containing $v$. Let $P$ be $C$ with $e$ deleted. Then $P$ is a path from $u$ to $v$. Since $P$ connects $T_1$ and $T_2$, there exists $g \in E(P)$ such that $g \cap V(T_1)$ and $g \cap V(T_2)$ are non-empty. Add $g$ to the union of $T_1$ and $T_2$, and denote the new graph by $\hat{T}$. Then, $\hat{T}$ is connected and by construction it has $n-1$ edges. Since $V(\hat{T}) = V(T_1) \cup V(T_2) = V(T) = V(G)$, $\hat{T}$ is a spanning tree of $G$. Then, $w(\hat{T}) - w(T) = w(g) - w(e) < 0$, contradiction. Hence, the set of weights of edges in $T$ is a subset of the set of weights of edges in $T'$. By the symmetry of this problem, they have to be the same set.

Suppose that the edges of $G$ all have distinct weights. Let $T$ and $T'$ be minimum spanning trees of $G$. Let $e \in E(T)$. Then the weight $w(e)$ of $e$ is in the set of weights of the edges in $T$, and therefore in the set of weights of the edges in $T'$. Thus, $e \in E(T')$ because no other edge has the same weight. Therefore, $E(T) \subseteq E(T')$, and they are then equal since they have the same cardinality. Also, since $T$ and $T'$ are spanning trees, $V(T) = V(T') = V(G)$. Thus, $T = T'$. Hence, $G$ has a unique minimum spanning tree.

✍ **Exercise 2.14(Average Number of Leaves)** How many trees on vertices labelled $1, 2, \ldots, n$ have the vertex labelled $k$ as a leaf? Use this to compute the average number of leaves of a tree with vertices labelled $1, 2, \ldots, n$ and show that this number is approximately $n/e$ for large values of $n$.

**Proof** By Cayley's theorem, there are $(n-1)^{n-3}$ trees on $n-1$ vertices labelled $1, \ldots, k-1, k+1, \ldots, n$. For each of those labelled trees, adding a vertex labelled $k$ and an edge $\{k, u\}$ for any $u \in \{1, \ldots, k-1, k+1, \ldots, n\}$ creates a distinct tree on vertices labelled $1, 2, \ldots, n$ where the vertex labelled $k$ is a leaf. This construction covers all such trees since removing the vertex labelled $k$ gives a tree on vertices labelled $1, \ldots, k-1, k+1, \ldots, n$. Therefore, the number of such trees equals to the number of ways to produce distinct trees in this construction, which is $(n-1)^{n-3}(n-1) = (n-1)^{n-2}$.

Let $\mathcal{A}$ be the set of all trees with vertices labelled $1, 2, \ldots, n$. Then, by Cayley's theorem, $|\mathcal{A}| = n^{n-2}$. For each tree $T \in \mathcal{A}$, let leaf$(T)$ be the set of labels of the leaves in $T$. For each $i \in \{1, 2, \ldots, n\}$, let $T(i)$ be the set of all trees on vertices labelled $1, 2, \ldots, n$ that have the vertex labelled $i$ as a leaf. Then, by the last part, $|T(i)| = (n-1)^{n-2}$ for all $i$. For each pair $T \in \mathcal{A}$ and $i \in \{1, 2, \ldots, n\}$, define $\delta(T, i)$ to be 1 if $T$ has the vertex labelled $i$ as a leaf, and 0 otherwise. Finally, let $\mu$ be the average number of leaves of a tree with vertices labelled $1, 2, \ldots, n$. Then,

$$\begin{aligned}
\mu &= \frac{1}{|\mathcal{A}|}\sum_{T\in\mathcal{A}} |\text{leaf}(T)| = \frac{1}{|\mathcal{A}|}\sum_{T\in\mathcal{A}}\sum_{i\in\text{leaf}(T)} 1 = \frac{1}{|\mathcal{A}|}\sum_{T\in\mathcal{A}}\sum_{i=1}^{n}\delta(T,i) = \frac{1}{|\mathcal{A}|}\sum_{i=1}^{n}\sum_{T\in\mathcal{A}}\delta(T,i) \\
&= \frac{1}{|\mathcal{A}|}\sum_{i=1}^{n}\sum_{T\in T(i)} 1 = \frac{1}{|\mathcal{A}|}\sum_{i=1}^{n}|T(i)| = \frac{1}{|\mathcal{A}|}\sum_{i=1}^{n}(n-1)^{n-2} \\
&= \frac{n(n-1)^{n-2}}{n^{n-2}} = n\left(\frac{n-1}{n}\right)^{n-2},
\end{aligned}$$

and

$$\lim_{n\to\infty}\frac{\mu}{n} = \lim_{n\to\infty}\left(\frac{n-1}{n}\right)^{n-2} = \lim_{n\to\infty}\left(\frac{n}{n-1}\right)^{2}\lim_{n\to\infty}\left(1-\frac{1}{n}\right)^{n} = e^{-1}.$$

Hence, $\mu$ is approximately $n/e$ for large $n$.

**Exercise 2.15(Last Caley Edges)** Recall that in our proof of Caley's Theorem that when converting a tree on vertices labelled $1, 2, \ldots, n$ that we stopped removing leaves when there were two vertices left. Prove that these vertices are the one labelled $n$ and the neighbor of this vertex along the unique path from the vertex labelled $n$ and the vertex labelled $n-1$.

**Proof**

**Lemma 2.5**

1. *If $T$ is a tree, then $T$ has at least 2 leaves.*
2. *If $T$ is a tree with exactly 2 leaves $u$ and $v$, then $T$ is a path from $u$ to $v$.*

**Proof** Let $T$ be a tree. Let $n$ be its number of vertices and $m$ be its number of leaves. Then, by handshake lemma,

$$2(n-1) = 2|E(T)| = \sum_{v \in V(T)} d_T(v) = \sum_{v \text{ is a leaf}} d_T(v) + \sum_{v \text{ is not a leaf}} d_T(v) \geq m + 2(n-m).$$

Thus, $m \geq 2$.

Suppose that $T$ has exactly 2 leaves $u$ and $v$. By the equation above, $m = 2$ implies that $d_T(w) = 2$ for any $w$ that is not a leaf. Let $P$ be a path from $u$ to $v$ in $T$, and assume for contradiction that there is a vertex $w$ that is not on this path. Then, there is a path $Q$ from $w$ to $v$. Let $x$ be the closet vertex to $u$ on among the vertex on $P$ that is also a vertex of $Q$. Then, $x$ is in an edge $e \in E(Q) \setminus E(P)$. Since $u$ and $v$ are of degree 1, they are not $x$. Thus, $x$ has two edges in $E(P) \setminus E(Q)$. Thus, $x$ has degree at least 3, contradiction. Hence, $P$ contains all vertices in $T$, so it has $n-1$ edges. Therefore, $T = P$ is a path from $u$ to $v$.

Since all trees have at least 2 leaves, the vertex labelled $n$ will never be deleted because either it is not a leaf or there is another leaf with smaller label. Consider the vertex labelled $n-1$. Either it is removed in the algorithm or it is not. If it is never removed, then the lemma follows. Suppose it is removed. Then, at the stage where the vertex labelled $n-1$ is about to be removed, all vertices with label smaller than $n-1$ are not leaves. Thus, by lemma 1, the vertices labelled $n$ and $n-1$ are the only leaves in the remaining tree. By lemma 2, the remaining tree at that stage is a path from the vertex labelled $n-1$ to that labelled $n$, which is the unique one in the original tree. Suppose the path is $(v_0, v_1, \ldots, v_m)$ where $v_0$ is the vertex labelled $n-1$ and $v_m$ is the one labelled $n$. Then, following the algorithm, we will remove, in order, $v_0$, $v_1$, $\ldots$, $v_{m-2}$. Thus, the vertices remained are $v_{m-1}$ and $v_m$, where $v_{m-1}$ is the neighbor of the vertex labelled $n$ along the unique path from the vertex labelled $n$ and the vertex labelled $n-1$.

**Exercise 2.16** Prove the following statements:

(a) A graph with $n$ vertices and $m$ edges has at least $n-m$ connected components.

**Proof** *This problem tests your understanding of basic graph theory definitions (components).*

The basic idea is that every edge added decreases the amount of connected components by at most one. To make this fully rigorous, we proceed by induction on the number $m$ of edges.

- Base case $m = 0$: When $m = 0$, we have a graph with $n$ vertices and no edges, so each vertex is its own connected components, therefore, the total number of connected components is $n = n - 0 = n - m$.
- Induction step $m \geq 1$: Let $G$ be a graph with $n$ vertices and $m$ edges. Delete any edge of $G$ to create a subgraph $G'$ with $n$ vertices and $m-1$ edges. Applying our inductive hypothesis to $G'$ shows that $G'$ has at least $n-(m-1)$ connected components. The deleted edge can connect at most two of the connected components in $G'$ together, reducing the total number of connected components by at most one. Therefore, the number of connected components of $G$ is at least $n-(m-1)-1 = n-m$, completing

the induction step.

(b) Every $n$-vertex graph with at least $n$ edges contains a cycle.

**Proof** *This problem tests your understanding of basic graph theory definitions (cycles, trees and forests).*

The contrapositive of this statement is that any acyclic graph (i.e. any forest) with $n$ vertices has at most $n-1$ edges, so it suffices to prove this statement. We can do so by showing the stronger statement, that any forest with $k$ connected components has exactly $n-k$ edges, since each forest has at least 1 connected components. To prove this, recall from class that any tree $T$ has exactly $|V(T)|-1$ edges. Let $F$ be a forest with $n$ vertices and $k$ connected components, $T_1, \dots, T_k$. Each connected component is a tree, so we have

$$|E(F)| = \sum_{i=1}^{k} |E(T_i)| = \sum_{i=1}^{k} (|V(T_i)|-1) = \left(\sum_{i=1}^{k} |V(T_i)|\right) - k = n-k.$$

✍ **Exercise 2.17** A vertex with degree one is called a leaf. A forest is an acyclic graph, i.e., whose components are trees.

(a) Prove that deleting a leaf from a tree $T$ produces a tree with $|V(T)|-1$ vertices.

**Proof** *This problem tests your understanding of basic graph theory definitions (trees).*

Let $T$ be a tree and let $\ell$ be a leaf of $T$. Let $T'$ be the graph obtained from $T$ by deleting vertex $\ell$. We need to show $T'$ is a tree. To do so we need to show that $T'$ is connected and acyclic.

To show that $T'$ is connected, let $u$ and $v$ be distinct vertices in $T'$. Since $T$ is connected there exists a path in $T$ from $u$ to $v$. Since the interior vertices of a path must all have degree $\geq 2$ in $T$, they cannot be vertex $\ell$. So this path is also present in $T'$, and therefore $T'$ is connected.

To show it $T'$ acyclic, suppose for contradiction that there exists a cycle in $T'$. Since $T'$ is a subgraph of $T$, this cycle would also be present in $T$, which is a contradiction since $T$ is a tree. Therefore $T'$ is acyclic.

(b) Prove that deleting a vertex $v \in V(T)$ from a tree $T$ produces a forest consisting of $d_T(v)$ components.

**Proof** *This problem tests your understanding of basic graph theory definitions (trees and forests).*

Let $T$ be a tree, and let $v$ be a vertex of $T$. Let $F$ be the graph obtained by deleting vertex $v$. Since $T$ is acyclic and $F$ is a subgraph of $T$, $F$ is acyclic so it is a forest. Since deleting vertex $v$ deleted $k$ edges, and every tree has one fewer edges than vertices,

$$|E(F)| = |E(T)| - k = |V(T)| - 1 - k = |V(F)| - k$$

By the argument from Problem 1(b), it therefore follows that $F$ consists of $k$ connected components.

(c) Prove that for any tree $T$ with $|V(T)| \geq 2$ vertices the following is true: if $\Delta(T) \leq 2$, then $T$ is a path.

**Proof** *This problem trains applying induction (a classical proof technique) in the context of graph theory.*

We proceed by induction on $n = |V(T)|$, the number of vertices.

- Base case $n = 2$: Suppose $n = 2$. The only tree with two vertices is two vertices connected by a single edge, this is a path on two vertices.
- Induction step $n \geq 2$. Let $T$ be a tree with $n$ vertices and $\Delta(T) \leq 2$. Let $v$ be a leaf of $T$ (which we know exists, see also Problem 3), and let $T' = T \setminus v$ be the graph obtained by deleting $v$ from $T$. Then $T'$ is also a tree by Problem 2(a), with $\Delta(T') \leq \Delta(T) \leq 2$. So by our inductive hypothesis, $T'$ is a path. Since the maximum degree of $T$ was at most $\Delta(T) \leq 2$, the leaf vertex can only be attached to a leaf vertex of $T'$, and it follows that $T$ is also a path, completing the induction step.

✍ **Exercise 2.18** Give at least two different proofs for the fact that any tree $T$ contains at least $\Delta(T)$ leaves.

**Proof** *This problem trains applying different proof techniques in the context of graph theory.*

For each of the following proofs, $\ell(T)$ denotes the number of leaves of a tree $T$.

Induction (on $n = |V(T)$, the number of vertices):

- Base Case $1 \le n \le 3$: There is (up to relabelling of the vertices) only one possible tree $T$ for each number of vertices $|V(T)| \in \{1, 2, 3\}$, and for each of them it is straightforward to check (by hand) that it contains at least as many leaves as the maximum degree.
- Induction step $n \ge 4$: Let $T$ be a tree with $n$ vertices, and let $v$ be a leaf of $T$. Let $T' = T \setminus v$ be the graph obtained by deleting $v$ from $T$. By Problem 2(a) we konw that $T'$ is a tree, so by our inductive hypothesis $T'$ has at least $\Delta(T')$ many leaves. Now we consider two cases:
  - Case 1: $v$ is attached to a leaf of $T'$. Then $\ell(T) = \ell(T')$, since adding on $v$ makes the vertex it is attached to no longer a leaf in $T$, but $v$ itself becomes a leaf. Also, since $v$ was attached to a leaf, the max degree cannot change, and $\Delta(T) = \Delta(T')$. So $\ell(T) = \ell(T') \ge \Delta(T') = \Delta(T)$
  - Case 2: $v$ is not attached to a leaf of $T'$. Then $\ell(T) = \ell(T') + 1$, since all leaves of $T'$ are still leaves of $T$, and we have added the new leaf $v$. Also, $\Delta(T) \le \Delta(T') + 1$, since we have added only one edge to go from $T'$ to $T$. So $\ell(T) = \ell(T') + 1 \ge \Delta(T') + 1 \ge \Delta(T)$.

  In either case, $T$ has at least $\Delta(T)$ many leaves, completing the induction step.

Vertex deletion:

Let $v$ be a vertex of $T$ with maximum degree. Let $T'$ be the graph $T \setminus v$. By problem 2b, $T'$ is a forest with $\Delta(T)$ components. Each connected component of $T'$ is a tree and is either a single vertex or has two leaves. In either case, it contributes at least one leaf to $T$. So $T$ has at least $\Delta(T)$ leaves.

Handshake Lemma based:

The number of edges in a tree with $n$ vertices is $n - 1$, so by the Handshake Lemma,

$$\sum_{v \in V} d(V) = 2n - 2.$$

We also know that

$$\sum_{v \in V} d(v) \ge \Delta(T) + 2 \cdot (\#\text{nonleaf vertices} - 1) + 1 \cdot (\#\text{leaf vertices}) = \Delta(T) + 2\big(n - \ell(T) - 1\big) + \ell(T).$$

Putting these together gives

$$2n - 2 \ge \Delta(T) + 2\big(n - \ell(T) - 1\big) + \ell(T),$$

which simplifies to

$$\ell(T) \ge \Delta(T).$$

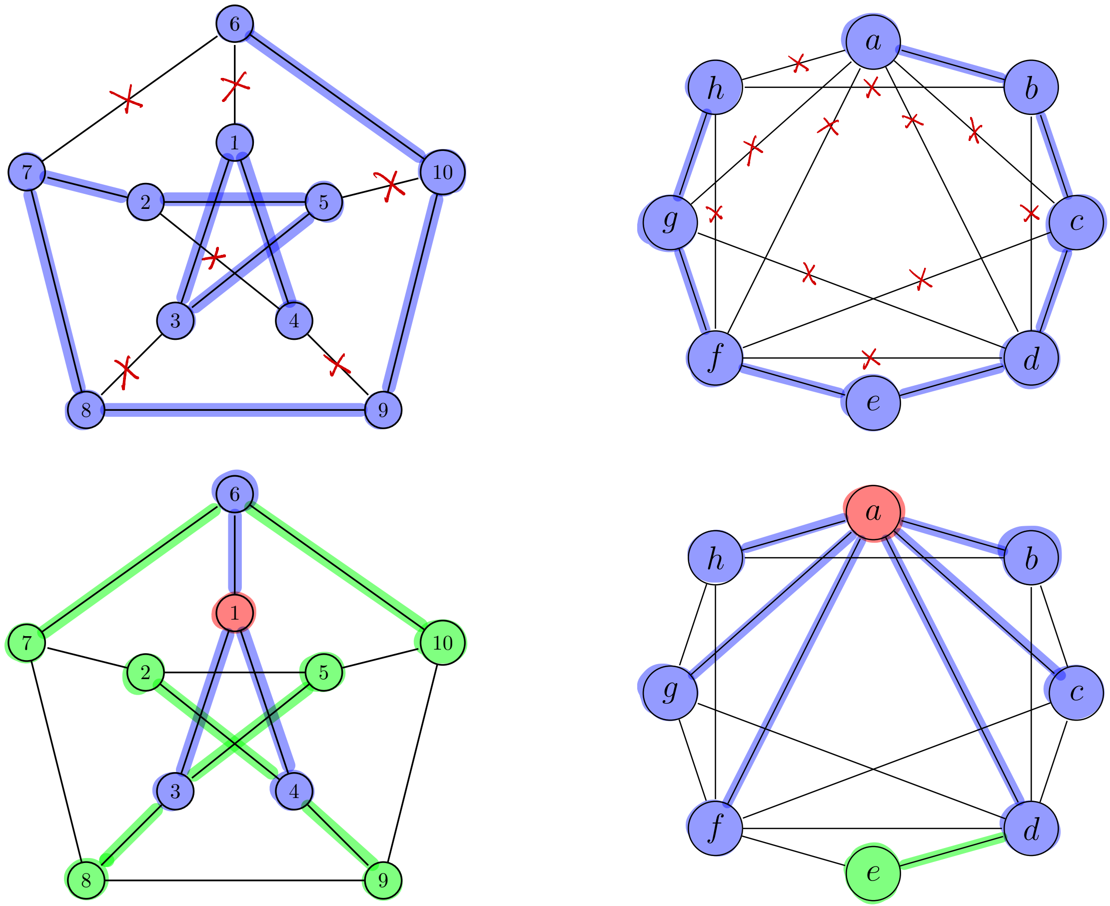
**Exercise 2.19** This problem illustrates algorithmic approaches for finding a spanning tree $T \subseteq G$ of a connected graph $G$.

(a) In class we learned one algorithmic way to create a spanning tree $T \subseteq G$ with $V(T) = V(G)$: iteratively delete one edge that is part of a cycle until no cyles are left. Using this cycle-removal algorithm, find spanning trees $T$ of the two graphs depicted below.

(b) Read Section 3.2 in the Verstraete lecture notes, and understand how to construct a spanning tree using breadth-first-search (which we only implicitly touched upon in class). Using this breadth-first algorithm, find spanning trees $T$ of the two graphs depicted below (starting with vertex 1 and $a$, respectively; if there are ties, prioritize vertices with smaller labels).

**Proof**

(a) The edges marked with red x are deleted from the original graphs (many valid solutions exist).
(b) Edges with colors are selected to form the spanning tree following the breadth-first search algorithm.




✍ **Exercise 2.20** A graph is $k$-regular if all its vertices have degree $k$. Prove the following statements:

(a) Any $n$-vertex graph with $m \geq n$ edges has at least $m - n + 1$ cycles.

(b) For $k \geq 1$, any $k$-regular bipartite graph has the same number of vertices in each partition class.

(c) For any connected graph $G$ and any $1 \leq k \leq |V(G)|$, $G$ has a connected subgraph with exactly $k$ vertices.

**Proof**

(a) Let $C_1, C_2, \cdots, C_k$ denote the connected components of $G$. Writing $T_i$ for the spanning tree of $C_i$, we define the forest $F := \cup_{i=1}^{k} T_i$. From Problem 1(a) of HW2, it follows that $|E(F)| = n - k$. We now add the remaining $m - n + k$ edges of $G$ to $F$. A new cycle will be completed each time when an edge is added to the previous graph, and these cycles are distinct since each one contains an edge that was not previously included. Since $k \geq 1$, the number of cycles is thus at least $m - n + 1$.

(b) Let $A$ and $B$ be the two partition classes of the bipartite graph $G$ (so in $G$ there are no edge within $A$ and $B$, respectively). If we count the edges of $G$ according to their endpoints $A$, then $e(G) = \sum_{a \in V(A)} d(v) = k|V(A)|$. Similarly, $e(G) = \sum_{a \in V(B)} d(v) = k|V(B)|$. Diving by $k \geq 1$ thus yields $|V(A)| = |V(B)|$.

(c) Let $T \subseteq G$ be a spanning tree of $G$ with $|V(T)| = |V(G)|$. By Problem 2(a) of HW2, by deleting one leaf of $T$ we obtain a tree $T' \subseteq T$ with $|V(G)| - 1$ vertices. We repeat this process of deleting leaves until there are $k$ vertices left and obtain $H$, where $H$ is a connected subgraph (since it a tree by construction).

✍ **Exercise 2.21** For $n \geq 3$, let $G$ be an $n$-vertex graph such that every graph obtained by deleting one vertex is a tree. Determine the number of edges $e(G) = |E(G)|$ of $G$, and then argue that $G$ is a cycle $C_n$ on $n$ vertices.

**Proof** Let $G_i = G \setminus \{i\}$ be the graph obtained by deleting vertex $i$. Then $|V(G_i)| = n - 1$ and $e(G_i) =$

$n-2 = e(G) - d_G(i)$ for each vertex $i \in V(G)$. By Handshaking lemma, it follows that

$$n(n-2) = \sum_{1\le i\le n}^{n} e(G_i) = \sum_{1\le i\le n} [e(G) - d_G(i)] = ne(G) - \sum_{1\le i\le n} d_G(i) = (n-2)e(G).$$

Thus $e(G) = n$, and $d_G(i) = 2$ for each $i \in V(G)$. Now there are at least two ways to conclude that $G = C_n$:
(1) Since the degree of each vertex $i$ of $G$ is two, every component of $G$ must be a cycle. Since $G_i = G \setminus \{i\}$ is a tree and thus connected, it follows that $G$ consists of only one component. Hence $G = C_n$.
(2) Using $|V(G)| = n$ and $e(G) = n$, by Problem 1(b) of HW2 we infer that $H$ contains a cycle. Since $G_i = G \setminus \{i\}$ is a tree and thus has no cycle, every cycle in $G$ must contain $i$. Since this is true for all vertices $i$, it follows that every cycle of $G$ must contain all vertices. Hence $G$ contains a spanning cycle $C_n$, i.e., one containing all $n$ vertices. But $G$ has no additional edges beyond $C_n$ (since $e(G) = e(C_n) = n$), so $G = C_n$.

✍ **Exercise 2.22** (a) Prove that every graph $G$ has a bipartite subgraph $H \subseteq G$ with $e(H) \ge e(G)/2$ many edges (where we use the shorthand $e(F) = |E(F)|$ for the number of edges, as usual).

(Hint: In fact, there is a bipartite subgraph $H \subseteq G$ where the degrees satisfy $d_H(v) \ge d_G(v)/2$ for all $v \in V(H) = V(G)$.)

(b) Deduce that if $G$ has $n$ vertices and at least $2n-1$ edges, then $G$ contains an even cycle.

**Proof**

(a) Proof by extremality. Let $H$ be a bipartite subgraph of $G$ with $V(H) = V(G) = A \cup B$, which contains the largest number of edges. If $d_H(v) \ge d_G(v)/2$ for each vertex $v \in V(H)$, then $e(H) \ge e(G)$ by Handshaking lemma. If otherwise there is a vertex $v \in A$ with $d_H(v) < d_G(v)/2$, then we switch $v$ to the other partition-class to form $B' = B \cup \{v\}$ and $A' = A \setminus \{v\}$. In the resulting the new bipartite graph $H'$ with partition-classes $A'$ and $B'$ we have $d_{H'}(v) = d_G(v) - d_H(v) > d_G(v)/2$ and thus $e(H') = e(H) - d_H(v) + d_{H'}(v) > e(H)$ holds (to see this note that you can go from $H$ and $H'$ by first removing $d_H(v)$ edges adjacent to $v$ and then adding $d_{H'}(v)$ edges adjacent to $v$). Hence $e(H') > e(H)$, contracting that $H$ has the largest number of edges.

Proof by induction on $|V(G)|$. Base case $|V(G)| = 1$: $H = G$ works as then trivially $e(G)/2 = 0 = e(H)$.

Induction step $|V(G)| \ge 2$: choose any vertex $v \in V(G)$ and delete $v$ to construct $G' = G \setminus \{v\}$. By the induction hypothesis, $G'$ contains a bipartite subgraph $H'$ with $e(H') \ge e(G')/2$. Let $A'$ and $B'$ be the two partition-classes of $H'$. We now want to add $v$ to $H'$ to form $H$ with $H' \subseteq H \subseteq G$ and $V(H) = V(H') \cup \{v\}$. If $v$ has more neighbors in $A'$, then it will be assigned to $B'$; otherwise $A'$. By this construction, at least half of edges containing $v$ will be added to $H'$. It therefore follows that

$$e(H) = e(H') + d_H(v) \ge \frac{1}{2}e(G') + \frac{1}{2}d_G(v) = \frac{1}{2}\Big(e(G \setminus \{v\}) + d_G(v)\Big) = \frac{1}{2}e(G).$$

(b) By part (a) there is a bipartite subgraph $H \subseteq G$ with $e(H) \ge e(G) \ge n - 1/2$, so that $e(H) \ge n$ (since the number of edges is an integer). Using $|V(H)| = n$ and $e(H) \ge n$, by Problem 1(b) of HW2 we infer that $H$ contains a cycle. Bipartite graphs contain no odd cycles, so the cycle in $H$ must be an even cycle.

## 2.3 Eulerian and Hamiltonian Graphs

### 2.3.1 Eulerian Graphs

**Definition 2.12 (Eulerian circuit/trail)**

*Let $G = (V, E)$ be a graph with finite vertices.*

1. *An **Eulerian circuit** is a circuit that uses every edge of a graph exactly once.*
2. *An **Eulerian trail** similarly uses each edge exactly once, but does not start and end at the same vertex.*
3. *Graph G is **Eulerian** if it contains an Eulerian circuit.*
4. *Graph G is **semi-Eulerian** if it contains an Eulerian trail, but not a circuit.*

♣

**Remark** The (semi-)Eulerian graphs must be connected, except for isolated vertices, since there is no Eulerian circuit/trail that connects both of $u$ and $v$ if there is no path from u to v.

**Theorem 2.6 (Criteria of Eulerian circuit/trial)**

*Let G be a connected multigraph, then*

1. *G has an Eulerian circuit if and only if all of the vertices of G have even degree.*
2. *G has an Eulerian trail if and only if exactly two vertices of G have odd degree. Furthermore, those 2 vertices will be the ends of any Eulerian trial.*

♡

**Corollary 2.1**

*If graph G contains a $uv$-Eulerian trial, then $G + \mathrm{edge}(u, v)$ is a Eulerian circuit.*

♡

**Lemma 2.6**

*If $e$ is a bridge in a finite graph G, then there is at least one vertex of odd degree on each side of it.*

♡

### 2.3.2 Hamiltonian Graphs

**Definition 2.13 (Hamiltonian path/cycle)**

*Let $G = (V, E)$ be a graph with finite vertices.*

1. *A **Hamiltonian path/cycle** in G is a path/cycle that uses every vertex of G exactly once.*
2. *A graph is **traceable/Hamiltonian** if it has a Hamiltonian path/cycle.*

♣

**Lemma 2.7**

*If a graph G on $n$ vertices has minimum degree $\delta(G) \geq (n-1)/2$, then G is connected.*

♡

**Definition 2.14**

*The closure of an $n$-vertex graph G, denoted $C(G)$, consists in adding edges between any two non-adjacent vertices whose sum of degrees is at least $n$.*

♣

**Theorem 2.7 (Dirac)**

*If G is a graph with n vertices, where $n \geq 3$ and $\deg(v) \geq n/2$, for every vertex $v \in G$, then G is Hamiltonian. (Not necessary)*

♡

**Theorem 2.8 (Ore)**

*Let $G$ be a graph with $n$ vertices and let $u$ and $v$ be non-adjacent vertices in $G$ such that* $\deg(u)+\deg(v) \geq n$. *Let $G+(u,v)$ denote the super graph of $G$ obtained by joining $u$ and $v$ by an edge. Then $G$ is Hamiltonian if and only if $G+(u,v)$ is Hamiltonian.* ♡

### 2.3.3 Exercises

**Exercise 2.23(Infinite Semi-Eulerian Graphs)** Define a semi-infinite trail on a graph $G$ to be a sequence of vertices $v_1, v_2, v_3, \ldots$ so that for each $i$ there is an edge in $G$ between $v_i$ and $v_{i+1}$ and so that these edges are all distinct. Call $G$ semi-Eulerian if there exists such a trail that uses every edge of $G$ exactly once. For this problem, we will consider graphs $G$ with infinitely many vertices and edges for which each vertex is incident on only finitely many edges (and thus has a well defined degree).

(a) Show that if $G$ has a semi-infinite Eulerian trail starting at a vertex $v$, then $\deg(v)$ is odd and all other degrees in $G$ are even, and that $G$ is connected except for isolated vertices.

(b) Unfortunately, the conditions in (a) are not sufficient. Give an example of a graph $G$ with a vertex $v$ of odd degree and all other degrees even so that $G$ does not have a semi-Eulerian trail.

(c) Show that any graph $G$ with a semi-infinite Eulerian trail must satisfy the following property: If $G'$ is any subgraph of $G$ obtained by removing finitely many edges of $G$, then $G'$ contains only one connected component of infinite size.

(d) It turns out that the conditions in (a) and (c) together are sufficient. The proof relies on the following **Lemma**: Suppose $G$ is a graph with only one vertex $v$ of odd degree and satisfying the condition in (c). If $P$ is any finite path starting at $v$, then the edges of all of the finite connected components of $G-P$ can be partitioned into circuits. Prove this lemma.

**Proof**

(a) Let $P$ denote the semi-infinite trail starting at vertex $v$. Consider vertex $u \neq v$, which is not the starting point. When there is an edge in, there must be an edge out, since the trail would not stop at $u$ and every edge at $u$ is used her, then the out degree always equals in degree, hence

$$\deg_{out}(u) = \deg_{in}(u) \implies \deg(u) = \deg_{out}(u) + \deg_{in}(u) = 2\ \deg_{in}(u)\,. \tag{2.3}$$

However, for the starting point $v$, the argument above holds true except for the first step out, then difference between out degree and in degree should be 1, hence

$$\deg_{out}(v) = \deg_{in}(v) + 1 \implies \deg(v) = \deg_{out}(v) + \deg_{in}(v) = 2\ \deg_{in}(v) + 1\,. \tag{2.4}$$

(b) The counter example can be seen in Figure 2.13, where the trail starting at $v$ can't be semi-infinite Eulerian, since a trail starting at $v$ can explore only one branch and it would never come back to go through vertices in other branches.

(c) Let $\{f_1, f_2, \cdots, f_k\}$ denote the set of finitely many edges removed from $G$, i.e., $E(G') = E(G) \setminus \{f_1, f_2, \cdots, f_k\}$. We assume that there are 2 connected components of infinite size in $G'$.

Note that $G$ contains a semi-infinite Eulerian trail $(e_1, e_2, e_3, \cdots)$, which contains all edges in $G$, then each $f_j$ is some $e_{i_j}$. Let $N = \max_{1\leq j\leq k} i_j$ and consider the finite trail $P = (e_1, e_2, \cdots, e_N)$. Define $G'' := G - P$, which contains a semi-infinite Eulerian trail $(e_{N+1}, e_{N+2}, \cdots)$, thus $G''$ is connected except for isolated vertices.

On the other hand, $G''$ is a subgraph of $G'$, obtained by removing finitely many edges $\{e_1, e_2, \cdots, e_N\} \setminus \{f_1, \cdots, f_k\}$. However, by our assumption, $G'$ contains 2 infinite components. Neither of them can be

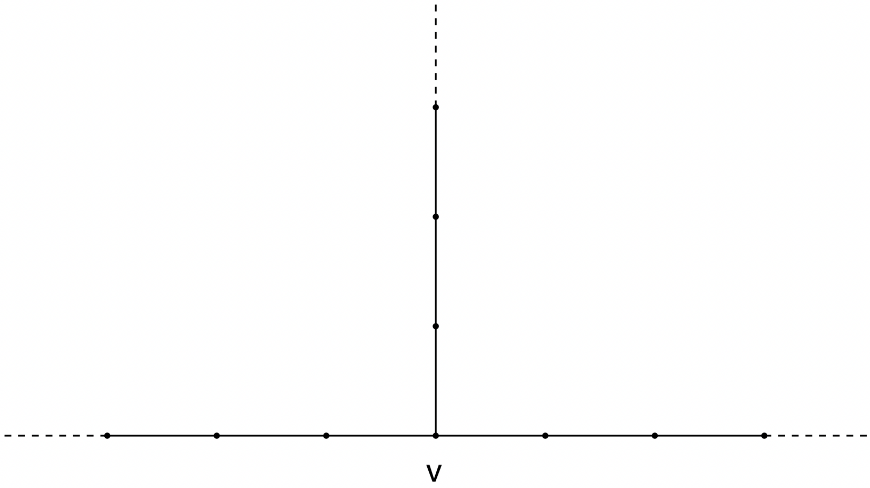



**Figure 2.13:** Counter example for part $(b)$.

eliminated by removing finitely many edges. Then by this construction, $G''$ must have 2 disjoint connected components, indicating that $G''$ is disconnected, which contradicts to the conclusion above.

(d) Let $P = (v, \cdots, u)$ denote the finite path starting at $v$, with $u$ being the endpoint. When $P$ is removed from $G$, there is only one vertex in $G - P$ with odd degree, which turns out to be $u$ due to conditions in $(a)$. The connected component containing $u$, denoted by $G'$, should have infinite size, otherwise it would contradict to the *HandShake* Lemma since the number of vertices with odd degrees is odd. Let $H \subset (G - P)$ be a connected component of $G - P$, which is disjoint from $G'$. By discussion above, any vertex $w \in H$ should have even degree. Therefore $H$ is Eulerian, which can be partitioned into circuits. An example can be seen in Figure 2.14.

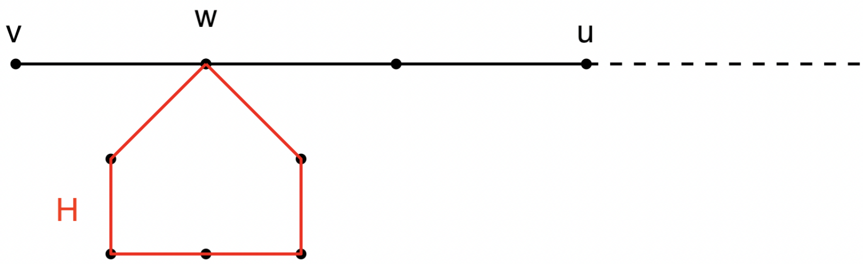



**Figure 2.14:** An illustrating example for part $(d)$.

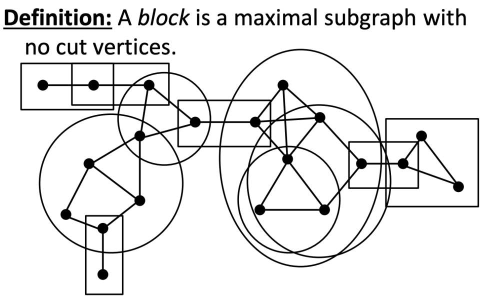
✍ **Exercise 2.24(Toughness of Hamiltonian Graphs)** Let $G$ be a Hamiltonian graph and let $\{v_1, v_2, \ldots, v_k\}$ be a set of $k$ vertices of $G$ for some $k \geq 1$. Let $G'$ be the subgraph of $G$ obtained by removing all of the vertices $v_i$ and all of their edges from $G$. Prove that $G'$ has at most $k$ connected components.

**Proof** Let $C$ be the Hamiltonian cycle in graph $G = (V, E)$ with $|V| = n$, then $V(C) = V(G) = n$. The cycle $C$ with $k$ vertices removed is always a disjoint union of at most $k$ paths, which can be proved by following induction.

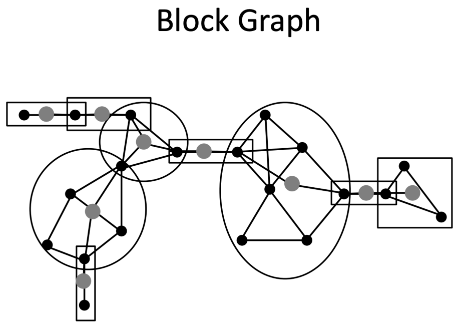
1. $k = 1$: By removing vertex $v_1$ and its incident edges in $C$, we would obtain a path $P = (v_2, \cdots, v_n)$.
2. $k = 2$: If we remove vertex $v_i$ with $3 \le i \le n-1$, then we get 2 connected components $(v_2, \cdots, v_{i-1})$ and $(v_{i+1}, \cdots, v_n)$; otherwise we would obtain only one connected component, either $(v_3, \cdots, v_n)$ or $(v_2, \cdots, v_{n-1})$.

   ...
3. $k = m+1$: Suppose that we can obtain at most $m$ disjoint paths if $m$ vertices are removed. When we are going to remove the $(m+1)^{\text{th}}$ vertex and its incident edges, the number of connected components will not increase when we remove isolated vertex or ends of some path. The only way to increase the number is to remove a middle vertex in some path to break the path into 2 pieces, which will only increase the number by $1$. Therefore we can obtain at most $m+1$ disjoint paths when $m+1$ vertices are removed, which finishes the induction argument.

Now, the cycle $C$ is partitioned into at most $k$ disjoint pieces $P_1, P_2, \cdots, P_k$ when $k$ vertices and their incident edges are removed, then $V(G') = V(P_1) \cup V(P_2) \cup \cdots \cup V(P_k)$. The subgraph $G'$ can be obtained from $P_1, P_2, \cdots, P_k$ by adding edges outside the Hamiltonian cycle $C$, which will not produce new connected component. Therefore, $G'$ has at most $k$ connected components.




**(a)** Blocks




**(b)** Block Graph

## 2.4 Structure of Connected Graphs

**Definition 2.15**

*Cut vertex and Block*

1. *A **cut vertex** in a connected graph $G$ is a vertex $v$ so that subgraph $G - v$ is disconnected.*
2. *A **block** is a maximal subgraph without cut vertices. See fig. 2.15a.*
3. *A **block graph** $B$ of the given connected graph $G$ is defined by*
   - *Vertices of $B$ are either cut vertices of $G$, or blocks of $G$.*
   - *There is an edge between a cut vertex $v$, and a block $B$ if and only if $v \in B$. See fig. 2.15b.*
4. ***Ear**: given a subgraph $H$ of $G$, an ear is a path $P$ in $G$ starting and ending at vertices $u, v \in H$, but intermediate vertices are outside $H$.*
5. *A **theta** in a graph is a pair of vertices with three vertex-disjoint paths between them.*
6. ***Cut Numbers**: Given a connected graph $G$ and vertices $u$ and $v$, $\kappa(u, v)$ is the minimum number of vertices that one needs to remove from $G$ to disconnect $u$ and $v$.*

♣

**Remark**

1. A cut vertex always separates at least two blocks.
2. Adding an ear can't create cut vertices.

**Lemma 2.8**

1. *Every edge is contained in some block.*
2. *The intersection of two blocks is either empty, or containing a single cut vertex.*
3. *Any block without a theta as a subgraph is either a $K_2$ or a cycle.*

♡

**Theorem 2.9 (Block Tree)**

*The block graph is always a tree, namely, connected and acyclic.*

♡

**Theorem 2.10 (Decomposition)**

*Any block is either a $K_2$ or can be obtained by starting with a cycle and adding ears.*

♡

**Proposition 2.2**

*Any connected, theta-less graph $G$ is a tree of cycles and $K_2$*

♠

**Theorem 2.11**

*For a finite graph $G$ with at least 3 vertices, the following are equivalent:*

1. *$G$ is a single block.($G$ has no cut vertices)*
2. *Any two edges of $G$ are in a common cycle.(Any two edges share a cycle)*
3. *Any two vertices of $G$ are in a common cycle.(Any two vertices share a cycle.)*

**Theorem 2.12 (Menger's Theorem)**

*There exist $\kappa(u, v)$ **vertex-disjoint** paths between $u$ and $v$ for any pair of vertices in any graph $G$.*

### 2.4.1 Exercises

**Exercise 2.25(Block Structure of Trees)** Let $T$ be a finite tree. Show that $T$ is the block graph of some other tree $T'$ if and only if the following condition holds: $T$ is a bipartite graph where all of the leaves are in the same part and the vertices in that part all have degree 1 or 2.

**Proof** "$\Longrightarrow$" Assume that $T$ is the block graph of tree $T'$. By definition, $T$ has some vertices that correspond to blocks in $T'$ and the other vertices correspond to cut vertices in $T'$, and all edges in $T$ connect one to the other. So we can identify $T$ as the bipartite graph where one part is all the vertices correspond to blocks and the other part is all the vertices correspond to cut vertices (See Figure 1). The leaves in $T$ correspond to blocks, since a cut vertex would separates blocks. Hence all of the leaves are in the same part.

Now we prove that all the vertices correspond to blocks have degree at most 2. Since $T'$ is a tree, each block is a $K_2$, either a leaf and its neighbor, or two adjacent non-leaf vertices. Hence in $T$ the vertices correspond to blocks can either be a leaf or be connected to 2 vertices correspond to cut vertices. So the degree is at most 2.

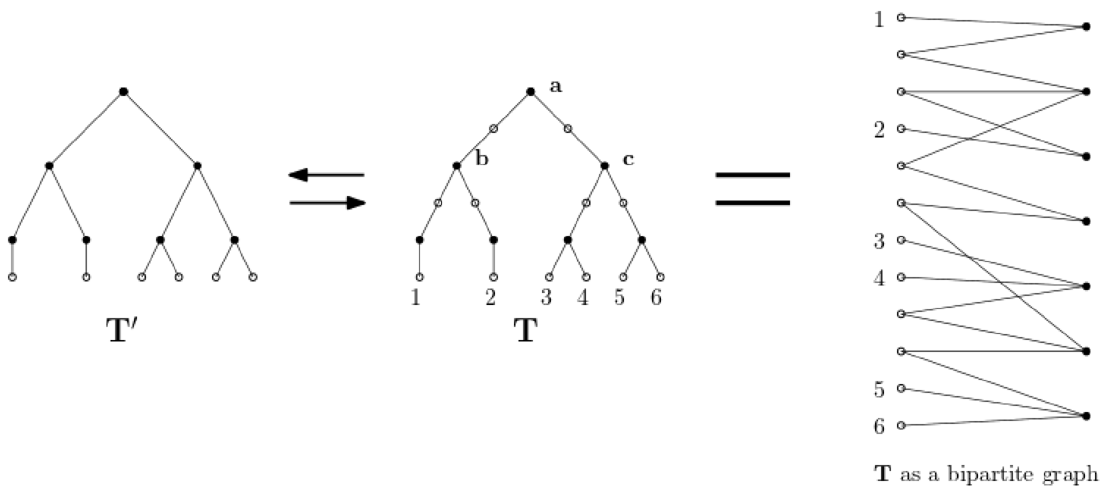



**Figure 2.16:** An example of a tree $T'$ and its block tree $T$.

"$\Longleftarrow$" Assume that $T$ is a bipartite graph where all of the leaves are in the same part and the vertices in that part all have degree 1 or 2. We can construct $T'$ in the following way:

The vertices of T' are the leaves of T and the vertices of T in the other part of the bipartition.

The edges of T' connect a leaf to its adjacent vertex or two cut vertices that have a single block vertex between them.

Inserting a block vertex between any pairs of adjacent cut vertices of $T'$, we get $T$, hence $T$ is the block tree of $T'$.

✍ **Exercise 2.26(Distinct Paths in a Block)** Let $G$ be a finite graph that consists of a single block with more than 2 vertices. Let $v$ and $w$ be two different vertices of $G$.

(a) Show that for any two distinct vertices $s$ and $t$ it is either possible to find an $s-v$ path and a $t-w$ path that do not share any vertices or it is possible to find a $t-v$ path and an $s-w$ path that do not share any vertices. [20 points]

(b) Show that if $G$ has $m$ edges and $n$ vertices that there are at least $m-n+2$ different paths from $v$ to $w$. Hint: Use induction on the number of ears in an ear decomposition of $G$. [20 points]

**Proof**

(a) By theorem 4.2.4 proven in class, vertex $v$ and $s$ share a cycle $C$. Let $p$ be the $t-w$ path. If $p$ doesn't intersect with $C$, then we can take an $s-v$ path in $C$ and it doesn't share any vertices with $p$, we are done. If $p$ intersects with $C$, let $x$ be the vertex where $p$ intersects $C$ at the first time, and $y$ be the vertex where $p$ intersects $C$ at the last time.

- Case 1, if $x$ and $y$ are on the same side of $C$, then there is an $x-y$ path $p'$ on $C$ that doesn't contain $s$ or $v$. We can connect $s$ and $v$ by the other side of cycle, and connect $t$ and $w$ by the $t-x-p'-y-w$ path. By construction these two paths doesn't share vertices.
- Case 2, if $x$ and $y$ are on different sides of $C$, then on $C$ there is an $s-y$ path and an $x-v$ path such that they don't share vertices. We can connect $s$ and $w$ by the $s-y-w$ path and connect $t$ and $v$ by the $t-x-v$ path.
- Case 3, $x$ or $y$ coincide with $s$ or $v$, i.e., $p$ passes through $s$ or $v$. By theorem 4.2.4, $t$ and $w$ share a cycle, so there is another $t-w$ path $p'$ which doesn't share vertices with $p$ except $t$ and $w$. If $p$ passes through both $s$ and $v$, then $p'$ cannot pass through either $s$ or $v$, so we go back to previous cases. If $p$ passes through $s$, and $p'$ passes through $v$, then $p$ contains an $s-t$ path, and $p'$ contains a $v-w$ path, by construction they don't share vertices.

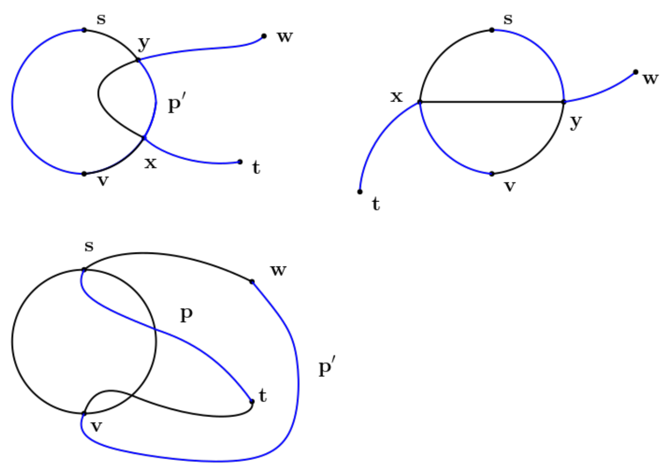



**Figure 2.17:** Constructing $s-v$ and $t-w$ paths in different cases.

(b) Note that $G$ is a cycle plus ears, we can do induction on the number of ears $k$.
When $k=0$, $G$ is a cycle, $m-n+2=2$, and there are two $v-w$ paths.
When $G$ has $k-1$ ears, since by adding one ear we add a path plus two edges, we increase $m-n$ by 1, so $m-n+2=k-1+2=k+1$. As inductive hypothesis, we assume that there are at least $k+1$ paths that connect $v$ and $w$. When $G$ has $k$ ears, $m-n+2=k+2$. Let $s,t$ be the start and end vertices

of the ear $P$ which doesn't contain $w$ or $v$. From part (a) we know that we have an $s - v$ path and a $t - w$ path that do not share any vertices, or we have a $t - v$ path and an $s - w$ path that do not share vertices. Hence we can define a new $v - w$ path as $v - s - P - t - w$ or $v - t - P - s - w$ which is different from the $k + 1$ paths that connect $v$ and $w$ in $G - P$, since only the new path contains $P$. So we have $k + 2$ different $v - w$ paths in $G$. By induction, we have shown that $G$ has $m - n + 2$ different $v - w$ paths.

## 2.5 Planar Graphs

### 2.5.1 Planarity and Euler Formula

**Definition 2.16**

*A **planar embedding** of a graph $G$ is a drawing of $G$ so that*

1. *Each vertex of $G$ corresponds to a point in the plane.*
2. *Each edge of $G$ corresponds to a curve connecting its endpoints.*
3. *No two edge-curves cross except at endpoints.*

*A graph $G$ is **planar** if it has a planar embedding.* ♣

**Remark** Applications.

- A circuit can be thought of as a graph with gates as vertices and wires as edges. With some technologies it is important to lay out circuits with few or no crossed wires.
- Given a map with simply connected regions, the adjacency graph on regions is planar.

**Lemma 2.9**

*Let $G$ be a planar graph. Then every subgraph of $G$ is also planar.* ♡

**Definition 2.17 (Faces)**

*A planar embedding of a graph divides the plane into regions, which are called **faces**. The number of **sides** of the face is the number of edges in this cycle, while every face has at least 3 sides.* ♣

**Theorem 2.13 (Euler's Formula)**

*For any planar embedding of a connected graph $G$ with $v$ vertices, $e$ edges and $f$ faces (including the infinite face)*

$$v - e + f = 2\,. \tag{2.5}$$

♡

**Proof** We begin by proving our result for trees, where $e = v - 1$, so we need only show $f = 1$. Use induction on $v$. For $v > 1$, contracting a leaf into the tree doesn't change number of faces.

For general graphs, use induction on $e$. Base case, $G$ is a tree. Otherwise, $G$ has a cycle, while cycle separates plane into inside and outside. Remove edge of cycle, decreases $f$ by 1, which implies $v - (e - 1) + (f - 1) = 2$.

**Lemma 2.10 (Dual Handshake Lemma)**

*For a connected, planar graph*

$$\sum_{\textit{Faces } f} \textit{Sides}\,(f) = 2|E|\,. \tag{2.6}$$

♡

**Proof** Count the number of pairs of an edge as a side of a face (be careful to count edges that are double sides twice). Each edge on two faces and each face $f$ has Sides $(f)$ edges.

**Remark** If $G$ only has faces with at least $k$ sides then, $e \geq kf/2$.

**Theorem 2.14 (Edge Bound)**

*If $G$ is a connected planar graph with $|V| \geq 3$, then*

$$|E| \leq 3|V| - 6\,. \tag{2.7}$$

**Proof** Combine $v - e + f = 2$ and $e \geq 3f/2$. The equality holds if and only if all faces are triangles.

**Remark** If each face has at least $k$ sides, the maximum number of edges is $(v-2)/(1-2/k)$.

**Theorem 2.15 (Minimum Degree)**

*If $G$ is a finite, connected planar graph, its vertices have minimum degree at most 5.*

**Proof** Otherwise, each vertex has degree 6 or more. The Handshake lemma implies $2e \geq 6v$, then $3v - 6 \geq e \geq 3v$. Contradiction.

**Lemma 2.11 (Triangulations)**

*For any planar embedding of a graph $G$, there is a way to add more edges to $G$ to get a new planar graph $G'$ in which all faces are triangles.*

**Theorem 2.16 (Fary's Theorem)**

*Any finite (simple) planar graph $G$ has a plane embedding where all of the edges are straight line segments(with the same topology).*

**Proof** Proof Strategy.

- Induct on $v$. The case $v \leq 3$ is trivial.
- Assume $G$ is connected (otherwise draw each component separately.)
- Find a vertex $v$ with $deg(v) \leq 5$.
- Remove $v$ and triangulate $G_1 = G - v$.
- Add $v$ back.

**Lemma 2.12**

*Given any polygon $P$ in the plane with at most 5 sides, there is a point $v$ inside of $P$ with straight line paths to each of $P$'s vertices.*

**Proof** [proof of lemma] Consider cases based on the locations of the non-convex (i.e. $>$ 180 degrees) angles. There are at most 2 non-convex angles in a polygon, and other cases follow by enumeration. (The lemma doesn't hold for hexagons.)

### 2.5.2 Non-planarity and subdivision

**Theorem 2.17**

*Non-planar graphs*

1. *$K_5$ is non-planar. If it were, we would have $e \leq 3v - 6 = 9$. However $e = 10$ for $K_5$.*
2. *$K_{3,3}$ is non-planar. It has no odd cycles. Therefore, if $K_{3,3}$ is planar, any face of it has at least 4 sides, and we would have $e \leq (v-2)/(1-2/k) = 8$ edges. However, $e = 9$ for $K_{3,3}$.*

**Definition 2.18**

*A **subdivision** of a graph G is obtained by placing vertices in the middle of some of its edges.* ♣

**Lemma 2.13**

*If $G'$ is a subdivision of $G$, then $G'$ is planar if and only if $G$ is.* ♡

**Remark** The act of subdividing a graph does not change the planarity of the graph at all, since the fundamental shape (the topological shape) has not changed.

**Theorem 2.18 (Kuratowski's Theorem)**

*A finite graph $G$ is planar if and only if it has no subdivision of a $K_5$ or $K_{3,3}$ as a subgraph.* ♡

### 2.5.3 Polytope and convexity

**Definition 2.19 (Polygon, polyhedron and polytope)**

*In geometry,*

- *a **polygon** is a 2-dimensional figure that is described by a finite number of straight line segments connected to form a closed polygonal chain.*
- *a **polyhedron** is a 3-dimensional figure bounded by finitely many flat faces. Two faces meet at an edge and edges meet at vertices.*
- *a **polytope** is a generalization in any number of dimensions of the 3-dimensional polyhedron. The sides of a $(n+1)$-polytope consist of $n$-polytopes, which may have $(n-1)$-polytopes in common.* ♣

**Definition 2.20 (Convexity)**

*A set $K \subset \mathbb{R}^n$ is convex if, for each pair of distinct points $a, b \in K$, the closed segment with endpoints $a$ and $b$ is contained within $K$ (a compact convex set with a finite number of extreme points, vertex representation).* ♣

**Definition 2.21**

*Geometrically, a polyhedron is **convex** if it contains every line segment connecting two of its points. Algebraically, a convex $n$-polytope may be defined as an intersection of a finite number of half-spaces(half-space representation), which can be written as the set of solutions to a system of linear inequalities*

$$\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\,, \tag{2.8}$$

*where $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{x} \in \mathbb{R}^n$ and $\boldsymbol{b} \in \mathbb{R}^m$.* ♣

**Remark** A more subtle distinction between convex and non-convex polyhedron is given by their **Euler characteristic**, which combines the numbers of vertices $V$, edges $E$, and faces $F$ of a polyhedron into a single number $\chi$ defined by the formula $\chi = V - E + F$. Any convex polyhedron's surface has Euler characteristic $\chi = V - E + F = 2$. However, the surfaces of non-convex polyhedra can have various Euler characteristics.

**Definition 2.22 (Polyhedral Graphs)**

*Given a convex polyhedron, it can be turned into a planar graph by projecting vertices/edges onto a sphere, which can then be flattened onto a plane by a linear map.* ♣

**Definition 2.23 (Regular Polyhedra)**

*A regular polyhedron is a highly symmetric polyhedron, such as cubes, which has the following properties:*

- *all edges are the same length.*
- *all faces are regular polygons with the same number $s$ of sides.*
- *The same number of faces, $d$, meet at each vertex.*

♣

### 2.5.4 Exercises

**Exercise 2.27(Euler's Formula and Regular Tessellations)** Let $G$ be an infinite graph with a planar embedding. Say that the embedding is periodic if translating all of the edges and vertices one unit up or to the left yields the same collection of edges and vertices. Say that the embedding is locally finite if any disk intersects only finitely many edges and vertices.

(a) For $G$ a locally finite, periodic planar graph, say that two vertices are equivalent if you can obtain one from the other by translating it an integer number of units up and an integer number of units to the left. Let $v$ be the number of equivalence classes of vertices. Similarly, let $e$ and $f$ be the number of equivalence classes of edges and faces of $G$. Prove that $v - e + f = 0$. [20 points]

(b) Suppose that such a graph $G$ is $d$-regular and has exactly $s$ sides per face. What are all possible values of $d$ and $s$ for which this can happen? [20 points]

**Proof**

(a) Since $G$ is locally finite, the intersection of $G$ and a disk is a finite subgraph $H$. Let $V, E, F$ be the number of vertices, edges and faces of $H$, since $H$ has a planar embedding, by Euler formula we have

$$V - E + F = 2. \tag{2.9}$$

To relate $V, E, F$ with $v, e, f$, we assume $H$ contains $N_1$ units, defined as the minimal subgraph that generates $G$ by translation (see Figure 3[1]) and intersects with $N_2$ units. Obviously $N_1 \leqslant N_2$.

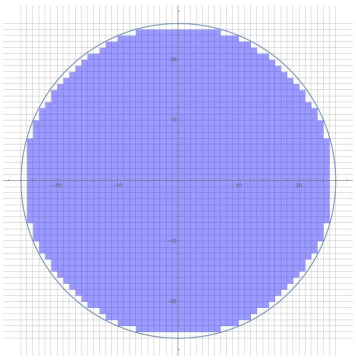


**Figure 2.18:** the shape in blue is the $N_1$ units contained in the disk.

Since $H$ has more edges, vertices and the infinite face than $N_1$ units, we have

$$V \geqslant N_1 v, E \geqslant N_1 e, F \geqslant N_1 f, \tag{2.10}$$

on the other hand, $H$ is contained in $N_2$ units, we have

$$V \leqslant N_2 v, E \leqslant N_2 e, F \leqslant N_2 f, \tag{2.11}$$

[1]Source: http://www.maths.surrey.ac.uk/hosted-sites/R.Knott/Numbers/year2016.html

hence

$$2 = V - E + F \geqslant N_1 v - N_2 e + N_1 f, \tag{2.12}$$

rearranging, we have

$$v - \frac{N_2}{N_1} e + f \leqslant \frac{2}{N_1} \tag{2.13}$$

Note that this is true for any size of disk. If we let the radius of the disk go to infinity, then $N_1$ and $N_2$ will go to infinity, but $\frac{N_2}{N_1}$ will go to 1, since for any finite graph, when the disk is large enough to cover the entire graph, we have $N_1 = N_2$. Hence we have

$$v - e + f \leqslant 0. \tag{2.14}$$

Meanwhile,

$$2 = V - E + F \leqslant N_2 v - N_1 e + N_2 f, \tag{2.15}$$

rearranging, we have

$$v - \frac{N_1}{N_2} e + f \geqslant \frac{2}{N_2}, \tag{2.16}$$

let the radius of the disk go to infinity, using the same argument, we have

$$v - e + f \geqslant 0, \tag{2.17}$$

hence we have

$$v - e + f = 0. \tag{2.18}$$

(b) Using the same technique as (a), we consider the finite subgraph $H$ which is the intersection of $G$ and a disk. By applying the handshake lemma and the dual handshake lemma to $H$, we have

$$2E = dV, \quad 2E = sF, \tag{2.19}$$

and from (2)(3) we have

$$2N_1 e \leqslant 2E = dV \leqslant dN_2 v, \tag{2.20}$$

hence

$$2e \leqslant \frac{N_2}{N_1} dv, \tag{2.21}$$

let the radius of the disk go to infinity, in (a) we have shown that $\frac{N_2}{N_1}$ will go to 1, so we have

$$2e \leqslant dv. \tag{2.22}$$

Meanwhile

$$2N_2 e \geqslant 2E = dV \geqslant dN_1 v, \tag{2.23}$$

hence

$$2e \geqslant \frac{N_1}{N_2} dv, \tag{2.24}$$

let the radius of the disk go to infinity we have

$$2e \geqslant dv, \tag{2.25}$$

so from (14)(17) we have

$$2e = dv. \tag{2.26}$$

Similarly, from the dual handshake lemma in (11), we have

$$\frac{N_1}{N_2} sf \leqslant 2e \leqslant \frac{N_2}{N_1} sf, \tag{2.27}$$

let the radius of the disk go to infinity we have

$$2e = sf. \tag{2.28}$$

Plug equation (18)(20) in (10) we have

$$\frac{2e}{d} - e + \frac{2e}{f} = 0, \tag{2.29}$$

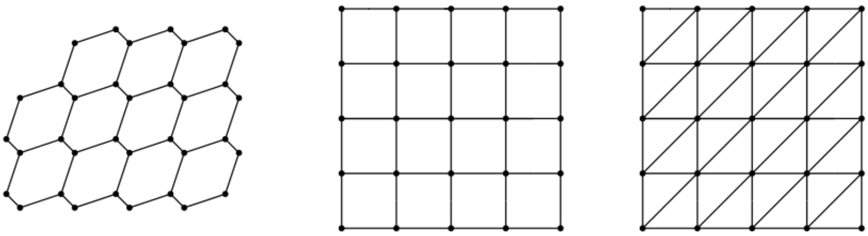
rearranging, since $e \neq 0$, $d, s$ are the solution of

$$\frac{1}{d} + \frac{1}{s} = \frac{1}{2}, \tag{2.30}$$

and we have

$$d = \frac{2s}{s-2} = 2 + \frac{4}{s-2}. \tag{2.31}$$

Since $d, s$ are positive integers, $s - 2$ must divide 4, hence $s$ can only be $3, 4$ or 6. So all possible values of $d$ and $s$ are $\{d = 3, s = 6\}$, $\{d = 4, s = 4\}$ and $\{d = 6, s = 3\}$. See Figure 4.



$\{\mathbf{d = 3, s = 6}\}$ $\{\mathbf{d = 4, s = 4}\}$ $\{\mathbf{d = 6, s = 3}\}$

**Figure 2.19:** The regular tessellations.

✍ **Exercise 2.28(Coloring Almost Planar Graphs)** Call a simple graph $G$ almost planar if it can be drawn in the plane with at most one pair of crossing edges. Show that any almost planar graph $G$ has $\chi(G) \leq 5$.

**Proof** Suppose that $G$ is an almost planar graph with an embedding where only the edges $(a, b)$ and $(c, d)$. Removing vertex a from the graph, we get a planar embedding of $G \setminus a$ (since no remaining edges cross). By the Four Color Theorem, we can color $G \setminus a$ with at most four colors. Coloring a with a fifth color, shows that $\chi(G) \leq 5$.

✍ **Exercise 2.29** Let $n$ and $k$ be positive integers. What is the greatest number of edges that a simple graph $G$ with $nk$ vertices and chromatic number at most $k$ can have?

**Proof** The answer is $(k-1)n^2/2$. This can be achieved by splitting the vertices into $k$ groups of size $n$ and drawing edges between each pair of vertices not in the same group. This graph is $k$-colorable by coloring each group a different color. To show that this is optimal, we note that if a $k$-coloring of $G$ has $n_i$ vertices of the $i$th color, then we have $\sum_i n_i = nk$. Furthermore, $G$ cannot have any edges between vertices of the same color. Thus, the greatest number of edges that $G$ can have is the number of pairs of vertices not in the same group. This is at most $(nk)(nk-1)/2 - \sum_i n_i(n_1 - 1)/2$. It is not hard to see that this is maximized when $n_i = n$ for all $1 \leq i \leq k$.

## 2.6 Coloring

### 2.6.1 Vertex Coloring

**Definition 2.24**

*Vertex coloring, max Degree, chromatic Number and clique number.*

- *A **vertex coloring** of a graph $G$ is an assignment of a color to each vertex of $G$, such that no two adjacent vertices have the same color. This is a $k$-coloring if only $k$ different colors are used.*
- *The **chromatic Number** $\chi(G)$ of a graph $G$ is the smallest number $k$ so that $G$ has a $k$-coloring.*
- *The **clique number** $\omega(G)$ of a graph $G$ is the largest $n$ so that $K_n$ is a subgraph of $G$.*
- *The **max Degree** $\Delta(G) := \max_v d(v)$ denotes the maximum degree of any vertex of $G$.*

♣

**Remark** For $G$ a graph on $n$ vertices, we have $\chi(G) = 1$ when $G$ consists of isolated vertices; $\chi(G) = 2$ if and only if $G$ is bipartite. Any tree $T$ is bipartite, so $\chi(T) = 2$. For cycles, $\chi(C_n) = 2$ for $n$ even, and 3 for $n$ odd. For complete graph, $\chi(K_n) = n$ since every 2 vertices are adjacent. Also, we have the following upper and lower bounds(far from tight)

$$\omega(G) \leq \chi(G) \leq n\,. \tag{2.32}$$

**Lemma 2.14**

*For any graph $G$, $\chi(G) \leq \Delta(G) + 1$. (Proved by Greedy Coloring.)*

♡

#### 2.6.1.0.1 Greedy Coloring Strategy

- Color one vertex at a time, giving each a color that doesn't conflict.
- If a vertex $v$ has $d(v)$ neighbors, it is enough to have $d(v) + 1$ colors to choose from.

**Remark** This upper bound is often far from tight. For example, $\Delta(K_{n,n}) = n$ however $\chi(K_{n,n}) = 2$.

**Theorem 2.19 (Brook's Theorem)**

*If a finite connected graph $G$ is neither an odd cycle nor a complete graph, then $\chi(G) \leq \Delta(G)$.*

♡

**Remark** We have $\chi(C_n) = 3$ and $\Delta(C_n) = 2$ for cycle when $n$ odd; and $\chi(K_n) = n$ while $\Delta(K_n) = n - 1$ for complete garph.

**Proof** Algorithm to color different type graphs.

- Not regular: greedy coloring ending at $v$.
- Not 3-connected: Break into parts, color parts inductively, recolor to make them match.
- 3-connected, find $v$ with non-adjacent neighbors $u$,$w$. Greedily color so that $u$,$w$ assigned same color first, $v$ assigned color last.

**Proposition 2.3 (Chromatic Numbers and Subgraphs)**

*Let $G$ be a finite graph with a complete subgraph $H$. Suppose that $G \setminus H$ is split into disjoint connected components $E$ and $F$. Prove that $\chi(G) = max(\chi(H \cup E), \chi(H \cup F))$, where $H \cup E$ and $H \cup F$ denote the induced subgraphs on the relevant set of vertices.*

♠

**Proof** Let $k = max(\chi(H \cup E)$. On the one hand, any coloring of $G$ restricts to colorings of $H \cup E$ and $H \cup F$, which means that $\chi(G) \geq k$. On the other hand, by assumption we have colorings of $H \cup E$ and $H \cup F$ with only $k$ colors each. Since $H$ is a complete subgraph, each coloring must assign the vertices in $H$ distinct colors. By renaming the colors in the coloring of $H \cup F$, we can make the colors of the vertices in $H$ agree with those in the coloring of $H \cup E$ and so that the other colors used are the same. Once these colorings use the same list of $k$ colors and agree on how they color $H$, we can combine them to make a $k$-coloring of $G$ simply by coloring each $v \in G$ the color that it would be assigned in either the coloring of $H \cup E$ or $H \cup F$. It is easy to check that this coloring works, and thus that $\chi(G) \leq k$.

### 2.6.2 Coloring Planar Graphs

**Theorem 2.20 (Kempe)**

*Every planar graph is* 5*-colorable.*

**Proof** See proof for Theorem 1.47 in *Combinatorics and Graph Theory, John M. Harris, Jeffry L. Hirst, Michael J. Mossinghoff.* Induction on number of vertices.

**Theorem 2.21 (The Four Color Theorem)**

*Every planar graph is* 4*-colorable.(Optimal)*

**Proof** Same idea. Thousands cases to be verified and computer assisted.

### 2.6.3 Edge Coloring

**Definition 2.25**

*An **edge coloring** of a graph is an assignment of a color to each edge so that no two edges incident on the same vertex have the same color.*

**Lemma 2.15**

*Any edge coloring of a graph $G$ requires at least $\Delta(G)$ colors.*

**Proof** At the maximum-degree vertex, all $\Delta(G)$ edges need to different colors.

**Lemma 2.16**

*There always is a coloring with at most $2\Delta(G) - 1$ colors.(Proved by greedy algorithm.)*

**Theorem 2.22 (Vizing's)**

*Any finite graph $G$ has an edge coloring with at most $\Delta(G) + 1$ colors.*

### 2.6.4 Exercises

✍ **Exercise 2.30(Uniqueness of Planar Embeddings)**

(a) Let $G = (V, E)$ be a connected, planar graph with $|E| = 3|V| - 6$. Show that any two planar embeddings of $G$ have the same set of faces (in particular, that if one embedding has a face whose sides consist of some collection of edges, then the other planar embedding will have a face with the same sides). Hint: Note that $G$ must be triangulated. Consider the faces including a given vertex $v$. [30 points]

(b) Show that this is no longer the case if we drop the condition $|E| = 3|V| - 6$. In particular, give a connected, planar graph $G$ with two different embeddings that have different faces. [10 points]

**Proof** By Euler's formula, $|V| = 2 + |E| - |F|$. Thus, $|E| = 3|V| - 6 = 3|E| - 3|F|$, and equivalently, $2|E| = 3|F|$. Since $|E| > 0$, $|V| \geq 3$, and thus every face has at least 3 sides. By the dual hand shake lemma, $3|F| = 2|E| = \sum_{f \in F} \text{sides}(f) \geq 3|F|$. Thus, the $\geq$ in the equation is actually an equality, which is true if and only if every face has exactly 3 sides. Hence, $G$ is triangulated.

Let $A, B : G \to \mathbb{R}^2$ be two embeddings of $G$, and $f$ be a face in $A$. Then, $f$ has 3 sides, say $e_1 = \{u, v\}$, $e_2 = \{v, w\}$ and $e_3 = \{w, u\}$. It suffices to show that in the embedding $B$, $e_1$, $e_2$ and $e_3$ are the sides of a face. Let $D$ be the closed region bounded by edges $e_1$, $e_2$ and $e_3$ in the embbeding $B$.

Case 1: If $D$ contains no other vertices, then $e_1$, $e_2$ and $e_3$ form a face.

Case 2: Suppose that $D$ contains some vertex $x$ other than $u, v$, and $w$. Suppose for contradiction that there exists a vertex $y$ such that $B(y) \notin D$. Then, since $G$ is connected, there exists a path $P$ from $y$ to $x$. The first intersection point of $B(P)$ and $D$ must be among $\{u, v, w\}$ since $B$ is a embedding of planar graph. By relabelling the vertices, we may assume without loss of generality that the first inter section point is $v$. Then, along the path $P$, there is a neighbour $v_1$ of $v$ that is outside of $D$ and a neighbour $v_2$ of $v$ that is inside of $D$. Consider the neighbours of $v$ in the embedding $A$. Since every face has only 3 sides, by tracing the edges between the neighbours of $v$, we have a cycle containing all neighbours of $v$, on which $u$ and $w$ are neighbours. Thus, there is a path $P'$ from $v_2$ to $v_1$ which does contain $u, v$ or $w$. However, by the same argument above, any path from $v_2$ (outside $D$) to $v_1$ (inside $D$) has to pass through $u$, $v$, or $w$, contradiction. Hence, there is no vertex $y$ with $B(y) \notin D$. Thus, $D$ contains all vertices, and $e_1$, $e_2$ and $e_3$ form a face.

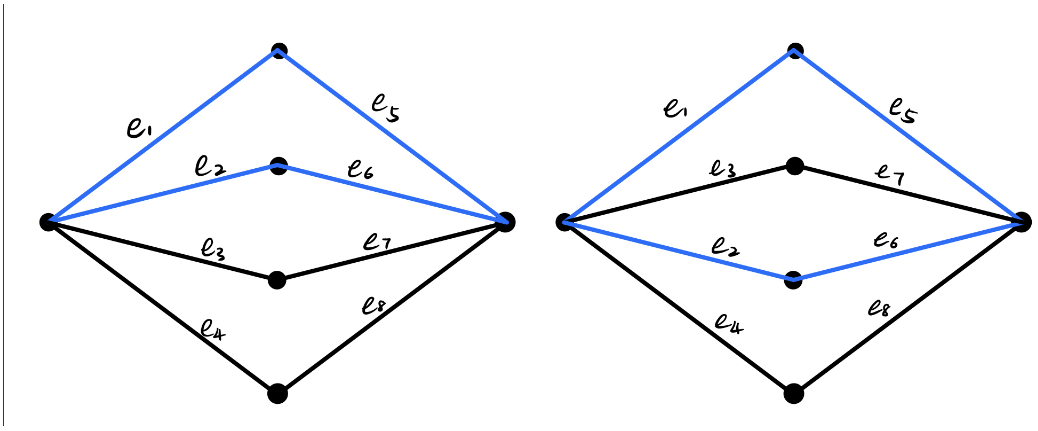



**Figure 2.20:** The graph G and two embeddings

**Proof** In the first embedding, there is a face with sides $\{e_1, e_2, e_5, e_6\}$, while in the second embedding, there is no face with sides $\{e_1, e_2, e_5, e_6\}$.

✍ **Exercise 2.31(Triangle-less Chromatic Number)** Give an example of a graph $G$ which contains no triangle and which has chromatic number at least 4.

**Proof** Let $G$ be the graph as in figure 2. It has no triangles. Suppose for contradiction that $G$ is 3-colorable by colors $\{a, b, c\}$. Suppose without loss of generality that vertex 1 is of color $a$. Since the outer cycle is odd, all three color will be used on the outer cycle, in particular, there is one vertex of color $a$. Suppose without loss of generality that vertex 7 has color $a$, and that vertex 2 is of color $b$. Then since vertex 2 has color $b$ and vertex 7 has color $a$, vertex 8 and 11 both have color $c$. Then, based on the colors on vertex 1, 8, and 11, vertex 5 and 4 should have color $b$. Thus, vertex 9 and 10 are both of color $a$, which is a contradiction to the fact that they are neighbours.

✍ **Exercise 2.32(Number of Colorings for Graphs with Small Maximum Degree)** Let $G$ be a connected graph with $m$ vertices and let $n \geq \Delta(G) + k$ for some positive integer $k$. Show that there are at least $(k+1)^{m-1}$

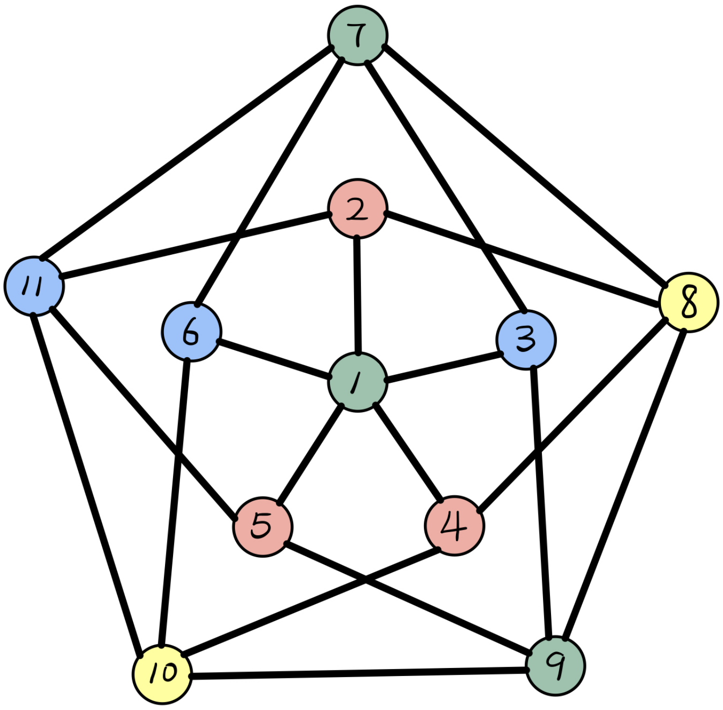



**Figure 2.21:** The graph G and a 4-coloring of G

ways to color $G$ with any set of $n$ colors.

**Proof** Run the BFS algorithm and let $(v_1, v_2, \ldots, v_m)$ be the sequence of all vertices in the order we discover them. Then for $i \geq 2$, $v_i$ has a neighbour $v_j$ for $j < i$. Thus, for $i \geq 2$, the number of neighbours of $v_i$ after itself is most $\Delta(G) - 1$.

Color the vertices using the greedy algorithm in the order $(v_m, \ldots, v_2, v_1)$. That is, choose an arbitrary color for $v_m$, and given the coloring of $(v_m, \ldots, v_{i+1})$, choose an arbitrary color for $v_i$ that is different from the colors of its neighbours that are already colored. For each edge $\{v_i, v_j\}$, suppose $i < j$, and then by the way we choose the color on $v_i$, its color is different from that of $v_j$. Then, for each $i \geq 2$, at the step where we choose the color of $v_i$, $v_i$ has at most $\Delta(G) - 1$ colored neighbours, and we have at least $n - (\Delta(G) - 1) = k + 1$ choices for the color of $v_i$. Thus, this procedure gives at least $(k+1)^{m-1}$ distinct colorings of $G$.

✍ **Exercise 2.33** Given an order $v_1, \ldots, v_n$ of the vertices, the greedy coloring algorithm assigns vertex $v_1$ color 1, and, in general, assigns $v_i$ the first color from $\{1, 2, \ldots\}$ that has not yet appeared on any neighbor $v_j$ of $v_i$ with $j < i$.

(a) Show that every graph $G$ has a vertex ordering for which the greedy coloring algorithm uses $\chi(G)$ colors.

**Proof** *The goal of this problem was to think about the behavior of the greedy coloring algorithm.*

Let $G$ be an $n$-vertex graph. Consider a coloring $c : V(G) \to \{1, \ldots, \chi(G)\}$ of $G$ using $\chi(G)$ colors. Order the vertices $v_1, \ldots, v_n$ of $G$ as follows: start with the vertices colored 1 in $c$, then those of color 2, etc.

We will show using induction on $1 \leq i \leq n$, that the greedy algorithm applied to this graph with this ordering will assign vertex $v_i$ color at most $c(v_i)$. For the base case $i = 1$, note that the greedy algorithm assigns color 1 to $v_1$. For the induction step $2 \leq i \leq n$, by the induction hypothesis we know that, for every $1 \leq j < i$, the greedy algorithm has assigned vertex $v_j$ color at most $c(v_j)$. Note that the only vertices $v_j$ with $c(v_j) = c(v_i)$ are those in the same color class with $v_i$ in the optimal coloring $c$, and those are not adjacent to $v_i$. Hence the colors used on earlier neighbors of $v_i$ are all in the set $\{1, \ldots, c(v_i) - 1\}$, and the greedy algorithm thus assigns vertex $v_i$ color at most $c(v_i)$, completing the proof of the induction step.

Therefore the greedy algorithm uses $\chi(G)$ colors (by the inductive argument at most $\chi(G)$ colors are used, which implies that exactly $\chi(G)$ colors are used, since no proper coloring can use fewer than $\chi(G)$ colors).

(b) For the path $P$ on 4 vertices (which clearly satisfies $\chi(P) = \Delta(P) = 2$), show that there is a vertex ordering for which the greedy coloring algorithm uses $\Delta(P) + 1 = 3$ colors.

**Proof** *The goal of this problem was to learn about suboptimality of the greedy coloring algorithm.*
If we order the endpoints first and second, they will be colored the same color. To color the two interior vertices, we would then need two additional colors, for a total of 3.

(c) For every $n \geq 2$, construct a bipartite graph with $2n$ vertices and an vertex ordering for which the greedy coloring algorithm uses $n$ colors rather than 2 colors.

**Proof** *The goal of this problem was to demonstrate suboptimality of the greedy coloring algorithm.*
Consider the bipartite graph $G = (A \cup B, E)$ where $A = \{a_1, \ldots, a_n\}$, $B = \{b_1, \ldots, b_n\}$ and there is an edge between $a_i$ and $b_j$ iff $i \neq j$. Order the vertices in the order $a_1, b_1, a_2, b_2, \ldots, a_n, b_n$.

We will show by induction on $1 \leq i \leq n$ that the greedy algorithm applied to this graph with this ordering will, for all $1 \leq j \leq i$, assign color $j$ to vertices $a_j$ and $b_j$. For the base case $i = 1$, note that the greedy algorithm assigns color 1 to $a_1$ and $b_1$ (here we exploit that there is no edge $a_1 b_1$). For the induction step $2 \leq i \leq n$, by the induction hypothesis we know that, for all $1 \leq j < i$, the greedy algorithm has assigned color $j$ to vertices $a_j$ and $b_j$. When assigning a color to $a_i$, $a_i$ is adjacent to $b_1, b_2, \ldots, b_{i-1}$, which have been assigned colors $1, 2, \ldots, i-1$ by our inductive hypothesis. So $a_i$ will be assigned color $i$. Similarly, $b_i$ will be assigned color $i$ (here we exploit that there is no edge $a_i b_i$), completing the proof of the induction step.

Therefore, since vertices $a_n$ and $b_n$ are assigned color $n$, the greedy algorithm needs $n$ colors.

✍ **Exercise 2.34** (a) Prove that $\chi'(G) \geq e(G)/\alpha'(G)$ for any graph $G$ with $e(G) \geq 1$ edges.

**Proof** *This goal of this problem was to review the proof of $\chi(G) \geq n(G)/\alpha(G)$ from class.*
Suppose we have an edge coloring of $G$ with $\chi'(G)$ colors. Let $E_i$ denote the set of edges colored with color $i$. Since the color sets partition the edges, we have $e(G) = \sum_{1 \leq i \leq \chi'(G)} |E_i|$. Since each set of same colored edges must be disjoint, they form a matching and are therefore smaller than the maximum matching. So

$$e(G) = \sum_{1 \leq i \leq \chi'(G)} |E_i| \leq \sum_{1 \leq i \leq \chi'(G)} \alpha'(G) = \alpha'(G) \cdot \chi'(G),$$

and the result follows by dividing by $\alpha'(G)$.

(b) Prove that $e(G) \geq \binom{\chi(G)}{2}$ for any graph $G$.

**Proof** *This problem tests your understanding of basic graph theory definitions (chromatic number).*
Suppose we have a proper coloring containing $\chi(G)$ colors. Then there must be an edge between every pair of colors. To see this, assume otherwise, that there exist two colors, e.g. red and blue, which contain no edges between them. Then, coloring all red vertices blue as well would give a proper coloring with one fewer colors, contradicting the minimality of $\chi(G)$. There are $\binom{\chi(G)}{2}$ pairs of colors, so there are at least that many edges.

✍ **Exercise 2.35** A graph $G$ is called $d$-degenerate if $\delta(H) \leq d$ for all induced subgraphs $H \subseteq G$. In class we proved that $d$-degenerate graphs $G$ satisfy $\chi(G) \leq 1 + d$ (see also Section 6.4 in the Verstraete lecture notes).

(a) Prove that any $n$-vertex $d$-degenerate graph $G$ contains at most $e(G) \leq dn$ edges.

**Proof** *This goal of this problem was to train using induction (a classical proof technique) in graph theory.*
We proceed by induction on the number of vertices $n$. For the case case $n = 1$, note that we trivially have $e(G) = 0 \leq dn$. For the inductive step $n \geq 2$, let $G$ be a $d$-degenerate graph with $n$ vertices. Since $G$ is $d$-degenerate and is an induced subgraph of itself, there must be a vertex $v$ with degree $d_G(v) = \delta(G) \leq d$.

Delete $v$ from $G$ to obtain an induced subgraph $G' := G - \{v\}$. Note that $G'$ is again $d$-degenerate, since any induced subgraph $H \subseteq G' \subseteq G$ is also an induced subgraph of $G$. Invoking our inductive hypothesis we thus infer $e(G') \leq d(n-1)$, so that by construction of $G'$ we obtain

$$e(G) \leq e(G') + d_G(v) \leq d(n-1) + d = dn,$$

completing the proof by induction.

(b) Prove that if the longest path in $G$ has at most $\ell$ vertices, then $\chi(G) \leq \ell$.

**Proof** *This goal of this problem was to train using extremality arguments in graph theory.*

By the result from class, it suffices to show that if the longest path in $G$ has at most $\ell$ vertices, then $G$ is $(\ell-1)$-degenerate. To prove this, consider any induced subgraph $H \subseteq G$. Consider the longest path in $H$, and let $v$ be one of its endpoints. Every edge incident to $v$ must connect back to the path, or the path could be extended. Since any path contains at most $\ell$ vertices, we infer that $\delta(H) \leq d_H(v) \leq \ell - 1$ (note that the vertex $v$ itself also counts as one vertex on the path). Since this holds for any induced subgraph $H \subseteq G$, we have established that $G$ is $(\ell-1)$-degenerate, as desired.

**Exercise 2.36** Given an integer $g \geq 3$, assume that $G$ is a connected planar graph which contains no cycles with less than $g$ vertices, and for which the number of vertices is at least $n(G) \geq g/2 + 1$. Using Euler's formula, prove that $G$ has at most $e(G) \leq \frac{g}{g-2}(n-2)$ edges.

**Proof** This is essentially Theorem 7.1.3 in the Verstraete notes, though there is one key difference: we do *not* assume that $G$ contains a cycle (which explains the case distinction below).

We first consider the case when $G$ contains no cycle: then $G$ is a tree (as it is connected and acyclic), and so

$$e(G) \leq n(G) - 1 \leq \frac{g}{g-2}\Big(n(G) - 2\Big),$$

where the last inequality is equivalent to the assumption $n(G) \geq g/2 + 1$ (by elementary algebra).

It remains to consider the case when $G$ contains a cycle, which by assumption contains at least $g$ vertices. In this case every face $F$ in graph $G$ has degree at least $g$. By the Handshake Lemma it follows that

$$2e(G) = \sum_{F \in F(G)} \deg(F) \geq g|F(G)|.$$

By Euler's formula it follows that

$$2 - n(G) + e(G) = |F(G)| \leq \frac{2}{g} e(G).$$

Solving this inequality for $e(G)$ yields

$$e(G) \leq \frac{g}{g-2}\Big(n(G) - 2\Big),$$

completing the proof of the lemmaed edge-bound.

**Exercise 2.37** *The goal of this problem is to become more familiar with the planarity concept, in particular to learn how edge-counts (and additional cycle-length information) can be used to prove that certain graphs are not planar.*

In the following parts (a)–(c) you may use Problem 1 above (or use any result from class):

(a) Answer Question 7.1 from the Verstraete Lecture notes.

**Proof**

**(1)** This graph is not planar. From Figure 2.22 we infer that $|V(G)| = 10$; furthermore, the length of each cycle is at least 5. Invoking Problem 1 with $g = 5$ then gives the edge bound $|E(G)| \leq \frac{5}{3}(|V(G)| - 2) = \frac{40}{3} \approx 13.3$, contracting the fact that $|E(G)| = 15$.

**(2)** This is a planar graph: a planar embedding is shown in Figure 2.22.

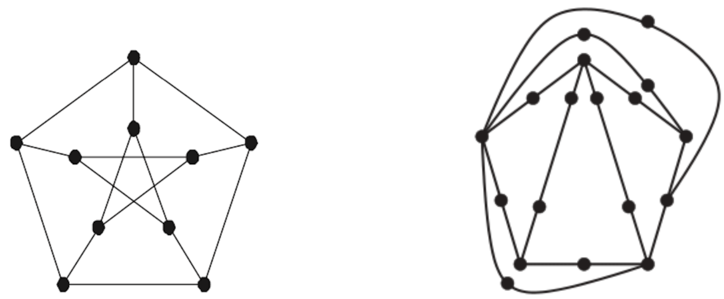


**Figure 2.22:** Problem 2(a): graph (1) on the left and graph (2) on the right.

**(3)** This graph is not planar. From Figure 2.23 we infer that $|V(G)| = 16$; furthermore, the graph is 4-regular graph with the shortest cycle having length 4. Invoking Problem 1 with $g = 4$ then gives the edge bound $|E(G)| \le 2|V(G)| - 4 = 28$, contradicting the fact that $|E(G)| = 4|V(G)|/2 = 32$ by Handshake lemma.

**(4)** This graph is not planar. From Figure 2.23 we infer that 2.23, $|V(G)| = 14$; furthermore, the shortest cycle has length 6. Invoking Problem 1 with $g = 7$ then gives the edge bound $|E(G)| \le 3|V(G)|/2 - 3 = 18$, contradicting the fact that $|E(G)| = 21$.

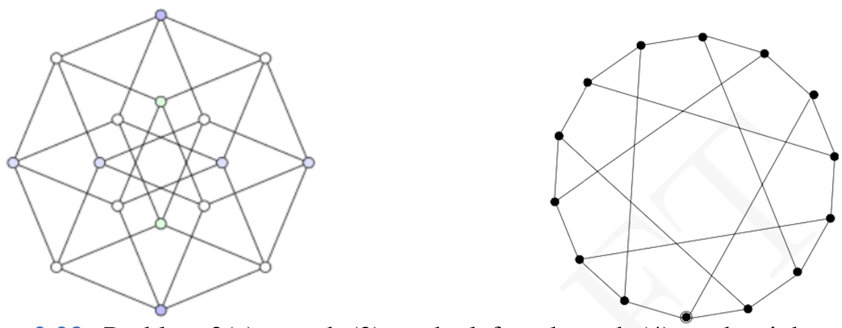


**Figure 2.23:** Problem 2(a): graph (3) on the left and graph (4) on the right.

(b) In class we showed that $K_r$ is planar for $1 \le r \le 4$ (by giving concrete planar embeddings), and not planar for $r \ge 5$ (by proving that $K_5$ is not planar). In this problem your task is to determine all positive integers $1 \le r \le s$ for which $K_{r,s}$ is planar; justify your answer.

**Proof** Since $K_{r,s}$ is bipartite it contains no odd cycle, which implies that $K_{r,s}$ contains no cycles of length at most 3. Invoking Problem 1 with $g = 4$ then gives the edge bound $|E(G)| \le 2|V(G)| - 4$. Note that $|V(K_{r,s})| = r + s$ and $|E(K_{r,s})| = rs$. So if $G = K_{r,s}$ is planar, then the inequality $rs \le 2(r+s) - 4$ must hold. For $r = s = 3$ this edge bound is not satisfied, thus $K_{3,3}$ is not planar. Since any $K_{r,s}$ with $3 \le r \le s$ contains $K_{3,3}$, it follows that these graphs are non-planar, too. The remaining graphs are of form $K_{1,s}$ or $K_{2,s}$ for some integer $s \in \mathbb{N}^+$ (which each satisfy the edge-bound mentioned above, but this does *not* yet ensure they they are planar: you have to provide a drawing without crossings). The star $K_{1,s}$ is obviously planar. Furthermore, $K_{2,s}$ has a planar embedding that is illustrated for $s = 4$ in Figure 2.24. To sum up: $K_{r,s}$ with $1 \le r \le s$ is non-planar when $r \ge 3$, and planar when $1 \le r \le 2$.

(c) Suppose three houses (A,B,C) each need to be connected to the water, gas, and electricity companies (W,G,E), with a separate line from each house to each company. Is there a way to make all nine connections without any of the lines crossing each other?

**Proof** If one draws the corresponding graph, for example as shown in Figure 2.24, then one sees that the question of whether one can make all these lines non-crossing is equivalent to the question whether $K_{3,3}$ is

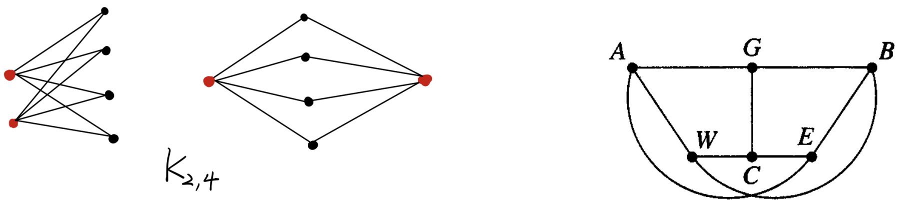



**Figure 2.24:** Problem 2(b) on the left and 2(c) on the right.

planar. By part (b) we know that $K_{3,3}$ is not planar (which followed from the edge-bound of Problem 1 with $g = 4$), so it is impossible to make all these lines non-crossing.

✍ **Exercise 2.38** *The goal of this problem is to demonstrate that planar graphs have fairly small chromatic number.* In the following parts (a)–(b) you may use Problem 7 above (or use any result from class):

(a) Show that every planar graph is 6-colorable.

**Proof** An important caveat here (compared to many proofs you can find in various books etc) is that we only proved Problem 1 for *connected* graphs. Indeed, in many other sources you will often find the argument that by the Handshake Lemma any planar graph $H$ has minimum degree at most 5 (if not, then the Handshake Lemma and Problem 1 with $g = 3$ (assuming that it applies for all graphs, not just connected ones) would imply that $6n(H) = \sum_{v \in V(H)} d_H(v) = 2|E(H)| \leq 6n(H) - 12$, which is a contradiction), which in turn implies that the graph $G$ is 5-degenerate and thus 6-colorable. *To justify this textbook argument one would need to prove that Problem 1 extends to connected graphs* (which requires a short and not too complicated argument).

So if we only want to use Problem 1 for connected graphs, then it perhaps is easiest to directly prove the desired result (that any planar graph $G$ is 6-colorable) by *induction* over the number $n$ of vertices:

1. The base case $1 \leq n \leq 6$ is trivial, since any graph with $1 \leq n \leq 6$ vertices is clearly 6-colorable (by giving each vertex its own color).
2. For the induction step $n \geq 7$ we simply use a case distinction. Either $G$ is not connected, in which case we apply the induction hypothesis to each component of $G$ (note that each component has at most $n - 1$ vertices), so we know that each component of $G$ can be colored by at most 6 colors, which in turn implies that $G$ can be colored using at most 6 colors. Otherwise $G$ is connected, in which case we invoke Problem 1 with $g = 3$ to deduce that $G$ contains a vertex $v$ with minimum degree at most 5 (reaching a contradiction as above: if all vertices of $G$ had degree at least 6, then $6n(G) = \sum_{v \in V(G)} d_G(v) = 2|E(G)| \leq 6n(G) - 12$ would yield a contradiction); we then take $v$ out of $G$ and use the induction hypothesis to color $G - \{v\}$ with at most 6 colors, which we can then readily extend to a proper 6 coloring of $G$ ($v$ is connected with at most 5 neighbors in $G - \{v\}$, so we can always find a color for $u$ that is different from all its neighbors).

Remark: as yet another alternative (as usual for proofs, there are many different ways to reach the goal!), one can also use the above proof-framework to inductively prove that every planar graph has minimum degree at most 5: the base case $1 \leq n \leq 6$ is trivial, and for the induction step the same case distinction as above works (either $G$ is disconnected, and then in one of the components we have minimum degree at most five by the induction hypothesis; or $G$ is connected, and then the above Problem 1 + Handshaking argument works).

(b) Show that every triangle-free planar graph is 4-colorable.

**Proof** This is essentially the same argument as for (a), the key difference being now that due to triangle-freeness there are no cycles of length three, so we can apply Problem 1 with $g = 4$ (note that if $G$ is triangle-free, then any subgraph $H \subseteq G$ is also triangle-free) and then focus in our argument on vertices with minimum degree at most 3 (if all vertices of the triangle-free graph $G$ had degree at least 4, then the Handshake Lemma and

Problem 1 with $g = 4$ would imply that $4n(G) = \sum_{v\in V(G)} d_G(v) = 2|E(G)| \leq 4n(G) - 8$, which is a contradiction), which then translates into 4-colorability of $G$.

**Exercise 2.39** Are the following statements about the chromatic number true? Prove or disprove:

(a) Every graph G has a $\chi(G)$-vertex-coloring where $\alpha(G)$ vertices get the same color.

**Proof** *This problem tests your understanding of basic graph theory definitions (chromatic number).*
False. See Figure 2.25 for a counterexample: this is a 2-colorable graph $G$ with the largest independent set of size 4, but (one can check that) in any 2-coloring of $G$ at most 3 vertices are assigned the same color.

(b) If $G$ and $H$ are graphs on the same vertex-set, then $\chi(G \cup H) \leq \chi(G) + \chi(H)$.

**Proof** *This problem tests your understanding of basic graph theory definitions (chromatic number).*
False. See Figure 2.25 for a counterexample. Let $G = K_{3,3}$ and $H$ be a union of two disjoint $K_3$. Then $\chi(G) = 2$, $\chi(H) = 3$. However, $G \cup H = K_6$, and $\chi(G \cup H) = 6 > \chi(G) + \chi(H) = 5$.

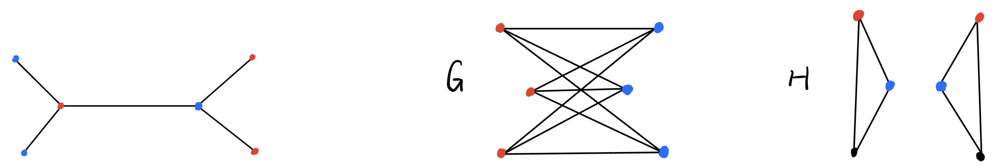



**Figure 2.25:** Counterexamples for Problem 4(a) and 4(b).

(c) Any $n$-vertex graph $G$ with degree-sequence $d_1 \geq d_2 \geq \cdots \geq d_n$ (so vertex $v_1$ has degree $d_1$, vertex $v_2$ has degree $d_2$, and so on) satisfies $\chi(G) \leq 1 + \max_{1\leq i\leq n} \min\{d_i, i-1\}$.
**Proof** *This goal of this problem was to review the proof of $\chi(G) \leq 1 + \Delta(G)$ from class.*
True. The upper bound is obtained from greedy coloring algorithm, where one tries to color vertices in the order $v_1, \cdots, v_n$ and assign to $v_i$ the smallest-indexed color not previously used on its lower-indexed neighbors. When coloring vertex $v_i$, the number of previous colored neighbors of $v_i$ is at most the minimum between $d_i$ and $i-1$. After assigning a color to $v_i$, the number of colors that has appeared so far is at most $1 + \min\{d_i, i-1\}$. Since this holds after each step, the desired upper bound is obtained by taking the maximum over $1 \leq i \leq n$, i.e., $\chi(G) \leq \max_{1\leq i\leq n}(1 + \min\{d_i, i-1\}) = 1 + \max_{1\leq i\leq n} \min\{d_i, i-1\}$.

## 2.7 Matching

**Definition 2.26 (Maching)**

*A **matching** in a graph G is a set of independent edges. That is, it is a set of edges in which no pair shares a vertex. Given a mating M in graph G,*

- *the vertices belonging to the edges of M are said to be M-**saturated(unmatched)**, otherwise M-**unsaturated(ununmatched)**.*
- *M is said to be a **perfect** matching if it saturates every vertex of G.*
- *the **size** of M is the number of edges inside M.*
- *a **maximum** matching in a graph is a matching that has the largest possible cardinality;*
- *a **maximal** matching is a matching that cannot be enlarged by the addition of any edge. (A maximum matching is always maximal, but not vice versa.*
- *an M-**alternating** path is a path in G where the edges alternate between M-edges and non-M-edges.*
- *an M-**augmenting** path is an M-alternating path where both end vertices are M-unsaturated.* ♣

**Theorem 2.23 (Berge's Theorem)**

*Let M be a matching in a graph G. M is maximum if and only if G contains no M-augmenting paths.* ♡

### 2.7.1 Independent Set and Vertex Cover

**Definition 2.27 (Independent set)**

*An **independent set** in a graph G is a set $X \subset V(G)$ of vertices such that no pair of which forms an edge of G, i.e., the subgraph $G[X]$ induced by $X$ has no edges.*

1. *The maximum size of an independent set in a graph G is denoted $\alpha(G)$.*
2. *The maximum size of a matching in a graph G is denoted $\alpha'(G)$.* ♣

**Definition 2.28 (Vertex cover)**

*A **vertex cover** of G is a set of vertices $X \subset V(G)$ such that every edge is incident on some vertex of $X$, i.e., $e \cap X \neq \emptyset$ for every $e \in E(G)$. The minimum size of a vertex cover of G is denoted $\beta(G)$.* ♣

**Lemma 2.17**

*The size of the maximum matching is at most the size of the minimum vertex cover.* ♡

**Proof** Each edge of $M$ uses different vertex of $C$.

**Lemma 2.18**

*For any graph G, $\alpha(G) + \beta(G) = |V(G)|$.* ♡

**Proof** If $I$ is an independent set of vertices in G, then $V(G) \setminus I$ is a vertex cover: every edge of G has at least one end in $V(G) \setminus I$ since no edges have both ends in I. Conversely, if $C$ is a vertex cover, then every edge is incident with $C$ so no edges have both ends in $V(G) \setminus C$. Therefore $V(G) \setminus C$ is an independent set of $G$. We conclude $\alpha(G) + \beta(G) = |V(G)|$.

**Definition 2.29 (Edge cover)**

*A **edge cover** of $G$ is a set of edges $F \subset E(G)$ covering every vertex of $G$, i.e., $v \cap F \neq \emptyset$ for every $v \in V(G)$. The minimum size of a vertex cover of $G$ is denoted $\beta'(G)$.* ♣

**Lemma 2.19 (Gallai's Lemma)**

*Let $G$ be a connected graph without isolated vertices, $\alpha'(G) + \beta'(G) = |V(G)|$.* ♡

### 2.7.2 Matchings in Bipartite Graphs

**Theorem 2.24 (Hall's Theorem)**

*Let $G$ be a bipartite graph with partite sets $X$ and $Y$. $X$ can be matched into $Y$ if and only if $|N(S)| \geq |S|$ for all subsets $S$ of $X$, where $N(S)$ denotes the set of neighbors of elements of $S$.* ♡

**Proof** Proof gives an inductive algorithm.

- Consider current Maximum Matching
- If not all of X matched, try to find and add an augmenting path.
- Consider the set of all vertices reached now.
- Gives an S with larger size.

**Proposition 2.4**

*Any regular, bipartite graph has a **perfect** matching (i.e. a matching that uses all of the vertices).* ♠

**Proof** Handshake Lemma tells us $\sum_{v \in X} d(v) = |E| = \sum_{v \in Y} d(v)$ for bipartite graph, and $|X| = |Y|$ for regular bipartite graph, which is enough to find matching using all of $X$.

**Theorem 2.25**

*Every finite bipartite graph $G$ can be edge colored with $\Delta(G)$ colors.* ♡

**Proof** 1. Find a matching M which includes all of the vertices of $G$ of maximum degree.

2. Color all edges in M one color, inductively color $G - M$. Since M includes an edge from each vertex of max degree $\Delta(G - M) = \Delta(G) - 1$, can be colored with $\Delta(G) - 1$ colors.

**Theorem 2.26 (Konig's, generalization of Hall's Theorem)**

*For any finite bipartite graph $G$, the size of the maximum matching equals the size of the minimum vertex cover.* ♡

**Definition 2.30 (System of distinct representatives)**

*Let $S_1, S_2, \cdots, S_n$ be sets. Then the sets have a **system of distinct representatives** or **transversal** if we can select $s_1 \in S_1, s_2 \in S_2, \cdots, s_n \in S_n$ such that $s_1, s_2, ..., s_n$ are all different.* ♣

The problem of determining whether sets $S_1, S_2, \cdots, S_n$ have a system of distinct representatives can be solved via Hall's Theorem, as follows.

Let $G$ be a bipartite graph with parts $A = S_1, S_2, \cdots, S_n$ and $B = S_1 \cup S_2 \cup S_n$, and where $(a, b) \in E(G)$ with $a \in A$ and $b \in B$ if $b \in a$. In other words, join a set to all the elements it contains. Then $G$ has a matching covering A if and only if $S_1, S_2, \cdots, S_n$ have a system of distinct representatives: the edges of the matching tell us which set each element is a representative for.

Halls' Theorem gives a necessary and sufficient condition for distinct representatives:

**Theorem 2.27**

*Sets $S_1, S_2, \cdots, S_n$ have a system of distinct representatives if for every set $I \subset \{1, \cdots, n\}$, $|\cup_{i \in I} S_i| \geq |I|$.* ♡

### 2.7.3 Matchings in Graphs That are not Bipartite

**Theorem 2.28 (Tutte's)**

*Let $G$ be a finite graph and $\Omega(G)$ denote the number of connected components of $G$ with an odd number of vertices. If $|S| \geq \Omega(G - S)$ for every set $S$ of vertices, then $G$ has a perfect matching.* ♡

**Theorem 2.29 (Petersen's)**

*Any bridgeless, 3-regular graph has a perfect matching.* ♡

### 2.7.4 Exercises

✍ **Exercise 2.40** (a) Let $G = (A \cup B, E)$ be a bipartite graph. Prove that the maximum matching of $G$ has size $|A| - d^*$, where in $d^* := \max_S\{|S| - |N_G(S)|, 0\}$ the maximum is over all non-empty $S \subseteq A$. (You may use the following result from class: if $|N_G(X)| \geq |X| - d$ for all non-empty $X \subseteq A$, then $G$ contains a matching $M$ of size $|M| \geq |A| - d$.)

**Proof** *This goal of this problem was to review the necessary condition in Hall's Theorem.*
We separately show that the maximum size $\alpha'(G)$ of a matching of $G$ satisfies $\alpha'(G) \geq |A| - d^*$ and $\alpha'(G) \leq |A| - d^*$, which together establishes the desired equality $\alpha'(G) = |A| - d^*$.

We start with $\alpha'(G) \leq |A| - d^*$, by mimicking the 'necessary' argument from Hall's Theorem. This is trivially true when $d^* = 0$, so we henceforth assume $d^* \geq 1$. Fix $S \subseteq A$ satisfying $|S| - |N(S)| = d^*$ (this exists when $d^* > 0$). In any matching in $G$, at most $|N(S)|$ elements of $S$ can be matched, leaving at least $|S| - |N(S)| = d^*$ elements unmatched. So no matching can have size larger than $|A| - d^*$, i.e., $\alpha'(G) \leq |A| - d^*$.

We now turn to $\alpha'(G) \geq |A| - d^*$, by using the hint, i.e., the result from class. Note that for any non-empty $S \subseteq A$, by definition we have $d^* \geq |S| - |N_G(S)|$ and thus $|N_G(S)| \geq |S| - d^*$, so by the result from class (mentioned in the hint) we infer that $G$ contains a matching of size $|A| - d^*$, so that $\alpha'(G) \geq |A| - d^*$.

(b) Let $G = (A \cup B, E)$ be a bipartite graph where $|N(X)| > |X|$ for every non-empty $X \subsetneq A$. Prove that every edge of $G$ belongs to some matching of size $|A|$.

**Proof** *This goal of this problem was to review the proof of Hall's Theorem.*
Fix an arbitrary edge $e = av \in E(G)$ of $G$, where the vertices satisfy $a \in A$ and $b \in B$, say. Let $G' := G \backslash \{a, b\}$ be the graph where we remove the vertices $\{a, b\}$ from $G$ (which means that automatically we remove all edges containing those vertices). As in the proof of the first case of the induction step of Hall's Theorem, it suffices to show that there is a matching $M' \subseteq E$ of size $|M'| := |A| - 1$ in $G'$, since adding the edge $e$ to $M$ then gives a matching $M := M' \cup \{e\}$ of size $|M| = |A|$ in $G$.
To show that such a matching exists in $G'$, it suffices to verify Hall's condition for $G'$: for any non-empty $S \subseteq A \setminus \{a\}$, note that the neighbors of $S$ in $G'$ consist of all the neighbors of $S$ in $G$, except for possibly the vertex $b$, implying that $|N_{G'}(S)| \geq |N_G(S)| - 1 \geq |S|$. Hence Hall's condition holds, so by Hall's

Theorem there is a matching of size $|A|-1$ in $G'$, which completes the proof, as discussed.

**Exercise 2.41** Prove that every tree has at most one perfect matching.

**Proof** *This goal of this problem was to train using induction (a classical proof technique) in graph theory.*

Proof by induction on the number $n$ of vertices (what the instructor had in mind).
*Base case* $1 \leq n \leq 2$*:* There is only one tree on $n=1$ vertices, namely a single vertex: this tree has no perfect matching (and thus at most one perfect matching). Similarly, there is only one tree on $n=2$ vertices, namely a single edge: this tree has exactly one perfect matching.
*Induction step* $n \geq 3$*:* Let $T$ be any tree with $n \geq 3$ vertices. Note that $T$ contains a leaf $v$. Every perfect matching of $T$ must contain the edge $e$ which matches this leaf $v$ to its unique neighbor $w$ in $T$. Delete the leaf $v$ and its neighbor $w$ to obtain the graph $T' := T \setminus \{v,w\}$. Note that $T'$ will be a forest, and each connected component will be a tree with at most $n-2$ vertices (note that we earlier used the base case $1 \leq n \leq 2$ in order to here ensure that $1 \leq n-2$ holds). Using the induction hypothesis we thus infer that each connected component of $T'$ has at most one perfect matching, which implies that $T = T \setminus \{v,w\}$ itself must have at most one perfect matching. Recalling that every perfect matching must contain the edge $e = vw$, it follows that $T$ must also have at most one perfect matching, completing the proof of the induction step.

Proof by contradiction (an alternative solution). Assume that there exist two distinct perfect matchings $M_1, M_2$ in $T$. Let $T'$ denote the subgraph of $T$ consisting only of edges in exactly one of the two perfect matchings $M_1, M_2$. Since each vertex is in exactly one edge of each perfect matching, $T'$ is a 2-regular subgraph. Therefore $T'$ contains a cycle, which contradicts the fact that $T$ was a tree (since trees are acyclic).

**Exercise 2.42**

(a) A school with 20 professors forms 10 committees, each containing 6 professors, such that every professor is on exactly 3 committees. Prove that it is possible to select a distinct representative from each committee.

**Proof** *This goal of this problem was to train the usage of Hall's Theorem via suitable auxiliary graphs.*
Let $G$ be a bipartite graph with vertex-classes $V(G) = P \cup C$, where $P$ is the set of professors and $C$ is the set of committees, with an edge whenever a professor is on a committee. Picking a distinct representative from each committee is equivalent to finding a matching of size $|C| = 10$ in $G$.
To show that such a matching exists in $G$, it suffices to verify Hall's condition for $G$: for any non-empty subset $S \subseteq C$ we proceed by counting the number of edges between $S$ and $N_G(S)$, similar as in class (the proof of matchings in $k$-regular graphs). To this end note that by construction, each committee vertex $c \in S$ has 6 neighbors in $P$, and each professor vertex $p \in N_G(S)$ has at most 3 neighbors in $S$ (since it has 3 neighbors in $C$). Hence $6|S| = e(S, N_G(S)) \leq 3|N_G(S)|$, and thus $|N_G(S)| \geq 2|S|$. Therefore Hall's condition holds, so by Hall's Theorem there is a matching of size $|C|$ in $G$, which completes the proof, as discussed.

(b) A class of 100 is participating in an oral exam. The committee consists of 25 members. Each student is interviewed by one member of the committee. It is known that each student likes at least 10 committee members. Prove that we can arrange the exam schedule such that each student is interviewed by one of the committee members that he likes, and each committee member interviews at most 10 students.

**Proof** *This goal of this problem was to train the usage of Hall's Theorem via suitable auxiliary graphs.*
Proof via suitable auxiliary graphs (what the instructor had in mind). Consider a bipartite graph $G$ with vertex-set $V(G) = C \cup S$, where the set $C$ contains ten vertices corresponding to each committee member (you can think of $C$ as all possible time-slots of the committee members), and the set $S$ which consists of all 100 students. It contains an edge between $s \in S$ and $c \in C$ whenever the student $s$ likes the committee member corresponding to the vertex $c$. Picking a matching of size $|S| = 100$ in $G$ is equivalent to picking desired exam assignment of

students to committees (where each student is interviewed by one of the committee members that he likes, and each committee member interviews at most 10 students).

To show that such a matching exists in $G$, it suffices to verify Hall's condition for $G$: for any non-empty subset $X \subseteq S$ of students, note that there are at least $|N_G(X)| \geq 10 \cdot 10 = 100 \geq |S| \geq |X|$ neighbors of $X$ in $C$ (since every such set contains at least one student $s \in X$, which in turn likes at least 10 different professors, each of which correspond to 10 different vertices in $C$). Therefore Hall's condition holds, so by Hall's Theorem there is a matching of size $|S|$ in $G$, which completes the proof, as discussed.

Proof via ad-hoc reasoning (alternative solution). Split the 100 students into 10 groups consisting of 10 students each. For each group, assign the students in that group each to a different committee member they are interested in: this is possible by Hall's Theorem because each student likes at least 10 members (so Hall's condition holds, similar as in the auxiliary graph solution). This ad-hoc construction ensures that each committee member will be assigned to at most one student from each group, i.e. at most 10 students.

✍ **Exercise 2.43** (a) King Kong and his friends are sharing some bananas. Each gorilla has a particular subset of the bananas that he/she is interested in. Coincidentally, it turns out that for each set of $S$ gorillas, the collection of bananas that any of them are interested in is of size at least $2|S|$. Prove that there is a way to give two bananas to each gorilla, so that every gorilla receives only bananas that he/she was interested in.

**Proof** *This goal of this problem was to train the usage of Hall's Theorem via suitable auxiliary graphs.*

Proof via suitable auxiliary graphs (what the instructor had in mind). Consider a bipartite graph $H$ with vertex-set $V(H) = G \cup B$, where the set $G$ contains two vertices corresponding to each gorilla, and $B$ is the set of all bananas. It contains an edge between $g \in G$ and $b \in B$ whenever the gorilla corresponding to vertex $g$ is interested in the banana $b$. Picking a matching of size $|G|$ in $H$ is equivalent to desired assignment of bananas to gorilla (where every gorilla receives two bananas that he/she is interested in).

To show that such a matching exists in $H$, it suffices to verify Hall's condition for $H$: for any non-empty subset $X \subseteq G$, note that $N_H(X)$ consists of all bananas that a gorilla corresponding to some vertex in $X$ is interested in, and that there are at least $s_X := |X|/2$ gorillas corresponding to vertices in $X$ (since we introduced two vertices for each gorilla). By assumption each set $S$ of gorillas is interested in at least $2|S|$ bananas, so it follows that $|N_H(S)| \geq 2 \cdot s_X = |X|$. Therefore Hall's condition holds, so by Hall's Theorem there is a matching of size $|G|$ in $H$, which completes the proof, as discussed.

Proof via ad-hoc reasoning (alternative solution). Let $F$ be a bipartite graph with vertex-classes $V(F) = G \cup B$, where $G$ is the set of gorillas and $B$ is the set of bananas, with an edge from gorilla $g$ to banana $b$ whenever $g$ is interested in $b$. Hall's condition holds in $F$ since for any non-empty $S \subseteq A$ we have $|N_F(S)| \geq 2|S| \geq |S|$. By Hall's Theorem there thus exists a matching $M_1$ in $F$. Delete all of the edges in $M_1$ from $F$ to obtain the graph $F' := F - M$. Hall's condition also holds in $F'$, since for any non-empty $S \subseteq A$ we have $|N_{F'}(S)| \geq |N_F(S)| - |S| \geq 2|S| - |S| = |S|$, since the matching $M_1$ removes at most $|S|$ vertices from the neighbors of $S$ in $F$. By Hall's Theorem there thus exists a matching $M_2$ in $F'$. Combining the two matchings $M_1$ and $M_2$ yields a collection $M := M_1 \cup M_2 \subseteq E(F)$ of edges from $F$ that is equivalent to desired assignment of bananas to gorilla (where every gorilla receives two bananas that he/she is interested in).

(b) A group of people are planning their summer vacations. Each person likes some of the trips $t_1, \ldots, t_n$, but will travel on at most one of them. Each trip $t_i$ has capacity $c_i$. In terms of which people like trips, derive a necessary and sufficient condition for filling all trips (to capacity) with people.

**Proof** *This goal of this problem was to train the usage of Hall's Theorem via suitable auxiliary graphs.*

Proof via suitable auxiliary graphs (what the instructor had in mind). The idea is to generalize the auxiliary graph argument from part (a), which is based on 'vertex duplication' (in one of the partition classes multiple vertices are created for one subject/object). To this end we first consider the bipartite graph $G$ with vertex-set $V(G) := T \cup P$, where $T$ contains all trips $t_i$, and $P$ is the set of all people. It contains an edge between $t \in T$ and $p \in P$ whenever the person $p$ is likes the trip $t$.

Mimicking the idea from Hall's Theorem (leading to Hall's condition), an obvious *necessary condition* for finding a desired assignment of people to trips (so that each trip is filled to capacity $c_i$) is the following:

$$|N_G(S)| \geq \textstyle\sum_{t_i \in S} c_i \qquad \text{for all non-empty set } S \subseteq T := \{t_1, t_2, \ldots\} \text{ of trips} \tag{2.33}$$

i.e., that the number of people interested in going on the trips in $S$ is at least the sum $\sum_{t_i \in S} c_i$ of their capacities (otherwise there are not enough people interested in those trips to fill them to capacity).

To show the *condition* (2.33) *is also sufficient*, we construct an auxiliary graph $H$ similar to part (a). Let $H$ be a bipartite graph with vertex-set $V(H) := T^+ \cup P$, where $T^+$ contains $c_i$ vertices corresponding to each trip $t_i \in T$, and $P$ contains a vertex corresponding to each person. It contains an edge between $t \in T^+$ and $p \in P$ whenever the person $p$ is interested in the trip corresponding to the vertex $t$. Picking a matching of size $|T^+| = \sum_{t_i \in T} c_i$ in $G$ is equivalent to the desired assignment of people to trips (so that every trip $t_i$ is filled to capacity $c_i$). To show that such a matching exists in $H$, it suffices to verify Hall's condition for $H$: for any non-empty subset $X \subseteq T^+$, note that $N_H(X)$ equals $N_G(T_X)$ where $T_X \subseteq T$ consists of all trips corresponding to a vertex in $X$, so using the necessary condition (2.33) we obtain

$$|N_H(X)| = |N_G(T_X)| \geq \textstyle\sum_{t_i \in T_X} c_i \geq |X|,$$

where the last inequality follows by noting that every trip $t_i$ appears at most $c_i$ times in $X$. By Hall's Theorem there thus is a matching of size $|T^+|$ in $H$, which completes the proof, as discussed.

Proof via ad-hoc inductive reasoning (alternative solution). Note that the capacities $c_i$ are integer valued. Hence we may assume $c_i \geq 1$ (since deleting/ignoring trips with capacity $c_i = 0$ does not change any valid assignment of people to trips). Using induction on the largest capacity, we shall now prove that the *condition* (2.33) *is sufficient*, by generalizing the idea of the ad-hoc argument from part (a).

The *base case*, when the maximum capacity is 1, is exactly given by Hall's theorem (since then, by our $c_i \geq 1$ assumption, the capacity of all trips equals 1, so that $\sum_{t_i \in S} c_i = |S|$ holds).

For the *induction step*, we create a bipartite auxiliary graph $F$ with vertex-set $V(F) = A \cup B$, where $T := \{t_1, t_2, \ldots\}$ is the set of trips (without duplications) and $P$ is the set of people, with an edge whenever someone is interested in a trip (analogous to the auxiliary graph solution above). For any non-empty $S \subseteq T$, using $c_i \geq 1$ and the necessary condition (2.33) we infer that

$$|N_F(S)| \geq \textstyle\sum_{t_i \in S} c_i \geq \sum_{t_i \in S} 1 = |S|.$$

Therefore Hall's condition holds, so by Hall's Theorem there is a matching $M \subseteq E(F)$ of size $|M| = |T|$ in $F$. Delete each edge of from this matching $M$ and any trip that was capacity $c_i = 1$, to get a new graph $F'$ with vertex-set $V(F') := T' \cup P$ and $T' \subseteq T$ (we deleted all trips with capacity 1 since these are already filled up by one person assigned via $M$). The point is that if we can fill every trip $t_i \in T'$ in $F'$ with $c_i - 1$ people, then by adding back the assignments from the matching $M$ we can fill up every trip $t_i \in T$ in $F$ to capacity $c_i$ (to clarify: all trips $t_i \in T \setminus T'$ have capacity $c_i = 1$, and so these are already filled up to capacity by $M$, as discussed). Similar to part (a), for any non-empty $S \subseteq T'$ we obtain using the necessary condition (2.33) that

$$|N_{F'}(S)| \geq |N_F(S)| - |S| \geq \left(\textstyle\sum_{t_i \in S} c_i\right) - |S| = \sum_{t_i \in S} (c_i - 1).$$

Invoking the induction hypothesis, in $G'$ we can thus fill all trips $t_i \in T'$ with $c_i - 1$ people, which completes the proof of the induction step, as discussed.

✍ **Exercise 2.44** Let $G$ be a regular graph. Prove that every bridge of $G$ is in every perfect matching of $G$.

**Proof** If the graph is not regular, it is not true. Take a simple path with four vertices 1,2,3,4 and three edges 12, 23, 34. The 23 edge is a bridge and the 12, 34 edges form a perfect matching.

Let $G$ be a regular graph of degree $d$ and $e = (u, v)$ be a bridge. Let a perfect matching does not contain edge $e$. Then graph $G \setminus e$ has two components $G_1$ and $G_2$, where $u \in V(G_1)$ and $v \in V(G_2)$. The number of vertices of each component is even, since there is no more edge between $G_1$ and $G_2$, and perfect matching is within each component. On the other hand, by the handshaking lemma we have for the graph $G_1$

$$2|E(G_1)| = \sum_{v \in V(G_1)} d(v) = d(|V(G_1)| - 1) + (d - 1) = d|V(G_1)| - 1,$$

where the left hand side is even, however the right hand side is odd. Contradiction.

**Exercise 2.45** (a) A standard deck of 52 playing cards is shuffled and then dealt into 13 piles of 4 cards each. Prove that, regardless of how the deck is shuffled, there is always a way to select one card from each pile so that each of the 13 possible ranks (2, 3, ..., 10, Jack, King, Queen, Ace) occurs once.

**Proof** *This goal of this problem was to train the usage of Hall's Theorem via suitable auxiliary graphs.* We consider the bipartite graph $G = (R \cup P, E)$, where $R = \{2, 3, 4, 5, 6, 7, 8, 9, 10, \text{Jack}, \text{Queen}, \text{King}, \text{Ace}\}$ contains all ranks, and $P = \{p_1, \cdots, p_{13}\}$ represents the 13 piles of 4 cards. It contains an edge between $r \in R$ and $p \in P$ whenever pile $p$ contains at least one card of rank $r$. Since $G$ is not 4-regular contrary to what one might think (as a rank may appear multiple times in a pile), we need to be slightly more careful than usual when verifying Hall's Condition. Fix a non-empty $S \subseteq R$. Note that there are exactly $4|S|$ cards with ranks from $S$ (since there are 4 cards of each rank). Furthermore, $N_G(S)$ contains all piles that contain cards with ranks from $S$, and that those piles together contain at most $4|N_G(S)|$ such cards (since each pile contains 4 cards). Hence $4|S| \leq 4|N_G(S)|$ follows, i.e., that $|N_G(S)| \geq |S|$. Hence Hall's condition holds, so by Hall's Theorem there is a matching $M_4 \subseteq E$ of size $|M_4| = |A| = 13$ in $G$, which in turn represents a way of selecting cards from each pile, such that the group of selected cards contains each rank.

(b) Prove that any $k$-regular bipartite graph $G = (A \cup B, E)$ on $n \geq 2$ vertices ($k \geq 1$) has the following property: its edge set $E$ can be written as the union of $k$ edge-disjoint perfect matchings.

(Hint: in class we proved that such a $k$-regular bipartite graph contains a perfect matching.)

**Proof** *This goal of this problem was to train using induction (a classical proof technique) in graph theory.* We proceed by induction on $k \geq 1$. *The base case* $k = 1$ corresponds to the result from class mentioned in the hint. For *the induction step* $k \geq 2$, using the result from class mentioned in the hint we find a perfect matching $M_k$ in $G$. Deleting all the edges of $M_k$ from $G$, we obtain the graph $G' := G \setminus M_k$, which is $(k-1)$-regular (since every vertex is the endpoint of exactly one edge from the perfect matching $M_k$). Invoking the induction hypothesis, it follows that the edge-set of $G'$ can be written as the union of $(k - 1)$ edge-disjoint perfect matching, say $E(G') = M_1 \cup \cdots \cup M_{k-1}$. Since $E(G') = E(G) \setminus M_k$ and $M_k$ are edge-disjoint, it follows that the edge set $E(G) = M_1 \cup \cdots M_{k-1} \cup M_k$ can be written as the union of $k$ edge-disjoint perfect matchings, completing the proof of the induction step.

(c) Prove that in part (a) we can in fact go one step further: we can divide the cards into 4 groups of 13 cards each, where each group has one card from every pile, and one card from every rank

**Proof** *This goal of this problem was to train the usage of Hall's Theorem via suitable auxiliary graphs.* The idea is to remove a matching and apply induction, similar to part (b), though we find it is easier to describe the argument iteratively. Writing $G_4 = (R \cup P, E)$ for the auxiliary bipartite graph from part $(a)$, recall that in part (a) we found a matching $M_4 \subseteq E$ of size $|M_4| = |R| = 13$, which presented a suitable way of selecting cards from each pile (such that the group of selected cards contains each rank). We now remove all those cards selected by $M_4$, and define the resulting new auxiliary bipartite graph by $G_3 = (R \cup P, E(G_4) \setminus M_4)$. Since $M_4$ is a matching, note that there are only 3 cards left in each pile, and for each rank there are also 3 cards left. Hence we can repeat the reasoning used in part (a) to infer that $3|S| \leq 3|N_{G_3}(S)|$, i.e., that $|N_{G_3}(S)| \geq |S|$. So by Hall's Theorem there is a matching $M_3 \subseteq E(G_3) = E(G_4) \setminus M_4 = E \setminus M_4$ of size $|M_3| = |A| = 13$, which in turn represents a suitable way of selecting cards from each pile. Iterating this argument, we can similarly find a matching $M_2 \subseteq E(G_2) = E(G_3) \setminus M_3 = E \setminus (M_3 \cup M_4)$ of size $|M_2| = |A| = 13$ in $G_3 = (R \cup P, E(G_3) \setminus M_3)$, and also a matching $M_1 \subseteq E(G_1) = E(G_2) \setminus M_2 = E \setminus (M_2 \cup M_3 \cup M_4)$ of size $|M_1| = |A| = 13$ in $G_1 = (R \cup P, E(G_2) \setminus M_2)$. This way we obtain 4 edge-disjoint matchings $M_1, \ldots, M_4$, each of which represents a suitable way of selecting cards from each pile (such that the group of selected cards contains each rank).

| | | | |
|---|---|---|---|
| 3♣ | 4♥ | J♠ | 9♣ |
| J♥ | 2♣ | 4♦ | 5♣ |
| 6♦ | 6♥ | 6♣ | Q♦ |
| 7♥ | K♣ | 5♠ | 6♠ |
| 9♠ | 8♥ | 10♠ | 9♦ |
| Q♣ | 7♦ | 3♠ | 9♥ |
| 5♦ | A♦ | J♣ | 5♥ |
| A♥ | 7♠ | 2♠ | K♦ |
| K♥ | 4♠ | 3♦ | 4♣ |
| 3♥ | A♠ | 10♣ | Q♥ |
| A♣ | 7♣ | J♦ | 2♥ |
| 10♥ | Q♠ | 10♦ | K♠ |
| 8♦ | 2♦ | 8♣ | 8♠ |

| | | | |
|---|---|---|---|
| 3♣ | 4♥ | J♠ | 9♣ |
| J♥ | 2♣ | 4♦ | 5♣ |
| 6♦ | 6♥ | 6♣ | Q♦ |
| 7♥ | K♣ | 5♠ | 6♠ |
| 9♠ | 8♥ | 10♠ | 9♦ |
| Q♣ | 7♦ | 3♠ | 9♥ |
| 5♦ | A♦ | J♣ | 5♥ |
| A♥ | 7♠ | 2♠ | K♦ |
| K♥ | 4♠ | 3♦ | 4♣ |
| 3♥ | A♠ | 10♣ | Q♥ |
| A♣ | 7♣ | J♦ | 2♥ |
| 10♥ | Q♠ | 10♦ | K♠ |
| 8♦ | 2♦ | 8♣ | 8♠ |

| | | | |
|---|---|---|---|
| 3♣ | 4♥ | J♠ | 9♣ |
| J♥ | 2♣ | 4♦ | 5♣ |
| 6♦ | 6♥ | 6♣ | Q♦ |
| 7♥ | K♣ | 5♠ | 6♠ |
| 9♠ | 8♥ | 10♠ | 9♦ |
| Q♣ | 7♦ | 3♠ | 9♥ |
| 5♦ | A♦ | J♣ | 5♥ |
| A♥ | 7♠ | 2♠ | K♦ |
| K♥ | 4♠ | 3♦ | 4♣ |
| 3♥ | A♠ | 10♣ | Q♥ |
| A♣ | 7♣ | J♦ | 2♥ |
| 10♥ | Q♠ | 10♦ | K♠ |
| 8♦ | 2♦ | 8♣ | 8♠ |

**Figure 2.26:** One example of problem 2(a) and 2(c), courtesy of Prof. Tesler's lecture notes in winter 2022.

**Exercise 2.46** Let $G$ be a connected graph on $2n$ vertices. Suppose that for any vertices $u, v, w$ of $G$ with no two of them connected by an edge, $d(u) + d(v) + d(w) \geq 3n - 2$. Show that $G$ has a perfect matching.

**Proof** Suppose for sake of contradiction that $G$ does not have a perfect matching, then by the contrapositive statement of Tutte's theorem, there must be a set $S$ with $|S| < \Omega(G - S)$. Note that $G$ is connected with $2n$ vertices and $|S| \geq 1$, then $|\Omega(G - S)| \geq 3$. It can't be $|\Omega(G - S)| = 2$ when $|S = 1|$ since $|G|$ would be odd.

Let $u, v$ and $w$ be from the three smallest components of $G \setminus S$ (there must be at least 3 since $|\Omega(G \setminus S)| \geq 3$). We note that $u, v$ and $w$ can only have neighbors either in $S$ or as other elements of their component. If $|S| = m < n$, then the contribution of neighbors in $S$ to the sum of $u, v$ and $w$ degrees is at most $3m$. The sum of the sizes of their components is at most $\lfloor 3(2n-2)/(m+2) \rfloor$, since there are at least $m+2$ components using up the remaining $2n - m$ vertices(note that we picked the smallest 3 connected components). Therefore we have

$$d(u) + d(v) + d(w) \leq 3m + \lfloor 3(2n - m)/(m + 2) \rfloor - 3 \leq 3n - 3$$

where the last inequality holds since $3m + 3(2n - m)/(m + 2)$ is maximized when $m$ takes one of the extreme values $m = 1$ or $m = n - 1$.

## 2.8 Flow and Cut

**Definition 2.31 (Network, flow)**

*Let $\vec{G} = (V, \vec{E})$ be a digraph and $s, t \in V$. We shall refer to $s$ as the source vertex and $t$ as the sink vertex in what follows.*

1. *A **network** is a directed graph $\vec{G}$ with designated source vertex $s$ and sink vertex $t$.*
2. *A **flow** is a subgraph of $\vec{G}$ so that for each vertex $v$ other than $s$ and $t$ $d_{in}(v) = d_{out}(v)$.*
3. *The **size** of a flow is $d_{out}(s) - d_{in}(s) = d_{in}(t) - d_{out}(t)$ restricted to flow $F$. The total flow out of $s$ equals the total flow into $t$.*
4. *Given $\vec{G}$ and flow $F$, an **augmenting path** is an $s - t$ path that uses either edges of $G$ unused by $F$ in the forwards direction, or edges used by $F$ in the backwards direction.*

♣

**Lemma 2.20**

*Given an augmenting path, you can add it to $F$ to get a path with 1 more unit of flow.* ♡

**Definition 2.32 (Cut)**

*A **cut** is a partition of the vertices into two sets $S$ and $T$, which contain $s$ and $t$, respectively.*
*The **size** of a cut is the total number of edges from vertices in $S$ to vertices in $T$.* ♣

**Lemma 2.21**

*For a network $\vec{G}$, a flow $F$ and a cut$(S, T)$ it is the case that*

$$Size(F) = \{\#\text{of edges in } F \text{ from } S \text{ to } T\} - \{\#\text{of edges in } F \text{ from } T \text{ to } S\}.$$

♡

**Proof** Consider the sum over all $v$ in $S$ of $d_{out}(v) - d_{in}(v)$. On the one hand this is 0 except for $v = s$, where it is Size(F). On the other hand, each edge contributes to one in degree and one out degree. This makes its total contribution 0 unless it crosses the cut. This gives 1 for each edge from $S$ to $T$ and $-1$ for each edge from $T$ to $S$.

**Theorem 2.30 (Maxflow-Mincut)**

*For any network $\vec{G}$, the size of a maximum flow in $G$ is the same as the size of a minimum cut.* ♡

**Proof** Let $F$ be a flow and $C(S, T)$ be a cut.

1. (Maxflow $\le$ Mincut): By previous Lemma,

$$\text{Size}(F) = \{\#\text{of edges in } F \text{ from } S \text{ to } T\} - \{\#\text{of edges in } F \text{ from } T \text{ to } S\} \le \text{Size}(C).$$

   Any flow is smaller than any cut, so the maximum flow size is at most than the minimum cut size.
2. (Maxflow $\ge$ Mincut): Let $F$ be the maximum flow. Then there is no augmenting path. Let $S$ be the set of vertices $v$ you can reach from $s$ using unused forward edges or used backwards edges. By previous Lemma, $Size(F) = \{\#\text{of edges in } F \text{ from } S \text{ to } T\} = Size(C)$, thus Maxflow $\ge$ Mincut.

**Remark** Recall that the Cut number $\kappa(s, t)$ is defined as the minimum number of vertices to remove from graph $G$ to disconnect $s$ and $t$. Menger's Theorem indicates that there exists $\kappa(s, t)$ vertex disjoint paths between $s$ and $t$.

Maxflow-Mincut is the edge version of Menger's Theorem. It says that the minimum number of edges you need to remove to disconnect $s$ from $t$ (the smallest cut size), is the maximum number of edge-disjoint paths (the maximum flow size).

## 2.9 Ramsey Theory

Now, we consider coloring edges of the complete graph $K_n$.

**Theorem 2.31**

*If we color the edges of $K_6$ arbitrarily with red and blue, we can always find a red $K_3$ as a subgraph.* ♡

**Theorem 2.32**

*For any positive integers $p$ and $q$, there exists an integer $N$ such that whenever $n \geq N$, any red-blue coloring of the edges of a $K_n$, there is either a red $K_p$ or a blue $K_q$.* ♡

**Definition 2.33**

*The smallest such number N is called the Ramsey Number, $R(p, q)$.* ♣

We have the following recurrence relation.

$$R(p, q) \leq R(p-1, q) + R(p, q-1)\,. \tag{2.34}$$

**Remark** Ramsey Theory more generally studies these kinds of patterns and when certain types of structures must exist within sufficiently much noise.

**Theorem 2.33 (Upper bound)**

$R(p, q) \leq 2^{p+q}$. ♡

**Proof** Proved by induction. The base case for $p = 1$ and $q = 1$ is trivial. Then for $p, q > 1$,

$$R(p, q) \leq R(p-1, q) + R(p, q-1) \leq 2^{p-1+q} + 2^{p+q-1} = 2^{p+q}\,. \tag{2.35}$$

**Theorem 2.34 (Lower bound)**

*If $n \geq 3$, then $R(n, n) \geq 2^{n/2}$.* ♡

**Proof** [Probabilistic Proof] Let $N = R(n, n)$. We would get $N^n$ many collections of $n$ vertices, each has $2^{-n(n-1)/2}$ probability of being monochromatic. The expected number of monochromatic $K_n$ is roughly $N^n/2^{-n(n-1)/2} = [N/2^{-(n-1)/2}]^n$. If $N \leq 2^n$, then the expectation is smaller than 1, which means there are some coloring without monochromatic $K_n$.

**Remark** $R(p, q) \geq 2^{\min\{p,q\}}$. Combined with the upper bound, this says that symmetric Ramsey numbers are exponentially large.

**Definition 2.34 (Graph Ramsey Numbers)**

*For finite graphs $G$ and $H$, we define the **graph Ramsey number** $R(G, H)$ to be the minimum integer $n$ so that any red-blue coloring of $K_n$ has either a red copy of $G$ or a blue copy of $H$ as a subgraph of $K_n$.* ♣

**Theorem 2.35**

$R(G, H) \leq R(|V(G)|, |V(H)|)$. ♡

**Proof** Let $m = R(|V(G)|, |V(H)|)$. Any red-blue coloring of $K_m$ has either a monochromatic complete red graph on $|V(G)|$ or monochromatic blue complete graph on $|V(H)|$. These contain a red copy of $G$ or blue copy of $H$.

**Theorem 2.36**

*If $m$ and $n$ are integers with $m-1$ dividing $n-1$ and $T_m$ is a tree with $m$ vertices then $R(T_m, K_{1,n}) = m+n-1$.* ♡

**Lemma 2.22**

*Let $T$ be any tree on $k$ vertices and $G$ a graph with $\delta(G) \geq k-1$. The $G$ contains a copy of $T$.* ♡

### 2.9.1 Exercises

**Exercise 2.47** Prove for positive integers $n$, $m$ and $k$ that $R(n, m+k-1) \geq R(n,m) + R(n,k) - 1$. Hint: Start with a coloring of the complete graph on $R(n,m)-1$ vertices with no red $K_n$ or blue $K_m$ and a coloring of the complete graph on $R(n,k)-1$ vertices with no red $K_n$ or blue $K_k$. Combine them to get a coloring of the complete graph on $R(n,m)+R(n,k)-2$ vertices with no red $K_n$ nor blue $K_{m+k-1}$.

**Proof** Let $M = R(n,m)$ and $K = R(n,k)$. By the definition of Ramsey numbers, there is a red-blue coloring of the edges of $K_{M-1}$ so that there is neither a red $K_n$ nor a blue $K_m$. Similarly, there is a red-blue coloring of the edges of $K_{K-1}$ so that there is neither a red $K_n$ nor a blue $K_k$. Consider the following red-blue coloring of the edges of $K_{M+K-2}$. Split the vertices into a $K_{M-1}$, a $K_{K-1}$, and the connected edges. Color the edges of the $K_{M-1}$ as specified above, the edges of the $K_{K-1}$ as specified above and all of the connecting edges blue. We lemma that this doesn't have a red $K_n$ nor a blue $K_{m+k-1}$, showing that $R(n, m+k-1) \geq R(n,m) + R(n,k) - 1$. For red $K_n$, note that since all connecting vertices are blue, a red $K_n$ would need to be contained either in the $K_{M-1}$ or the $K_{K-1}$, which by assumption don't have it. For the blue $K_{m+k-1}$, it must have either at least m vertices in the $K_{M-1}$ (which implies that the $K_{M-1}$ has a blue Km), or at least $k$ vertices in the $K_{K-1}$ (which implies that the $K_{K-1}$ has a blue $K_k$). Since neither of these can happen, there is no blue $K_{m+k}$. This completes the proof

## 2.10 Extremal Graph Theory

**Definition 2.35 (F-free)**

*A graph $G$ is $F$-free if it has no subgraphs isomorphic to $F$.*

**Definition 2.36**

$ex(n, F)$ *is the greatest number of edges that a single, $F$-free graph with $n$ vertices can have.*

**Theorem 2.37 (Turan)**

*The $K_{r+1}$-free graph on $n$ vertices with the greatest number of vertices is the complete $r$-partite graph with parts as close to each other as possible.*

**Theorem 2.38 (Kovari-Sos-Turan)**

$$ex(n, K_{r,s}) \leq \left(\frac{s-1}{2}\right)^{1/r} n^{2-1/r} + \frac{1}{2}(r-1)n\,. \tag{2.36}$$