Combinatorics

Joseph R. Mileti

May 9, 2014

1 Introduction

1.1 Sets and Sequences

Recall that a set is a collection of elements without regard to repetition and order. Intuitively, a set is a box and the only thing that matters are the things that are inside it, and furthermore the box does not have more than 1 of any given item. For example, $\{2, 5\}$ is a set with 2 elements. Since all that matters are the elements, we define two sets to be equals if they have the same elements, regardless of how the sets themselves are defined. For example, the set $\{n \in \mathbb{N} : n \text{ is an even prime}\}$ equals the set $\{n \in \mathbb{N} : 3 < n^2 < 8\}$ even though the descriptions are very different. Similarly, since order doesn't matter, we have $\{3, 7\} = \{7, 3\}$ and $\{1, 2, 3\} = \{3, 1, 2\}$. Although we typically would not even write something like $\{2, 5, 5\}$, if we choose to do so then we would have $\{2, 5, 5\} = \{2, 5\}$.

We use $\in$ to represent membership. Thus, we have $2 \in \{2, 5\}$ and $3 \notin \{2, 5\}$. Since sets are mathematical objects, they may be elements of other sets. For example, we can form the set $S = \{1, \{2, 3\}\}$. Notice that we have $1 \in S$ and $\{2, 3\} \in S$, but $2 \notin S$ and $3 \notin S$. As a result, S has only 2 elements, namely 1 and $\{2, 3\}$. Thinking of a set as a box, one element of S is the number 1, and the other is a different box. The empty set is the unique set with no elements. We can write it as $\{\}$, but instead we typically denote it by $\emptyset$. There is only *one* empty set, because if both A and B have no elements, they they have the exactly the same elements for vacuous reasons, and hence $A = B$.

Sets can be either finite or infinite. Some examples of infinite sets are the standard "universes" of numbers.

Notation 1.1. *We use the following notation for the basic sets of numbers.*

- $\mathbb{N} = \{0, 1, 2, 3, \dots\}$ *is the set of natural numbers.*
- $\mathbb{N}^+ = \{1, 2, 3, \dots\}$ *is the set of positive natural numbers.*
- $\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$ *is the set of integers.*
- $\mathbb{Q} = \{\frac{a}{b} : a, b \in \mathbb{Z} \text{ and } b \neq 0\}$ *is the set of rational numbers.*
- $\mathbb{R}$ *is the set of real numbers.*

Beyond these fundamental sets, there are various ways to define new sets. In some cases, we can simply the list the elements as we did above. Although this often works for small finite sets, it is almost never a good idea to list the elements of a set with 20 or more elements, and it rarely works for infinite sets (unless there is an obvious pattern like $\{5, 10, 15, 20, \dots\}$). One of the standard ways to define a set S is to carve it out of some bigger set A by describing a certain property that may or may not be satisfied by an element of A. We did this above with $\{n \in \mathbb{N} : 3 < n^2 < 8\}$, but let's look at a simpler example. Suppose that we define.

$$S = \{n \in \mathbb{N} : 5 < n < 13\}.$$

In this case, we are taking $A = \mathbb{N}$, and forming a set S by carving out those elements of A that satisfy the condition that $5 < n < 13$. Thus, we have

$$S = \{6, 7, 8, 9, 10, 11, 12\}.$$

In this case, it is important that we put the "$\mathbb{N}$" in the above, because if we wrote $\{n : 5 < n < 13\}$ then it would be unclear what n we should consider. For example, should $\frac{11}{2}$ be in this set? How about $\sqrt{11}$? Sometimes the "universe" of numbers (or other mathematical objects) that we are working within is clear, but typically it is best to write the global set that you are picking elements from in order to avoid such ambiguity. Notice that when we define a set, there is no guarantee that it has any elements. For example, $\{n \in \mathbb{N} : n^2 = 2\} = \emptyset$. Keep in mind that we can also use words in our description of sets, as we did with $\{n \in \mathbb{N} : n \text{ is an even prime}\}$.

Another way to describe a set is through a "parametric" description. Rather than carving out a certain subset of a given set by describing a property that the elements must satisfy, we can instead form all the elements one obtains by varying a value through a particular set. For example, consider the following description of a set:

$$S = \{3x^2 + 1 : x \in \mathbb{R}\}$$

Although the notation looks quite similar to the above (in both case we have curly braces, with a : in the middle), this set is described differently. Notice that the "formula" or "description" appears on the left of the :, while the set that the variable is coming from appears on the right. The idea here is that instead of carving out a subset of $\mathbb{R}$ by using a property, we instead let x vary through all real numbers, plug each of these real numbers x into $3x^2 + 1$, and form the set of all possible outputs. For example, we have $4 \in S$ because $4 = 3 \cdot 1^2 + 1$. In other words, when $x = 1$, the left hand side gives the value 4, so we should put $4 \in S$. Notice also that $4 = 3 \cdot (-1)^2 + 1$, so we can also see that $4 \in S$ because of the "witness" -1. Of course, we are forming a set, so we do not repeat the number 4. We also have $1 \in S$ because $1 = 3 \cdot 0^2 + 1$, and we have $76 \in S$ because $76 = 3 \cdot 5^2 + 1$. Notice also that $7 \in S$ because $7 = 3 \cdot (\sqrt{2})^2 + 1$.

Now it is possible and indeed straightforward to turn any parametric description of a set into one where we carve out a subset by a property. In our case of $S = \{3x^2 + 1 : x \in \mathbb{R}\}$ above, we can alternatively write it as

$$S = \{y \in \mathbb{R} : \text{There exists } x \in \mathbb{R} \text{ with } y = 3x^2 + 1\}$$

Notice how we flipped the way we described by the set by introducing a "there exists" quantifier. This is always possible for a parametric description. For example, we have

$$\{5n + 4 : n \in \mathbb{N}\} = \{m \in \mathbb{N} : \text{There exists } n \in \mathbb{N} \text{ with } m = 5n + 4\}$$

Thus, these parametric descriptions are not essentially new ways to describe sets, but they can be more concise and hence clear.

By the way, we can use multiple parameters in our description. For example, consider the set

$$S = \{18m + 33n : m, n \in \mathbb{Z}\}$$

Now we are simply letting m and n vary through all possible values in $\mathbb{Z}$ and collecting all of the values $18m + 33n$ that result. For example, we have $15 \in S$ because $15 = 18 \cdot (-1) + 33 \cdot 1$. We also have $102 \in S$ because $102 = 18 \cdot 2 + 33 \cdot 2$. Notice that we are varying m and n independently, so they might take different values, or the same value (as in the case of $m = n = 2$). Don't be fooled by the fact that we used different letters! As above, we can flip this description around as above by writing

$$S = \{k \in \mathbb{Z} : \text{There exists } m, n \in \mathbb{Z} \text{ with } k = 18m + 33n\}$$

1.2 Subsets and Set Equality

We write $A \subseteq B$ to mean that every element of A is an element of B. More formally, $A \subseteq B$ means that for all x, if $x \in A$, then $x \in B$. Written more succinctly, $A \subseteq B$ means that for all $a \in A$, we have that $a \in B$. Since two sets are equal exactly when they have the same elements, notice that $A = B$ if and only if both $A \subseteq B$ and $B \subseteq A$.

To prove that $A \subseteq B$, one takes a completely arbitrary $a \in A$, and argues that $a \in B$. For example, let $A = \{6n : n \in \mathbb{Z}\}$ and let $B = \{2n : n \in \mathbb{Z}\}$. Since both of these sets are infinite, we can't show that $A \subseteq B$ by taking each element of A in term and showing that it is an element of B. Instead, we take an *arbitrary* $a \in A$, and show that $a \in B$. Here's the proof.

Proposition 1.2. *Let $A = \{6n : n \in \mathbb{Z}\}$ and $B = \{2n : n \in \mathbb{Z}\}$. We have $A \subseteq B$.*

Proof. Let $a \in A$ be arbitrary. By definition of A, this means that we can fix an $m \in \mathbb{Z}$ with $a = 6m$. Notice then that $a = 2 \cdot (3m)$. Since $3m \in \mathbb{Z}$, it follows that $a \in B$. Since $a \in A$ we arbitrary, we conclude that $A \subseteq B$. □

Make sure that you understand that logic of the argument above. First, we took an arbitrary element a from the set A. Now in the definition of $A = \{6n : n \in \mathbb{Z}\}$, the n is varying over all integers. Since $a \in A$, there must be one fixed integer value of n that puts a into the set A. In our proof, we chose to call that one fixed integer m. This was an arbitrary choice of name, and we could have chosen almost any other name for it. We could have called it ℓ, k, b, x, or α. The only really awful choice would be to call it a, because we have already given the letter a a meaning (namely as our arbitrary element of a). We could even have called it n, and in the future we will likely do this. However, to avoid confusion in our first arguments, we've chosen to use a different letter so that we keep straight in our mind the varying n in the definition of the set and the particular integer m that puts our one fixed (but arbitrary) a into the set A.

Now we have our a and we've fixed an integer m with $a = 6m$. To show that $a \in B$, we need to come up with an integer $n \in \mathbb{Z}$ such that $a = 2n$. Notice that this n might have little or nothing to do with the m that put a into the set A. There is absolutely no reason at all to think that the n that works might be our special m. In fact, in our case, that is certainly too much to hope for because $12 \in A$ by witness $n = 2$, and $12 \in B$ by witness $n = 6$ (and of course $2 \neq 6$). So how do we go about finding an n such that $a = 2n$? Well, the only thing that we know about our a is that $a = 6m$. Using this knowledge, we think about how we can insert a 2 into the mix. Of course, we know that $6 = 2 \cdot 3$. Thus, since $a = 6m$, we can also write $a = (2 \cdot 3) \cdot m = 2 \cdot (3m)$. Don't panic from the fact that we have $3m$ rather than a single "letter" like n here! After all, the n is just a dummy variable whose name doesn't matter! We had an honest fixed $m \in \mathbb{Z}$ with $a = 6m$, and we succeeded in finding an integer n with $a = 2m$, namely $n = 3m$. The only thing we need check is that $3m$ really is an integer, but that is the case because m is an integer. Therefore, we were successful in show that $a \in B$ because we have found an integer n such that $a = 2n$.

What would go wrong if we tried to prove that $B \subseteq A$? Let's try it. Let $b \in B$ be arbitrary. Since $b \in B$, we can fix $m \in \mathbb{Z}$ with $b = 2m$. Now our goal is to try to prove that we can find $n \in \mathbb{Z}$ with $b = 6n$. It's not obvious how to obtain a 6 from that 2, but perhaps we come up with the following idea. Since $b = 2m$ and $2 = \frac{6}{3}$, we can write $b = 6 \cdot \frac{m}{3}$. At this point we celebrate because we have found an n with $b = 6n$. Woohoo! Pause for a minute and think about what the problem is. We have indeed found a number n such that $b = 6n$. However, we have not checked that this n is an integer. In general, dividing an integer by 3 does not result in an integer, so this argument currently has a hole in it.

Although that argument has a problem, we can not immediately conclude that $B \not\subseteq A$. Our failure to find an argument does not mean that an argument does not exist. So how can we show that $B \not\subseteq A$. All that we need to do is find just *one example* of an element of B that is not an element of A. We choose as our example to take 2. However, we need to convince everybody that this choice works. So let's do it! First, notice that $2 = 2 \cdot 1$, so $2 \in B$ because $1 \in \mathbb{Z}$. We now need show that $2 \notin A$, and we'll do this using a proof by contradiction. Suppose instead that $2 \in A$. Then, by definition, we can fix an $m \in \mathbb{Z}$ with $2 = 6m$. We then have that $m = \frac{2}{6} = \frac{1}{3}$. However, this a contradiction because $\frac{1}{3} \notin \mathbb{Z}$. Since our assumption that $2 \in A$ led to a contradiction, we conclude that $2 \notin A$. We found an example of an element that is in B but not in A, so we conclude that $B \not\subseteq A$.

Since $A = B$ if and only both $A \subseteq B$ and $B \subseteq A$, we can prove that two sets are equal by doing two proofs like the above. Such a strategy is called a proof by "double containment". You will have some examples of this on the homework.

1.3 Ordered Pairs and Sequences

In contrast to sets, we define *ordered pairs* in such a way that order and repetition *do* matter. We denote an ordered pair using normal parentheses rather than curly braces. For example, we let $(2, 5)$ be the ordered pair whose first element is 2 and whose second element is 5. Notice that we have $(2, 5) \neq (5, 2)$ despite the fact that $\{2, 5\} = \{5, 2\}$. Make sure to keep a clear distinction between the ordered pair $(2, 5)$ and the set $\{2, 5\}$. We *do* allow the possibility of creating something like $(2, 2)$, and here the repetition of 2's is meaningful. Furthermore, we do not use $\in$ in ordered pairs, so we would **not** write $2 \in (2, 5)$. We'll talk about ways to refer to the two elements of an ordered pair later.

We can generalize ordered pairs to the possibility of having more than 2 elements. In this case, we have an ordered list of n elements, so something like $(5, 4, 5, -2)$. We use call such an object an n-tuple, a list with n elements, or a finite sequence of length n. Thus, for example, we could call $(5, 4, 5, -2)$ a 4-tuple. It is also possible to have infinite sequences (i.e. infinite lists), but we will wait to discuss these when the time comes.

1.4 Operations on Sets and Sequences

Aside from listing elements, carving out subsets of a given set using a given property, and giving a parametric description (which as mentioned above is just a special case of the previous type), there are other ways to build sets.

Definition 1.3. *Given two sets A and B, we define*

$$A \cup B = \{x : x \in A \textit{ or } x \in B\}$$

and call this set the union *of A and B.*

Here, as in mathematics generally, we use *or* to mean "inclusive or". In other words, if x is an element of both A and B, then we still put x into $A \cup B$. Here are a few examples (we leave the proofs of the latter results until we have more theory):

- $\{1, 2, 7\} \cup \{4, 9\} = \{1, 2, 4, 7, 9\}$.
- $\{1, 2, 3\} \cup \{2, 3, 5\} = \{1, 2, 3, 5\}$.
- $\{2n : n \in \mathbb{N}\} \cup \{2n + 1 : n \in \mathbb{N}\} = \mathbb{N}$.
- $\{2n : n \in \mathbb{N}^+\} \cup \{2n + 1 : n \in \mathbb{N}^+\} = \{2, 3, 4, \dots\}$.
- $\{2n : n \in \mathbb{N}^+\} \cup \{2n - 1 : n \in \mathbb{N}^+\} = \{1, 2, 3, 4, \dots\} = \mathbb{N}^+$.
- $A \cup \emptyset = A$ for every set A.

Definition 1.4. *Given two sets A and B, we define*

$$A \cap B = \{x : x \in A \textit{ and } x \in B\}$$

and call this set the intersection *of A and B.*

Here are a few examples (again we leave some proofs until later):

- $\{1,2,7\} \cap \{4,9\} = \emptyset$.
- $\{1,2,3\} \cap \{2,3,5\} = \{2,3\}$.
- $\{1,\{2,3\}\} \cap \{1,2,3\} = \{1\}$.
- $\{2n : n \in \mathbb{Z}\} \cap \{3n : n \in \mathbb{Z}\} = \{6n : n \in \mathbb{Z}\}$.
- $\{3n+1 : n \in \mathbb{N}^+\} \cap \{3n+2 : n \in \mathbb{N}^+\} = \emptyset$.
- $A \cap \emptyset = \emptyset$ for every set A.

Definition 1.5. *Given two sets A and B, we define*

$$A \backslash B = \{x : x \in A \text{ and } x \notin B\}$$

and call this set the (relative) complement *of B (in A).*

In many cases, we have $B \subseteq A$, but we occasionally us it generally. Here are a few examples:

- $\{5,6,7,8,9\} \backslash \{5,6,8\} = \{7,9\}$.
- $\{1,2,7\} \backslash \{4,9\} = \{1,2,7\}$.
- $\{1,2,3\} \cap \{2,3,5\} = \{1\}$.
- $\{2n : n \in \mathbb{Z}\} \backslash \{4n : n \in \mathbb{Z}\} = \{4n+2 : n \in \mathbb{Z}\}$.
- $A \backslash \emptyset = A$ for every set A.
- $A \backslash A = \emptyset$ for every set A.

Definition 1.6. *Given a set A, we let $\mathcal{P}(A)$ be the set of all subsets of A, and we call $\mathcal{P}(A)$ the* power set *of A.*

For example, we have

$$\mathcal{P}(\{1,2\}) = \{\emptyset, \{1\}, \{2\}, \{1,2\}\}$$

and

$$\mathcal{P}(\{4,5,7\}) = \{\emptyset, \{4\}, \{5\}, \{7\}, \{4,5\}, \{4,7\}, \{5,7\}, \{4,5,7\}\}$$

Definition 1.7. *Given two sets A and B, we let $A \times B$ be the set of all ordered pairs (a,b) such that $a \in A$ and $b \in B$, and we call this set the* Cartesian product *of A and B.*

For example, we have

$$\{1,2,3\} \times \{6,8\} = \{(1,6),(1,8),(2,6),(2,8),(3,6),(3,8)\}$$

and

$$\mathbb{N} \times \mathbb{N} = \{(0,0),(0,1),(1,0),(2,0),\dots,(4,7),\dots\}$$

Notice that elements of $\mathbb{R} \times \mathbb{R}$ correspond to points in the plane.

We can also generalize the concept of a Cartesian product to more than 2 sets. If we are given n sets $A_1, A_2, \dots, A_n$, we let $A_1 \times A_2 \times \cdots \times A_n$ be the set of all n-tuples $(a_1, a_2, \dots, a_n)$ such that $a_i \in A_i$ for each i. For example, we have

$$\{1,2\} \times \{3\} \times \{4,5\} = \{(1,3,4),(1,3,5),(2,3,4),(2,3,5)\}$$

In the special case when $A_1, A_2, \dots, A_n$ are all the same set A, we use the notation A^n to denote the set $A \times A \times \cdots \times A$ (where we have n copies of A). Thus, A^n is the set of all finite sequences of elements of A of length n. For example, $\{0,1\}^n$ is the set of all finite sequences of 0's and 1's of length n. Notice that this notation fits in with the notation $\mathbb{R}^n$ that we are used to in Calculus and Linear Algebra.

Definition 1.8. *Given a set A, we let A^* be the set of all finite sequences of elements of A of any length, including the empty sequence (the unique sequence of length 0).*

Thus, for example, the set $\{0,1\}^*$ is the set of all finite sequences of 0's and 1's. If we use λ to denote the empty sequence and write things like 010 in place of the more precise $(0,1,0)$, then we have

$$\{0,1\}^* = \{\lambda, 0, 1, 00, 01, 10, 11, 000, 001, \dots\}$$

Notice that if $A \neq \emptyset$, then A^* is an infinite set.

Definition 1.9. *Given two finite sequences σ and τ, we let $\sigma\tau$ be the* concatenation *of σ and τ, i.e. if $\sigma = (a_1, a_2, \dots, a_m)$ and $\tau = (b_1, b_2, \dots, b_n)$, then $\sigma\tau = (a_1, a_2, \dots, a_m, b_1, b_2, \dots, b_n)$.*

1.5 The Cardinality of Sets

We will spend a significant amount of time trying to count the number of elements in certain sets. For now, we will study some simple properties that will become extremely useful later when employed in clever ways.

Definition 1.10. *Given a set A, we let $|A|$ be the number of elements of A, and we call $|A|$ the cardinality of A. If A is infinite, then we write $|A| = \infty$.*

Of course, if we list the elements of a set A, then it's usually quite easy to determine $|A|$. For example, we trivially have $|\{1, \sqrt{2}, \frac{5}{2}, 18\}| = 4$. However, it can be very hard to determine the cardinality of a set. For example, consider the set

$$A = \{(x,y) \in \mathbb{Z}^2 : x^3 = y^2 + 1\}$$

Determining the elements of A is nontrivial. It's easy to see that $(1,0) \in A$, but it's not clear whether there are any other elements. Using some nontrivial number theory, it is possible to show that $A = \{(1,0)\}$, and hence $|A| = 1$.

We start with one of the most basic, yet important, rules about the cardinality of sets.

Definition 1.11. *We say that two sets A and B are disjoint if $A \cap B = \emptyset$.*

Fact 1.12 (Sum Rule)**.** *If A and B are finite disjoint sets, then $|A \cup B| = |A| + |B|$.*

We won't give a formal proof of this fact, because it is so basic that it's hard to know what to assume (although if one goes through the trouble of axiomatizing math with something like set theory, then it's possible to give a formal proof by induction on $|B|$). At any rate, the key fact is that since A and B are disjoint, they have no elements in common. Therefore, each element of $A \cup B$ is in exactly one of A or B. Notice that the assumption that A and B are disjoint is essential. If $A = \{1,2\}$ and $B = \{2,3\}$, then $|A| = 2 = |B|$, but $|A \cup B| = 3$ because $A \cup B = \{1,2,3\}$.

Although the next result is again very intuitive, we show how to prove it using the Sum Rule.

Proposition 1.13 (Complement Rule)**.** *If A and B are finite sets and $B \subseteq A$, then $|A \backslash B| = |A| - |B|$.*

Proof. Notice that $A\backslash B$ and B are disjoint sets and that $(A\backslash B) \cup B = A$. Using the Sum Rule, we conclude that $|A\backslash B| + |B| = |A|$. The result follows. □

We can now easily generalize this to the case where B may not be a subset of A.

Proposition 1.14 (General Complement Rule)**.** *If A and B are finite sets, then $|A\backslash B| = |A| - |A \cap B|$.*

Proof. We have $A\backslash B = A\backslash(A \cap B)$. Since $A \cap B \subseteq A$, we can now apply the Complement Rule. □

We can generalize the Sum Rule to the following.

Definition 1.15. *A collection of sets $A_1, A_2, \dots, A_n$ is* pairwise disjoint *if $A_i \cap A_j = \emptyset$ whenever $i \neq j$.*

Fact 1.16 (General Sum Rule)**.** *If $A_1, A_2, \dots, A_n$ are finite sets that are pairwise disjoint sets, then $|A_1 \cup A_2 \cup \cdots \cup A_n| = |A_1| + |A_2| + \cdots + |A_n|$.*

Again, we won't give a formal proof of this fact (although it it possible to do so from the Sum Rule by induction on n). Notice that as above the pairwise disjoint assumption is key, and it's not even enough to assume that $A_1 \cap A_2 \cap \cdots \cap A_n = \emptyset$ (see the homework).

Proposition 1.17. *If A and B are finite sets, we have $|A \cup B| = |A| + |B| - |A \cap B|$.*

Proof. Consider the three sets $A\backslash B$, $B\backslash A$, and $A \cap B$. These three sets are pairwise disjoint, and their union is $A \cup B$. Using the General Sum Rule, we conclude that

$$|A \cup B| = |A\backslash B| + |B\backslash A| + |A \cap B|$$

Now $|A\backslash B| = |A| - |A \cap B|$ and $|B\backslash A| = |B| - |A \cap B|$ by the General Complement Rule. Plugging these in, we conclude that

$$|A \cup B| = |A| - |A \cap B| + |B| - |A \cap B| + |A \cap B|$$

and hence

$$|A \cup B| = |A| + |B| - |A \cap B|$$

□

Proposition 1.18 (Product Rule)**.** *If A and B are finite sets, then $|A \times B| = |A| \cdot |B|$.*

Proof. Let $n = |A|$ and let $m = |B|$. List the elements of A so that $A = \{a_1, a_2, \dots, a_n\}$. Similarly, list the elements of B so that $B = \{b_1, b_2, \dots, b_m\}$. For each i, let

$$A_i = \{(a_i, b_j) : 1 \leq j \leq m\} = \{(a_i, b_1), (a_i, b_2), \dots, (a_i, b_m)\}$$

Thus, A_i is the subset of $A \times B$ consisting only of those pairs whose first element is a_i. Notice that the sets $A_1, A_2, \dots, A_n$ are pairwise disjoint and that

$$A \times B = A_1 \cup A_2 \cup \cdots \cup A_n$$

Furthermore, we have that $|A_i| = m$ for all i. Using the General Sum Rule, we conclude that

$$\begin{aligned} |A \times B| &= |A_1| + |A_2| + \cdots + |A_n| \\ &= m + m + \cdots + m \\ &= n \cdot m \\ &= |A| \cdot |B| \end{aligned}$$

The result follows. □

Using induction (see below), one can prove the following generalization.

Proposition 1.19 (General Product Rule)**.** *If $A_1, A_2, \dots, A_n$ are finite sets, then $|A_1 \times A_2 \times \cdots \times A_n| = |A_1| \cdot |A_2| \cdots |A_n|$.*

Corollary 1.20. *If A is a finite set and $n \in \mathbb{N}^+$, then $|A^n| = |A|^n$.*

Corollary 1.21. *For any $n \in \mathbb{N}^+$, we have that $|\{0,1\}^n| = 2^n$, i.e. there are 2^n many sequences of 0's and 1's of length n.*

1.6 Relations

Definition 1.22. *Let A and B be sets. A* (binary) relation *between A and B is a subset $R \subseteq A \times B$. If $A = B$, then we call a subset of $A \times A$ a* (binary) relation on A.

For example, let $A = \{1, 2, 3\}$ and $B = \{6, 8\}$ as above. Let

$$R = \{(1, 6), (1, 8), (3, 8)\}$$

We then have that R is a relation between A and B, although certainly not a very interesting one. However, we'll use it to illustrate a few facts. First, in a relation, it's possible for an element of A to be related to multiple elements of B, as in the case for $1 \in A$ in our example R. Also, it's possible that an element of A is related to no elements of B, as in the case of $2 \in A$ in our example R.

For a more interesting example, consider the binary relation on $\mathbb{Z}$ defined by $R = \{(a, b) \in \mathbb{Z}^2 : a < b\}$. Notice that $(4, 7) \in R$ and $(5, 5) \notin R$.

By definition, relations are sets. However, it is typically cumbersome to use set notation to write things like $(1, 6) \in R$. Instead, it usually makes much more sense to use infix notation and write $1R6$. Moreover, we can use better notation for the relation by using a symbol like $\sim$ instead of R. In this case, we would write $1 \sim 6$ instead of $(1, 6) \in \sim$ or $2 \not\sim 8$ instead of $(2, 8) \notin \sim$.

With this new notation, we give a few examples of binary relations on $\mathbb{R}$:

- Given $x, y \in \mathbb{R}$, we let $x \sim y$ if $x^2 + y^2 = 1$.
- Given $x, y \in \mathbb{R}$, we let $x \sim y$ if $x^2 + y^2 \leq 1$.
- Given $x, y \in \mathbb{R}$, we let $x \sim y$ if $x = \sin y$.
- Given $x, y \in \mathbb{R}$, we let $x \sim y$ if $y = \sin x$.

Again, notice from these examples that given $x \in \mathbb{R}$, there many 0, 1, 2, or even infinitely many $y \in \mathbb{R}$ with $x \sim y$.

If we let $A = \{0, 1\}^*$ be the set of all finite sequences of 0's and 1's, then the following are binary relations on A:

- Given $\sigma, \tau \in A$, we let $\sigma \sim \tau$ if σ and τ have the same number of 1's.
- Given $\sigma, \tau \in A$, we let $\sigma \sim \tau$ if σ occurs as a consecutive subsequence of τ (for example, we have $010 \sim 001101011$ because 010 appears in positions 5-6-7 of 001101011).

For a final example, let A be the set consisting of the 50 states. Let R be the subset of $A \times A$ consisting of those pairs of states whose second letter of their postal codes are equal. For example, we have (Iowa,California) $\in R$ and and (Iowa, Virginia) $\in R$ because the postal codes of these sets are IA, CA, VA. We also have (Minnesota, Tennessee) $\in R$ because of the postal codes MN and TN. Now (Texas, Texas) $\in R$, but there is no $a \in A$ with $a \neq$ Texas such that (Texas, a) $\in R$ because no other state has X as the second letter of its postal code. Texas stands alone.

1.7 Equivalence Relations

Definition 1.23. *An* equivalence relation *on a set A is a binary relation $\sim$ on A having the following three properties:*

- *$\sim$ is reflexive: $a \sim a$ for all $a \in A$.*
- *$\sim$ is symmetric: Whenever $a, b \in A$ satisfy $a \sim b$, we have $b \sim a$.*

- *$\sim$ is transitive: Whenever $a, b, c \in A$ satisfy $a \sim b$ and $b \sim c$, we have $a \sim c$.*

Consider the binary relation $\sim$ on $\mathbb{Z}$ where $a \sim b$ means that $a \leq b$. Notice that $\sim$ is reflexive because $a \leq a$ for all $a \in \mathbb{Z}$. Also, $\sim$ is transitive because if $a \leq b$ and $b \leq c$, then $a \leq c$. However, $\sim$ is not symmetric because $3 \sim 4$ but $4 \not\sim 3$. Thus, although $\sim$ satisfies two out of the three requirements, it is not an equivalence relation.

A simple example of an equivalence relation is where $A = \mathbb{R}$ and $a \sim b$ means that $|a| = |b|$. In this case, it is straightforward to check that $\sim$ is an equivalence relation. We now move on to some more interesting examples which we treat more carefully.

Example 1.24. *Let A be the set of all $n \times n$ matrices with real entries. Let $M \sim N$ mean that there exists an invertible $n \times n$ matrix P such that $M = PNP^{-1}$. We then have that $\sim$ is an equivalence relation on A.*

Proof. We need to check the three properties.

- Reflexive: Let $M \in A$. The $n \times n$ identity matrix I is invertible and satisfies $I^{-1} = I$, so we have $M = IMI^{-1}$. Therefore, $\sim$ is reflexive.
- Symmetric: Let $M, N \in A$ with $M \sim N$. Fix a $n \times n$ invertible matrix P with $M = PNP^{-1}$. Multiplying on the left by P^{-1} we get $P^{-1}M = NP^{-1}$, and now multiplying on the right by P we conclude that $P^{-1}MP = N$. We know from linear algebra that P^{-1} is also invertible and $(P^{-1})^{-1} = P$, so $N = P^{-1}M(P^{-1})^{-1}$ and hence $N \sim M$.
- Transitive: Let $L, M, N \in A$ with $L \sim M$ and $M \sim N$. Since $L \sim M$, we may fix a $n \times n$ invertible matrix P with $L = PMP^{-1}$. Since $M \sim N$, we may fix a $n \times n$ invertible matrix Q with $M = QNQ^{-1}$. We then have
$$L = PMP^{-1} = P(QNQ^{-1})P^{-1} = (PQ)N(Q^{-1}P^{-1})$$
Now by linear algebra, we know that the product of two invertible matrices is invertible, so PQ is invertible and furthermore we know that $(PQ)^{-1} = Q^{-1}P^{-1}$. Therefore, we have
$$L = (PQ)N(PQ)^{-1}$$
so $L \sim N$.

Putting it all together, we conclude that $\sim$ is an equivalence relation on A. □

Example 1.25. *Let A be the set $\mathbb{Z} \times (\mathbb{Z}\backslash\{0\})$, i.e. A is the set of all pairs $(a, b) \in \mathbb{Z}^2$ with $b \neq 0$. Define a relation $\sim$ on A as follows. Given $a, b, c, d \in \mathbb{Z}$ with $b, d \neq 0$, we let $(a, b) \sim (c, d)$ mean $ad = bc$. We then have that $\sim$ is an equivalence relation on A.*

Proof. We check the three properties.

- Reflexive: Let $a, b \in \mathbb{Z}$ with $b \neq 0$. Since $ab = ba$, it follows that $(a, b) \sim (a, b)$.
- Symmetric: Let $a, b, c, d \in \mathbb{Z}$ with $b, d \neq 0$, and $(a, b) \sim (c, d)$. We then have that $ad = bc$. From this, we conclude that $cb = da$ so $(c, d) \sim (a, b)$.
- Transitive: Let $a, b, c, d, e, f \in \mathbb{Z}$ with $b, d, f \neq 0$ where $(a, b) \sim (c, d)$ and $(c, d) \sim (e, f)$. We then have that $ad = bc$ and $cf = de$. Multiplying the first equation by f we see that $adf = bcf$. Multiplying the second equation by b gives $bcf = bde$. Therefore, we know that $adf = bde$. Now $d \neq 0$ by assumption, so we may cancel it to conclude that $af = be$. It follows that $(a, b) \sim (e, f)$

Therefore, $\sim$ is an equivalence relation on A. □

Let's analyze the above situation more carefully. We have $(1,2) \sim (2,4)$, $(1,2) \sim (4,8)$, $(1,2) \sim (-5,-10)$, etc. If we think of (a,b) as representing the fraction $\frac{a}{b}$, then the relation $(a,b) \sim (c,d)$ is saying exactly that the fractions $\frac{a}{b}$ and $\frac{c}{d}$ are equal. You may never have thought about equality of fractions as the result of imposing an equivalence relation on pairs of integers, but that is exactly what it is. We will be more precise about this below.

Definition 1.26. *Let $\sim$ be an equivalence relation on a set A. Given $a \in A$, we let*

$$\overline{a} = \{b \in A : a \sim b\}$$

The set $\overline{a}$ is called the equivalence class *of a.*

Some sources use the notation $[a]$ instead of $\overline{a}$. This notation helps emphasize that the equivalence class of a is a *subset* of A rather than an element of A. However, it is cumbersome notation when we begin working with equivalence classes. We will stick with our notation, although it might take a little time to get used to. Notice that by the reflexive property of $\sim$, we have that $a \in \overline{a}$ for all $a \in A$.

For example, let's return to where A is the set consisting of the 50 states and R is the subset of $A \times A$ consisting of those pairs of states whose second letter of their postal codes are equal. It's straightforward to show that R is an equivalence relation on A. We have

$$\overline{\text{Iowa}} = \{\text{California}, \text{Georgia}, \text{Iowa}, \text{Louisiana}, \text{Massachusetts}, \text{Pennsylvania}, \text{Virginia}, \text{Washington}\}$$

while

$$\overline{\text{Minnesota}} = \{\text{Indiana}, \text{Minnesota}, \text{Tennessee}\}$$

and

$$\overline{\text{Texas}} = \{\text{Texas}\}$$

Notice that each of these are sets, even in the case of $\overline{\text{Texas}}$.

For another example, suppose we are working with $A = \mathbb{Z} \times (\mathbb{Z}\backslash\{0\})$ where $(a,b) \sim (c,d)$ means that $ad = bc$. As discussed above, some elements of $\overline{(1,2)}$ are $(1,2)$, $(2,4)$, $(4,8)$, $(-5,-10)$, etc. So

$$\overline{(1,2)} = \{(1,2), (2,4), (4,8), (-5,-10), \dots\}$$

Again, I want to emphasize that $\overline{(a,b)}$ is a subset of A.

The following proposition is hugely fundamental. It says that if two equivalence classes overlap, then they must in fact be equal. In other words, if $\sim$ is an equivalence on A, then the equivalence classes *partition* the set A into pieces.

Proposition 1.27. *Let $\sim$ be an equivalence relation on a set A and let $a, b \in A$. If $\overline{a} \cap \overline{b} \neq \emptyset$, then $\overline{a} = \overline{b}$.*

Proof. Suppose that $\overline{a} \cap \overline{b} \neq \emptyset$. Fix $c \in \overline{a} \cap \overline{b}$. We then have $a \sim c$ and $b \sim c$. By symmetry, we know that $c \sim b$, and using transitivity we get that $a \sim b$. Using symmetry again, we conclude that $b \sim a$.

We first show that $\overline{a} \subseteq \overline{b}$. Let $x \in \overline{a}$. We then have that $a \sim x$. Since $b \sim a$, we can use transitivity to conclude that $b \sim x$, hence $x \in \overline{b}$.

We next show that $\overline{b} \subseteq \overline{a}$. Let $x \in \overline{b}$. We then have that $b \sim x$. Since $a \sim b$, we can use transitivity to conclude that $a \sim x$, hence $x \in \overline{a}$.

Putting this together, we get that $\overline{a} = \overline{b}$. □

With that proposition in hand, we are ready for the foundational theorem about equivalence relations.

Theorem 1.28. *Let $\sim$ be an equivalence relation on a set A and let $a, b \in A$.*

1. *$a \sim b$ if and only if $\overline{a} = \overline{b}$.*

2. $a \not\sim b$ *if and only if* $\overline{a} \cap \overline{b} = \emptyset$.

Proof. We first prove 1. Suppose first that $a \sim b$. We then have that $b \in \overline{a}$. Now we know that $b \sim b$ because $\sim$ is reflexive, so $b \in \overline{b}$. Thus, $b \in \overline{a} \cap \overline{b}$, so $\overline{a} \cap \overline{b} \neq \emptyset$. By the previous proposition, we conclude that $\overline{a} = \overline{b}$.

Suppose conversely that $\overline{a} = \overline{b}$. Since $b \sim b$ because $\sim$ is reflexive, we have that $b \in \overline{b}$. Therefore, $b \in \overline{a}$ and hence $a \sim b$.

We now use everything we've shown to get 2 with little effort. Suppose that $a \not\sim b$. Since we just proved 1, it follows that $\overline{a} \neq \overline{b}$, so by the previous proposition we must have $\overline{a} \cap \overline{b} = \emptyset$. Suppose conversely that $\overline{a} \cap \overline{b} = \emptyset$. We then have $\overline{a} \neq \overline{b}$ (because $a \in \overline{a}$ so $\overline{a} \neq \emptyset$), so $a \not\sim b$ by part 1. □

Therefore, given an equivalence relation $\sim$ on a set A, the equivalence classes partition A into pieces. Working out the details in our postal code example, one can show that $\sim$ has 1 equivalence class of size 8 (namely $\overline{\text{Iowa}}$, which is the same set as $\overline{\text{California}}$ and 6 others), 3 equivalence classes of size 4, 4 equivalence classes of size 3, 7 equivalence classes of size 2, and 4 equivalence classes of size 1.

Let's revisit the example of $A = \mathbb{Z} \times (\mathbb{Z} \backslash \{0\})$ where $(a, b) \sim (c, d)$ means $ad = bc$. The equivalence class of $(1, 2)$, namely the set $\overline{(1,2)}$ is the set of all pairs of integers which are ways of representing the fraction $\frac{1}{2}$. In fact, this is how once can "construct" the rational numbers from the integers. We simply *define* the rational numbers to be the set of equivalence classes of A under $\sim$. In other words, we let

$$\frac{a}{b} = \overline{(a,b)}$$

So when we write something like

$$\frac{1}{2} = \frac{4}{8}$$

we are simply saying that

$$\overline{(1,2)} = \overline{(4,8)}$$

which is true because $(1, 2) \sim (4, 8)$.

1.8 Functions

Intuitively, given two sets A and B, a function $f : A \to B$ is a input-output "mechanism" that produces a *unique* output $b \in B$ for any given input $a \in A$. Up through calculus, the vast majority of functions that we encounter are given by simple formulas, so this "mechanism" was typically interpreted in an algorithmic and computational sense. However, some functions such as $f(x) = \sin x$, $f(x) = \ln x$, or integral functions like $f(x) = \int_a^x g(t)\ dt$ (given a continuous function $g(t)$ and a fixed $a \in \mathbb{R}$) were defined in more interesting ways where it was not at all obvious how to compute them. We are now in a position to define functions as relations that satisfy a certain property. Thinking about functions from this more abstract point of view eliminates the vague "mechanism" concept because they will simply be certain types of sets. With this perspective, we'll see that functions can be defined in any way that a set can be defined. This approach both clarifies the concept of a function as well as providing us with some much needed flexibility in defining functions in more interesting ways.

Definition 1.29. *Let A and B be sets. A* function *from A to B is relation f between A and B such that for each $a \in A$, there is a unique $b \in B$ with $(a, b) \in f$.*

For example, let $A = \{c, q, w, y\}$ and let $B = \mathbb{N} = \{0, 1, 2, 3, 4, \dots\}$. An example of a function from A to B is the set

$$f = \{(c, 71), (q, 4), (w, 9382), (y, 4)\}.$$

Notice that in the definition of a function from A to B, we know that for every $a \in A$, there is a unique $b \in B$ such that $(a, b) \in f$. However, as this example shows, it may not be the case that for every $b \in B$, there is a unique $a \in A$ with $(a, b) \in f$. Be careful with the order of quantifiers!

Thinking of functions as special types of relations, and in particular as special types of sets, is occasionally helpful (see below), but is often awkward in practice. For example, writing $(c, 71) \in f$ to mean that f sends c to 71 gets annoying very quickly. Using infix notation like $c\ f\ 71$ is not much better. Thus, we introduce some new notation matching up with our old experience with functions.

Notation 1.30. *Let A and B be sets.*

- *Instead of writing "f is a function from A to B", we typically use the shorthand notation "$f \colon A \to B$".*
- *If $f \colon A \to B$ and $a \in A$, we write $f(a)$ to mean the unique $b \in B$ such that $(a, b) \in f$.*

Therefore, in the above example of f, we have

$$\begin{aligned} f(c) &= 71 \\ f(q) &= 4 \\ f(w) &= 9382 \\ f(y) &= 4 \end{aligned}$$

Definition 1.31. *Let $f \colon A \to B$ be a function. We define the following.*

- *We call A the* domain *of f.*
- *We call B the* codomain *of f.*
- *We define $\text{range}(f) = \{b \in B : \text{There exists } a \in A \text{ with } f(a) = b\}$.*

Notice that given a function $f \colon A \to B$, we have $\text{range}(f) \subseteq B$, but it is possible that $\text{range}(f) \neq B$. For example, in the above case, we have that the codomain of f is $\mathbb{N}$, but $\text{range}(f) = \{4, 71, 9382\}$.

In general, given a function $f \colon A \to B$, it may be very difficult to determine $\text{range}(f)$ because we may need to search through all $a \in A$. For example, fix $n \in \mathbb{N}^+$ and define $f \colon \mathbb{N} \to \{0, 1, 2, \dots, n-1\}$ by letting $f(a)$ be the remainder when dividing a^2 by n. This simple but strange looking function has many interesting properties. Given a large number n, computing whether a given number in $\{0, 1, 2, \dots, n-1\}$ is an element of $\text{range}(f)$ is thought to be extremely hard. In fact, it is widely believed that there is no efficient algorithm to compute this when n is the product of two primes, and this is the basis for some cryptosystems and pseudo-random number generators.

One nice feature of our definition of a function is that we immediately obtain a nice definition for when two functions $f \colon A \to B$ and $g \colon A \to B$ are equal because we have defined when two sets are equal. Unwrapping this definition, we see that $f = g$ exactly when f and g have the same elements, which is precisely the same thing as saying that $f(a) = g(a)$ for all $a \in A$. In particular, the *manner* in which we describe functions does not matter so long as the functions behave the same on all inputs. For example, if we define $f \colon \mathbb{R} \to \mathbb{R}$ and $g \colon \mathbb{R} \to \mathbb{R}$ by letting $f(x) = \sin^2 x + \cos^2 x$ and $g(x) = 1$, then we have that $f = g$. Just like sets (because after all functions are sets!), the definitions do not matter as long as the elements are the same.

Definition 1.32. *Suppose that $f \colon A \to B$ and $g \colon B \to C$ are functions. The* composition *of g and f, denoted $g \circ f$, is the function $g \circ f \colon A \to C$ defined by $(g \circ f)(a) = g(f(a))$ for all $a \in A$.*

Instead of defining $g \circ f$ in function language, one can also define function composition directly in terms of the sets f and g. Suppose that $f \colon A \to B$ and $g \colon B \to C$ are functions. Define a new set

$$R = \{(a, c) \in A \times C : \text{There exists } b \in B \text{ with } (a, b) \in f \text{ and } (b, c) \in g\}$$

Now R is a relation, and one can check that it is a function (using the assumption that f and g are both functions). We define $g \circ f$ to be this set.

Proposition 1.33. *Let A, B, C, D be sets. Suppose that $f\colon A \to B$, that $g\colon B \to C$, and that $h\colon C \to D$ are functions. We then have that $(h \circ g) \circ f = h \circ (g \circ f)$. Stated more simply, function composition is associative whenever it is defined.*

Proof. Let $a \in A$ be arbitrary. We then have

$$\begin{aligned} ((h \circ g) \circ f)(a) &= (h \circ g)(f(a)) \\ &= h(g(f(a))) \\ &= h((g \circ f)(a)) \\ &= (h \circ (g \circ f))(a) \end{aligned}$$

Therefore $((h \circ g) \circ f)(a) = (h \circ (g \circ f))(a)$ for all $a \in A$. It follows that $(h \circ g) \circ f = h \circ (g \circ f)$. □

Notice that in general we have $f \circ g \neq g \circ f$ even when both are defined! For example, if $f\colon \mathbb{R} \to \mathbb{R}$ is $f(x) = x + 1$ and $g\colon \mathbb{R} \to \mathbb{R}$ is $g(x) = x^2$, then

$$(f \circ g)(x) = f(g(x)) = f(x^2) = x^2 + 1$$

while

$$(g \circ f)(x) = g(f(x)) = g(x + 1) = (x + 1)^2 = x^2 + 2x + 1$$

For example, we have $(f \circ g)(1) = 1^2 + 1 = 2$ while $(g \circ f)(1) = 1^2 + 2 \cdot 1 + 1 = 4$. Since we have found one example of an x with $(f \circ g)(x) \neq (f \circ g)(x)$, we conclude that $f \circ g \neq g \circ f$. It does not matter that there do exist some values of x with $(f \circ g)(x) = (f \circ g)(x)$ (for example, this is true when $x = 0$). Remember that two functions are equal precisely when they agree on *all* inputs, so to show that the two functions are not equal it suffices to find just one value where they disagree.

Definition 1.34. *Let A be a set. The function $id_A\colon A \to A$ defined by $id_A(a) = a$ for all $a \in A$ is called the* identity function *on A.*

The identity function does leave other functions alone when we compose with it. However, we have to be careful that we compose with the identity function on the correct set and the correct side.

Proposition 1.35. *For any function $f\colon A \to B$, we have $f \circ id_A = f$ and $id_B \circ f = f$.*

Proof. Let $f\colon A \to B$. For any $a \in A$, we have

$$(f \circ id_A)(a) = f(id_A(a)) = f(a)$$

Since $a \in A$ was arbitrary, it follows that $f \circ id_A = f$. For any $b \in B$, we have

$$(id_B \circ f)(a) = id_B(f(a)) = f(a)$$

because $f(a)$ is some element in B. Since $b \in B$ was arbitrary, it follows that $id_B \circ f = f$. □

Definition 1.36. *Let $f\colon A \to B$ be a function.*

- *We say that f is injective (or one-to-one) if whenever $f(a_1) = f(a_2)$ we have $a_1 = a_2$.*
- *We say that f is surjective (or onto) if for all $b \in B$ there exists $a \in A$ such that $f(a) = b$. In other words, f is surjective if $range(f) = B$.*
- *We say that f is bijective if both f is injective and surjective.*

An equivalent condition for f to be injective is obtained by simply taking the contrapositive, i.e. $f\colon A \to B$ is injective if and only if whenever $a_1 \neq a_2$, we have $f(a_1) \neq f(a_2)$. Stated in more colloquial language, f is injective if every element of B is hit by at most one element of A via f. In this manner, f is surjective if every element of B is hit by at least one element of a via f, and f is bijective if every element of B is hit by exactly one element of a via f.

If we want to prove that a function $f\colon A \to B$ is injective, it is usually better to use our official definition than the contrapositive one with negations. Thus, we want to start by assuming that we are given arbitrary $a_1, a_2 \in A$ that satisfy $f(a_1) = f(a_2)$, and using this assumption we want to prove that $a_1 = a_2$. The reason why this approach is often preferable is because it is typically easier to work with and manipulate a statement involving equality than it is to derive statements from a non-equality.

Example 1.37. *We have the following examples:*

- $f\colon \mathbb{R} \to \mathbb{R}$ *defined by* $f(x) = 2x$ *is both injective and surjective, so is bijective.*
- $f\colon \mathbb{Z} \to \mathbb{Z}$ *defined by* $f(n) = 2n$ *is injective but not surjective.*
- $f\colon \{0,1\}^* \to \mathbb{N}$ *defined by* $f(\sigma) = |\sigma|$ *is surjective but not injective.*
- $f\colon \{0,1\}^* \to \mathbb{Z}$ *defined by* $f(\sigma) = |\sigma|$ *is neither surjective nor injective.*
- $f\colon \mathbb{R} \to \mathbb{R}$ *defined by* $f(x) = \sin x$ *is neither injective nor surjective.*
- $f\colon \mathbb{N}^+ \to \mathbb{N}^+$ *defined by letting* $f(n)$ *be the number of positive divisors of* n *is surjective (not easy), but it is not injective.*
- $f\colon \mathbb{Q} \to \mathbb{Z}$ *defined by* $f(\frac{a}{b}) = a$ *is not even a function because we would need both* $f(\frac{1}{2}) = 1$ *and* $f(\frac{2}{4}) = 2$*, but this contradicts the definition of a function because* $\frac{1}{2} = \frac{2}{4}$*.*

Proposition 1.38. *Let* A, B, C *be sets and let* $f\colon A \to B$ *and* $g\colon B \to C$ *be functions*

1. *If* f *and* g *are both injective, then* $g \circ f$ *is injective.*
2. *If* f *and* g *are both surjective, then* $g \circ f$ *is surjective.*
3. *If* f *and* g *are both bijective, then* $g \circ f$ *is bijective.*
4. *If* $g \circ f$ *is injective, then* f *is injective.*
5. *If* $g \circ f$ *is surjective, then* g *is surjective.*

Proof.

1. Suppose that f and g are both injective. Let $a_1, a_2 \in A$ be arbitrary with $(g \circ f)(a_1) = (g\circ)(a_2)$. By definition of composition, we then have $g(f(a_1)) = g(f(a_2))$. Using the fact that g is injective, we conclude that $f(a_1) = f(a_2)$. Now we use the fact that f is injective to conclude that $a_1 = a_2$. Therefore, $g \circ f$ is injective.

2. Suppose that f and g are both surjective. Let $c \in C$ be arbitrary. Since g is surjective, we can fix $b \in B$ with $g(b) = c$. Since f is surjective, we can fix $a \in A$ with $f(a) = b$. We then have

$$\begin{aligned}(g \circ f)(a) &= g(f(a)) \\ &= g(b) \\ &= c\end{aligned}$$

Since $c \in C$ was arbitrary, we conclude that $g \circ f$ is surjective.

3. This follows from combining 1 and 2.

4. Suppose that $g \circ f$ is injective. Let $a_1, a_2 \in A$ be arbitrary with $f(a_1) = f(a_2)$. Applying g to both sides, we then have that $g(f(a_1)) = g(f(a_2))$, so $(g \circ f)(a_1) = (g \circ f)(a_2)$. Using the fact that $g \circ f$ is injective, it follows that $a_1 = a_2$. Therefore, f is injective.

5. Suppose that $g \circ f$ is surjective. Let $c \in C$ be arbitrary. Since $g \circ f$ is surjective, we can fix $a \in A$ with $(g \circ f)(a) = c$. By definition of composition, we then have $g(f(a)) = c$. Since $f(a) \in B$, we have succeeded in finding a b with $g(b) = c$ (namely $b = f(a)$). Since $c \in C$ was arbitrary, we conclude that g is surjective.

□

1.9 Divisibility

Definition 1.39. *Let $a, b \in \mathbb{Z}$. We say that a* divides *b, and write $a \mid b$, if there exists $m \in \mathbb{Z}$ with $b = am$.*

For example, we have $2 \mid 6$ because $2 \cdot 3 = 6$ and $3 \mid -21$ because $3 \cdot (-7) = 21$. On the other hand, we have $2 \nmid 5$. To see this, we argue as follows.

- We have $2 \cdot 0 = 0$, $2 \cdot 1 = 2$, and $2 \cdot 2 = 4$.
- For any $m \in \mathbb{Z}$ with $m > 2$, we have $m \geq 3$, so $2m \geq 6$.
- For any $m \in \mathbb{Z}$ with $m < 0$, we have $2m < 0$.

Therefore, for every $m \in \mathbb{Z}$, we have $2m \neq 5$. It follows that $2 \nmid 5$. We will see less painful ways to prove this later.

Notice that $a \mid 0$ for every $a \in \mathbb{Z}$ because $a \cdot 0 = 0$ for all $a \in \mathbb{Z}$. In particular, we have $0 \mid 0$ because as noted we have $0 \cdot 0 = 0$. Of course we also have $0 \cdot 3 = 0$ and in fact $0 \cdot m = 0$ for all $m \in \mathbb{Z}$, so every integer serves as a “witness” that $0 \mid 0$. Our definition says nothing about the $m \in \mathbb{Z}$ being unique.

Proposition 1.40. *Let $a, b, c \in \mathbb{Z}$. If $a \mid b$ and $b \mid c$, then $a \mid c$.*

Proof. Suppose that $a, b, c \in \mathbb{Z}$ are such that $a \mid b$ and $b \mid c$. Since $a \mid b$, we may fix $m \in \mathbb{Z}$ with $b = am$. Since $b \mid c$, we may fix $n \in \mathbb{Z}$ with $c = bn$. We then have

$$c = bn = (am)n = a(mn)$$

Since $mn \in \mathbb{Z}$, it follows that $a \mid c$. □

Proposition 1.41. *Let $a, b, c \in \mathbb{Z}$.*

1. *If $a \mid b$, then $a \mid bk$ for all $k \in \mathbb{Z}$.*
2. *If $a \mid b$ and $a \mid c$, then $a \mid (b + c)$.*
3. *If $a \mid b$ and $a \mid c$, then $a \mid (bk + c\ell)$ for all $k, \ell \in \mathbb{Z}$.*

Proof.

1. Suppose that $a \mid b$. Let $k \in \mathbb{Z}$ be arbitrary. Since $a \mid b$, we may fix $m \in \mathbb{Z}$ with $b = am$. We then have

$$bk = (am)k = a(mk)$$

Since $mk \in \mathbb{Z}$, it follows that $a \mid bk$. Since $k \in \mathbb{Z}$ was arbitrary, the result follows.

2. Suppose that $a \mid b$ and $a \mid c$. Since $a \mid b$, we may fix $m \in \mathbb{Z}$ with $b = am$. Since $a \mid c$, we may fix $n \in \mathbb{Z}$ with $c = an$. We then have
$$b + c = am + an = a(m + n)$$
Since $m + n \in \mathbb{Z}$, it follows that $a \mid b + c$.

3. This follows by combining 1 and 2 as follows. Suppose that $a \mid b$ and $a \mid c$. Let $m, n \in \mathbb{Z}$ be arbitrary. Since $a \mid b$, we conclude from part 1 that $a \mid bm$. Since $a \mid c$, we conclude from part 1 again that $a \mid cn$. Using part 2, it follows that $a \mid (bm + cn)$. Since $m, n \in \mathbb{Z}$ were arbitrary, the result follows. Alternatively, you should try to prove this directly without using the first two parts.

□

Proposition 1.42. *Suppose that $a, b \in \mathbb{Z}$. If $a \mid b$ and $b \neq 0$, then $|a| \leq |b|$.*

Proof. Suppose that $a \mid b$ and $b \neq 0$. Fix $d \in \mathbb{Z}$ with $ad = b$. Since $b \neq 0$, we have $d \neq 0$. Thus, $|d| \geq 1$, and so
$$|b| = |ad| = |a| \cdot |d| \geq |a| \cdot 1 = |a|$$
The result follows. □

Corollary 1.43. *Suppose that $a, b \in \mathbb{Z}$. If $a \mid b$ and $b \mid a$, then either $a = b$ or $a = -b$.*

Proof. Suppose first that $a \neq 0$ and $b \neq 0$. By the previous Proposition, we know that both $|a| \leq |b|$ and $|b| \leq |a|$. It follows that $|a| = |b|$, and hence either $a = b$ or $a = -b$.

Suppose now that $a = 0$. As above, since $a \mid b$, we may fix $m \in \mathbb{Z}$ with $b = am$. We then have $b = am = 0m = 0$ as well. Therefore, $a = b$.

Suppose fiinally that $b = 0$. Since $b \mid a$, we may fix $m \in \mathbb{Z}$ with $a = bm$. We then have $a = bm = 0m = 0$ as well. Therefore, $a = b$. □

Definition 1.44. *Let $a \in \mathbb{Z}$.*

- *We say that a is* even *if $2 \mid a$, i.e. if there exists $m \in \mathbb{Z}$ with $a = 2m$.*
- *We say that a is* odd *if there exists $m \in \mathbb{Z}$ with $a = 2m + 1$.*

Proposition 1.45. *No integer is both even and odd.*

Proof. Let $a \in \mathbb{Z}$. Suppose that a is both even and odd. Since a is even, we may fix $m \in \mathbb{Z}$ with $a = 2m$. Since a is odd, we may fix $n \in \mathbb{Z}$ with $a = 2n + 1$. We then have $2m = 2n + 1$, so $2(m - n) = 1$. Since $m - n \in \mathbb{Z}$, we conclude that $2 \mid 1$, which contradicts Proposition 1.42. This contradiction implies that a is not both even and odd. Since $a \in \mathbb{Z}$ was arbitrary, the result follows. □

We prove the following "obvious" fact below in much more generality. For now, we will simply assume it.

Proposition 1.46. *Every integer is either even or odd.*

Proposition 1.47. *Let $a \in \mathbb{Z}$. If a^2 is even, then a is even.*

Proof. We prove the contrapositive. That is, we show that whenever a is odd, then a^2 is odd. Suppose that a is odd. Fix $m \in \mathbb{Z}$ with $a = 2m + 1$. We then have
$$\begin{aligned} a^2 &= (2m + 1)^2 \\ &= 4m^2 + 4m + 1 \\ &= 2(2m^2 + 2m) + 1 \end{aligned}$$
so a^2 is odd. We have shown that if a is odd, then a^2 is odd. Therefore, if a^2 is even, then a is even. □

Theorem 1.48. $\sqrt{2}$ *is irrational.*

Proof. Suppose for the sake of obtaining a contradiction that $\sqrt{2}$ is irrational. Fix $a, b \in \mathbb{Z}$ with

$$\sqrt{2} = \frac{a}{b}$$

where $\frac{a}{b}$ is in lowest terms, i.e. a and b have no common divisor (this seems obviously possible, but we will formally justify it later). Squaring both sides, we then have

$$2 = \frac{a^2}{b^2}$$

so

$$2b^2 = a^2$$

Since $b^2 \in \mathbb{Z}$, we conclude that $2 \mid a^2$. Using the previous Proposition, it follows that a is even. Fix $m \in \mathbb{Z}$ with $a = 2m$. We then have

$$2b^2 = (2m)^2 = 4m^2.$$

Dividing each side by 2, we conclude that

$$b^2 = 2m^2.$$

Since $m^2 \in \mathbb{Z}$, it follows that $2 \mid b^2$. Using the previous Proposition again, we conclude that b is even. We have shown that both a and b are even, i.e. that both $2 \mid a$ and $2 \mid b$. This is a contradiction because a and b were assumed to have no common factors. Therefore, $\sqrt{2}$ is irrational. □

Proposition 1.49. *If $a \in \mathbb{Z}$ is odd, then a is the difference of two perfect squares, i.e. there exist $b, c \in \mathbb{Z}$ with $a = b^2 - c^2$.*

Proof. Let $a \in \mathbb{Z}$ be odd. Fix $m \in \mathbb{Z}$ with $a = 2m + 1$. Let $b = m + 1$ and $c = m$. Notice that $b, c \in \mathbb{Z}$ and that

$$\begin{aligned} b^2 - c^2 &= (m+1)^2 - m^2 \\ &= m^2 + 2m + 1 - m^2 \\ &= 2m + 1 \\ &= a \end{aligned}$$

Therefore, we shown the existence of b and c (namely $b = m + 1$ and $c = m$) for which $a = b^2 - c^2$. Since $a \in \mathbb{Z}$ was an arbitrary odd number, we conclude that every odd integer is the difference of two perfect squares. □

2 Fundamental Proof Techniques

2.1 Induction

Suppose that we want to prove that a certain statement is true for all natural numbers. In other words, we want to do the following:

- Prove that the statement is true for 0.
- Prove that the statement is true for 1.
- Prove that the statement is true for 2.

- Prove that the statement is true for 3.
- ....

Of course, since there are infinitely many natural numbers, going through each one in turn does not work because we will never handle them all this way. How can we get around this? Suppose that when we examine the first few proofs above that they look the same except that we replace 0 by 1 everywhere, or 0 by 2 everywhere, etc. In this case, one is tempted to say that "the pattern continues" or something similar, but that is not convincing because we can't be sure that the pattern does not break down when we reach 5413. However, one way to see that the "the pattern continues" and handle all of the infinitely many possibilities at once is to take an arbitrary natural number n, and prove that the statement is true for k using *only* the fact that n is a natural number (but *not* any particular natural number).

This method of taking an arbitrary $n \in \mathbb{N}$ and proving that the statement is true for n is that standard way of proving a statement involving a "for all" quantifier. This technique also works to prove that a statement is true for all real numbers or for all matrices, as long as we take an *arbitrary* such object. However, there is a different method one can use to prove that every natural number has a certain property, and this one does not carry over to other situations such as the real numbers. The key fact is that the natural numbers start with 0 and proceed in discrete steps forward. Consider what would if we can prove each of the following:

- Prove that the statement is true for 0.
- Prove that if the statement is true for 0, then the statement is true for 1.
- Prove that if the statement is true for 1, then the statement is true for 2.
- Prove that if the statement is true for 2, then the statement is true for 3.
- ....

Suppose that we are successful in doing this. From the first line, we then know that the statement is true for 0. Since we now know that it's true for 0, we can use the second line to conclude that the statement is true for 1. Since we now know that it's true for 1, we can use the second line to conclude that the statement is true for 2. And so on. In the end, we are able to conclude that the statement is true for all natural numbers.

Let's examine this situation more closely. On the fact of it, each line looks more complicated than the corresponding lines for proving a theorem directly. However, the key fact is that from the second line onward, we now have an additional assumption! Thus, instead of proving that the statement is true for 3 without any help, we can now use the assumption that the statement is true for 2 in that argument. Extra assumptions are always welcome because we have more that we can use in the actual argument.

Of course, as in our discussion at the beginning of this section, we can't hope to prove each of these infinitely many things one at a time. In an ideal world, the arguments from the second line onward all look exactly the same with the exception of replacing the number involved. Thus, the idea is to prove the following.

- Prove that the statement is true for 0.
- Prove that if the statement is true for n, then the statement is true for $n + 1$.

Notice that for the second line, we would need to prove that it is true for an arbitrary $n \in \mathbb{N}$, just like we would have to in a direct argument. An argument using these method is called a proof by (mathematical) *induction*, and it is an extremely useful and common technique in combinatorics. We now state this approach formally in terms of sets, which allows us to bypass the vague notion of "statement" that we used above.

Fact 2.1 (Principle of Mathematical Induction on $\mathbb{N}$)**.** *Let $X \subseteq \mathbb{N}$. Suppose that the following are true:*

- $0 \in X$ *(the* base case*)*

- *$n+1 \in X$ whenever $n \in X$ (the* inductive step*)*

We then have that $X = \mathbb{N}$.

Once again, here's the intuitive argument for why induction is valid. By the first assumption, we know that $0 \in X$. Since $0 \in X$, the second assumption tells us that $1 \in X$. Since $1 \in X$, the second assumption again tells us that $2 \in X$. By repeatedly applying the second assumption in this manner, each element of $\mathbb{N}$ is eventually determined to be in X. Notice that a similar argument works if we start with a different base case, i.e. if we start by proving that $3 \in X$ and then prove the inductive step, then it follows that $n \in X$ for all $n \in \mathbb{N}$ with $n \geq 3$.

We now give many examples of proofs by induction.

Proposition 2.2. *For any $n \in \mathbb{N}^+$, we have*

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}$$

We give two proofs. The first is a clever argument that avoids induction, while the second is a typical application of induction.

Proof 1. We first give a proof with induction. Let $n \in \mathbb{N}^+$ be arbitrary. Let $S = 1 + 2 + \cdots + (n-1) + n$. We also have $S = n + (n-1) + \cdots + 2 + 1$. Adding both of these we conclude that

$$2S = (n+1) + (n+1) + \cdots + (n+1) + (n+1)$$

and hence

$$2S = n(n+1).$$

Dividing both sides by 2, we conclude that

$$S = \frac{n(n+1)}{2}$$

so $1 + 2 + \cdots + (n-1) + n = \frac{n(n+1)}{2}$. Since $n \in \mathbb{N}^+$ was arbitrary, the result follows. □

Proof 2. We now give a proof using induction.

- *Base Case:* For $n = 1$, the statement is true because $\frac{1 \cdot 2}{2} = 1$.
- *Inductive Step:* Assume that the statement is true for some fixed $n \in \mathbb{N}^+$, i.e. suppose that n is a number for which we know that

  $$1 + 2 + \cdots + n = \frac{n(n+1)}{2}.$$

  We then have

  $$\begin{aligned} 1 + 2 + \cdots + n + (n+1) &= \frac{n(n+1)}{2} + (n+1) && \text{(by the inductive hypothesis)} \\ &= \frac{n^2 + n + 2n + 2}{2} \\ &= \frac{n^2 + 3n + 2}{2} \\ &= \frac{(n+1)(n+2)}{2} \\ &= \frac{(n+1)((n+1)+1)}{2}. \end{aligned}$$

  Thus, the statement is true for $n+1$.

By induction, we conclude that

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}$$

for all $n \in \mathbb{N}^+$. □

Theorem 2.3. *For any $n \in \mathbb{N}^+$, we have*

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

i.e.

$$1 + 3 + 5 + 7 + \cdots + (2n-1) = n^2$$

Proof. We give a proof by induction.

- *Base Case:* Suppose that $n = 1$. We have

$$\sum_{k=1}^{1} (2k-1) = 2 \cdot 1 - 1 = 1$$

  so the left hand-side is 1. The right-hand side is $1^2 = 1$. Thus, the statement is true when $n = 1$.

- *Inductive Step:* Assume that the statement is true for some fixed $n \in \mathbb{N}^+$, i.e. suppose that n is a number for which we know that

$$\sum_{k=1}^{n} (2k-1) = n^2.$$

  Notice that $2(n+1) - 1 = 2n + 2 - 1 = 2n + 1$, hence

$$\begin{aligned}
\sum_{k=1}^{n+1} (2k-1) &= [\sum_{k=1}^{n} (2k-1)] + [2(n+1) - 1] \\
&= [\sum_{k=1}^{n} (2k-1)] + (2n+1) \\
&= n^2 + (2n+1) && \text{(by induction)} \\
&= (n+1)^2
\end{aligned}$$

  Thus, the statement is true for $n + 1$.

By induction, we conclude that

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

for all $n \in \mathbb{N}^+$. □

Proposition 2.4. *For all $n \in \mathbb{N}$, we have $3 \mid (4^n - 1)$.*

Proof. We give a proof by induction.

- *Base Case:* Suppose that $n = 0$. We have $4^0 - 1 = 1 - 1 = 0$, hence $3 \mid (4^0 - 1)$ because $3 \cdot 0 = 0$. Thus, the statement is true when $n = 0$.

- *Inductive Step:* Assume that the statement is true for some fixed $n \in \mathbb{N}^+$, i.e. suppose that n is a number for which we know that $3 \mid (4^n - 1)$. Fix $k \in \mathbb{Z}$ with $3k = 4^n - 1$. We then have

$$\begin{aligned} 4^{n+1} - 1 &= 4 \cdot 4^n - 1 \\ &= 4 \cdot (3k+1) - 1 \\ &= 12k - 3 \\ &= 3 \cdot (4k - 1) \end{aligned}$$

  Since $4k - 1 \in \mathbb{Z}$, we conclude that $3 \mid (4^{n+1} - 1)$. Thus, the statement is true for $n + 1$.

By induction, we conclude that $3 \mid (4^n - 1)$ for all $n \in \mathbb{N}$. □

Proposition 2.5. *We have $2n + 1 < n^2$ for all $n \in \mathbb{N}$ with $n \geq 3$.*

Proof. We give a proof by induction.

- *Base Case:* Suppose that $n = 3$. We have $2 \cdot 3 + 1 = 7$ and $3^2 = 9$, so $2 \cdot 3 + 1 < 3^2$. Thus, the statement is true when $n = 3$.

- *Inductive Step:* Assume that the statement is true for some fixed $n \in \mathbb{N}$ with $n \geq 3$, i.e. suppose that $n \geq 3$ is a number for which we know that $2n + 1 < n^2$. Since $2n + 1 \geq 2 \cdot 3 + 1 = 7 > 2$, we then have

$$\begin{aligned} 2(n+1) + 1 &= 2n + 3 \\ &= (2n+1) + 2 \\ &= n^2 + 2 \\ &< n^2 + 2n + 1 \\ &= (n+1)^2 \end{aligned}$$

  Thus, the statement is true for $n + 1$.

By induction, we conclude that $2n + 1 < n^2$ for all $n \in \mathbb{N}$ with $n \geq 3$. □

Proposition 2.6. *We have $n^2 < 2^n$ for all $n \geq 5$.*

Proof. We give a proof by induction.

- *Base Case:* Suppose that $n = 5$. We have $5^2 = 25$ and $2^5 = 32$, so $5^2 < 2^5$. Thus, the statement is true when $n = 5$.

- *Inductive Step:* Assume that the statement is true for some fixed $n \in \mathbb{N}$ with $n \geq 5$, i.e. suppose that $n \geq 5$ is a number for which we know that $n^2 < 2^n$. Since $n^2 = n \cdot n \geq 3n = 2n + n > 2n + 1$, we have then have

$$\begin{aligned} (n+1)^2 &= n^2 + 2n + 1 \\ &< n^2 + n^2 \\ &= 2n^2 \\ &< 2 \cdot 2^n \\ &= 2^{n+1} \end{aligned}$$

  Thus, the statement is true for $n + 1$.

By induction, we conclude that $n^2 < 2^n$ for all $n \geq 5$. □

Theorem 2.7. *For all $x \in \mathbb{R}$ with $x \geq -1$ and all $n \in \mathbb{N}^+$, we hve $(1+x)^n \geq 1 + nx$.*

Proof. On the face of it, this looks a little different because it we are also quantifying over infinitely many real numbers x. Since x is coming from $\mathbb{R}$, we can't induct on x. However, we *can* take an arbitrary $x \in \mathbb{R}$ with $x \geq -1$, and then induct on n for this particular x. We now carry out that argument.

Let $x \in \mathbb{R}$ with $x \geq -1$. For this x, we show that $(1+x)^n \geq 1 + nx$ for all $n \in \mathbb{N}^+$ by induction.

- *Base Case:* Suppose that $n = 1$. We then have that $(1+x)^1 = 1 + x = 1 + 1x$, so certainly $(1+x)^1 \geq 1 + 1x$.

- *Inductive Step:* Assume that the statement is true for some fixed $n \in \mathbb{N}^+$, i.e. suppose that n is a number for which we know that $(1+x)^n \geq 1 + nx$. Since $x \geq -1$, we have $1 + x \geq 0$, so we can multiply both sides of this inequality by $(1+x)$ to conclude that

$$(1+x)^n \cdot (1+x) \geq (1+nx) \cdot (1+x)$$

We then have

$$\begin{aligned}
(1+x)^{n+1} &= (1+x)^n \cdot (1+x) \\
&\geq (1+nx) \cdot (1+x) && \text{(from above)} \\
&= 1 + nx + x + nx^2 \\
&= 1 + (n+1)x + nx^2 \\
&\geq 1 + (n+1)x. && \text{(since } nx^2 \geq 0\text{)}
\end{aligned}$$

Hence, we have shown that $(1+x)^{n+1} \geq 1 + (n+1)x$, i.e. that the statement is true for $n+1$.

By induction, we conclude that $(1+x)^n \geq 1 + nx$ for all $n \in \mathbb{N}^+$. Since $x \in \mathbb{R}$ with $x \geq -1$ was arbitrary, the statement follows. □

Proposition 2.8. *For all $n \in \mathbb{N}^+$, we have*

$$\sum_{k=1}^{n} \frac{1}{k^2} \leq 2 - \frac{1}{n}.$$

Proof. We prove the statement by induction.

- *Base Case:* Suppose that $n = 1$. In this case, we have

$$\sum_{k=1}^{1} \frac{1}{k^2} = \frac{1}{1^2} = 1$$

and

$$2 - \frac{1}{1} = 2 - 1 = 1$$

hence

$$\sum_{k=1}^{1} \frac{1}{k^2} \leq 2 - \frac{1}{1}$$

- *Inductive Step:* Assume that the statement is true for some fixed $n \in \mathbb{N}^+$, i.e. suppose that n is a number for which we know that

$$\sum_{k=1}^{n} \frac{1}{k^2} \leq 2 - \frac{1}{n}.$$

We then have

$$\begin{aligned}\sum_{k=1}^{n+1}\frac{1}{k^2} &= \left(\sum_{k=1}^{n}\frac{1}{k^2}\right)+\frac{1}{(n+1)^2}\\ &\leq 2-\frac{1}{n}+\frac{1}{(n+1)^2}\\ &= 2-\left(\frac{1}{n}-\frac{1}{(n+1)^2}\right)\\ &= 2-\frac{(n+1)^2-n}{n(n+1)^2}\\ &= 2-\frac{n^2+n+1}{n(n+1)^2}\\ &\leq 2-\frac{n^2+n}{n(n+1)^2}\\ &= 2-\frac{n(n+1)}{n(n+1)^2}\\ &= 2-\frac{1}{n+1}.\end{aligned}$$

Thus, the statement is true for $n+1$.

By induction, we conclude we conclude that

$$\sum_{k=1}^{n}\frac{1}{k^2}\leq 2-\frac{1}{n}$$

for all $n \in \mathbb{N}^+$. □

Theorem 2.9. *Let $a, b \in \mathbb{N}$ with $b \neq 0$. There exist unique $q, r \in \mathbb{N}$ such that $a = qb + r$ and $0 \leq r < b$. Uniqueness here means that if $a = q_1 b + r_1$ with $0 \leq r_1 < b$ and $a = q_2 b + r_2$ with $0 \leq r_2 < b$, then $q_1 = q_2$ and $r_1 = r_2$.*

Proof. We first prove existence. Since we want to prove something for all $a, b \in \mathbb{N}$, it might at first seem unclear how to apply induction to both a and b. The answer in this case is not to induct on both, but to fix b and induct on a. In other words, let $b \in \mathbb{N}$ with $b > 0$ be arbitrary, and for this fixed b, we prove the existence of q, r for all $a \in \mathbb{N}$ by induction on a. That is, for this fixed b, we define

$$X = \{a \in \mathbb{N} : \text{ There exist } q, r \in \mathbb{N} \text{ with } a = qb + r\}$$

and show that $X = \mathbb{N}$ by induction.

- *Base Case:* Suppose that $a = 0$. We then have $a = 0 \cdot b + 0$ and since $0 < b$, we may take $q = 0$ and $r = 0$.

- *Inductive Step:* Assume that the statement is true for some fixed $a \in \mathbb{N}$. Fix $q, r \in \mathbb{Z}$ with $0 \leq r < b$ such that $a = qb + r$. We then have $a + 1 = qb + (r + 1)$. Since $r, b \in \mathbb{N}$ with $r < b$, we know that $r + 1 \leq b$. If $r + 1 < b$, then we are done. Otherwise, we have $r + 1 = b$, hence

$$\begin{aligned}a+1 &= qb+(r+1)\\ &= qb+b\\ &= (q+1)b\\ &= (q+1)b+0\end{aligned}$$

and so we may take $q + 1$ and 0. Thus, the statement is true for $a + 1$.

Therefore, the existence part of the theorem follows by induction.

We now prove uniqueness. Let $a, b \in \mathbb{N}$ with $b \neq 0$. Suppose that $q_1, q_2, r_1, r_2 \in \mathbb{N}$ are such that

$$q_1 b + r_1 = a = q_2 b + r_2$$

and both $0 \leq r_1 < b$ and $0 \leq r_2 < b$. We then have

$$b(q_2 - q_1) = r_1 - r_2$$

hence $b \mid (r_2 - r_1)$. Now $-b < -r_1 \leq 0$, so adding this to $0 \leq r_2 < b$, we conclude that

$$-b < r_2 - r_1 < b$$

and therefore

$$|r_2 - r_1| < b$$

Now if $r_2 - r_1 \neq 0$, then since $b \mid (r_2 - r_1)$, we would conclude from Proposition 1.42 that $|b| \leq |r_2 - r_1|$, a contradiction. It follows that $r_2 - r_1 = 0$, and hence $r_1 = r_2$. Since

$$q_1 b + r_1 = q_2 b + r_2$$

and $r_1 = r_2$, we conclude that $q_1 b = q_2 b$. Now $b \neq 0$, so it follows that $q_1 = q_2$. □

Proposition 2.10. *Let $a, b \in \mathbb{N}$ with $b \neq 0$. Write $a = qb + r$ for the unique choice of $q, r \in \mathbb{N}$ with $0 \leq r < b$. We then have that $b \mid a$ if and only if $r = 0$.*

Proof. If $r = 0$, then $a = qb + r = bq$, so $b \mid a$. Suppose conversely that $b \mid a$ and fix $m \in \mathbb{Z}$ with $a = bm$. Notice that since $a \geq 0$ and $b \geq 0$, we must have that $m \geq 0$. We then have that both $a = mb + 0$ and $a = qb + r$, so by the uniqueness part of the above theorem, we must have $r = 0$. □

2.2 Strong Induction

Remember our original model for induction:

- Prove that the statement is true for 0.
- Prove that if the statement is true for 0, then the statement is true for 1.
- Prove that if the statement is true for 1, then the statement is true for 2.
- Prove that if the statement is true for 2, then the statement is true for 3.
- Prove that if the statement is true for 3, then the statement is true for 4.
- ....

In the previous section, we argued why this model was sound and gave many examples. However, upon closer inspection, it appears that we can assume more. In the second line, when proving that the statement is true for 1 we are allowed to assume that the statement is true for 0. Now in the third line, when proving that the statement is true for 2, we only assume that it is true for 1. If we are knocking down the natural numbers in order, then we've already proved that it's true for 0, so why can't we assume that as well? The answer is that we can indeed assume it, and in general when working to prove that the statement is true for a natural number n, we can assume that we know it is true for all smaller values. In other words, we do the following:

- Prove that the statement is true for 0.
- Prove that if the statement is true for 0, then the statement is true for 1.
- Prove that if the statement is true for 0 and 1, then the statement is true for 2.
- Prove that if the statement is true for 0, 1, and 2, then the statement is true for 3.
- Prove that if the statement is true for 0, 1, 2, and 3, then the statement is true for 4.
- ....

Suppose that we are successful in doing this. From the first line, we then know that the statement is true for 0. Since we now know that it's true for 0, we can use the second line to conclude that the statement is true for 1. Since we now know that it's true for both 0 and 1, we can use the second line to conclude that the statement is true for 2. And so on. In the end, we are able to conclude that the statement is true for all natural numbers.

As usual, we can't hope to prove each of these infinitely many things one at a time. In an ideal world, the arguments from the second line onward all look exactly the same with the exception of replacing the number involved. Thus, the idea is to prove the following.

- Prove that the statement is true for 0.
- Prove that if the statement is true for each of $0, 1, 2, \dots, n$, then the statement is true for $n + 1$.

Alternatively, we can state this as follows:

- Prove that the statement is true for 0.
- Prove that if the statement is true for each of $0, 1, 2, \dots, n-1$, then the statement is true for n (for $n \geq 1$).

An argument using these method is called a proof by *strong induction*. As we will see in the examples below, sometimes we need to modify this clean structure to include several base cases to get the argument going. Rather than going through a theoretical discussion of how and why one would do this, it's easier to illustrate the technique by example.

Proposition 2.11. *Define a sequence a_n recursively by letting $a_0 = 0$, $a_1 = 1$, and*

$$a_n = 3a_{n-1} - 2a_{n-2}$$

for $n \geq 2$. Show that $a_n = 2^n - 1$ for all $n \in \mathbb{N}$.

Proof. We prove that $a_n = 2^n - 1$ for all $n \in \mathbb{N}$ by strong induction.

- *Base Case:* We handle two bases where $n = 0$ and $n = 1$ because our inductive step will use the result for two steps back. When $n = 0$, we have $a_0 = 0$ and $2^0 - 1 = 1 - 1 = 0$, so $a_0 = 2^0 - 1$. When $n = 1$, we have $a_1 = 1$ and $2^1 - 1 = 2 - 1 = 1$, so $a_1 = 2^1 - 1$.
- *Inductive Step:* Let $n \geq 2$ and assume that the statement is true for $0, 1, 2, \dots, n-1$, i.e. assume that $a_m = 2^m - 1$ for all $m \in \{0, 1, 2, \dots, n-1\}$. We prove that the statement is true for n. Notice that

since $n \geq 2$, we have $0 \leq n-1 < n$ and $0 \leq n-2 < n$, so we know that $a_{n-1} = 2^{n-1} - 1$ and $a_{n-2} = 2^{n-2} - 1$. Now

$$\begin{aligned}
a_n &= 3a_{n-1} - 2a_{n-1} && \text{(by definition since } n \geq 2\text{)} \\
&= 3 \cdot (2^{n-1} - 1) - 2 \cdot (2^{n-2} - 1) && \text{(by the inductive hypothesis)} \\
&= 3 \cdot 2^{n-1} - 3 - 2 \cdot 2^{n-2} + 2 \\
&= 3 \cdot 2^{n-1} - 2^{n-1} - 1 \\
&= (3 - 1) \cdot 2^{n-1} - 1 \\
&= 2 \cdot 2^{n-1} - 1 \\
&= 2^n - 1
\end{aligned}$$

Thus, $a_n = 2^n - 1$ and so the statement is true for n.

Using strong induction, we conclude that $a_n = 2^n - 1$ for all $n \in \mathbb{N}$. □

Proposition 2.12. *If $n \in \mathbb{N}$ and $n \geq 12$, then there exist $k, \ell \in \mathbb{N}$ with $n = 4k + 5\ell$.*

Proof. We give a proof by strong induction.

- *Base Case:* We first prove that the statement is true for $n \in \{12, 13, 14, 15\}$ (we will see why we need so many base cases below). We have
  - $12 = 4 \cdot 3 + 5 \cdot 0$
  - $13 = 4 \cdot 2 + 5 \cdot 1$
  - $14 = 4 \cdot 1 + 5 \cdot 2$
  - $15 = 4 \cdot 0 + 5 \cdot 3$

  Thus, the statement is true for $n \in \{12, 13, 14, 15\}$.

- *Inductive Step:* Let $n \geq 16$ and assume that the statement is true for $12, 13, 14, \dots, n-1$. We prove that the statement is true for n. Since $n \geq 16$, we have $12 \leq n - 4 < n$. Since $12 \leq n - 4 < 4$, we know that there exists $k, \ell \in \mathbb{N}$ with
  $$n - 4 = 4k + 5\ell.$$
  Adding 4 to both sides, we conclude that
  $$n = 4k + 5\ell + 4 = 4(k+1) + 5\ell$$
  Since $k + 1, \ell \in \mathbb{N}$, we conclude that the statement is true for n.

By (strong) induction, we conclude that for all $n \in \mathbb{N}$ with $n \geq 12$, there exist $k, \ell \in \mathbb{N}$ with $n = 4k + 5\ell$. □

Theorem 2.13. *Let $b \in \mathbb{N}$ with $b \geq 2$. For all $n \in \mathbb{N}^+$, there exists $a_i \in \mathbb{N}$ with $0 \leq a_i < b$ such that*

$$n = a_k b^k + a_{k-1} b^{k-1} + \cdots + a_1 b + a_0$$

Proof. Let $b \in \mathbb{N}$ with $b \geq 2$ be arbitrary. With this fixed b, we prove the result by strong induction on n.

- *Base Case:* Let $n = 1$. We may take $k = 1$ and $a_0 = 1 < b$.

- *Inductive Step:* Let $n \geq 2$ and assume that the statement is true for $1, 2, \dots, n-1$. Fix $q, r \in \mathbb{N}$ with $n = qb + r$ and $0 \leq r < b$. Notice that $q < n$ because $q \geq n$ would imply that

  $$n = qb + r \geq qb \geq nb \geq 2n > n$$

  a contradiction. Therefore, since $0 \leq q < n$, we may use strong induction to conclude that we can fix $a_i \in \mathbb{N}$ with $0 \leq a_i < b$ such that

  $$q = a_k b^k + a_{k-1} b^{k-1} + \cdots + a_1 b + a_0.$$

  We then have

  $$\begin{aligned} n &= qb + r \\ &= (a_k b^k + a_{k-1} b^{k-1} + \cdots + a_1 b + a_0) b + r \\ &= a_k b^{k+1} + a_{k-1} b^k + \cdots + a_1 b^2 + a_0 b + r \end{aligned}$$

  Since $0 \leq r < b$, we have shown that the statement is true for n.

The result follows by induction. □

Definition 2.14. *Suppose that $a, b \in \mathbb{Z}$. We say that $d \in \mathbb{Z}$ is a* common divisor *of a and b if both $d \mid a$ and $d \mid b$.*

The common divisors of 120 and 84 are $\{\pm 1, \pm 2, \pm 3, \pm 4, \pm 6, \pm 12\}$ (we will see a careful argument below). The common divisors of 10 and 0 are $\{\pm 1, \pm 2, \pm 5, \pm 10\}$. Every element of $\mathbb{Z}$ is a common divisor of 0 and 0. The following little proposition is fundamental to this entire section.

Proposition 2.15. *Suppose that $a, b, q, r \in \mathbb{Z}$ and $a = qb + r$ (we need not have $0 \leq r < |b|$). For any $d \in \mathbb{Z}$, we have that d is a common divisor of a and b if and only if d is a common divisor of b and r, i.e.*

$$\{d \in \mathbb{Z} : d \textit{ is a common divisor of } a \textit{ and } b\} = \{d \in \mathbb{Z} : d \textit{ is a common divisor of } b \textit{ and } r\}.$$

Proof. Suppose first that d is a common divisor of b and r. Since $d \mid b$, $d \mid r$, and $a = qb + r = bq + r1$, we may use Proposition 1.41 to conclude that $d \mid a$.

Conversely, suppose that d is a common divisor of a and b. Since $d \mid a$, $d \mid b$, and $r = a - qb = a1 + b(-q)$, we may use Proposition 1.41 to conclude that $d \mid r$. □

For example, suppose that we are trying to find the set of common divisors of 120 and 84 (we wrote them above, but now want to justify it). We repeatedly do division to reduce the problem as follows:

$$\begin{aligned} 120 &= 1 \cdot 84 + 36 \\ 84 &= 2 \cdot 36 + 12 \\ 36 &= 3 \cdot 12 + 0 \end{aligned}$$

The first line tells us that the set of common divisors of 120 and 84 equals the set of common divisors of 84 and 36. The next line tells us that the set of common divisors of 84 and 36 equals the set of common divisors of 36 and 12. The last line tells us that the set of common divisors of 36 and 12 equals the set of common divisors of 12 and 0. Now the set of common divisors of 12 and 0 is simply the set of divisors of 12 (because every number divides 0). Putting it all together, we conclude that the set of common divisors of 120 and 84 equals the set of divisors of 12.

Definition 2.16. *Let $a, b \in \mathbb{Z}$. We say that an element $d \in \mathbb{Z}$ is a* greatest common divisor *of a and b if:*

- $d \geq 0$
- *d is a common divisor of a and b.*
- *Whenever $c \in \mathbb{Z}$ is a common divisor of a and b, we have $c \mid d$.*

Notice that we are *not* defining the greatest common divisor of a and b to be the largest divisor of a and b. The primary reason we do not is because this description fails to capture the most fundamental property (namely that of being divisible by all other divisors, not just larger than them). Furthermore, if we were to take that definition, then 0 and 0 would fail to have a greatest common divisor because every integer is a common divisor of 0 and 0. With this definition however, it is a straightforward matter to check that 0 satisfies the above three conditions.

Since we require more of a greatest common divisor than just picking the largest, we first need to check that they do indeed exist. The proof is an inductive formulation of the above method of calculation.

Theorem 2.17. *Every pair of integers $a, b \in \mathbb{Z}$ has a unique greatest common divisor.*

We first sketch the idea of the proof in the case where $a, b \in \mathbb{N}$. If $b = 0$, we are done because it is simple to verify that a is a greatest common divisor of a and 0. Suppose then that $b \neq 0$. Fix $q, r \in \mathbb{N}$ with $a = qb + r$ and $0 \leq r < b$. Now the idea is to assert inductively the existence of a greatest common divisor of b and r because this pair is "smaller" than the pair a and b. The only issue is how to make this intuitive idea of "smaller" precise. There are several ways to do this, but perhaps the most straightforward is to only induct on b. Thus, our base case handles all pairs of form $(a, 0)$. Next, we handle all pairs of the form $(a, 1)$ and in doing this we can use the fact the we know the result for all pairs of the form $(a', 0)$. Notice that we can we even change the value of the first coordinate here which is why we used a'. Then, we handle all pairs of the form $(a, 2)$ and in doing this we can use the fact that we know the result for all pairs of the form $(a', 0)$ and $(a', 1)$. We now begin the formal argument.

Proof. We begin by proving existence only in the special case where $a, b \in \mathbb{N}$. We use (strong) induction on b to prove the result. That is, we let

$$X = \{b \in \mathbb{N} : \text{For all } a \in \mathbb{N}, \text{ there exists a greatest common divisor of } a \text{ and } b\}$$

and prove that $X = \mathbb{N}$ by strong induction.

- *Base Case:* Suppose that $b = 0$. Let $a \in \mathbb{N}$ be arbitrary. We then have that the set of common divisors of a and b equals the set of divisors of a (because every integer divides 0), so a satisfies the requirement of a greatest common divisor of a and 0. Since $a \in \mathbb{N}$ was arbitrary, we showed that there exists a greatest common divisor of a and 0 for every $a \in \mathbb{N}$, hence $0 \in X$.

- *Inductive Step:* Suppose then that $b \in \mathbb{N}^+$ and we know the result for all smaller natural numbers. In other words, we are assuming that $c \in X$ whenever $0 \leq c < b$. We prove that $b \in X$. Let $a \in \mathbb{N}$ be arbitrary. From above, we may fix $q, r \in \mathbb{Z}$ with $a = qb + r$ and $0 \leq r < b$. Since $0 \leq r < b$, we know by strong induction that $r \in X$, hence b and r have a greatest common divisor d. By Proposition 2.15, the set of common divisors of a and b equals the set of common divisors of b and r. It follows that d is a greatest common divisor of a and b. Since $a \in \mathbb{N}$ was arbitrary, we showed that there exists a greatest common divisor of a and b for every $a \in \mathbb{N}$, hence $b \in X$.

Therefore, we have shown that $X = \mathbb{N}$, which implies that whenever $a, b \in \mathbb{N}$, there exists a greatest common divisor of a and b.

To turn the argument into a proof for all $a, b \in \mathbb{Z}$, we simply note the set of divisors of an element $m \in \mathbb{Z}$ equals the set of divisors of $-m$. So, for example, if $a < 0$ but $b \geq 0$ we can simply take a greatest common divisor of $-a$ and b (which exists since $-a, b \in \mathbb{N}$) and note that it will also be a greatest common divisor of a and b. A similar argument works if $a \geq 0$ and $b < 0$, or if both $a < 0$ and $b < 0$.

For uniqueness, suppose that c and d are both greatest common divisors of a and b. Since d is a greatest common divisor and c is a common divisor, we know by the last condition that $c \mid d$. Similarly, since c is a greatest common divisor and d is a common divisor, we know by the last condition that $d \mid c$. Therefore, either $c = d$ or $c = -d$. Using the first requirement that a greatest common divisor must be nonnegative, we must have $c = d$. □

Definition 2.18. *Let $a, b \in \mathbb{Z}$. We let $\gcd(a, b)$ be the unique greatest common divisor of a and b.*

For example we have $\gcd(120, 84) = 12$ and $\gcd(0, 0) = 0$. The following corollary is immediate from Proposition 2.15.

Corollary 2.19. *Suppose that $a, b, q, r \in \mathbb{Z}$ and $a = qb + r$. We have $\gcd(a, b) = \gcd(b, r)$.*

The method of using repeated division and this corollary to reduce the problem of calculating greatest common divisors is known as the *Euclidean Algorithm*. We saw it in action of above with 120 and 84. Here is another example where we are trying to compute $\gcd(525, 182)$. We have

$$\begin{aligned}
525 &= 2 \cdot 182 + 161 \\
182 &= 1 \cdot 161 + 21 \\
161 &= 7 \cdot 21 + 14 \\
21 &= 1 \cdot 14 + 7 \\
14 &= 2 \cdot 7 + 0
\end{aligned}$$

Therefore, $\gcd(525, 182) = \gcd(7, 0) = 7$.

Theorem 2.20. *For all $a, b \in \mathbb{Z}$, there exist $k, \ell \in \mathbb{Z}$ with $\gcd(a, b) = ka + \ell b$.*

Proof. We begin by proving existence in the special case where $a, b \in \mathbb{N}$. We use induction on b to prove the result. That is, we let

$$X = \{b \in \mathbb{N} : \text{For all } a \in \mathbb{N}, \text{ there exist } k, \ell \in \mathbb{Z} \text{ with } \gcd(a, b) = ka + \ell b\}$$

and prove that $X = \mathbb{N}$ by strong induction.

- *Base Case:* Suppose that $b = 0$. Let $a \in \mathbb{N}$ be arbitrary. We then have that

  $$\gcd(a, b) = \gcd(a, 0) = a$$

  Since $a = 1 \cdot a + 0 \cdot b$, so we may let $k = 1$ and $\ell = 0$. Since $a \in \mathbb{N}$ was arbitrary, we conclude that $0 \in X$.

- *Inductive Step:* Suppose then that $b \in \mathbb{N}^+$ and we know the result for all smaller nonnegative values. In other words, we are assuming that $c \in X$ whenever $0 \leq c < b$. We prove that $b \in X$. Let $a \in \mathbb{N}$ be arbitrary. From above, we may fix $q, r \in \mathbb{Z}$ with $a = qb + r$ and $0 \leq r < b$. We also know from above that $\gcd(a, b) = \gcd(b, r)$. Since $0 \leq r < b$, we know by strong induction that $r \in X$, hence there exist $k, \ell \in \mathbb{Z}$ with

  $$\gcd(b, r) = kb + \ell r$$

  Now $r = a - qb$, so

  $$\begin{aligned}
  \gcd(a, b) &= \gcd(b, r) \\
  &= kb + \ell r \\
  &= kb + \ell(a - qb) \\
  &= kb + \ell a - qb\ell \\
  &= \ell a + (k - q\ell)b
  \end{aligned}$$

  Since $a \in \mathbb{N}$ was arbitrary, we conclude that $b \in X$.

Therefore, we have shown that $X = \mathbb{N}$, which implies that whenever $a, b \in \mathbb{N}$, there exists $k, \ell \in \mathbb{Z}$ with $\gcd(a, b) = ka + \ell b$. □

Given $a, b \in \mathbb{Z}$, we can explicitly calculate $k, \ell \in \mathbb{Z}$ by "winding up" the work created from the Euclidean Algorithm. For example, we saw above that $\gcd(525, 182) = 7$ by calculating

$$\begin{aligned}
525 &= 2 \cdot 182 + 161 \\
182 &= 1 \cdot 161 + 21 \\
161 &= 7 \cdot 21 + 14 \\
21 &= 1 \cdot 14 + 7 \\
14 &= 2 \cdot 7 + 0
\end{aligned}$$

We now use these steps in reverse to calculate:

$$\begin{aligned}
7 &= 1 \cdot 7 + 0 \cdot 0 \\
&= 1 \cdot 7 + 0 \cdot (14 - 2 \cdot 7) \\
&= 0 \cdot 14 + 1 \cdot 7 \\
&= 0 \cdot 14 + 1 \cdot (21 - 1 \cdot 14) \\
&= 1 \cdot 21 + (-1) \cdot 14 \\
&= 1 \cdot 21 + (-1) \cdot (161 - 7 \cdot 21) \\
&= (-1) \cdot 161 + 8 \cdot 21 \\
&= (-1) \cdot 161 + 8 \cdot (182 - 1 \cdot 161) \\
&= 8 \cdot 182 + (-9) \cdot 161 \\
&= 8 \cdot 182 + (-9) \cdot (525 - 2 \cdot 182) \\
&= (-9) \cdot 525 + 26 \cdot 182
\end{aligned}$$

This wraps everything up perfectly, but it is easier to simply start at the fifth line.

We end this section with a useful result.

Definition 2.21. *Two elements $a, b \in \mathbb{Z}$ are* relatively prime *if* $\gcd(a, b) = 1$.

Proposition 2.22. *Let $a, b, c \in \mathbb{Z}$. If $a \mid bc$ and $\gcd(a, b) = 1$, then $a \mid c$.*

Proof. Since $a \mid bc$, we may fix $m \in \mathbb{Z}$ with $bc = am$. Since $\gcd(a, b) = 1$, we may fix $k, \ell \in \mathbb{Z}$ with $ak + b\ell = 1$. Multiplying this last equation through by c we conclude that $akc + b\ell c = c$, so

$$\begin{aligned}
c &= akc + \ell(bc) \\
&= akc + na\ell \\
&= a(kc + n\ell)
\end{aligned}$$

It follows that $a \mid c$. □

2.3 Using Functions to Count

The following fact is intuitively clear. If $f\colon A \to B$ is injective, then every element of B is hit by at most one element of A, so the there must be at least as many elements in B as their are in A. The others can be argued similarly. One can do a formal proof by induction on the cardinalities of A and B, but just as for the Sum Rule we will avoid being so formal.

Fact 2.23. *Let A and B be finite sets and let $f\colon A \to B$ be a function.*

- *If f is injective, then $|A| \leq |B|$.*
- *If f is surjective, then $|B| \leq |A|$.*
- *If f is bijective, then $|A| = |B|$.*

The last of these is often very helpful when to trying to determine the cardinality of a set, and is sometimes called the "Bijection Principle". In fact, we've already used this type of argument informally. Recall on the homework that if A is a set with $|A| = n$ and $D = \{(a,a) : a \in A\}$, then we asserted that $|D| = n$. Formally, one can argue this by showing that there is a bijection between A and D. Namely, define $f\colon A \to A^2$ by letting $f(a) = (a,a)$. We then have f is injective and range$(f) = D$, so we get a bijection $f\colon A \to D$. It follows that $|A| = |D|$, and hence $|D| = n$.

In general, if we have a set A and want to know $|A|$, then the idea is to build a set B and a bijection $f\colon A \to B$ where $|B|$ is much easier to determine. The most fundamental example of this is the following:

Proposition 2.24. *Given a set A with $|A| = n$, we have $|\mathcal{P}(A)| = |\{0,1\}^n|$.*

Proof. Let $A = \{a_1, a_2, \dots, a_n\}$ where the a_i are distinct. Define a function $f\colon \{0,1\}^n \to \mathcal{P}(A)$ by letting $f(b_1, b_2, \dots, b_n) = \{a_i : b_i = 1\}$. In other words, given a finite sequence $(b_1, b_2, \dots, b_n)$ of 0's and 1's, we send it to the subset of A obtained by including a_i precisely when the i^{th} element of the sequence is a 1. Notice that if $(b_1, b_2, \dots, b_n) \neq (c_1, c_2, \dots, c_n)$, then we can fix an i with $b_i \neq c_i$, and in this case we have $f(b_1, b_2, \dots, b_n) \neq f(c_1, c_2, \dots, c_n)$ because a_i is one of the sets but not the other. Furthermore, given any $S \subseteq A$, if we let $(b_1, b_2, \dots, b_n) \in \{0,1\}^n$ be defined by letting

$$b_i = \begin{cases} 1 & \text{if } a_i \in S \\ 0 & \text{if } a_i \notin S \end{cases}$$

then $f(b_1, b_2, \dots, b_n) = S$, so f is surjective. Therefore, f is a bijection, and hence $|\{0,1\}^n| = |\mathcal{P}(A)|$. □

Corollary 2.25. *If $|A| = n \in \mathbb{N}^+$, then $|\mathcal{P}(A)| = 2^n$.*

Proof. This is immediate from the bijection principle and Corollary 1.20. □

Since this result is so fundamental, we give another proof that uses both induction and the bijection principle.

Proof 2 of Corollary 2.25. We prove the result by induction on $n \in \mathbb{N}^+$.

- *Base Case:* Suppose that $n = 1$. Let A be a set with $|A| = 1$, say $A = \{a\}$. We then have that $\mathcal{P}(A) = \{\emptyset, \{a\}\}$, so $|\mathcal{P}(A)| = 2 = 2^1$.
- *Induction Step:* Assume that the statement is true for some fixed $n \in \mathbb{N}^+$, i.e. assume that for some fixed $n \in \mathbb{N}^+$, we know that $|\mathcal{P}(A)| = 2^n$ for all sets A with $|A| = n$. Consider an arbitrary set A with $|A| = n + 1$. Fix some (any) element $a_0 \in A$. Let $\mathcal{S} \subseteq \mathcal{P}(A)$ be the collection of subsets of A not having a_0 as an element, and let $\mathcal{T} \subseteq \mathcal{P}(A)$ be the collection of subsets of A having a_0 as an element. Notice then that $\mathcal{S}$ and $\mathcal{T}$ are disjoint sets with $\mathcal{P}(A) = \mathcal{S} \cup \mathcal{T}$, so by the Sum Rule we know that

  $$|\mathcal{P}(A)| = |\mathcal{S}| + |\mathcal{T}|.$$

  Now consider the function $f\colon \mathcal{S} \to \mathcal{T}$ defined by letting $f(B) = B \cup \{a_0\}$, i.e. given $B \in \mathcal{S}$, we have that B is a subset of A not having a_0 as an element, and we send to the subset of A obtained by throwing a_0 in as a new element. Notice that f is a bijection, so $|\mathcal{S}| = |\mathcal{T}|$. Therefore, we have

  $$|\mathcal{P}(A)| = |\mathcal{S}| + |\mathcal{S}|.$$

Finally, notice that $\mathcal{S} = \mathcal{P}(A\backslash\{a_0\})$, so since $|A\backslash\{a_0\}| = n$, we can use induction to conclude that $|A\backslash\{a_0\}| = 2^n$. Therefore,

$$|\mathcal{P}(A)| = 2^n + 2^n = 2 \cdot 2^n = 2^{n+1}.$$

Thus, the statement is true for $n + 1$.

By induction, we conclude that if $|A| = n \in \mathbb{N}^+$, then $|\mathcal{P}(A)| = 2^n$. □

By the way, Corollary 2.25 is also true in the case $n = 0$. When $n = 0$, we have $A = \emptyset$ and $\mathcal{P}(\emptyset) = \{\emptyset\}$, so $|\mathcal{P}(\emptyset) = 1 = 2^0$.

Proposition 2.26. *Let A be a set with $|A| = n \in \mathbb{N}^+$ and let $k \in \mathbb{N}$ be such that $0 \leq k \leq n$. The number of subsets of A having cardinality k equals the number of subsets of A having cardinality $n - k$.*

Proof. Let $\mathcal{S}$ be the collection of all subsets of A having cardinality k, and let $\mathcal{T}$ be the collection of all subsets of A having cardinality $n - k$. Define $f \colon \mathcal{S} \to \mathcal{T}$ by letting $f(B) = A\backslash B$, i.e. given $B \subseteq A$ with $|B| = k$, send it to the complement of B in A (notice that if $|B| = k$, then $|A\backslash B| = n - k$ by the complement rule). Notice that f is a bijection (it is surjective because if $C \subseteq A$ is such that $|C| = n - k$, then $|A\backslash C| = k$ and $f(A\backslash C) = C$). Therefore, $|\mathcal{S}| = |\mathcal{T}|$. □

Thus, despite the fact that we do not (yet) have a formula for the number of subsets of a certain size, we know that the number of subsets of size k must equal the number of subsets of size $n - k$ even without this knowledge.

2.4 The Pigeonhole Principle

We know that if A and B are finite sets and $f \colon A \to B$ is an injective function, then $|A| \leq |B|$. Taking the contrapositive of this fact, we obtain the following.

Corollary 2.27 (Pigeonhole Principle)**.** *If A and B are finite sets with $|A| > |B|$, and $f \colon A \to B$ is a function, then there exist $a_1, a_2 \in A$ with $a_1 \neq a_2$ such that $f(a_1) = f(a_2)$. Informally, if $n > k$ and we have placed n balls into k boxes, then (at least) one box will contain at least 2 balls.*

For a very simple example, in any group of 13 people, there must exist (at least) 2 people in the group who were born in the same month.

Proposition 2.28. *Given $n + 1$ integers, it is always possible to find two whose difference is divisible by n.*

Proof. Let A be a set of $n + 1$ integers, so $A = \{a_0, a_1, \dots, a_n\}$. For each i, write

$$a_i = nq_i + r_i$$

where $0 \leq r_i < n$, so $r_i \in \{0, 1, 2, \dots, n - 1\}$. Define $f \colon A \to \{0, 1, 2, \dots, n - 1\}$ by letting $f(a_i) = r_i$ for each i. Since $|A| = n + 1$ and $|\{0, 1, 2, \dots, n - 1\}| = n$, we know by the Pigeonhole Principle that f is not injective. Fix $i \neq j$ with $r_i = r_j$. We then have

$$\begin{aligned} a_i - a_j &= (nq_i + r_i) - (nq_j + r_j) \\ &= n(q_i - q_j) + (r_i - r_j) \\ &= n(q_i - q_j) && \text{(since } r_i - r_j = 0\text{)} \end{aligned}$$

so $n \mid (a_i - a_j)$. □

Proposition 2.29. *Let $a_n = 777 \cdots 7$ where there are n many 7's. There exists an $n \leq 2014$ such that $2013 \mid a_n$.*

Proof. For each n with $1 \leq n \leq 2014$, write

$$a_i = 2011q_i + r_i$$

where $0 \leq r_i < 2013$. Since we have 2013 many possible distinct r_i, it follows that there exists distinct $i < j$ with $r_i = r_j$. We then have

$$2013 \mid (a_j - a_i)$$

as above. The problem is that $a_j - a_i$ does not equal any of the a_n. However

$$a_j - a_i = 777 \cdots 700 \cdots 0 = a_{j-i} \cdot 10^i$$

since $\gcd(2013, 10^i) = 1$ (because the only prime divisors of 10^i are 2 and 5, but these do not divide 2013), we can use Proposition 2.22 to conclude that $2013 \mid a_{j-i}$. □

Proposition 2.30. *Suppose we have a gathering of $n \geq 2$ people, and at the beginning of the gathering some pairs of people shake hands. There always must exist (at least) two people who have shaken the same number of hands.*

Proof. Label the people with the numbers $1, 2, 3, \dots, n$. We can then define a function $f \colon \{1, 2, 3, \dots, n\} \to \{0, 1, 2, \dots, n-1\}$ by letting $f(k)$ is the number of people that person k shook hands with. On the face of it, this looks bad because both sets have n elements. However, it is impossible that both 0 and $n-1$ are elements of range(f) because if somebody shook hands with all of the other $n-1$ people, then everybody shook hands with a least one person, so $0 \notin \text{range}(f)$. Thus, we can either view f as a function $f \colon \{1, 2, 3, \dots, n\} \to \{0, 1, 2, \dots, n-2\}$ or as a function $f \colon \{1, 2, 3, \dots, n\} \to \{1, 2, \dots, n-1\}$. In either case, f is not injective by the Pigeonhole Principle, so there exist two people who have shaken the same number of hands. □

Proposition 2.31. *Let $f \colon \{0,1\}^* \to \{0,1\}^*$ be injective. For every $n \in \mathbb{N}^+$, there exists $\sigma \in \{0,1\}^n$ with $|f(\sigma)| \geq |\sigma|$ (here $|\tau|$ is the length of the finite sequence τ).*

Proof. Let $f \colon \{0,1\}^* \to \{0,1\}^*$ be injective. Let $n \in \mathbb{N}^+$ be arbitrary. Suppose instead that $|f(\sigma)| < |\sigma|$ for all $\sigma \in \{0,1\}^n$. Notice that $|\{0,1\}|^n = 2^n$ and the number of sequences of length strictly less than n is $1 + 2 + 4 + \cdots + 2^{n-1}$ because we can write it as the union $\{0,1\}^0 \cup \{0,1\}^1 \cup \{0,1\}^2 \cup \cdots \cup \{0,1\}^{n-1}$ where the sets are pairwise disjoint. Now the key fact is that

$$1 + 2 + 2^2 + \cdots + 2^{n-1} = 2^n - 1$$

which one can prove either by induction or by noting that

$$\begin{aligned}
1 + 2 + 2^2 + \cdots + 2^{n-1} &= (1 + 2 + 2^2 + \cdots + 2^{n-1}) \cdot 1 \\
&= (1 + 2 + 2^2 + \cdots + 2^{n-1}) \cdot (2 - 1) \\
&= (2 + 2^2 + 2^3 + \cdots + 2^n) - (1 + 2 + 2^2 + \cdots + 2^{n-1}) \\
&= 2^n - 1.
\end{aligned}$$

Since $|\{0,1\}^n| = 2^n$ and the set of sequences of length strictly less than n is $2^n - 1$, we may use the Pigeonhole Principle to conclude that there exists distinct $\sigma_1, \sigma_2 \in \{0,1\}^n$ with $f(\sigma_1) = f(\sigma_2)$, which contradicts the fact that f is injective. Therefore, there must exist $\sigma \in \{0,1\}^n$ with $|f(\sigma)| \geq |\sigma|$. □

We can interpret the previous proposition as follows. Suppose that we have a compression algorithm, i.e. a program that takes a sequence of 0's and 1's and tries to compress it down to a shorter sequence (think of any standard zip program). If we look at how the function behaves on every input, we obtain a function $f \colon \{0,1\}^* \to \{0,1\}^*$. Of course, for this compression algorithm to be at all useful, we would need to be able to uncompress any file back to its original. In order to do this, the function f must be injective (otherwise, if two files compress to the same thing, we would have no way to know which file to return). This proposition says that any purported compression scheme must in fact fail to actually shrink the size of some file, and in fact for *every* length n, there is a file of length n that is not actually made smaller.

Proposition 2.32. *Let $n \in \mathbb{N}^+$. Given a set $S \subseteq \{1, 2, 3, \ldots, 2n\}$ with $|S| \geq n + 1$, there always exists a pair of distinct elements $a, b \in S$ with $a \mid b$.*

Before proving this proposition, we examine some special cases in order to get some intuition. First, consider the case when $n = 5$ so that $2n = 10$. We want to prove that whenever we have at least 6 numbers from the set $\{1, 2, 3, \ldots, 10\}$, we can find two distinct numbers a and b such that $a \mid b$. The idea is to build five "boxes" of numbers with the following properties:

- Every number from $\{1, 2, 3, \ldots, 10\}$ is in a box.
- Given any two distinct numbers from the same box, one divides the other.

Suppose that we are successful in doing this. Then given any set of at least six numbers, we can find two of the numbers in the same box (because we only five boxes), and then we will be done. So let's build five boxes with the above properties in the case where $n = 5$:

- Box 1: $\{1, 2, 4, 8\}$
- Box 2: $\{3, 6\}$
- Box 3: $\{5, 10\}$
- Box 4: $\{7\}$
- Box 5: $\{9\}$

We now want to generalize this argument. The key idea behind the above boxes was as follows: Given a natural number, keep dividing by 2 until we reach an odd number, and put two numbers in the same box if we arrive at the same odd number. In order to formalize this, we prove the following lemma.

Lemma 2.33. *Let $n \in \mathbb{N}^+$. There exist unique $k, \ell \in \mathbb{N}$ such that ℓ is odd and $n = 2^k \ell$.*

Proof. We first prove the existence of k and ℓ by strong induction on n.

- When $n = 1$, we can write $1 = 2^0 \cdot 1$, so we can take $k = 0$ and $\ell = 1$.
- Let $n \in \mathbb{N}$ with $n \geq 2$, and assume that we know the existence part is true for all m with $1 \leq m < n$. We prove it for n. First, notice that if n is odd, then we can simply write $n = 2^0 n$, and we are done. Suppose then that n is even. Fix $m \in \mathbb{Z}$ with $n = 2m$ and notice that $1 \leq m < n$. By induction, we can fix $k, \ell \in \mathbb{N}$ such that ℓ is odd and $m = 2^k \ell$. We then have $n = 2m = 2^{k+1}\ell$, hence the result holds for n.

The existence of k and ℓ for all n follows by induction.

We now prove uniqueness. Suppose that $k_1, k_2, \ell_1, \ell_2 \in \mathbb{N}$ are such that ℓ_1 and ℓ_2 are both odd and $2^{k_1}\ell_1 = 2^{k_2}\ell_2$. If $k_1 < k_2$, then dividing both sides by 2^{k_1}, we would be able to conclude that $\ell_1 = 2^{k_2 - k_1}\ell_2$, which contradicts the fact that ℓ_1 is odd (since $k_2 - k_1 \geq 1$). A similar contradiction occurs if $k_1 > k_2$. Therefore, we must have that $k_1 = k_2$. Diving both sides by $2^{k_1} = 2^{k_2}$, we then conclude that $\ell_1 = \ell_2$. This gives uniqueness. □

Proof of Proposition 2.32. $S \subseteq \{1, 2, 3, \ldots, 2n\}$ with $|S| \geq n + 1$ be arbitrary. Let X be the set of all odd integers ℓ with $1 \leq \ell \leq 2n$, and notice that $|X| = n$. Define a function $f \colon S \to X$ as follows. Given $a \in S$, write $a = 2^k \ell$ for the unique k and ℓ from the previous lemma, and define $f(a) = \ell$ (notice that $\ell \leq 2n$ because $a \leq 2n$). Intuitively, associate to each given $n \in S$ the unique odd number obtained by repeatedly dividing by 2 until we reach an odd number. Since $|S| \geq n + 1$ and $|X| = n$, the Pigeonhole Principle tells us that we can find distinct $a, b \in S$ with $a < b$ such that $f(a) = f(b)$. Call this unique value ℓ, i.e. let $\ell = f(a) = f(b)$, and fix $k_1, k_2 \in \mathbb{N}$ with $a = 2^{k_1}\ell$ and $b = 2^{k_2}\ell$. Since $a < b$, we have $k_1 < k_2$. Now

$$b = 2^{k_2}\ell = 2^{k_2-k_1} \cdot 2^{k_1} \cdot \ell = 2^{k_2-k_1} \cdot a$$

so $a \mid b$. This completes the proof. □

Suppose that we have a finite sequence of (possibly real) numbers. For example, consider the following sequence of 10 numbers:

$$3 \quad 1 \quad 6 \quad 9 \quad 0 \quad 2 \quad 8 \quad 5 \quad 7 \quad 4$$

Although these numbers are not sorted in any sense, one can find a decently long decreasing subsequence by pulling out the $9, 8, 7, 4$. It turns out that no matter what sequence of length n one looks at, it always possible to pull out an increasing or decreasing subsequence of length about $\sqrt{n}$.

Definition 2.34. *Suppose that $a_1, a_2, \dots, a_n$ is a finite sequence of real numbers. Suppose that we have a sequence of indices with $1 \leq i_1 < i_2 < \cdots < i_k \leq n$. We then call $a_{i_1}, a_{i_2}, \dots, a_{i_k}$* a subsequence *of $a_1, a_2, \dots, a_n$.*

Definition 2.35. *Suppose that $a_1, a_2, \dots, a_n$ is a finite sequence of real numbers.*

- *We call the sequence* increasing *if $a_1 \leq a_2 \leq \cdots \leq a_n$.*
- *We call the sequence* decreasing *if $a_1 \geq a_2 \geq \cdots \geq a_n$.*
- *We call the sequence* monotonic *if it is either increasing or decreasing.*

For example, suppose that $a_1, a_2, \dots, a_{10}$ is our original sequence

$$3 \quad 1 \quad 6 \quad 9 \quad 0 \quad 2 \quad 8 \quad 5 \quad 7 \quad 4$$

Notice that $9, 8, 7, 4$ is a decreasing subsequence of this sequence (with $i_1 = 4$, $i_2 = 7$, $i_3 = 9$, and $i_4 = 10$).

Theorem 2.36. *Let $n \in \mathbb{N}^+$. Given a sequence of $(n-1)^2 + 1$ real numbers, it is always possible to find a monotonic subsequence of length n.*

Proof. Consider an arbitrary sequence

$$a_1, a_2, a_3, \dots, a_{(n-1)^2+1}$$

of $(n-1)^2 + 1$ many real numbers. Associate to each i the pair $(k, \ell) \in \mathbb{N}^+ \times \mathbb{N}^+$ where k is the length of the longest increasing subsequence ending with (and including) a_i and ℓ is the length of the longest decreasing subsequence ending with (and including) a_i. If any one of these pairs has a coordinate that is at least n, then we are done. Otherwise, every pair (k, ℓ) is such that $1 \leq k \leq n-1$ and $1 \leq \ell \leq n-1$. There are only $(n-1)^2$ many possible pairs, so since we have $(n-1)^2 + 1$ many numbers some pair must be repeated by the Pigeonhole Principle. Fix $i < j$ with $(k_i, \ell_i) = (k_j, \ell_j)$. Now if $a_j \geq a_i$, then we can add a_j onto the end of the longest increasing subsequence ending in a_i to form an increasing subsequence of length $k_i + 1 > k_j$, a contradiction. Similarly, if $a_j \leq a_i$, then we can add a_j onto the end of the longest decreasing subsequence ending in a_i to form an idecreasing subsequence of length $\ell_i + 1 > \ell_j$, a contradiction. □

For example, for our sequence

$$3 \quad 1 \quad 6 \quad 9 \quad 0 \quad 2 \quad 8 \quad 5 \quad 7 \quad 4$$

we would assign the values

$$(1,1) \quad (1,2) \quad (2,1) \quad (3,1) \quad (1,3) \quad (2,2) \quad (3,2) \quad (3,3) \quad (4,3) \quad (3,4)$$

Thus, we either take either $0, 2, 5, 7$ as an increasing subsequence or $9, 8, 7, 4$ as a decreasing subsequence.

3 Counting

3.1 Arrangements, Permutations, and Combinations

Let A be a finite set with $|A| = n$. Given $k \in \mathbb{N}^+$, the set A^k is the set of all finite sequences of length n whose elements are all from A. Occasionally, especially in computer science, such a finite sequence is called a *string* over A of length k. We already know that $|A^k| = |A|^k = n^k$, so we can count the number of finite sequences of length k. For example, if $A = \{a, b, c, d\}$, then there are exactly $4^2 = 16$ many two letter strings over A. There are exactly $128^2 = 16,384$ many two character long ASCII sequences, and there are 10^7 many potential phone numbers.

Notice that in a finite sequences, we might have repetition. For example, if $A = \{1, 2, 3\}$, then $(1, 1, 3) \in A^3$ and $(3, 1, 2, 3) \in A^4$. Suppose that A is a set with $|A| = n$, and we want to count the number of sequences of length 2 where there is no repetition, i.e. we want to determine the cardinality of the set

$$B = \{(a, b) \in A^2 : a \neq b\}$$

There are (at least) two straightforward ways to do this.

- *Method 1:* As on the first homework, we use the complement rule. Let $D = \{(a, a) : a \in A\}$ and notice that $|D| = n$ because $|A| = n$. Since $B = A^2 \backslash D$, it follows that $|B| = |A^2| - |D| = n^2 - n = n(n-1)$.
- *Method 2:* We use a modified version of the product rule as follows. Think about constructed an element of B in two stages. First, we need to pick the first coordinate of our pair, and we have n choices here. Now once we fix the first coordinate of our pair, we have $n - 1$ choices for the second coordinate because we can choose any element of A other than the one that we chose in the first round. By making these two choices in succession, we determine an element of B, and furthermore, every element of B is obtained via a unique sequence of such choices. Therefore, we have $|B| = n(n - 1)$.

Notice that in the argument for Method 2 above, we are not directly using the Product Rule. The issue is that we can not write B in the form $B = X \times Y$ where $|X| = n$ and $|Y| = n-1$ because the choice of second coordinates depends upon the choice of first component. For example, if $A = \{1, 2, 3\}$, then if we choose 1 as our first coordinate, then we can choose any element of $\{2, 3\}$ for the second, while if we choose 3 as our first coordinate, then we can choose any element of $\{1, 2\}$ for the second. However, the key fact is that the *number* of choices for the second coordinate is the same no matter what we choose for the first.

Suppose more generally that we are building a set of objects in stages, in such a way that a sequence of choices throughout the stages determines a unique object, and no two distinct sequences determine the same object. Suppose also that we have the following number of choices at each stage:

- There are n_1 many choices at the first stage.
- For each choice in the first stage, there are n_2 many objects to pair with it in the second stage.
- For each pair of choices in the first two stages, there are n_3 many choices to append at the third stage.
- ...
- For each sequence of choices in the first $k - 1$ stages, there are n_k many choices to append at the k^{th} stage.

In this situation, there are $n_1 n_2 n_3 \cdots n_k$ many total objects in the set. The argument is similar to the argument for the Product Rule, and again we will omit a formal proof.

With this new rule in hand, we can count a new type of object.

Definition 3.1. *Let A be a finite set with $|A| = n$. A* permutation *of A is an element of A^n without repeated elements, i.e. it's a linear ordering of all n elements of A without repetition.*

For example, consider $A = \{1,2,3\}$. One example of permutation of A is $(3,1,2)$. The set of all permutations of A is:

$$\{(1,2,3),(1,3,2),(2,1,3),(2,3,1),(3,1,2),(3,2,1)\}$$

Thus, there are 6 permutations of the set $\{1,2,3\}$. In order to count the number of permutations of a set with n element, we use our new technique.

Proposition 3.2. *If A is a finite set with $n \in \mathbb{N}^+$ elements, then there are $n!$ many permutations of A.*

Proof. We can build a permutation of A through a sequence of choices.

- We being by choosing the first element, and we have n choices.
- Once we've chosen the first element, we have $n-1$ choices for the second because we can choose any element of A other than the one chosen in the first stage.
- Next, we have $n-2$ many choices for the third element.
- ...
- At stage $n-1$, we have chosen $n-2$ distinct elements so far, so we have 2 choices here.
- Finally, we have only choice remaining for the last position.

Since every such sequence of choices determines a permutation of A, and distinct choices given distinct permutations, it follows that there are $n(n-1)(n-2)\cdots 2 \cdot 1 = n!$ many permutations of A. □

Alternatively, we can give a recursive description of the number of permutations of a set with n-elements, and use that to derive the above result. Define $f\colon \mathbb{N}^+ \to \mathbb{N}^+$ by letting $f(n)$ be the number of permutations of $\{1,2,\dots,n\}$. Notice that $f(1) = 1$. Suppose that we know the value of $f(n)$. We show how to build all permutations of $\{1,2,\dots,n+1\}$ from the $f(n)$ many permutations of $\{1,2,\dots,n\}$ along with an element of the set $\{1,2,\dots,n,n+1\}$. Given a permutation of $\{1,2,\dots,n\}$ together with a number k with $1 \leq k \leq n+1$, we form a permutation of $\{1,2,\dots,n,n+1\}$ by taking our permutation of $\{1,2,\dots,n\}$, and inserting $n+1$ into the sequence in position k (and then shifting all later numbers to the right). For example, if $n = 4$ and we have the permutation $(4,1,3,2)$, and we have $k = 2$, then we insert 5 into the second position to form the permutation $(4,5,1,3,2)$.

In this way, we form all permutations of $\{1,2,\dots,n,n+1\}$ in a unique way. More formally, if we let $\mathcal{R}_n$ is the set of all permutation $\{1,2,\dots,n\}$, then this rule provides a bijection from $\mathcal{R}_n \times \{1,2,\dots,n,n+1\}$ to $\mathcal{R}_{n+1}$. Therefore, we have $f(n+1) = (n+1)\cdot f(n)$ for all $n \in \mathbb{N}^+$. Combining this with the fact that $f(1) = 1$, we conclude that $f(n) = n!$ for all $n \in \mathbb{N}^+$.

Definition 3.3. *Let A be a finite set with $|A| = n$, and let $k \in \mathbb{N}$ with $1 \leq k \leq n$. A* partial permutation *of A of length k is an element of A^k with no repeated element. A partial permutation of length k is also called a k-*permutation *of A.*

Proposition 3.4. *If A is a finite set with $n \in \mathbb{N}^+$ elements and $k \in \mathbb{N}^+$ is such that $1 \leq k \leq n$, then there are*

$$n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!}$$

many k-permutations of A.

Proof. The proof is the same as for permutations, except we stop after k stages. Notice that the last term in the product, corresponding to the number of choices at stage k, is $n-(k-1)=n-k+1$ because at stage k we have chosen the first $k-1$ many element. Finally, notice that

$$\begin{aligned}\frac{n!}{(n-k)!} &= \frac{n(n-1)(n-2)\cdot(n-k+1)(n-k)(n-k-1)\cdots 1}{(n-k)(n-k-1)\cdots 1} \\ &= n(n-1)(n-2)\cdots(n-k+1)\end{aligned}$$

giving the last equality. □

For example, using the standard 26-letter alphabet, there are $26\cdot 25\cdot 24=\frac{26!}{23!}=15,600$ many three-letter strings of letters having no repetition.

Notation 3.5. *If $k,n\in\mathbb{N}^+$ with $1\leq k\leq n$, we use the notation $(n)_k$ or $P(n,k)$ for the number of k-permutations of a set with n elements, i.e. we define*

$$(n)_k=P(n,k)=\frac{n!}{(n-k)!}$$

Suppose that A and B are finite sets and $|A|=m$ and $|B|=n$.

- We claim that number of functions f with domain A and codomain B equals n^m. To see this, list the elements of A in some order as $a_1,a_2,\dots,a_m$. A function assigns a unique value in B to each a_i, so we go through that a_i in order. For a_1, we have n possible images because we can choose any element of B. Once we've chosen this, we now have n possible images for a_2. As we go along, we always have n possible images for each of the a_i. Therefore, the number of functions from A to B is $n\cdot n\cdots n=n^m$.

- Notice that if $n<m$, then there are no injective functions $f\colon A\to B$ by the Pigeonhole Principle. Suppose instead that $n\leq m$. We claim that the number of injective functions f with domain A and codomain B equals $P(n,m)=\frac{n!}{(n-m)!}$. The argument is similar to the one for general functions, but we get fewer choices as we progress through A. As above, list the elements of A in some order as $a_1,a_2,\dots,a_m$. A function assigns a unique value in B to each a_i, so we go through that a_i in order. For a_1, we have n possible images because we can choose any element of B. Once we've chosen this, we now have $n-1$ possible images for a_2 because we can choose any value of B other than the one we sent a_1 to. Then we have $n-2$ many choices for a_3, etc. Once we arrive at a_m, we have already used up $m-1$ many elements of B, so we have $n-(m-1)=n-m+1$ many choices for where to send a_m. Therefore, the number of functions from A to B is

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

  which is $P(n,m)$.

Suppose we ask the following question: Let $A=\{1,2,3,4,5,6,7\}$. How many element of A^4 contain the number 7 at least once? In other words, how many three digit numbers are there such that each digit is between 1 and 7 (inclusive), and 7 occurs at least once? A natural guess is that the answer is $4\cdot 7^3$ for the follow reason:

- First, pick one of 4 positions to place the 7.
- Now we have three positions open. Going through them in order, we have 7 choices for what to put in each of these three positions.

This all looks great, but unfortunately, there is a problem. It is indeed true that such a sequence of four choices does create one of the number we are looking for. If we choose the sequence $3, 1, 5, 1$, then we obtain the number 1571. However, the sequence of choices $2, 7, 3, 4$ and the sequence of choices $1, 7, 3, 4$ both produce the same string, namely 7734.

The key idea is to count the complement. Instead of counting the number of elements of A^4 that *do* contain the number 7 at least once, we count the number of elements of A^4 that *do not* contain the number 7 at all, and subtract this amount from the total number of elements in A^4. Now since $|A| = 7$, we have that $|A^4| = 7^4$ because we have 7 choices for each of the 4 spots. To count the number of elements of A^4 that do not contain a 7, we simply notice that we have 6 choices for each of the 4 spots, so there are 6^4 of these. Therefore, by the Complement Rule, the number of elements of A^4 that do contain the number 7 at least once is $7^4 - 6^4$.

We next move on to a fundamental question that will guide a lot of our later work. Let $n \in \mathbb{N}^+$. We know that there are 2^n many subsets of $\{1, 2, \dots, n\}$. However, what if we ask how many subsets there are of a certain size? For instance, how many subsets are there of $\{1, 2, 3, 4, 5\}$ having exactly 3 elements? The intuitive idea is to make 3 choices: First, pick one of the 5 elements to go into our set. Next, put one of the 4 remaining elements to add to it. Finally, finish off the process by picking one of the 3 remaining elements. For example, if we choose the number $1, 3, 5$ then we get the set $\{1, 3, 5\}$. Thus, a natural guess is that there are $5 \cdot 4 \cdot 3$ many subsets with 3 elements. However, recall that a set has neither repetition nor order, so just as in the previous example we count the same set multiple times. For example, picking the sequence $3, 5, 1$ would also give the set $\{1, 3, 5\}$. In fact, we arrive at the set $\{1, 3, 5\}$ in the following six ways:

$$\begin{array}{ccc} 1,3,5 & 1,5,3 & 3,1,5 \\ 3,5,1 & 5,1,3 & 5,3,1 \end{array}$$

At this point, we may be tempted to throw our hands in the air as we did above. However, there is one crucial difference. In our previous example, some sequences of 4 numbers including a 7 were counted once (like 1571), some were counted twice (like 7712), and others were counted three or four times. However, in our current situation, *every* subset is counted exactly 6 times because given a set with 3 elements, we know that there are $3! = 6$ many permutations of that set (i.e. ways to arrange the elements of the set in order). The fact that we count each element 6 times means that the total number of subsets of $\{1, 2, 3, 4, 5\}$ having exactly 3 elements equals $\frac{5 \cdot 4 \cdot 3}{6} = 10$. The general principle that we are applying is the following:

Proposition 3.6 (Quotient Rule)**.** *Suppose that A is a finite set with $|A| = n$. Suppose that $\sim$ is an equivalence relation on A, and that every equivalence class has exactly k elements. In this case there are $\frac{n}{k}$ many equivalence classes.*

Proof. Let ℓ be the number of equivalence classes. To obtain an element of A, we can first pick one of the ℓ equivalence classes, and then pick one of the k many elements from that class. Since the equivalence classes partition A, it follows that this sequence of choices produces each element of A in a unique way. Thus, $n = k \cdot \ell$ by the Product Rule. It follows that $\ell = \frac{n}{k}$. □

Proposition 3.7. *Let $n, k \in \mathbb{N}^+$ and with $1 \leq k \leq n$. Suppose that A is a finite set with $|A| = n$. The number of subsets of A having exactly k elements equals:*

$$\frac{n(n-1)(n-2)\cdots(n-k+1)}{k!} = \frac{n!}{k! \cdot (n-k)!}$$

Proof. We generalize the above argument. We know that the number of k-permutations of A equals

$$n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!}.$$

Now a k-permutation of A picks k distinct elements of A, put also assigns an order to the elements. Now every subset of A of size k is coded by exactly $k!$ many such k-permutations because we can order the subset in $k!$ many ways. Therefore, by the Quotient Rule, the number of subsets of A having exactly k elements equals

$$\frac{n(n-1)(n-2)\cdots(n-k+1)}{k!} = \frac{n!}{k!\cdot(n-k)!}$$

□

Notice also that if $k = 0$, then there is one subset of any set having zero elements (namely $\emptyset$). Thus, by defining $0! = 1$, the above formula works in the case when $k = 0$ as well.

Definition 3.8. *Let $n, k \in \mathbb{N}$ and with $0 \leq k \leq n$. We define the notations $\binom{n}{k}$ and $C(n,k)$ by*

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

We call this the number of k-combinations *of an n-element set, and pronounce $\binom{n}{k}$ as "n choose k".*

For example, the number of 5-card poker hands from a standard 52-card deck is:

$$\binom{52}{5} = \frac{52!}{5!\cdot 48!} = 2,598,960.$$

We now give a number of example of counting problems:

- Over the standard 26-letter alphabet, how many "words" of length 8 have exactly 5 consonants and 3 vowels? We build every such word in a unique way via a sequence of choices.
  - First, we pick out a subset of 3 of the 8 positions to house the vowels, and we have $\binom{8}{3}$ many possibilities.
  - Next, we pick 3 vowels in order allowing repetition to fill in these positions. Since we have 5 vowels, there are 5^3 many possibilities.
  - Finally, we pick 5 consonants in order allowing repetition to fill in remaining 5 positions. Since we have 21 consonants, there are 21^5 many possibilities.

  Since every word is uniquely determined by this sequence of choices, the number of such words is

  $$\binom{8}{3}\cdot 5^3 \cdot 21^6 = 56 \cdot 5^3 \cdot 21^5$$

- How many ways are there to seat n people around a circular table (so the only thing that matters is the relative position of people with respect to each other)? To count this, we use the Quotient Rule. We first consider each of the chairs as distinct. List the people in some order, and notice that we have n choices for where to seat the first person, then $n-1$ for where to seat the second, then $n-2$ for the third, and so forth. Thus, if the seats are distinct, then we have $n!$ many ways to seat the people. However, two such seating arrangements are equivalent if we can get one from the other via a rotation of the seats. Since there are n possible rotations, each seating arrangement occurs n times in this count, so the total number of such seatings is $\frac{n!}{n} = (n-1)!$.

  More formally, we can think about this as following. Consider all permutations of an n-element set (the people): we know that there are $n!$ of these. Now given two permutations, which are just sequences of length n without repetition, we consider two of these sequences equivalent exactly when every pair of numbers is the same distance apart where we allow "wrap around" (since the seating is circular). We then have that two such sequences are equivalent precisely when one is a cyclic shift of the other. Thus, every equivalence class has exactly n elements, and hence there are $\frac{n!}{n} = (n-1)!$ many circular arrangements.

- Suppose that we are in a city where all streets are straight and either east-west or north-south. Suppose that we are at one corner, and want to travel to a corner that is m blocks east and n blocks north, but we want to do it efficiently. More formally, we want to count the number of ways to get from the point $(0,0)$ to the point (m,n) where at each stage we either increase the x-coordinate by 1 or we increase the y-coordinate by 1. At first sight, it appears that we at each intersection, we have 2 choices: Either go east or go north. However, this is not really the case, because if we can east m times, then we are forced to go north the rest of the way. The idea for how to count this is that such a path is uniquely determined by a sequence of $m+n$ many E's and N's (representing east and north) having exactly m many E's. To determine such a sequence, we need only choose the positions of the m many E's, and there are

  $$\binom{m+n}{m}$$

  man choices. Of course, we could instead choose the positions of n many N's to count it as

  $$\binom{m+n}{n}$$

  which is the same number.

- How many anagrams (i.e. rearrangements of the letters) are there of MISSISSIPPI? Here is one approach. Notice that MISSISSIPPI has one M, four I's, four S's, and two P's, for a total of eleven letters. First pick the position of the M and notice that we have 11 choices. Once that is done, pick the position of the four I's and notice that this amount to picking a 4 element subset of the remaining 10 positions. There are $\binom{10}{4}$ many such choices. Once that is done, pick the position of the four S's and notice that this amount to picking a 4 element subset of the remaining 6 positions. There are $\binom{6}{4}$ many such choices. Once this is done, the position of the two P's is fixed. This gives a total number of anagrams equal to

  $$11\binom{10}{4}\binom{6}{4} = 11 \cdot \frac{10!}{4! \cdot 6!} \cdot \frac{6!}{4! \cdot 2!} = \frac{11!}{4! \cdot 4! \cdot 2!} = 34,650$$

  Another argument is as follows. Think of distinguishing common letters with different colors. We then have 11! many ways to rearrange the letters, but this number overcounts the numbers of anagrams. Each actual anagram comes about in $4! \cdot 4! \cdot 2!$ many ways because we can permute the currently distinct four I's amongst each other in 4! ways, we can permute the currently distinct four S's amongst each other in 4! ways, and we can permute the the currently distinct two P's amongst each other in 2! many ways. Thus, since each actual anagram is counted $4! \cdot 4! \cdot 2!$ many times in the 11! count, it follows that there are

  $$\frac{11!}{4! \cdot 4! \cdot 2!} = 34,650$$

  many anagrams of MISSISSIPPI.

As mentioned above, there are a total of

$$\binom{52}{5} = 2,598,960$$

many (unordered) 5-card poker hands from a standard 52-card deck. Using this, we now count the number of special hands of each type. We use the fact that each card has one of four suits (clubs, diamonds, hearts, and spades) and one of thirteen ranks (2, 3, 4, 5, 6, 7, 8, 9,10, jack, queen, king, ace). We follow the common practice of allowing the ace to be either a low card or a high card for a straight, but we do not allow "wrap around" straights such as king, ace, 2, 3, 4.

- Straight Flush: There are
$$4 \cdot 10 = 40$$
many of these because they are determined by a choice of suits and the rank of the lowest card (from ace through 10). The probability is about .00154%.

- Four of a kind: There are
$$13 \cdot 48 = 624$$
of these because we choose a rank (and take all four cards of that rank), and the choose one of the remaining 48 cards. The probability is about .0256%.

- Full House: There are
$$13 \cdot \binom{4}{3} \cdot 12 \cdot \binom{4}{2} = 3,744$$
many, which can be seen by making the following sequence of choices:

  - Choose one of the 13 ranks for the three of a kind.
  - Choose 3 of the 3 suits for the three of a kind.
  - Choose one of the 12 remaining ranks for the pair.
  - Choose 2 of the 4 suits for the pair.

  The probability is about .14406%.

- Flush: There are
$$4 \cdot \binom{13}{5} = 5,148$$
many because we need to choose 1 of the 4 suits, and then 5 of the 13 ranks. However, 40 of these are actually straight flushes, so we really have $5,108$ many flushes that are not stronger hands. The probability is about .19654%

- Straight: There are
$$10 \cdot 4^5 = 10,240$$
many because we need to choose the rank of the lowest card, and the suits for the five cards in increasing order of rank. However, we again have that 40 of these are straight flushes, so we really have $10,200$ many straights that are not stronger hands. The probability is about .39246%.

- Three of a kind: There are
$$13 \cdot \binom{4}{3} \cdot \binom{12}{2} \cdot 4^2 = 54,912$$
many, which can be seen by making the following sequence of choices:

  - Choose one of the 13 ranks for the three of a kind.
  - Choose 3 of the 3 suits for the three of a kind.

  - Choose two of the other ranks for the remaining two cards (they are different because we do not want to include full houses).
  - Choose the suit of the lower ranked card not in the three of a kind.
  - Choose the suit of the higher ranked card not in the three of a kind.

  (Alternatively, we can choose the last two cards in different ranks in $48 \cdot 44$ many ways, but then we need to divide by 2 because the order of choosing these does not matter.) The probability is about 2.1128%.

- Two Pair: There are
$$\binom{13}{2} \cdot \binom{4}{2}^2 \cdot 44 = 123,552$$
  many, which can be seen by making the following sequence of choices:
  - Choose the two ranks for the two pairs.
  - Choose the two suits for the lower ranked pair.
  - Choose the two suits for the higher ranked pair.
  - Choose one of the 44 cards not in these two ranks.

  The probability is about 4.7539%.

- One pair: There are
$$13 \cdot \binom{4}{2} \cdot \binom{12}{3} \cdot 4^3 = 1,098,240$$
  many, which can be seen by making the following sequence of choices:
  - Choose the rank for the pair.
  - Choose the two suits for the pair.
  - Choose three distinct ranks for the other three cards (which are not the same rank as the pair).
  - Choose the suit of the lowest ranked card not in the pair.
  - Choose the suit of the middle ranked card not in the pair.
  - Choose the suit of the highest ranked card not in the pair.

  The probability is about 42.257%.

3.2 The Binomial Theorem and Properties of Binomial Coefficients

Recall that if $n, k \in \mathbb{N}$ with $k \leq n$, then we defined
$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$$

Notice that when $k = n = 0$, then $\binom{n}{k} = 1$ because we define $0! = 1$, and indeed there is a unique subset of $\emptyset$ having 0 elements, namely $\emptyset$. When $n, k \in \mathbb{N}$ with $n < k$, then we define
$$\binom{n}{k} = 0$$

because there are no subsets of an n-element set with cardinality k (notice that the above formula doesn't make sense because $n - k < 0$.

Using Proposition 2.26, we know that whenever $k, n \in \mathbb{N}$ are such that $k \leq n$, then

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

because the function that takes the relative complement is a bijection between subsets of cardinality k and subsets of cardinality $n-k$. Of course, one can prove this directly from the formulas because

$$\begin{aligned}\binom{n}{n-k} &= \frac{n!}{(n-k)! \cdot (n-(n-k))!} \\ &= \frac{n!}{(n-k)! \cdot k!} \\ &= \frac{n!}{k! \cdot (n-k)!} \\ &= \binom{n}{k}\end{aligned}$$

Although the algebraic manipulations here are easy, the bijective proof feels more satisfying because it "explains" the formula. Proving that two numbers are equal by showing that the both count the numbers of elements in one common set, or by proving that there is a bijection between a set counted by the first number and a set counted by the second, is called either a *combinatorial proof* or a *bijective proof.*

Proposition 3.9. *Let $n, k \in \mathbb{N}^+$ with $0 < k < n$. We have*

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

Proof. One extremely unenlightening proof is to expand out the formula on the right and do terrible algebraic manipulations on it. If you haven't done so, I encourage you to do it. However, we use the combinatorial description of $\binom{n}{k}$ to give a more meaningful combinatorial argument. Let $n, k \in \mathbb{N}$ with $k \leq n$. Consider a set A with n many elements. To determine $\binom{n}{k}$, we need to count the number of subsets of A of size k. We do this as follows. Fix an arbitrary $a \in A$. Now an arbitrary subset of A of size k fits into exactly one of the following types.

- The subset has a as an element. In this case, to completely determine the subset, we need to pick the remaining $k-1$ elements of the subset from $A \backslash \{a\}$. Since $A \backslash \{a\}$ has $n-1$ elements, the number of ways to do this is $\binom{n-1}{k-1}$.
- The subset does not have a as an element. In this case, to completely determine the subset, we need to pick all k elements of the subset from $A \backslash \{a\}$. Since $A \backslash \{a\}$ has $n-1$ elements, the number of ways to do this is $\binom{n-1}{k}$.

Putting this together, we conclude that the number of subsets of A of size k equals $\binom{n-1}{k-1} + \binom{n-1}{k}$. □

Using this proposition, together with the fact that

$$\binom{n}{0} = 1 \qquad \text{and} \qquad \binom{n}{n} = 1$$

for all $n \in \mathbb{N}$, we can compute $\binom{n}{k}$ recursively to obtain the following table. The rows are labeled by n and the columns by k. To determine the number that belongs in a given square, we simply add the number above it and the number above and to the left. This table is known as *Pascal's Triangle*:

$\binom{n}{k}$	0	1	2	3	4	5	6	7
0	1	0	0	0	0	0	0	0
1	1	1	0	0	0	0	0	0
2	1	2	1	0	0	0	0	0
3	1	3	3	1	0	0	0	0
4	1	4	6	4	1	0	0	0
5	1	5	10	10	5	1	0	0
6	1	6	15	20	15	6	1	0
7	1	7	21	35	35	21	7	1

There are many curious properties of Pascal's Triangle that we will discover in time. On of the first things to note is that these numbers seem to appear in other places. For example, if $x, y \in \mathbb{R}$, then we have:

- $(x+y)^1 = x + y$
- $(x+y)^2 = x^2 + 2xy + y^2$
- $(x+y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$
- $(x+y)^4 = x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4$

Looking at these, it appears that the coefficients are exactly the corresponding elements of Pascal's Triangle. What is the connection here? Notice that if we do not use commutativity and do not collect like terms, we have

$$\begin{aligned}(x+y)^2 &= (x+y)(x+y)\\ &= x(x+y) + y(x+y)\\ &= xx + xy + yx + yy\end{aligned}$$

and so

$$\begin{aligned}(x+y)^3 &= (x+y)(x+y)^2\\ &= (x+y)(xx + xy + yx + yy)\\ &= x(xx + xy + yx + yy) + y(xx + xy + yx + yy)\\ &= xxx + xxy + xyx + xyy + yxx + yxy + yyx + yyy.\end{aligned}$$

In other words, it looks like when we fully expand $(x+y)^n$, without using commutativity or collecting x's and y's, that we are getting a sum of all sequences of x's and y's of length n. Thus, if we want to know the coefficient of $x^{n-k}y^k$, then we need only ask how many such sequences have exactly k many y's (or equivalently exactly $n-k$ many x's), and the answer is $\binom{n}{k} = \binom{n}{n-k}$ because we need only pick out the position of the y's (or the x's). More formally, we can prove this by induction.

Theorem 3.10 (Binomial Theorem)**.** *Let $x, y \in \mathbb{R}$ and let $n \in \mathbb{N}^+$. We have*

$$\begin{aligned}(x+y)^n &= \binom{n}{0}x^n + \binom{n}{1}x^{n-1}y + \cdots + \binom{n}{n-1}xy^{n-1} + \binom{n}{n}y^n\\ &= \sum_{k=0}^{n}\binom{n}{k}x^{n-k}y^k\\ &= \sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k}\end{aligned}$$

Proof. We prove the result by induction. When $n = 1$, we trivially have

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

Suppose then that we have an $n \in \mathbb{N}^+$ for which we know that the statement is true. We then have

$$\begin{aligned}
(x+y)^{n+1} &= (x+y)^n \cdot (x+y) \\
&= \left(\binom{n}{0}x^n + \binom{n}{1}x^{n-1}y + \cdots + \binom{n}{n-1}xy^{n-1} + \binom{n}{n}y^n\right) \cdot (x+y) \\
&= \left(\binom{n}{0}x^n + \binom{n}{1}x^{n-1}y + \cdots + \binom{n}{n-1}xy^{n-1} + \binom{n}{n}y^n\right) \cdot x \\
&\qquad \left(\binom{n}{0}x^n + \binom{n}{1}x^{n-1}y + \cdots + \binom{n}{n-1}xy^{n-1} + \binom{n}{n}y^n\right) \cdot y \\
&= \binom{n}{0}x^{n+1} + \binom{n}{1}x^n y + \binom{n}{2}x^{n-1}y^2 + \cdots + \binom{n}{n-1}x^2y^{n-1} + \binom{n}{n}xy^n \\
&\qquad + \binom{n}{0}x^n y + \binom{n}{1}x^{n-1}y^2 + \cdots + \binom{n}{n-2}x^2y^{n-1} + \binom{n}{n-1}xy^n + \binom{n}{n}y^{n+1} \\
&= x^{n+1} + \left(\binom{n}{1} + \binom{n}{0}\right) \cdot x^n y + \left(\binom{n}{2} + \binom{n}{1}\right) \cdot x^{n-1}y^2 + \cdots + \left(\binom{n}{n} + \binom{n}{n-1}\right) \cdot xy^n + y^{n+1} \\
&= \binom{n+1}{0}x^{n+1} + \binom{n+1}{1}x^n y + \binom{n+1}{2}x^{n-1}y^2 + \cdots + \binom{n+1}{n}xy^n + \binom{n+1}{n+1}y^{n+1}
\end{aligned}$$

where we have used Proposition 3.9 to combine each of the sums to get the last line. □

Corollary 3.11. *For any $n \in \mathbb{N}^+$, we have*

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n} = 2^n$$

Proof 1. We use the Binomial Theorem in the special case where $x = 1$ and $y = 1$ to obtain

$$\begin{aligned}
2^n &= (1+1)^n \\
&= \sum_{k=0}^{n} \binom{n}{k} \cdot 1^{n-k} \cdot 1^k \\
&= \sum_{k=0}^{n} \binom{n}{k} \\
&= \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n}.
\end{aligned}$$

This completes the proof. □

Proof 2. Let $n \in \mathbb{N}^+$ be arbitrary. We give a combinatorial proof by arguing that both sides count the number of subsets of an n-element set. Suppose then that A is a set with $|A| = n$. On the one hand, we know that $|\mathcal{P}(A)| = 2^n$ by Corollary 2.25.

We know argue that

$$|\mathcal{P}(A)| = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n}$$

For each $k \in \mathbb{N}$ with $0 \leq k \leq n$, let $\mathcal{P}_k(A)$ be the subset of $\mathcal{P}(A)$ consisting of those subsets of A having exactly k elements. We then have that

$$\mathcal{P}(A) = \mathcal{P}_0(A) \cup \mathcal{P}_1(A) \cup \mathcal{P}_2(A) \cup \cdots \cup \mathcal{P}_n(A)$$

and furthermore that the $\mathcal{P}_k(A)$ are pairwise disjoint (i.e. if $k \neq \ell$, then $\mathcal{P}_k(A) \cap \mathcal{P}_\ell(A) = \emptyset$). Therefore,

$$|\mathcal{P}(A)| = |\mathcal{P}_0(A)| + |\mathcal{P}_1(A)| + |\mathcal{P}_2(A)| + \cdots + |\mathcal{P}_n(A)|$$

Now for each k with $0 \leq k \leq n$, we know that

$$|\mathcal{P}_k(A)| = \binom{n}{k}$$

so it follows that

$$|\mathcal{P}(A)| = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n}.$$

Hence

$$2^n = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n}.$$

because both sides count the number of elements of $\mathcal{P}(A)$. □

Corollary 3.12. *For any $n \in \mathbb{N}^+$, we have*

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

Thus

$$\binom{n}{0} + \binom{n}{2} + \binom{n}{4} + \cdots = 2^{n-1} = \binom{n}{1} + \binom{n}{3} + \binom{n}{5} + \dots$$

Proof 1. We use the Binomial Theorem in the special case where $x = 1$ and $y = -1$ to obtain

$$\begin{aligned} 0 &= 0^n \\ &= (1 + (-1))^n \\ &= \sum_{k=0}^{n} \binom{n}{k} \cdot 1^{n-k} \cdot (-1)^k \\ &= \sum_{k=0}^{n} (-1)^k \binom{n}{k} \\ &= \binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \cdots + (-1)^n \binom{n}{n}. \end{aligned}$$

This gives the first claim. Adding $\binom{n}{k}$ to both sides for each odd k, we conclude that

$$\binom{n}{0} + \binom{n}{2} + \binom{n}{4} + \cdots = \binom{n}{1} + \binom{n}{3} + \binom{n}{5} + \dots$$

Since

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n} = 2^n$$

by the previous result, it follows that

$$\binom{n}{0} + \binom{n}{2} + \binom{n}{4} + \dots = 2^{n-1} = \binom{n}{1} + \binom{n}{3} + \binom{n}{5} + \dots$$

□

Proof 2. Let $n \in \mathbb{N}^+$ be arbitrary. We begin by giving a combinatorial proof for the second claim. We first show that

$$\binom{n}{0} + \binom{n}{2} + \binom{n}{4} + \dots = 2^{n-1}$$

Let A be an arbitrary set with $|A| = n$, and list the elements of A as $A = \{a_1, a_2, \dots, a_n\}$. Recall that we know that $|\mathcal{P}(A)| = 2^n$ because for each i, we have 2 choices for whether or not to include a_i in our subset. Now in our case, the sum on the left

$$\binom{n}{0} + \binom{n}{2} + \binom{n}{4} + \dots$$

counts the numbers of subset of A having an even number of elements. We argue that 2^{n-1} also counts the number of subsets of A having an even number of elements. To build these subsets, we make the following sequence of choices:

- Determine whether to include a_1 in our subset: We have 2 choices.
- Determine whether to include a_2 in our subset: We have 2 choices.
- ...
- Determine whether to include a_{n-1} in our subset: We have 2 choices.
- Finally, examine the first $n-1$ choices, and determine whether we have included an even number of a_i. If so, do not include a_n in our subset. If not, include a_n in our subset.

Notice that in the last step, we do not make any choices, but do one of two things that are completely determined by the previous choices. Now no matter what sequence of choices we make, we end up with a subset of A having an even number of elements, and furthermore every subset with an even number of elements arrises in a unique way. Since there are 2 choices in each of the opening $n-1$ stages, it follows that there are 2^{n-1} many subsets of A with an even number of elements. Therefore,

$$\binom{n}{0} + \binom{n}{2} + \binom{n}{4} + \dots = 2^{n-1}$$

Now the proof that

$$\binom{n}{1} + \binom{n}{3} + \binom{n}{5} + \dots = 2^{n-1}$$

is completely analogous except for changing the last stage (or alternatively comes from the complement rule). Finally, since both of these sums equals 2^{n-1}, we conclude that

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

□

Proposition 3.13. *For any $n, k \in \mathbb{N}^+$ with $k \leq n$, we have*

$$k \cdot \binom{n}{k} = n \cdot \binom{n-1}{k-1}$$

hence

$$\binom{n}{k} = \frac{n}{k} \cdot \binom{n-1}{k-1}.$$

Proof. We claim that each side counts the number of ways of selecting a committee consisting of k people, including a distinguished president of the committee, from a group of n people. On the one hand, we can do this as follows:

- First pick the committee of k people from the total group of all n people. We have $\binom{n}{k}$ many ways to do this.
- Within this committee, choose one of the k people to serve as president. We have k options here.

Therefore, the number of possibilities is $k \cdot \binom{n}{k}$. On the other hand, we can count it as follows.

- First pick one of the n people to be the president.
- Next pick the remaining $k-1$ many people to serve on the committee amongst the remaining $n-1$ people. We have $\binom{n-1}{k-1}$ many ways to do this.

Therefore, the number of possibilities is $n \cdot \binom{n-1}{k-1}$.

Since each side counts the number of elements of one set, the values must be equal. Therefore,

$$k \cdot \binom{n}{k} = n \cdot \binom{n-1}{k-1}.$$

□

Proposition 3.14. *For any n, we have*

$$\sum_{k=1}^{n} k \cdot \binom{n}{k} = n \cdot 2^{n-1}.$$

Proof 1. We have

$$\begin{aligned}
\sum_{k=1}^{n} k \cdot \binom{n}{k} &= \sum_{k=1}^{n} n \cdot \binom{n-1}{k-1} && \text{(by Proposition 3.13)} \\
&= n \cdot \sum_{k=1}^{n} \binom{n-1}{k-1} \\
&= n \cdot \sum_{k=0}^{n-1} \binom{n-1}{k} \\
&= n \cdot 2^{n-1} && \text{(by Corollary 3.11)}
\end{aligned}$$

□

Proof 2. We give a direct combinatorial proof by arguing that both sides count the number of ways of building a committee, including a distinguished president of that committee, of any size from a group of n people.

One the one hand we can count the number of such committees as follows. We break up the situation into cases based on the size of the committee. For a committee of size k including a distinguished president, we know from Proposition 3.13 that there are $k \cdot \binom{n}{k}$ many ways to do this. Since we can break up the collection of all such committees into the pairwise disjoint union of those committees of size 1, those of size 2, etc. Therefore, by the sum Rule, the number of ways to do this is $\sum_{k=1}^{n} k \cdot \binom{n}{k}$.

On the other hand, we can count the number of such committees differently. First, pick the president of the committee, and notice that we have n choices. Once we pick the president, we need to pick the rest of the committee. Thus, we need to pick a subset (of any size) from the remaining $n-1$ people to fill out the committee, and we know that there are 2^{n-1} many subsets of a set of size $n-1$. Therefore, there are $n \cdot 2^{n-1}$ many such committees.

Since each side counts the number of elements of one set, the values must be equal. Therefore,

$$\sum_{k=1}^{n} k \cdot \binom{n}{k} = n \cdot 2^{n-1}.$$

□

Proof 3. We give another proof using the Binomial Theorem, which tells us that

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

for all $x, y \in \mathbb{R}$. Plugging in $y = 1$, we conclude that

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

for all $x \in \mathbb{R}$. Now each side is a function of the real variable x, so taking the derivative of each side, it follows that

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

for all $x \in \mathbb{R}$. Plugging in $x = 1$, we conclude that

$$n \cdot 2^{n-1} = \sum_{k=1}^{n} k \cdot \binom{n}{k}$$

This completes the proof. □

Proposition 3.15. *If $k \leq n$, then*

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

and since $\binom{m}{k} = 0$ if $m < k$, it follows that

$$\sum_{m=0}^{n} \binom{m}{k} = \binom{n+1}{k+1}$$

Proof. Using Proposition 3.9 repeatedly, we have:

$$\begin{aligned}
\binom{n+1}{k+1} &= \binom{n}{k} + \binom{n}{k+1} \\
&= \binom{n}{k} + \binom{n-1}{k} + \binom{n-1}{k+1} \\
&= \binom{n}{k} + \binom{n-1}{k} + \binom{n-2}{k} + \binom{n-2}{k+1} \\
&= \binom{n}{k} + \binom{n-1}{k} + \binom{n-2}{k} + \cdots + \binom{k+2}{k+1} \\
&= \binom{n}{k} + \binom{n-1}{k} + \binom{n-2}{k} + \cdots + \binom{k+1}{k} + \binom{k+1}{k+1} \\
&= \binom{n}{k} + \binom{n-1}{k} + \binom{n-2}{k} + \cdots + \binom{k+1}{k} + \binom{k}{k}.
\end{aligned}$$

where the last line follows from the fact that

$$\binom{k+1}{k+1} = 1 = \binom{k}{k}.$$

□

Plugging in $k = 1$, we get

$$\binom{1}{1} + \binom{2}{1} + \binom{3}{1} + \cdots + \binom{n}{1} = \binom{n+1}{2}.$$

for all $n \in \mathbb{N}^+$. Since $\binom{m}{1} = m$ for all $m \in \mathbb{N}^+$, it follows that

$$1 + 2 + 3 + \cdots + n = \binom{n+1}{2} = \frac{n(n+1)}{2}.$$

for all $n \in \mathbb{N}^+$. Notice that letting $k = 2$, we conclude that that

$$\binom{2}{2} + \binom{3}{2} + \binom{4}{2} + \cdots + \binom{n}{2} = \binom{n+1}{3}.$$

for all $n \in \mathbb{N}^+$. Since $\binom{1}{2} = 0$, we can also write this as

$$\binom{1}{2} + \binom{2}{2} + \binom{3}{2} + \cdots + \binom{n}{2} = \binom{n+1}{3}.$$

Now we can use these to find a formula for the sum of the first n squares:

$$1^2 + 2^2 + 3^2 + \cdots + n^2$$

The idea is to find $A, B \in \mathbb{R}$ such that

$$m^2 = A \cdot \binom{m}{1} + B \cdot \binom{m}{2}$$

is true for all $m \in \mathbb{N}^+$, because if we can do this, then we can use the above summation formulas for the two sums that appear on the right. Since $\binom{m}{1} = m$ for all $m \in \mathbb{N}^+$, and

$$\binom{m}{2} = \frac{m(m-1)}{2}$$

for all $m \in \mathbb{N}^+$ (even for $m = 1$ because then both sides are 0), we want to find A and B such that:

$$m^2 = A \cdot m + B \cdot \frac{m(m-1)}{2}$$

for all $m \in \mathbb{N}^+$. Now

$$\begin{aligned} A \cdot m + B \cdot \frac{m(m-1)}{2} &= A \cdot m + B \cdot \frac{m^2 - m}{2} \\ &= \left(A - \frac{B}{2}\right) \cdot m + \frac{B}{2} \cdot m^2 \end{aligned}$$

so equating coefficients with $m^2 = 0 \cdot m + 1 \cdot m^2$, we want to solve the linear system:

$$\begin{array}{ccccc} A & - & \frac{1}{2} \cdot B & = & 0 \\ & & \frac{1}{2} \cdot B & = & 1 \end{array}$$

Now $A = 1$ and $B = 2$ as the unique solution to this system, so it follows that

$$m^2 = \binom{m}{1} + 2 \cdot \binom{m}{2}$$

is true for all $m \in \mathbb{N}^+$. Thus, using Proposition 3.15, we conclude that

$$\begin{aligned} 1^2 + 2^2 + \cdots + n^2 &= \left[\binom{1}{1} + 2 \cdot \binom{1}{2}\right] + \left[\binom{2}{1} + 2 \cdot \binom{2}{2}\right] + \cdots + \left[\binom{n}{1} + 2 \cdot \binom{n}{2}\right] \\ &= \left[\binom{1}{1} + \binom{2}{1} + \cdots + \binom{n}{1}\right] + 2 \cdot \left[\binom{1}{2} + \binom{2}{2} + \cdots + \binom{n}{2}\right] \\ &= \binom{n+1}{2} + 2 \cdot \binom{n+1}{3} \\ &= \frac{(n+1)n}{2} + 2 \cdot \frac{(n+1)n(n-1)}{6} \\ &= \frac{3(n+1)n}{6} + \frac{2(n+1)n(n-1)}{6} \\ &= \frac{n(n+1)(3+2n-2)}{6} \\ &= \frac{n(n+1)(2n+1)}{6} \end{aligned}$$

One can generalize these techniques to get the sum of the first n cubes. Doing so would require finding $A, B, C \in \mathbb{R}$ such that

$$m^3 = A \cdot \binom{m}{1} + B \cdot \binom{m}{2} + C \cdot \binom{m}{3}$$

for all $m \in \mathbb{N}^+$. Although it's not too onerous to do the algebra in order to set up the linear system, and then solve for A, B, C, we will see more unified ways to determine these coefficients (along with for fourth powers, etc.) soon.

Suppose that we want to pick out 5 days from the month of February (having 28 days) in such a way that we do not pick two consecutive days. How can we count it? Although we want to pick out an unordered subset, one idea is to first count the number of *ordered* choices, and then divide by 5!. The idea then is to pick out one day, and we have 28 choices. Once we've picked that day out, we then pick out a second day. It may appear that we have 25 choices here because we've eliminated one day and it's two neighbors. However, that it is only true if we did not pick out the first or last days of February in our first choice. Thus, the number of options in round two depends on our choice from round one. You might think about counting those sets including the first and/or last days of February as special cases, but this doesn't solve all of the problems. For example, if we choose 11 and 18 in our first two rounds, then we've eliminated 6 days and have 22 choices for the third round. However, if we choose 11 and 13 in our first two rounds, then we've only eliminated 5 days and so have 23 choices for the third round. In other words, we need a new way to count this.

Let's attack the problem from a different angle. Instead of trying to avoid bad configurations directly, we think about picking out an arbitrary subset of 5 days and "spreading" them to guarantee that the result will not have any consecutive days. To do this, we will leave the lowest numbered day alone, but add 1 to the second lowest day (to ensure we have a "gap" between the first two), and then add 2 to the middle day, etc. More formally, given an arbitrary subset $\{a_1, a_2, a_3, a_4, a_5\}$ of $[28]$ with $a_1 < a_2 < a_3 < a_4 < a_5$, we turn it into the subset $\{a_1, a_2 + 1, a_3 + 2, a_4 + 3, a_5 + 4\}$ which does not have any consecutive days. The only problem is that now we might "overflow". For example, although

$$\{3, 4, 15, 16, 21\} \mapsto \{4, 6, 17, 19, 25\}$$

works out just fine, we also have

$$\{1, 10, 21, 26, 27\} \mapsto \{1, 11, 23, 30, 31\}$$

which is not allowed. However, there's an easy fix. Instead of picking our original subset from [28], we pick it from [24], for a total of $\binom{24}{5}$ many possibilities. In general, we have the following:

Proposition 3.16. *The number of subsets of* $[n] = \{1, 2, 3, \ldots, n\}$ *of size* k *having no two consecutive numbers equals* $\binom{n-k+1}{k}$.

Proof. We establish a bijection between the k-element subsets of $[n - k + 1]$ and the sets we want. Given a subset $\{a_1, a_2, a_3, \ldots, a_k\}$ of $[n - k + 1]$ with $a_1 < a_2 < a_3 < \cdots < a_k$, we map it to the set $\{a_1, a_2 + 1, a_3 + 2, \ldots, a_k + (k - 1)\}$, i.e. the i^{th} element of the new set equals $a_i + (i - 1)$. Now since $a_i < a_{i+1}$ for each i, we have that $a_{i+1} - a_i \geq 1$ for each i, and hence

$$\begin{aligned} a_{i+1} + ((i+1) - 1) - (a_i + (i - 1)) &= a_{i+1} + i - a_i - i + 1 \\ &= a_{i+1} - a_i + 1 \\ &\geq 1 + 1 \\ &= 2 \end{aligned}$$

for i, so there are no consecutive elements in the resulting set. Furthermore, since $a_k \leq n - k + 1$, we have $a_k + (k-1) \leq n - k + 1 + (k - 1) = n$, so the resulting subset is indeed a subset of $[n]$ of size k having no two consecutive elements. Notice that this function is injective because if $\{a_1, a_2 + 1, a_3 + 2, \ldots, a_k + (k - 1)\} = \{b_1, b_2 + 1, b_3 + 2, \ldots, b_k + (k-1)\}$, then $a_i + (i-1) = b_i + (i-1)$ for all i, hence $a_i = b_i$ for all i. Furthermore, given a subset $\{c_1, c_2, c_3, \ldots, c_k\}$ of $[n]$ with $c_1 < c_2 < c_3 < \cdots < c_n$ and $c_{i+1} - c_i \geq 2$ for all i, we have that $\{c_1, c_2 - 1, c_3 - 2, \ldots, c_k - (k - 1)\}$ is a subset of $[n - k + 1]$ that maps to $\{c_1, c_2, c_3, \ldots, c_k\}$, so it is surjective. The result follows. □

What if we just wanted to count the number number of subsets of $[n]$ having no two consecutive numbers, without any size restrictions? One approach is to sum over all possible sizes to obtain:

$$\sum_{k=0}^{n} \binom{n-k+1}{k} = \binom{n+1}{0} + \binom{n}{1} + \binom{n-1}{2} + \binom{n-2}{3} + \cdots + \binom{1}{n}$$

Of course, many of the terms on the right equal 0 because if $k > n-k+1$, i.e. if $k > \frac{n+1}{2}$, then $\binom{n-k+1}{k} = 0$. Thus, if we let $\lfloor m \rfloor$ be the greatest integers less than or equal to m, then we have

$$\sum_{k=0}^{\lfloor \frac{n+1}{2} \rfloor} \binom{n-k+1}{k}$$

For example, the number of subsets of $[6] = \{1,2,3,4,5,6\}$ having no two consecutive numbers is

$$\begin{aligned} \sum_{k=0}^{3} \binom{7-k}{k} &= \binom{7}{0} + \binom{6}{1} + \binom{5}{2} + \binom{4}{3} \\ &= 1 + 6 + 10 + 4 \\ &= 21 \end{aligned}$$

while the number of subsets of $[7] = \{1,2,3,4,5,6,7\}$ having no two consecutive numbers is

$$\begin{aligned} \sum_{k=0}^{4} \binom{8-k}{k} &= \binom{8}{0} + \binom{7}{1} + \binom{6}{2} + \binom{5}{3} + \binom{4}{4} \\ &= 1 + 7 + 15 + 10 + 1 \\ &= 34. \end{aligned}$$

Notice that we are summing up diagonals of Pascal's triangle, and we are seeing Fibonacci numbers. You will prove that this holds true generally on the homework.

Returning to the Binomial Theorem, what happens if we look powers of $x+y+z$ instead of $x+y$? For example, we have

$$\begin{aligned} (x+y+z)^2 &= (x+y+z)(x+y+z) \\ &= x(x+y+z) + y(x+y+z) + z(x+y+z) \\ &= xx + xy + xz + yx + yy + yz + zx + zy + zz \end{aligned}$$

Thus, we obtain a sum of $9 = 3 \cdot 3$ terms, where each term is an ordered product of two elements (with repetition) from $\{x, y, z\}$. If we work out $(x+y+z)^3$, we see a sum of $27 = 3 \cdot 3 \cdot 3$ terms, where each possible ordered sequence of 3 elements (with repetition) from $\{x, y, z\}$ appears exactly once. In general, one expects that we expand $(x+y+z)^n$, then we obtain a sequence of 3^n many terms where each possible ordered sequence of n elements (with repetition) from $\{x, y, z\}$ appears exactly once. What happens when we use collapse these sums by using commutativity, so write $xxz + xzx + zxx$ as $3x^2z$? In general, we are asking what the coefficient of $x^a y^b z^c$ will be in the result? Notice that we need only examine the coefficients where $a + b + c = n$ because each term involves a product of n of the variables. Suppose then that $a + b + c = n$. To know the coefficient of $x^a y^b z^c$, we want to know the number of sequences of x's, y's, and z's of length n having exactly a many x's, b many y's, and c many z's. To count these, we can first pick out that a positions in which to place the x's in $\binom{n}{a}$ many ways. Next, we have $n - a$ open positions, and need to pick out b positions to place the y's in $\binom{n-a}{b}$ many ways. Finally, we have $n - a - b = c$ many positions for the c many

z's, so they are completely determined. Thus, if $a+b+c=n$, then the coefficient of $x^a y^b z^c$ in $(x+y+z)^n$ equals

$$\begin{aligned}\binom{n}{a}\cdot\binom{n-a}{b} &= \frac{n!}{a!\cdot(n-a)!}\cdot\frac{(n-a)!}{b!\cdot(n-a-b)!}\\ &= \frac{n!}{a!\cdot b!\cdot(n-a-b)!}\\ &= \frac{n!}{a!\cdot b!\cdot c!}.\end{aligned}$$

More generally, suppose that we expand $(x_1+x_2+\cdots+x_k)^n$. In the result, we will have a sum of term of the form $x_1^{a_1}x_2^{a_2}\cdots x_k^{a_k}$ where the $a_i\in\mathbb{N}$ and $a_1+a_2+\cdots+a_k=n$. To determine the coefficient of such a term, we need only determine the number of sequences of x_i of length n such that there are exactly a_1 many x_1's, exactly a_2 many x_2's, ..., and exactly a_k many x_k's. Following the above template, the number of such sequences equals

$$\begin{aligned}&\binom{n}{a_1}\cdot\binom{n-a_1}{a_2}\cdot\binom{n-a_1-a_2}{a_3}\cdots\binom{n-a_1-a_2-\cdots-a_{k-2}}{a_{k-1}}\cdot\binom{n-a_1-a_2-\cdots-a_{k-1}}{a_k}\\ &\quad=\binom{n}{a_1}\cdot\binom{n-a_1}{a_2}\cdot\binom{n-a_1-a_2}{a_3}\cdots\binom{n-a_1-a_2-\cdots-a_{k-2}}{a_{k-1}}\cdot\binom{a_k}{a_k}\\ &\quad=\frac{n!}{a_1!\cdot(n-a_1)!}\cdot\frac{(n-a_1)!}{a_2!\cdot(n-a_1-a_2)!}\cdot\frac{(n-a_1-a_2)!}{a_3!\cdot(n-a_1-a_2-a_3)!}\cdots\frac{(n-a_1-a_2-\cdots-a_{k-1})!}{a_{k-1}!\cdot(n-a_1-a_2-\cdots-a_{k-2}-a_{k-1})!}\\ &\quad=\frac{n!}{a_1!\cdot a_2!\cdot a_3!\cdots a_{k-1}!\cdot(n-a_1-a_2-\cdots-a_{k-2}-a_{k-1})!}\\ &\quad=\frac{n!}{a_1!\cdot a_2!\cdot a_3!\cdots a_{k-1}!\cdot a_k!}\end{aligned}$$

Notice that this is just like our problem with anagrams of MISSISSIPPI. Instead of doing the above count, we could have treated all the x_1 as different (and x_2 as different, etc.), rearranged them in $n!$ many ways, and then divided by the overcount from the permuting the x_i within themselves in $a_1!$ ways, the x_2 within themselves in $a_2!$ many ways, etc.

Definition 3.17. *If $n, a_1, a_2, \dots, a_k \in \mathbb{N}$ and $a_1+a_2+\cdots+a_k=n$, we define*

$$\binom{n}{a_1,a_2,\dots,a_k}=\frac{n!}{a_1!\cdot a_2!\cdots a_k!}$$

We call this a multinomial coefficient*.*

The above argument proves the generalization of the Binomial Theorem:

Theorem 3.18 (Multinomial Theorem)**.** *For all $n,k\in\mathbb{N}^+$, we have*

$$(x_1+x_2+\cdots+x_k)^n=\sum\binom{n}{a_1,a_2,\dots,a_k}x_1^{a_1}x_2^{a_2}\cdots x_k^{a_k}$$

where the sum is taken over all k-tuples of nonnegative integers $(a_1,a_2,\dots,a_k)$ such that $a_1+a_2+\cdots+a_k=n$.

3.3 Compositions and Partitions

3.3.1 Compositions

There are six different M&M colors: Red, Yellow, Blue, Green, Orange, Brown. Suppose that we want to pick out 13 total M&M's. How ways can you do it, if all that matters is how many of each color we take?

Notice that we can model this as follows: if we let a_i be the number that you choose with color i, then we need $a_1 + a_2 + a_3 + a_4 + a_5 + a_6 = 13$.

Definition 3.19. *Let $n, k \in \mathbb{N}$. A sequence of nonnegative integers $(a_1, a_2, \ldots, a_k)$ such that $a_1 + a_2 + \cdots + a_k = n$ is called a* weak composition *of n into k parts. If all the a_i are positive, then it is called a composition of k into k parts.*

For example $(1, 3, 5, 3)$ is a composition of 12 into 4 parts and $(2, 0, 5, 1, 0, 0)$ is a weak composition of 8 into 6 parts.

One can view the number of weak compositions of n into k parts as the number of ways to distribute n identical balls into k distinct boxes. In this interpretation, the value a_i is the number of balls that we put into box i. We are treating the balls as identical because all that matters are the number of balls in each box, but the boxes are distinct because $(5, 2, 1)$ is different than $(2, 5, 1)$.

We can also view these another way. Recall that a k-permutation of n distinct objects is a way to pick out k of those objects where order matters and repetition is not allowed. Also, a k-combination of n distinct objects is a way to pick out k of those objects where order does not matter and repetition is not allowed. A different way to interpret a weak composition of n into k parts is as a way to pick out n objects from k distinct objects where order doesn't matter but repetition *is* allowed (yes, the n and k have switched, and this is incredibly annoying). The value a_i is the number of times that we pick out object i. Due to the fact that order doesn't matter but repetition is allowed, some sources think about something they call *multisets*. The idea is to allow one to write something like "$\{1, 1, 4\}$" and think about it as different from "$\{1, 4\}$", but the same as "$\{1, 4, 1\}$". Since, by definition, two sets are equal exactly when they have the same elements, we should introduce new notation rather than $\{$ and $\}$ used in sets. Instead of dealing with all of these, we write $(2, 0, 0, 1)$ to represent that we picked the number 1 twice and the number 4 once.

The number of weak compositions of n into k parts is the number of nonnegative integer solutions to the equation

$$x_1 + x_2 + \cdots + x_k = n$$

while the number of compositions of n into k parts is the number of positive integer solutions to the equation

$$x_1 + x_2 + \cdots + x_k = n.$$

How do we count the number of weak compositions of n into k parts? In the M&M case, think about lining them up in order of color, so red first, then yellow, etc. If we eliminate the colors from the M&M's themselves, then we only need some kind of "marker" to distinguish when we change colors. If we represent the M&M's as dots, then we can place 5 bars to denote the dividing line as to when we switch colors. Since we have 5 bars and 13 M&M's that we have to put into a line, we have18 positions and need to choose the positions for the 5 bars. Therefore, there are $\binom{18}{5}$ many possibilities.

Proposition 3.20. *Let $n, k \in \mathbb{N}$. The number of weak compositions of n into k parts is*

$$\binom{n+k-1}{k-1} = \binom{n+k-1}{n}.$$

Proof. As above, there is a bijection between arrangements of n dots and $k-1$ bars into a line and weak compositions of n into k parts (the number of dots before the first bar is a_1, then number of dots between the first and second is a_2, etc.). Since we want to place $n+k-1$ many objects and need only choose the $k-1$ positions for the bars, or alternatively the n positions for the dots. Therefore, the number of weak compositions of n into k parts equals

$$\binom{n+k-1}{k-1} = \binom{n+k-1}{n}.$$

□

Another way to visualize this is as follows: Consider the following bijection between subsets of $[n+k-1]$ of size n and weak compositions of n into k parts: Given a subset $\{a_1, a_2, \ldots, a_n\}$ of $[n+k-1]$ with $a_1 < a_2 < \cdots < a_n$, consider the multiset "$\{a_1, a_2 - 1, a_3 - 2, \ldots, a_n - (n-1)\}$" and form the corresponding weak composition. For example if $k = 5$ and $n = 3$, then $n + k - 1 = 7$ and we do the following:

$$\begin{aligned}
\{1,2,3\} &\mapsto \text{“}\{1,1,1\}\text{”} \mapsto (3,0,0,0,0)\\
\{1,3,7\} &\mapsto \text{“}\{1,2,5\}\text{”} \mapsto (1,1,0,0,1)\\
\{3,4,6\} &\mapsto \text{“}\{3,3,4\}\text{”} \mapsto (0,0,2,1,0)
\end{aligned}$$

More formally, given a subset $\{a_1, a_2, \ldots, a_n\}$ of $[n+k-1]$ with $a_1 < a_2 < \cdots < a_n$, we send it the sequence $(b_1, b_2, \ldots, b_k)$ where b_ℓ equals the number of i such that $a_i - (i-1) = \ell$, i.e. the cardinality of the set $\{i : a_i = i + \ell - 1\}$.

Now that we've determined the number of *weak* compositions of n into k parts, we can answer the count the number of compositions of n into k parts. The idea is that if $k \leq n$, then the number of positive integer solutions to the equation

$$x_1 + x_2 + \cdots + x_k = n$$

equals to the number of nonnegative solutions to

$$x_1 + x_2 + \cdots + x_k = n - k$$

Corollary 3.21. *Let $n, k \in \mathbb{N}$ with $k \leq n$. The number of compositions of n into k parts equals*

$$\binom{n-1}{k-1} = \binom{n-1}{n-k}.$$

Proof. First distribute one ball to each of the k boxes . We now have $n-k$ balls to put into k boxes with no restrictions, and so we want to count the number of weak compositions of $n-k$ into k parts. The answer to this is:

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

Since $(n-1)-(k-1) = n-k$, this also equals

$$\binom{n-1}{n-k}.$$

More formally, given a weak composition $(a_1, a_2, \ldots, a_k)$ of $n-k$ into k parts, the sequence $(a_1+1, a_2+1, \ldots, a_k+1)$ is composition of n into k parts, and this mapping is a bijection. □

Another way to visualize the previous corollary with a direct bijection is as follows: Consider the function

$$(a_1, a_2, \ldots, a_k) \mapsto \{a_1, a_1 + a_2, \ldots, a_1 + a_2 + \cdots + a_{k-1}\}$$

from compositions of n into k parts to $(k-1)$-element subsets of $[n-1]$. For example if $n = 10$ and $k = 4$, then

$$\begin{aligned}
(1,2,3,4) &\mapsto \{1,3,6\}\\
(6,1,1,2) &\mapsto \{6,7,8\}\\
(2,1,1,6) &\mapsto \{2,3,4\}
\end{aligned}$$

Notice that since $a_i \geq 1$ for all i, we have $a_1 < a_1 + a_2 < \cdots < a_1 + a_2 + \cdots + a_{k-1}$. Now since $a_1 + a_2 + \cdots + a_k = n$ and $a_k \geq 1$, it follows that $a_1 + a_2 + \cdots + a_{k-1} \leq n - 1$, and hence the set on the right is an element of $[n-1]$. Finally, one must check that this is a bijection, but I'll leave that to you (since we already have a proof of the result).

What happens if we try to count *all* compositions of a number n without specifying the number of parts? For example, we have 4 compositions of 3 given by (3), (1, 2), (2, 1), and (1, 1, 1). The compositions of 4 are (4), (1, 3), (3, 1), (2, 2), (2, 1, 1), (1, 2, 1), (1, 1, 2), and (1, 1, 1, 1), so we have 8 of those.

Theorem 3.22. *The number of compositions of n is 2^{n-1}.*

Proof. We give two proofs. The first is to notice that a composition of n must be a composition of n into k parts for some unique k with $1 \leq k \leq n$. Therefore, the number of compositions of n equals

$$\begin{aligned}\sum_{k=1}^{n} \binom{n-1}{k-1} &= \binom{n-1}{0} + \binom{n-1}{1} + \binom{n-1}{2} + \cdots + \binom{n-1}{n-1} \\ &= \sum_{k=0}^{n-1} \binom{n-1}{k} \\ &= 2^{n-1}. && \text{(by Corollary 3.11)}\end{aligned}$$

Alternatively, we can give a direct combinatorial proof. Write down n dots. Notice that we can not put a bar before the first dot or after the last one, and we also can not put two bar in the same place because in a composition all numbers must be positive. Therefore, a composition arises by picking a subset of the $n-1$ spaces between the dots to serve as bars (i.e. the dividers). Since there are 2^{n-1} many subsets of a set with $n-1$ many elements, it follows that there are 2^{n-1} many compositions of n. □

3.3.2 Set Partitions

Above we considered the case where the balls were identical and the boxes were distinct. Now consider the case where the balls are distinct but the boxes are identical.

Definition 3.23. *A (set)* partition *of a set A is a set $\{B_1, B_2, \ldots, B_k\}$ where the B_i are nonempty pairwise disjoint subsets of A with*

$$A = B_1 \cup B_2 \cup \cdots \cup B_k$$

In this case, we call this a partition of A into k nonempty parts.

Definition 3.24. *Given $n, k \in \mathbb{N}^+$ with $k \leq n$, we define $S(n,k)$ to be the number of partitions of $[n]$ into k nonempty parts. The numbers $S(n,k)$ are called the* Stirling numbers of the second kind *and are sometimes denoted by:*

$$S(n,k) = \left\{ {n \atop k} \right\}.$$

We also define $S(0,0) = 1$, $S(n,0) = 0$ if $n \geq 1$, and $S(n,k) = 0$ if $k > n$.

For example, we have $S(3,2) = 3$ because the following are all possible partitions of $[3] = \{1,2,3\}$ into 2 parts:

- $\{\{1\},\{2,3\}\}$
- $\{\{2\},\{1,3\}\}$
- $\{\{3\},\{1,2\}\}$

Notice that these are all of them because if we partition [3] into 2 parts, then one must have size 1 and the other have size 2, so the partition is completely determined by the choice of the the element that is in its own block (and hence there are $\binom{3}{1} = 3$ many choices).

Here are few more examples:

- If $n \geq 1$, then
$$S(n, 1) = 1 = S(n, n)$$
because the only partition on $[n]$ into one part is $\{\{1, 2, 3, \ldots, n\}\}$ and the only partition into n parts is $\{\{1\}, \{2\}, \ldots, \{n\}\}$.
- We have $S(4, 3) = \binom{4}{2} = 6$ because such a partition must have one set of size 2 and the others of size 1, so we need only choose the subset of size 2.
- More generally, for any $n \geq 2$, we have
$$S(n, n-1) = \binom{n}{2}$$
because a partition of $[n]$ into $n - 1$ many blocks must have one block of size 2 and $n - 2$ of size 1, so we need to pick the two unique elements for the block of size 2.
- The number $S(4, 2)$ is more interesting. We can partition $\{1, 2, 3, 4\}$ into a set of size 3 and a set of size 1, or into two sets of size 2. There are $\binom{4}{1} = 4$ ways to do the former because we need only pick the element in the set of size 1. For the latter, there are 3 possibilities:
    - $\{\{1, 2\}, \{3, 4\}\}$
    - $\{\{1, 3\}, \{2, 4\}\}$
    - $\{\{1, 4\}, \{2, 3\}\}$

  Therefore, $S(4, 2) = 4 + 3 = 7$.

In general, the numbers $S(n, k)$ are difficult to compute. Recall that
$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$
whenever $k, n \in \mathbb{N}^+$. We get a similar recurrence here.

Theorem 3.25. *For all $k, n \in \mathbb{N}^+$ with $k \leq n$, we have*
$$S(n, k) = S(n-1, k-1) + k \cdot S(n-1, k).$$
In other words, if $k \leq n$, then
$$\left\{ n \atop k \right\} = \left\{ n-1 \atop k-1 \right\} + k \cdot \left\{ n-1 \atop k \right\}.$$

Proof. A partition of $[n]$ into k parts is of one of two possible types:

- *Case 1:* The number n is in a block by itself. If we remove the block $\{n\}$, then we are left with a partition of $[n-1]$ into $k - 1$ parts, so there are $S(n-1, k-1)$ many such possibilities. Notice that every partition of $[n]$ into k blocks having $\{n\}$ as one of the blocks arises in a unique way from such a partition of $[n-1]$ into $k - 1$ parts. Thus, there are $S(n-1, k-1)$ many partitions of this type.

- *Case 2:* The number n is not in its own block. If we remove n from its block, we then obtain a partition of $[n-1]$ into k parts, and there are $S(n-1,k)$ many possible outcomes. Notice that each of these outcomes arise in k many ways because given a partition of $[n-1]$ into k blocks, we can add n into any of the blocks to obtain a partition of $[n]$ into k parts. Therefore, there are $k \cdot S(n-1,k)$ many partitions of this type.

It follows that $S(n,k) = S(n-1,k-1) + k \cdot S(n-1,k)$. □

We now get a triangle like Pascal's triangle, but with $S(n,k) = \left\{ {n \atop k} \right\}$ in place of $C(n,k) = \binom{n}{k}$.

$\left\{ {n \atop k} \right\}$	0	1	2	3	4	5	6	7
0	1	0	0	0	0	0	0	0
1	0	1	0	0	0	0	0	0
2	0	1	1	0	0	0	0	0
3	0	1	3	1	0	0	0	0
4	0	1	7	6	1	0	0	0
5	0	1	15	25	10	1	0	0
6	0	1	31	90	65	15	1	0
7	0	1	63	301	350	140	21	1

Given $n,k \in \mathbb{N}^+$, recall that we have the following:

- The number of functions $f\colon [n] \to [k]$ equals k^n because for each $i \in [n]$, we have k possibilities for the value of $f(i)$.
- If $k < n$, then there are no injective functions $f\colon [n] \to [k]$ by the Pigeonhole Principle.
- If $k > n$, then the number of injective functions $f\colon [n] \to [k]$ equal $k(k-1)(k-2)\cdots(k-n+1) = (k)_n = \frac{k!}{(k-n)!}$ because we have k choices for the value of $f(1)$, then $k-1$ for the value of $f(2)$, $\dots$, and finally $k-(n-1)$ for the value of $f(n)$.

Proposition 3.26. *Given $n,k \in \mathbb{N}^+$, there are exactly $k! \cdot S(n,k)$ many surjective functions $f\colon [n] \to [k]$.*

Proof. If $k > n$, then there are no surjective functions $f\colon [n] \to [k]$, and $k! \cdot S(n,k) = k! \cdot 0 = 0$. Suppose then that $k < n$. Consider a surjective $f\colon [n] \to [k]$. For each $c \in [k]$, let $B_c = \{a \in [n] : f(a) = c\}$, i.e. B_c is the set of all elements of $[n]$ than map to c. Since f is surjective, we know that $B_c \neq \emptyset$ for all $c \in [k]$. Furthermore, since f is a function, the sets $B_1, B_2, \dots, B_k$ are pairwise disjoint, and $[n] = B_1 \cup B_2 \cup \cdots \cup B_k$. Therefore, $\{B_1, B_2, \dots, B_k\}$ is a partition of $[n]$. Notice that each of these partitions arise in $k!$ many ways because we can reorder the B_i in terms of their outputs, i.e. if $n = 4$ and $k = 2$ then $\{\{1,4\},\{2,3\}\}$ is a partition arising from both the function

$$f(1) = 1 \quad f(2) = 2 \quad f(3) = 2 \quad f(4) = 1$$

and the function

$$f(1) = 2 \quad f(2) = 1 \quad f(3) = 1 \quad f(4) = 2$$

In other words, every surjective function arises uniquely from a partition of $[n]$ together with a permutation of $[k]$. Therefore, the number of surjective functions $f\colon [n] \to [k]$ equals $k! \cdot S(n,k)$. □

Theorem 3.27. *For all $m,n \in \mathbb{N}^+$, we have*

$$m^n = \sum_{k=1}^{n} k! \cdot S(n,k) \cdot \binom{m}{k}$$

i.e.

$$m^n = \sum_{k=1}^{n} k! \cdot \left\{ {n \atop k} \right\} \cdot \binom{m}{k}$$

Proof. The left-hand side m^n is simply the number of functions from $[n]$ to $[m]$. The key fact is that given any function $f \colon A \to B$, if we let $C = \text{range}(f)$, then we can view f as a surjective function $f \colon A \to C$. Thus, every function $f \colon [n] \to [m]$ can be viewed as a surjective function onto some nonempty subset of $[m]$. Now given a subset $X \subseteq [m]$ with $|X| = k$, we know from the previous proposition that there are $k! \cdot S(n,k)$ many surjections from $[n]$ to X. For a fixed k, there are $\binom{m}{k}$ many subsets of $[m]$ of size k, so there are $\binom{m}{k} \cdot k! \cdot S(n,k)$ many functions from $[n]$ to $[m]$ whose range has size k. Summing over all possible sizes for the range, we conclude that the number of functions from $[n]$ to $[m]$ equals

$$\sum_{k=1}^{n} k! \cdot \left\{ {n \atop k} \right\} \cdot \binom{m}{k}.$$

Therefore,

$$m^n = \sum_{k=0}^{n} \binom{m}{k} \cdot k! \cdot S(n,k).$$

□

In particular, we have

$$\begin{aligned} m^2 &= 1 \cdot 1 \cdot \binom{m}{1} + 2 \cdot 1 \cdot \binom{m}{2} \\ &= 1 \cdot \binom{m}{1} + 2 \cdot \binom{m}{2} \end{aligned}$$

for all $m \in \mathbb{N}$ as we learned above. We also have

$$\begin{aligned} m^3 &= 1 \cdot 1 \cdot \binom{m}{1} + 2 \cdot 3 \cdot \binom{m}{2} + 6 \cdot 1 \cdot \binom{m}{3} \\ &= 1 \cdot \binom{m}{1} + 6 \cdot \binom{m}{2} + 6 \cdot \binom{m}{3} \end{aligned}$$

and

$$\begin{aligned} m^4 &= 1 \cdot 1 \cdot \binom{m}{1} + 2 \cdot 7 \cdot \binom{m}{2} + 6 \cdot 6 \cdot \binom{m}{3} + 24 \cdot 1 \cdot \binom{m}{4} \\ &= 1 \cdot \binom{m}{1} + 14 \cdot \binom{m}{2} + 36 \cdot \binom{m}{3} + 24 \cdot \binom{m}{4} \end{aligned}$$

for all $m \in \mathbb{N}$. Using these formulas together with Proposition 3.15, we can now develop formulas for the sum of the first n cubes, the first n fourth powers, etc. For example, we have

$$\begin{aligned}
\sum_{m=1}^{n} m^3 &= \sum_{m=1}^{n} \left(1 \cdot \binom{m}{1} + 6 \cdot \binom{m}{2} + 6 \cdot \binom{m}{3}\right) \\
&= \sum_{m=1}^{n} \binom{m}{1} + 6 \cdot \sum_{m=1}^{n} \binom{m}{2} + 6 \cdot \sum_{m=1}^{n} \binom{m}{3} \\
&= \binom{n+1}{2} + 6 \cdot \binom{n+1}{3} + 6 \cdot \binom{n+1}{4} \\
&= \frac{(n+1)n}{2} + 6 \cdot \frac{(n+1)n(n-1)}{6} + 6 \cdot \frac{(n+1)n(n-1)(n-2)}{24} \\
&= \frac{(n+1)n}{2} + (n+1)n(n-1) + \frac{(n+1)n(n-1)(n-2)}{4} \\
&= \frac{(n+1)n}{4} \cdot (2 + 4(n-1) + (n-1)(n-2)) \\
&= \frac{(n+1)n}{4} \cdot (2 + 4n - 4 + n^2 - 3n + 2) \\
&= \frac{(n+1)n}{4} \cdot (n^2 + n) \\
&= \frac{(n+1)^2 n^2}{4} \\
&= \left(\frac{n(n+1)}{2}\right)^2
\end{aligned}$$

Therefore, we obtain the surprising result that

$$\sum_{m=1}^{n} m^3 = \left(\sum_{m=1}^{n} m\right)^2$$

for all $n \in \mathbb{N}^+$.

Definition 3.28. *Let $n \in \mathbb{N}$. The number of all partitions of $[n]$ into nonempty parts is denoted by $B(n)$ and is called the n^{th} Bell number. We also define $B(0) = 0$. Notice that*

$$B(n) = \sum_{k=0}^{n} S(n,k) = \sum_{k=0}^{n} \left\{ {n \atop k} \right\}$$

for all $n \in \mathbb{N}$.

Recall than an equivalence relation on A induces a partition of A into nonempty parts through the equivalence classes. Conversely, it's not hard to show that if $\{B_1, B_2, \ldots, B_k\}$ is a partition of A with each $B_i \neq \emptyset$, then then the relation $a \sim b$ if there exists an i with $a, b \in B_i$ is an equivalence relation on A whose equivalence classes are the B_i. Therefore, $B(n)$ equals the number of equivalence relations on a set of size n.

Adding up the rows of the above table, we get

n	$B(n)$
0	1
1	1
2	2
3	5
4	15
5	52
6	203
7	877

Theorem 3.29. *For any $n \in \mathbb{N}$, we have*

$$B(n+1) = \sum_{k=0}^{n} \binom{n}{k} \cdot B(k)$$

Proof. We need to argue that the right-hand side counts the number of partitions of $[n+1]$. We look at the block containing $n+1$. We examine how many elements are *not* in the block containing $n+1$. If there are k such elements, then there are $\binom{n}{k}$ many ways to choose these elements (and hence choose the $n-k$ many elements of $[n]$ grouped with $n+1$) and then $B(k)$ many ways to partition them. Thus,

$$B(n+1) = \sum_{k=0}^{n} \binom{n}{k} \cdot B(k)$$

Alternatively, we can count as follows. If that block has has k, then there are $\binom{n}{k-1}$ many ways to choose the other elements in the block, and then $B(n+1-k)$ many ways to partition the rest. Thus

$$\begin{aligned} B(n+1) &= \sum_{k=1}^{n+1} \binom{n}{k-1} \cdot B(n+1-k) \\ &= \sum_{k=0}^{n} \binom{n}{k} \cdot B(n-k) \\ &= \sum_{k=0}^{n} \binom{n}{n-k} \cdot B(k) \\ &= \sum_{k=0}^{n} \binom{n}{k} \cdot B(k) \end{aligned}$$

□

3.3.3 Integer Partitions

We've seen that compositions correspond to ways to distribute n identical balls to k distinct boxes in such a way that each box receives at least one ball. Also, (set) partitions correspond to ways to distribute n distinct balls to k identical boxes in such a way that each box receives at least one ball. We now introduce (integer) partitions that correspond to ways to distribute n identical balls to k identical boxes in such a way that each box receives at least one ball.

Definition 3.30. *An* (integer) partition *of an $n \in \mathbb{N}$ into k parts is a composition $(a_1, a_2, \ldots, a_k)$ of n where $a_1 \geq a_2 \geq \cdots \geq a_k$. The number of partitions of n into k parts is denoted by $p(n,k)$. We also define $p(0,0) = 1$*

Notice that $p(n,0) = 0$ if $n \geq 1$ and $p(n,k) = 0$ if $k > n$. We have $p(4,2) = 2$ because $(2,2)$ and $(3,1)$ are the only partitions. Notice that $p(7,3) = 1$ because the partitions are $(5,1,1)$, $(4,2,1)$, $(3,3,1)$, and $(3,2,2)$.

Definition 3.31. *The number of partitions of n (into any number of parts) is denoted by $p(n)$, so*

$$p(n) = \sum_{k=0}^{n} p(n,k).$$

Theorem 3.32. *For all $n, k \in \mathbb{N}$ with $0 < k < n$, we have*

$$\begin{aligned} p(n,k) &= \sum_{i=1}^{k} p(n-k,i) \\ &= p(n-k,1) + p(n-k,2) + \cdots + p(n-k,k) \end{aligned}$$

Proof. Suppose that you have a partition of n into k parts. Remove 1 from each of these parts. This gives a partition of $n-k$ into some number of parts which is at most k. Since $k < n$, something must be left. □

Thus, we have

$$p(7,3) = p(4,1) + p(4,2) + p(4,3)$$

and

$$p(7,4) = p(3,1) + p(3,2) + p(3,3) + p(3,4) = p(3,1) + p(3,2) + p(3,3)$$

$*$	0	1	2	3	4	5	6	7
0	1	0	0	0	0	0	0	0
1	0	1	0	0	0	0	0	0
2	0	1	1	0	0	0	0	0
3	0	1	1	1	0	0	0	0
4	0	1	2	1	1	0	0	0
5	0	1	2	2	1	1	0	0
6	0	1	3	3	2	1	1	0
7	0	1	3	4	3	2	1	1

Adding up the rows, we get

n	$p(n)$
0	1
1	1
2	2
3	3
4	5
5	7
6	11
7	15

The question of how fast $p(n)$ grows is extremely interesting and subtle. It turns out that

$$p(n) \sim \frac{1}{4\sqrt{3}} \cdot \exp\left(\pi\sqrt{\frac{2n}{3}}\right)$$

where $\exp(x) = e^x$ and $f(n) \sim g(n)$ means that

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1.$$

3.4 Inclusion-Exclusion

Recall that if A and B are any finite sets, then

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

What about three sets, i.e. if we wanted to count $|A \cup B \cup C|$? A natural guess would be that we need to subtract off the various intersection, so one might hope that $|A \cup B \cup C|$ equals

$$|A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C|$$

Let's examine if this is correct. Notice that if $x \in A$, but $x \notin B$ and $x \notin C$, then x contributes 1 to $|A \cup B \cup C$ and in the formula it contributes

$$1 + 0 + 0 - 0 - 0 - 0 = 1.$$

Similar arguments words if x is in only B, or x is in only C. Let's examine what happens if x is in two of the sets, say $x \in A$, $x \in B$, but $x \notin C$. Again, x contributes 1 to $|A \cup B \cup C|$, and in the formula it contributes

$$1 + 1 + 0 - 1 - 0 - 0 = 1.$$

Again, everything looks good so far. Finally, suppose that x is an element of each of A, B, and C. As usual, x contributes 1 to $|A \cup B \cup C|$, but in the formula it contributes

$$1 + 1 + 1 - 1 - 1 - 1 = 0.$$

Thus, elements that are if $A \cap B \cap C$ are not counted at all on the right-hand side. To correct this, we need to add it back in. We then claim that the correct formula is

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|.$$

Working though each of the possibilities, one can check that this count is correct to matter where x lies in the Venn diagram of sets. How does this generalize? For four sets, one can show by working through all of the cases that

$$\begin{aligned} |A_1 \cup A_2 \cup A_3 \cup A_4| = |A_1| &+ |A_2| + |A_3| + |A_4| \\ &- |A_1 \cap A_2| - |A_1 \cap A_3| - |A_1 \cap A_4| - |A_2 \cap A_3| - |A_2 \cap A_4| - |A_3 \cap A_4| \\ &+ |A_1 \cap A_2 \cap A_3| + |A_1 \cap A_2 \cap A_4| + |A_1 \cap A_3 \cap A_4| + |A_2 \cap A_3 \cap A_4| \\ &- |A_1 \cap A_2 \cap A_3 \cap A_4|. \end{aligned}$$

It's tedious to check them all possibilities here, and so would like a way to prove that this works in general. We'll do that below, but first we'll demonstrate how to use these formulas to count an interesting set. For our example, we will count the number of primes less than or equal to 120. Before jumping into this, we prove a few small but important facts.

Proposition 3.33. *Let $n \in \mathbb{N}^+$ with $n \geq 2$. If n is not prime, then there is a prime p such that $p \mid n$ and $p \leq \sqrt{n}$.*

Proof. Suppose that n is not prime. Since n is not prime, we can fix $d \in \mathbb{N}$ with $1 < d < n$ such that $d \mid n$. Fix $c \in \mathbb{Z}$ with $cd = n$. Notice that $c > 0$ because both $d > 0$ and $n > 0$, and moreover we must have $1 < c < n$ (if $c = 1$ then $d = n$, and if $c = n$ then $d = 1$). Now at least one of $c \leq \sqrt{n}$ or $c \leq \sqrt{n}$ must be true because otherwise $n = cd > \sqrt{n} \cdot \sqrt{n} = n$. In either case, this number has a prime divisor less than or equal to it, so n has a prime prime divisor p with $p \leq \sqrt{n}$. □

Proposition 3.34. *If $p \in \mathbb{Z}$ is prime and $p \mid ab$, then either $p \mid a$ or $p \mid b$.*

Proof. Suppose that $p \mid ab$ and $p \nmid a$. Since $\gcd(a, p)$ divides p and we know that $p \nmid a$, we have $\gcd(a, p) \neq p$. The only other positive divisor of p is 1, so $\gcd(a, p) = 1$. Therefore, by the Proposition 2.22, we conclude that $p \mid b$. □

Now that we've handled the product of two numbers, we get the following corollary about finite products by a trivial induction.

Corollary 3.35. *If p is prime and $p \mid a_1 a_2 \cdots a_n$, then $p \mid a_i$ for some i.*

Proposition 3.36. *If $a, b, c \in \mathbb{Z}$ are such that $a \mid c$, $b \mid c$, and $\gcd(a, b) = 1$, then $ab \mid c$.*

Proof. Since $a \mid c$ and $b \mid c$, we may fix $m, n \in \mathbb{Z}$ with $c = am$ and $c = bn$. Since $\gcd(a, b) = 1$, we may fix $k, \ell \in \mathbb{Z}$ with $ak + b\ell = 1$. Multiplying this equation through by c, we have $cak + cb\ell = c$. Therefore

$$\begin{aligned} c &= cak + cb\ell \\ &= (bn)ak + (am)b\ell \\ &= ab(nk + m\ell) \end{aligned}$$

Since $nk + m\ell \in \mathbb{Z}$, it follows that $ab \mid n$. □

Proposition 3.37. *Let $a \in \mathbb{Z}$ and let $p_1, p_2, \ldots, p_k$ be distinct primes. If $p_i \mid a$ for all i, then $p_1 p_2 \cdots p_k \mid a$.*

Proof. We prove the result by induction on k. Notice that if $k = 1$, then the statement is trivial. Suppose that we know the statement is true for a fixed $k \in \mathbb{N}$. Let $p_1, p_2, \ldots, p_k, p_{k+1}$ be distinct primes with the property that $p_i \mid a$ for all i. By induction, we know that $p_1 p_2 \cdots p_k \mid a$. We also have that $p_{k+1} \mid a$ by assumption. Now if $p_{k+1} \mid p_1 p_2 \cdots p_k$ then we would have $p_{k+1} \mid p_i$ for some $i \leq k$ by Corollary 3.35, so either $p_{k+1} = 1$ or $p_{k+1} = p_i$, a contradiction (because the p_i are distinct primes). Since $p_{k+1} \nmid p_1 p_2 \cdots p_k$, we must have $\gcd(p_1 p_2 \cdots p_k, p_{k+1}) = 1$. Since both $p_1 p_2 \cdots p_k \mid a$ and $p_{k+1} \mid a$, we may use Proposition 3.36 to conclude that $p_1 p_2 \cdots p_k p_{k+1} \mid a$. This completes the induction. □

We now return to counting the number of primes in [120]. By Proposition 3.33, if $a \in [120]$ is not prime and $a \geq 2$, then a is divisible by prime less than or equal to $\sqrt{a} \leq \sqrt{120}$. Now the primes less than or equal to $\sqrt{120} < 11$ are 2, 3, 5, and 7. We thus let $p_1 = 2$, $p_2 = 3$, $p_3 = 5$, and $p_4 = 7$. For each i, let A_i be the set of numbers in [120] divisible by p_i. We count

$$|A_1 \cup A_2 \cup A_3 \cup A_4|$$

which is the number of elements of [120] that are divisible by at least one 2, 3, 5, or 7. We have

$$|A_1| = \frac{120}{2} = 60 \qquad |A_2| = \frac{120}{3} = 40 \qquad |A_3| = \frac{120}{5} = 24 \qquad |A_4| = \left\lfloor \frac{120}{7} \right\rfloor = 17$$

To determine the cardinalities of intersections, we use Proposition 3.37. For example, the numbers divisible by both 2 and 3 are the numbers divisible by 6. Working these out, we conclude that

$$|A_1 \cap A_2| = \frac{120}{6} = 20 \qquad |A_1 \cap A_3| = \frac{120}{10} = 12 \qquad |A_1 \cap A_4| = \left\lfloor \frac{120}{14} \right\rfloor = 8$$

$$|A_2 \cap A_3| = \frac{120}{15} = 8 \qquad |A_2 \cap A_4| = \left\lfloor \frac{120}{21} \right\rfloor = 5 \qquad |A_3 \cap A_4| = \left\lfloor \frac{120}{35} \right\rfloor = 3$$

Next we compute

$$|A_1 \cap A_2 \cap A_3| = \frac{120}{30} = 4 \qquad |A_1 \cap A_2 \cap A_4| = \left\lfloor \frac{120}{42} \right\rfloor = 2$$

$$|A_1 \cap A_3 \cap A_4| = \frac{120}{70} = 1 \qquad\qquad |A_2 \cap A_3 \cap A_4| = \left\lfloor \frac{120}{105} \right\rfloor = 1$$

and

$$|A_1 \cap A_2 \cap A_3 \cap A_4| = \left\lfloor \frac{120}{210} \right\rfloor = 0$$

Thus

$$|A_1 \cup A_2 \cup A_3 \cup A_4| = (60 + 40 + 24 + 17) - (20 + 12 + 8 + 8 + 5 + 3) + (4 + 2 + 1 + 1) - 0 = 93$$

By the Complement Rule, it follows that there are

$$120 - 93 = 27$$

many numbers in [120] that not divisible by any of 2, 3, 5, or 7. All of these except 1 are prime, so this gives 26 new primes in [120]. Adding back in the primes 2, 3, 5, and 7, we see that there are 30 primes in [120].

Theorem 3.38 (Inclusion-Exclusion). *Let $A_1, A_2, \dots, A_n$ be finite sets. We then have*

$$\begin{aligned} |A_1 \cup A_2 \cup \dots \cup A_n| &= \sum_{S \subseteq [n] \backslash \{\emptyset\}} (-1)^{|S|-1} \cdot |\bigcap_{i \in S} A_i| \\ &= \sum_{k=1}^{n} (-1)^{k-1} \sum_{S \subseteq [n], |S| = k} |\bigcap_{i \in S} A_i| \end{aligned}$$

Less formally, this says that

$$|A_1 \cup A_2 \cup \dots \cup A_n| = \sum_i |A_i| - \sum_{i<j} |A_i \cap A_j| + \sum_{i<j<k} |A_i \cap A_j \cap A_k| - \dots$$

Proof. Suppose that $x \in A_1 \cup A_2 \cup \dots \cup A_n$. Let $T = \{i \in [n] : x \in A_i\}$, i.e. T is the nonempty set of indices i such that $x \in A_i$. Let $k = |T|$ and notice that $k \geq 1$. We examine the number of times that x is counted on each side. On the left, x contributes 1 to the cardinality. On the right, it contributes

$$\binom{k}{1} - \binom{k}{2} + \binom{k}{3} - \dots + (-1)^{k-1}\binom{k}{k}$$

to the sum. Now from Corollary 3.12, we know that

$$\binom{k}{0} - \binom{k}{1} + \binom{k}{2} - \binom{k}{3} + \dots - (-1)^{k}\binom{k}{k} = 0.$$

Hence

$$\binom{k}{1} - \binom{k}{2} + \binom{k}{3} - \dots + (-1)^{k-1}\binom{k}{k} = \binom{k}{0} = 1$$

Therefore, every $x \in A_1 \cup A_2 \cup \dots \cup A_n$ contributes 1 to both sides. The result follows. □

We next count the number of surjections $f\colon [n] \to [k]$. Of course we know that the answer is $k! \cdot S(n,k)$ from by Proposition 3.26,, but we count it in a different way using Inclusion-Exclusion (from which we will be able to derive a formula for $S(n,k)$). We first illustrate the general argument in the special case where $n = 7$ and $k = 4$, i.e. we count the number of surjections $f\colon [7] \to [4]$. The idea is to count the complement. We know that there are 4^7 many total functions $f\colon [7] \to [4]$, so we count the number of functions that are *not* surjective. Now a function can fail to be a surjective by missing 1, missing 2, missing 3, or missing 4.

Thus, given $i \in [4]$, we let A_i be the set of functions $f\colon [7] \to [4]$ such that $i \notin \text{range}(f)$. Then the set of functions $f\colon [7] \to [4]$ that are not surjective equals $A_1 \cup A_2 \cup A_3 \cup A_4$. Now we know that:

$$\begin{aligned}|A_1 \cup A_2 \cup A_3 \cup A_4| = |A_1| &+ |A_2| + |A_3| + |A_4| \\ &- |A_1 \cap A_2| - |A_1 \cap A_3| - |A_1 \cap A_4| - |A_2 \cap A_3| - |A_2 \cap A_4| - |A_3 \cap A_4| \\ &+ |A_1 \cap A_2 \cap A_3| + |A_1 \cap A_2 \cap A_4| + |A_1 \cap A_3 \cap A_4| + |A_2 \cap A_3 \cap A_4| \\ &- |A_1 \cap A_2 \cap A_3 \cap A_4|.\end{aligned}$$

To count $|A_1|$, we need to count the number of functions $f\colon [7] \to [4]$ such that $1 \notin \text{range}(f)$. This is just the number of functions $f\colon [7] \to \{2,3,4\}$, which equals 3^7. Similarly, $|A_2| = |A_3| = |A_4| = 3^7$. To count $|A_1 \cap A_2|$, we just need to count the number of functions $f\colon [7] \to [4]$ such that $1, 2 \notin \text{range}(f)$. This is just the number of functions $f\colon [7] \to \{3,4\}$, which equals 2^7. Following through on this, we conclude that

$$\begin{aligned}|A_1 \cup A_2 \cup A_3 \cup A_4| &= 3^7 + 3^7 + 3^7 + 3^7 \\ &- 2^7 - 2^7 - 2^7 - 2^7 - 2^7 - 2^7 \\ &+ 1^7 + 1^7 + 1^7 + 1^7 + 1^7 \\ &- 0.\end{aligned}$$

so

$$|A_1 \cup A_2 \cup A_3 \cup A_4| = 4 \cdot 3^7 - 6 \cdot 2^7 + 4 \cdot 1^7$$

Notice that coefficients are $\binom{4}{1} = 4$, $\binom{4}{2} = 6$, and $\binom{4}{3} = 4$ because $\binom{4}{m}$ is the number of ways to pick out m elements from [4]. It follows that the number of surjective functions $f\colon [7] \to [4]$ equals

$$4^7 - (4 \cdot 3^7 - 6 \cdot 2^7 + 4 \cdot 1^7) = 4^7 - 4 \cdot 3^7 + 6 \cdot 2^7 - 4 \cdot 1^7 = 8,400.$$

We now generalize this argument.

Theorem 3.39. *Let $n, k \in \mathbb{N}^+$ with $k \leq n$. The number of surjections $f\colon [n] \to [k]$ is*

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

Proof. The total number of functions $f\colon [n] \to [k]$ is k^n. For each $i \in [k]$, let A_i be the set of all functions $f\colon [n] \to [k]$ such that $i \notin \text{range}(f)$. We then have that

$$A_1 \cup A_2 \cup \cdots \cup A_k$$

is the set of all functions which are *not* surjective, and we count

$$|A_1 \cup A_2 \cup \cdots \cup A_k|$$

using Inclusion-Exclusion. Suppose that $S \subseteq [k]$ with $|S| = m$. We then have that

$$\bigcap_{i \in S} A_i$$

is the set of functions whose range is contained in $[k] \backslash S$, so since $|[k] \backslash S| = k - m$, it follows that

$$|\bigcap_{i \in S} A_i| = (k - |S|)^n = (k-m)^n$$

Therefore

$$\begin{aligned}
|A_1 \cup A_2 \cup \cdots \cup A_k| &= \sum_{S \subseteq [k] \setminus \{\emptyset\}} (-1)^{|S|-1} \cdot |\bigcap_{i \in S} A_i| \\
&= \sum_{m=1}^{k} (-1)^{m-1} \sum_{S \subseteq [k], |S|=m} |\bigcap_{i \in S} A_i| \\
&= \sum_{m=1}^{k} (-1)^{m-1} \sum_{S \subseteq [k], |S|=m} (k - |S|)^n \\
&= \sum_{m=1}^{k} (-1)^{m-1} \binom{k}{m} (k-m)^n
\end{aligned}$$

where the last line follows from the fact that $\binom{k}{m}$ is the number of subsets of $[k]$ of cardinality m. Thus, the number of surjections $f\colon [n] \to [k]$ is

$$\begin{aligned}
k^n - \sum_{m=1}^{k} (-1)^{m-1} \binom{k}{m} (k-m)^n &= k^n + \sum_{m=1}^{k} (-1)^{m} \binom{k}{m} (k-m)^n \\
&= \sum_{m=0}^{k} (-1)^{m} \binom{k}{m} (k-m)^n
\end{aligned}$$

□

Corollary 3.40. *Let $n, k \in \mathbb{N}^+$ with $k \leq n$. We have*

$$\begin{aligned}
S(n,k) &= \frac{1}{k!} \sum_{m=0}^{k} (-1)^m \binom{k}{m} (k-m)^n \\
&= \sum_{m=0}^{k} (-1)^m \frac{(k-m)^n}{m! \cdot (k-m)!}
\end{aligned}$$

Proof. We know that the number of surjections $f\colon [n] \to [k]$ equals $k! \cdot S(n,k)$ by Proposition 3.26, and it also equals

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

by Theorem 3.39. Therefore,

$$k! \cdot S(n,k) = \sum_{m=0}^{k} (-1)^m \binom{k}{m} (k-m)^n$$

and hence

$$\begin{aligned}
S(n,k) &= \frac{1}{k!} \sum_{m=0}^{k} (-1)^m \binom{k}{m} (k-m)^n \\
&= \sum_{m=0}^{k} (-1)^m \frac{(k-m)^n}{m! \cdot (k-m)!}
\end{aligned}$$

□

For example, since

$$\sum_{m=0}^{4}(-1)^m\binom{4}{m}(4-m)^7 = 8,400$$

form above, we have

$$S(7,4) = \frac{8,400}{24} = 350.$$

Definition 3.41. *We define a function $\varphi\colon \mathbb{N}^+ \to \mathbb{N}^+$ as follows. For each $n \in \mathbb{N}^+$, we let*

$$\varphi(n) = |\{a \in [n] : \gcd(a,n) = 1\}|$$

The function φ is called the Euler φ-function or Euler totient function.

For example, we have the following:

- $\varphi(1) = 1$ because $\gcd(1,1) = 1$.
- $\varphi(4) = 2$ because 1 and 3 are the only elements in $[4]$ that are relatively prime with 4.
- $\varphi(5) = 4$ because $1,2,3,4$ are all relatively prime with 5, but $\gcd(5,5) \neq 1$.
- $\varphi(6) = 2$ because 1 and 5 are the only elements in $[6]$ that are relatively prime with 6.
- $\varphi(p) = p-1$ for all primes p because if $1 \leq a < p$, then $\gcd(a,p) = 1$.

Proposition 3.42. *If $p \in \mathbb{N}$ is prime and $k \in \mathbb{N}^+$, then*

$$\varphi(p^k) = p^k - p^{k-1} = p^{k-1}(p-1) = p^k \cdot \left(1 - \frac{1}{p}\right)$$

Proof. Let

$$A = \{a \in [p^k] : \gcd(a,p^k) = 1\}$$

By definition, we have that $\varphi(p^k) = |A|$, so we need to count how many elements are in A. To this, we count the complement. In other words, we determine the cardinality of

$$B = [p^k]\backslash A = \{a \in [p^k] : \gcd(a,p^k) \neq 1\}.$$

Our claim is that

$$B = \{pm : 1 \leq m \leq p^{k-1}\}$$

Suppose first that $a = pm$ where $1 \leq m \leq p^{k-1}$. Notice that since $1 \leq m \leq p^{k-1}$, we have $p \leq pm \leq pp^{k-1}$, which is to say that $p \leq a \leq p^k$, so $a \in [p^k]$. Now we clearly have that $p \mid a$ because $a = pm$, and we also have $p \mid p^k$ because $p = pp^{k-1}$ and $k-1 \geq 0$, so $\gcd(n,p^k) \neq 1$ because $p > 1$ is a common divisor. It follows that $a \in B$, and since a was arbitrary we conclude that $\{pm : 1 \leq m \leq p^{k-1}\} \subseteq B$.

Suppose conversely that $a \in B$. Let $d = \gcd(a,p^k)$, so since $a \in B$ we know that $d > 1$. Now the only positive divisors of p^k are $1, p, p^2, \dots, p^k$, so since $d \mid p$ and $d \neq 1$, we know that $d \in \{p, p^2, \dots, p^k\}$. Since p divides every element of this set, it follows that $p \mid d$. Now we also know that $d \mid a$, so by transitivity of the divisibility relation it follows that $p \mid a$. Thus, we fix $m \in \mathbb{Z}$ with $a = pm$. Notice that $m > 0$ because $a > 0$ and $p > 0$. Finally, we must have $m \leq p^{k-1}$ because otherwise $m > p^{k-1}$ and so $a = pm > p^k$, contradicting our assumption that $a \in B$. Therefore, $B \subseteq \{pm : 1 \leq m \leq p^{k-1}\}$.

We've shown that $B = \{pm : 1 \leq m \leq p^{k-1}\}$. Now the set on the right has p^{k-1} many elements (one for each choice of m), so $|B| = p^{k-1}$. It follows that

$$\varphi(p^k) = |[p^k]\backslash B| = p^k - p^{k-1}$$

The latter two formulas are not just simple algebra. □

Suppose that $n = pq$ where p and q are distinct primes. Notice that if $a \in [n]$ with $\gcd(a, n) \neq 1$, then a must be divisible by either p or q (or both). There are $\frac{pq}{p} = q$ many elements divisible by p, and $\frac{pq}{q} = p$ many elements divisible by q. Also, there is one element divisible by both. Hence, there are $p + q - 1$ many elements divisible by at least one of p or q, and hence

$$\begin{aligned}\varphi(pq) &= pq - (p + q - 1) \\ &= pq - p - q + 1 \\ &= (p - 1)(q - 1)\end{aligned}$$

Theorem 3.43. *Suppose that $n \in \mathbb{N}$ with $n \geq 2$. Write $n = p_1^{k_1} p_2^{k_2} \cdots p_\ell^{k_\ell}$ where the p_i are distinct primes. We then have*

$$\begin{aligned}\varphi(n) &= n \cdot \left(1 - \frac{1}{p_1}\right) \cdot \left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_\ell}\right) \\ &= p_1^{k_1}\left(1 - \frac{1}{p_1}\right) \cdot p_2^{k_2}\left(1 - \frac{1}{p_2}\right) \cdots p_\ell^{k_\ell}\left(1 - \frac{1}{p_\ell}\right) \\ &= p_1^{k_1}(p_1 - 1) \cdot p_2^{k_2}(p_2 - 1) \cdot p_\ell^{k_\ell}(p_\ell - 1) \\ &= \varphi(p_1^{k_1}) \cdot \varphi(p_2^{k_2}) \cdots \varphi(p_\ell^{k_\ell})\end{aligned}$$

Proof. For each $i \in [\ell]$, let

$$A_i = \{a \in [n] : p_i \mid a\}$$

We calculate

$$|A_1 \cup A_2 \cup \cdots \cup A_k|$$

and notice that this is the set of numbers which are *not* relatively prime to n. We have

$$|A_i| = \frac{n}{p_i}$$

For $i < j$ we have

$$|A_i \cap A_j| = \frac{n}{p_i p_j}$$

while for $i < j < k$ we have

$$|A_i \cap A_j \cap A_k| = \frac{n}{p_i p_j p_k}$$

and so on. Thus

$$|A_1 \cup A_2 \cup \cdots \cup A_k| = \sum_i \frac{n}{p_i} - \sum_{i<j} \frac{n}{p_i p_j} + \sum_{i<j<k} \frac{n}{p_i p_j p_k} - \dots$$

It follows that

$$\begin{aligned}\varphi(n) &= n - \sum_i \frac{n}{p_i} + \sum_{i<j} \frac{n}{p_i p_j} - \sum_{i<j<k} \frac{n}{p_i p_j p_k} + \dots \\ &= n \cdot \left(1 - \sum_i \frac{1}{p_i} + \sum_{i<j} \frac{1}{p_i p_j} - \sum_{i<j<k} \frac{1}{p_i p_j p_k} + \dots\right) \\ &= n \cdot \left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_\ell}\right)\end{aligned}$$

□

For example, we have

$$504 = 2^3 \cdot 3^2 \cdot 7$$

so

$$\begin{aligned}\varphi(504) &= \varphi(2^3) \cdot \varphi(3^2) \cdot \varphi(7) \\ &= 4(2-1) \cdot 3(3-1) \cdot (7-1) \\ &= 4 \cdot 6 \cdot 6 \\ &= 144\end{aligned}$$

Definition 3.44. *A* derangement *of $[n]$ is a permutation $(a_1, a_2, \dots, a_n)$ of $[n]$ such that $a_i \neq i$ for all i.*

For example, $(3, 1, 4, 2)$ is a derangement of $[4]$, but $(3, 2, 4, 1)$ is not (because $a_2 = 2$).

Theorem 3.45. *Let $n \in \mathbb{N}^+$. The number of derangements of $[n]$ is*

$$n! \cdot \sum_{k=0}^{n} \frac{(-1)^k}{k!}$$

Proof. We know that there are $n!$ many permutations of $[n]$. For each $i \in [n]$, let A_i be the set of all permutations $(a_1, a_2, \dots, a_n)$ of $[n]$ such that $a_i = i$. We then have that

$$A_1 \cup A_2 \cup \cdots \cup A_n$$

is the set of all functions which are *not* derangements. We count

$$|A_1 \cup A_2 \cup \cdots \cup A_n|$$

using Inclusion-Exclusion. Suppose that $S \subseteq [n]$ with $|S| = k$. We then have that

$$\bigcap_{i \in S} A_i$$

is the set of permutations of $[n]$ such that $a_i = i$ for all $i \in S$. To count this, notice that k of the elements are determined, and the remaining $n - k$ elements can be permuted in the remaining $n - k$ spots arbitrarily, so

$$|\bigcap_{i \in S} A_i| = (n - |S|)! = (n - k)!$$

We then have

$$\begin{aligned}|A_1 \cup A_2 \cup \cdots \cup A_n| &= \sum_{S \subseteq [n] \backslash \{\emptyset\}} (-1)^{|S|-1} \cdot |\bigcap_{i \in S} A_i| \\ &= \sum_{k=1}^{n} (-1)^{k-1} \sum_{S \subseteq [n], |S| = k} |\bigcap_{i \in S} A_i| \\ &= \sum_{k=1}^{n} (-1)^{k-1} \sum_{S \subseteq [n], |S| = k} (n-k)! \\ &= \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} (n-k)! \\ &= \sum_{k=1}^{n} (-1)^{k-1} \frac{n!}{k!} \\ &= n! \cdot \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k!}\end{aligned}$$

Thus, the number of derangements of $[n]$ is

$$n! - n! \cdot \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k!} = n! \cdot \sum_{k=0}^{n} \frac{(-1)^k}{k!}$$

□

Notice that the fraction of permutations that are derangements equals

$$\begin{aligned}\sum_{k=0}^{n} \frac{(-1)^k}{k!} &= 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^n}{n!} \\ &= \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^n}{n!}\end{aligned}$$

For example, when $n = 6$, we have

$$\frac{1}{2} - \frac{1}{6} + \frac{1}{24} - \frac{1}{120} + \frac{1}{720} = \frac{53}{144} \approx .36806$$

so approximately 36.8% of the permutations are derangements. When $n = 7$, we have

$$\frac{53}{144} - \frac{1}{5040} = \frac{1854}{5040} = \frac{103}{280} \approx .36786$$

so again about 36.8% of the permutations are derangements. Now if you've seen infinite series, then you know that

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} - \dots$$

for all $x \in \mathbb{R}$. In particular, when $x = -1$, we have

$$\begin{aligned}e^{-1} &= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \\ &= 1 - (-1) + \frac{(-1)^2}{2!} - \frac{(-1)^3}{3!} + \frac{(-1)^4}{4!} - \dots \\ &= \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \dots\end{aligned}$$

Therefore, as n gets large, the percentage of permutations of $[n]$ that are derangements approaches the number

$$1/e \approx .36788$$

3.5 Permutations

Recall that given a finite set A with $|A| = n$, we defined a permutation of A to be an element of A^n without repeated elements. Consider the case where $A = [n]$. In this situation, we can view a permutation of A differently. Instead of thinking about the finite sequence $(a_1, a_2, \dots, a_n)$, we can think about the function $\sigma\colon [n] \to [n]$ defined by letting $\sigma(i) = a_i$ for all i. For example, if $n = 6$, then we can think about the permutation $(5, 6, 3, 1, 4, 2)$ instead as the function $\sigma\colon [6] \to [6]$ defined by:

- $\sigma(1) = 5$
- $\sigma(2) = 6$

- $\sigma(3) = 3$
- $\sigma(4) = 1$
- $\sigma(5) = 4$
- $\sigma(6) = 2$

Notice that since a permutation of $[n]$ does not have any repeated elements, it follows that every element of $[n]$ appears exactly once as an output of the corresponding function, and hence the corresponding function is a bijection. Conversely, given a bijection $\sigma\colon [n] \to [n]$, the sequence $(\sigma(1), \sigma(2), \dots, \sigma(n))$ is a permutation of $[n]$. In other words, permutations of $[n]$ and bijections from $[n]$ to $[n]$ are really the same thing.

Rather than list out the values of the function as we did in our example above, we can instead write out the values in a table. For example, for our σ above, we can write it as:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 6 & 3 & 1 & 4 & 2 \end{pmatrix}$$

In this table, entries on the top row are input values and corresponding entries on the bottom row are the output values. Notice that bottom row is simply our original sequence. We call $(5, 6, 3, 1, 4, 2)$ (or 563142 if we want to be even more compact) the *one-line notation* of σ and we call the above table the *two-line notation* of σ.

At this point, you may wonder why we care about viewing permutations as functions or in two-line notation. The primary answer is that functions can be *composed*. Recall that the composition of two bijections is a bijection by Proposition 1.38, so the composition of two permutations of $[n]$ is again a permutation of $[n]$. For example, consider the following two permutations of $[6]$:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 6 & 3 & 1 & 4 & 2 \end{pmatrix} \qquad \tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 1 & 5 & 6 & 2 & 4 \end{pmatrix}$$

Let's compute $\sigma \circ \tau$. Remember that function composition happens from right to left. That is, the composition $\sigma \circ \tau$ is obtained by performing τ first and following after by performing σ. For example, we have

$$(\sigma \circ \tau)(2) = \sigma(\tau(2)) = \sigma(1) = 5$$

Working through the 6 inputs, we obtain:

$$\sigma \circ \tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 5 & 4 & 2 & 6 & 1 \end{pmatrix}$$

On the other hand, we have

$$\tau \circ \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 5 & 3 & 6 & 1 \end{pmatrix}$$

Notice that $\sigma \circ \tau \neq \tau \circ \sigma$. Remember that function composition is not commutative in general!

Given a permutation σ of $[n]$, we can define $\sigma^2 = \sigma \circ \sigma$, $\sigma^3 = \sigma \circ \sigma \circ \sigma$, etc. Notice that since function composition of associative by Proposition 1.33, we do not need to insert parentheses in things like σ^3 because we know that $\sigma \circ (\sigma \circ \sigma) = (\sigma \circ \sigma) \circ \sigma$. We now show that if we start with $i \in [n]$ and repeatedly apply σ, we eventually cycle back around to i.

Proposition 3.46. *Let $\sigma\colon [n] \to [n]$ be a permutation and let $i \in [n]$. There exists $k \in \mathbb{N}^+$ with $1 \leq k \leq n$ such that $\sigma^k(i) = i$. Moreover, if k is the least positive integer with $\sigma^k(i) = i$, then the numbers*

$$i \quad \sigma(i) \quad \sigma^2(i) \quad \sigma^3(i) \quad \dots \quad \sigma^{k-1}(i)$$

are distinct

Proof. We first show that there exists $k \in \mathbb{N}^+$ with $1 \leq k \leq n$ such that $\sigma^k(i) = i$. Consider the first $n+1$ many numbers

$$\sigma(i) \quad \sigma^2(i) \quad \sigma^3(i) \quad \dots \quad \sigma^n(i) \quad \sigma^{n+1}(i)$$

Since we have a list of $n+1$ numbers, and only n possible values for those numbers, then there must exist $\ell < m$ with $\sigma^\ell(i) = \sigma^m(i)$ by the Pigeonhole Principle. Since $\sigma^m(i) = \sigma^\ell(\sigma^{m-\ell}(i))$, it follows that $\sigma^\ell(\sigma^{m-\ell}(i)) = \sigma^\ell(i)$. Now using the fact that σ^ℓ is injective (because it is a permutation as mentioned above), it follows that $\sigma^{m-\ell}(i) = i$. Since $1 \leq m - \ell \leq n$, we have shown the existence of positive $k \in \mathbb{N}$ with $\sigma^k(i) = i$.

Suppose now that k is the least positive integer with $\sigma^k(i) = i$. Assume that there is a repeat in the list:

$$i \quad \sigma(i) \quad \sigma^2(i) \quad \sigma^3(i) \quad \dots \quad \sigma^{k-1}(i)$$

We may then fix $0 \leq \ell < m \leq k$ with $\sigma^\ell(i) = \sigma^m(i)$. As above, this implies that $\sigma^{m-\ell}(i) = i$. Since $0 < m - \ell < k$, this would contradict the minimality of k. Therefore, we must have that the above values are distinct. □

With this proposition in mind, we now develop a new notation to represent permutations called *cycle notation*. The basic idea is to take an element of $[n]$ and follow its path through σ. For example, let's work with our

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 6 & 3 & 1 & 4 & 2 \end{pmatrix}$$

We begin by starting with 1, and notice that $\sigma(1) = 5$. Now instead of moving on to deal with 2, let's continue this thread and determine the value $\sigma(5)$. Looking above, we see that $\sigma(5) = 4$. If we continue on this path to investigate 4, we see that $\sigma(4) = 1$, and we have found a "cycle" $1 \to 5 \to 4 \to 1$ hidden inside σ. We will denote this cycle with the notation $(1\ 5\ 4)$. Now that those numbers are taken care of, we start again with the smallest number not yet claimed, which in this case is 2. We have $\sigma(2) = 6$ and following up gives $\sigma(6) = 2$. Thus, we have found the cycle $2 \to 6 \to 2$ and we denote this by $(2\ 6)$. We have now claimed all numbers other than 3, and when we investigate 3 we see that $\sigma(3) = 3$, so we form the sad lonely cycle (3). Putting this all together, we write σ in cycle notation as

$$\sigma = (1\ 5\ 4)(2\ 6)(3).$$

Notice that Proposition 3.46 justifies why we never "stuck" when trying to build these cycles. When we start with 1 and follow the path, we can not repeat a number before coming back to 1. For example, we will never see $1 \to 3 \to 6 \to 2 \to 6$ because then the purported permutation must send both 3 and 2 to 6, which would violate the fact that the purported permutation is injective. Also, if we finish a few cycles and start up a new one, then it is not possible that our new cycle has any elements in common with previous ones. For example, if we already have the cycle $1 \to 3 \to 2 \to 1$ and we start with 4, we can't find $4 \to 5 \to 3$ because then both 1 and 5 would map to 3.

Our conclusion is that this process of writing down a permutation in cycle notation never gets stuck and results in writing the given permutation as a product of disjoint cycles. Working through the same process with the permutation

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 1 & 5 & 6 & 2 & 4 \end{pmatrix}$$

we see that in cycle notation we have

$$\tau = (1\ 3\ 5\ 2)(4\ 6).$$

Now we can determine $\sigma \circ \tau$ in cycle notation directly from the cycle notations of σ and τ. For example, suppose we want to calculate the following:

$$(1\ 2\ 4)(3\ 6)(5) \circ (1\ 6\ 2)(3\ 5\ 4)$$

We want to determine the cycle notation of the resulting function, so we first need to determine where it sends 1. Again, function composition happens from right to left. Looking at the function represented on the right, we see the cycle containing 1 is (1 6 2), so the right function sends 1 to 6. We then go to the function on the left and see where it sends 6 The cycle containing 6 there is (3 6), so it takes 6 and sends it to 3. Thus, the composition sends 1 to 3. Thus, our result starts out as

$$(1\ 3$$

Now we need to see what happens to 3. The function on the right sends 3 to 5, and the function on the left takes 5 and leave it alone, so we have

$$(1\ 3\ 5$$

When we move on to see what happens to 5, we notice that the right function sends it to 4 and then the left function takes 4 to 1. Since 1 is the first element the cycle we started, we now close the loop and have

$$(1\ 3\ 5)$$

We now pick up the least element not in the cycle and continue. Working it out, we end with:

$$(1\ 2\ 4)(3\ 6)(5) \circ (1\ 6\ 2)(3\ 5\ 4) = (1\ 3\ 5)(2)(4\ 6)$$

Notice that cycle notation is not unique. For example, if $n = 4$, then (1 2 3 4) and (2 3 4 1) both represent the same function, namely the function that sends 1 to 2, 2 to 3, 3 to 4, and 4 to 1. In general, we can always "cyclically shift" a cycle without changing the actual function. Also, notice that (1 2)(3 4) = (3 4)(1 2), so we can also swap the ordering of the disjoint cycles.

Let's examine the possible cycle types for permutations of [4], along with the number of permutations each types.

- One 4-cycle, such as (1 2 3 4): There are two ways to count the number of 4-cycles. One approach is to list the elements of [4] in order in 4! ways, but realize that we are over counting because we can cyclically shift each result in 4 ways to arrive at the same permutation. Thus, there are $\frac{4!}{4} = 6$ many 4-cycles. Alternatively, we can say that any 4-cycle can be shifted uniquely to put the 1 first, at which point we have $3! = 6$ many ways to arrange the three numbers after it.

- One 3-cycle and one 1-cycle, such as (1 2 3)(4): There are $4 \cdot 2 = 8$ many such permutations because we need to choose the unique element that is in the 1-cycle in 4 possible ways, and then choose the 3-cycle in $\frac{3!}{3} = 2$ ways as in the argument for 4-cycles.

- Two 2-cycles, such as (1 2)(3 4): This one is a bit tricky. We can pick two element to go in one of the cycles in $\binom{4}{2} = 6$ many ways, and once we pick this the other cycle is completely determined. However, notice that we count each of these permutations twice with this method, because if we pick $\{1, 2\}$ then we are describing the permutation (1 2)(3 4), while if we pick $\{3, 4\}$, then we are describing the permutation (3 4)(1 2) = (1 2)(3 4) as well. In other words, we can't pick the "first" 2-cycle because we can list the cycles in either order. Therefore, we need to divide by 2 to handle the overcount, and so there are 3 possibilities here. Alternatively, one can notice that such a permutation is completely determined by the element that is in the cycle with 1, and we have 3 choices.

- One 2-cycle and two 1-cycles, such as (1 2)(3)(4): In this case, we need only pick the two elements of the 2-cycle (noting that order does not matter), and there are $\binom{4}{2} = 6$ many possibilities.

- Four 1-cycles, such as (1)(2)(3)(4): There is only 1 of these.

Notice that

$$6 + 8 + 3 + 6 + 1 = 24$$

as we expect because there are $4! = 24$ many permutations of [4].

Definition 3.47. *Let $k, n \in \mathbb{N}$ with $k \leq n$. The number of permutations of $[n]$ with exactly k total cycles is denoted by $c(n,k)$ and is called the* signless (or unsigned) Stirling numbers of the first kind. *Alternatively, these numbers are sometimes denoted by:*

$$c(n,k) = \begin{bmatrix} n \\ k \end{bmatrix}$$

We also define $c(0,0) = 1$, $c(n,0) = 0$ if $n \geq 1$, and $c(n,k) = 0$ if $k > n$.

For example, our above calculations show the following:

- $c(4,1) = 6$.
- $c(4,2) = 8 + 3 = 11$.
- $c(4,3) = 6$.
- $c(4,4) = 1$.

In general, we have the following values:

- $c(n,n) = 1$ for all $n \in \mathbb{N}^+$ because the only permutation of $[n]$ with n many cycles is the one where all elements of $[n]$ are fixed.
- $c(n,1) = (n-1)!$ for all $n \in \mathbb{N}^+$ because a permutation of $[n]$ with only 1 cycles must be a cycle of length n, and we can count this by looking at all $n!$ many ways to list the elements, and then divide by n for the n many cyclic shifts. Alternatively, we can place 1 at the front of the cycle, and then order the other $n-1$ elements in all possible $(n-1)!$ many ways afterwards.
- $c(n, n-1) = \binom{n}{2}$ (which also equals $S(n, n-1)$) for all $n \geq 2$. To see this, simply notice that a permutation of $[n]$ has exactly $n-1$ many cycles if and only if it consists of $n-2$ many 1-cycles and 2-cycles. Such a permutation is completely determined by the 2 elements in the 2-cycle.

Although we were able to directly calculate $c(4,k)$ for each k, it becomes more difficult to compute values like $c(9,3)$ because such a permutation may have three 3-cycles, or one 7-cycles and two 1-cycles, or a 5-cycles and 2-cycles, or a 2-cycle, 3-cycle, and 4-cycles, etc. Rather than attempting to calculate these values directly by looking at all possible cases, we now develop a recurrence similar to the one for the binomial coefficients and Stirling numbers of the second kind.

Theorem 3.48. *Let $k, n \in \mathbb{N}^+$ with $k \leq n$. We have*

$$c(n,k) = c(n-1,k-1) + (n-1) \cdot c(n-1,k)$$

In other words, if $k \leq n$, then

$$\begin{bmatrix} n \\ k \end{bmatrix} = \begin{bmatrix} n-1 \\ k-1 \end{bmatrix} + (n-1) \cdot \begin{bmatrix} n-1 \\ k \end{bmatrix}$$

Proof. We need to show that $c(n-1,k-1) + (n-1) \cdot c(n-1,k)$ counts the number of permutations of $[n]$ with exactly k cycles. We do this by considering two cases.

- Consider those permutations of $[n]$ with exactly k cycles in which n forms a 1-cycle by itself, i.e. where n is a fixed point of the permutation. Since n forms its own cycle, if we remove it, then the rest of permutation must be a permutation of $[n-1]$ with exactly $k-1$ cycles. Furthermore, every permutation of $[n-1]$ with exactly $k-1$ cycles arises uniquely in this way. Therefore, the number of permutations of $[n]$ with exactly k cycles in which n forms a 1-cycle by itself is $c(n-1,k-1)$.

- Now consider those permutations of $[n]$ with exactly k cycles in which n does not form a 1-cycle by itself, i.e. where n is not a fixed point of the permutation. If we simply delete n from the cycle notation, we obtain a permutation of $[n-1]$ with exactly k cycles. For example, if $n=8$ and $k=4$, and we have the permutation
$$(1\ 3\ 4)(2\ 8)(5\ 7)(6)$$
then by deleting 8 we arrive at the permutation
$$(1\ 3\ 4)(2)(5\ 7)(6)$$
Notice that we can also arrive at this latter permutation by deleting 8 from
$$(1\ 3\ 8\ 4)(2)(5\ 7)(6).$$
The key fact is that that every permutation of $[n-1]$ into exactly k cycles arises in $n-1$ ways from this process, because given a permeation of $[n-1]$ into exactly k cycles, we can insert n into the permutation after any of the numbers in cycle notation. Therefore, the number of permutations of $[n]$ with exactly k cycles in which n does not form a 1-cycle by itself equals $(n-1)\cdot c(n-1,k)$.

Since we have broken up the set of all permutations of $[n]$ with exactly k cycles into the disjoint union of two sets, it follows that $c(n,k) = c(n-1,k-1) + (n-1)\cdot c(n-1,k)$. □

Using this recurrence, we can compute the following table of values:

$c(n,k)$	0	1	2	3	4	5	6	7
0	1	0	0	0	0	0	0	0
1	0	1	0	0	0	0	0	0
2	0	1	1	0	0	0	0	0
3	0	2	3	1	0	0	0	0
4	0	6	11	6	1	0	0	0
5	0	24	50	35	10	1	0	0
6	0	120	274	225	85	15	1	0
7	0	720	1764	1624	735	175	21	1

The recurrence does indeed allow us to compute the values of $c(n,k)$ quickly, but we need more work to compute the permutations of a certain "cycle structure". For example, suppose that we want to count how many permutations of [20] that consist of four 2-cycles and three 4-cycles. Notice that this value will occur as one summon in the term $c(20,7)$ (as will those permutations consisting of one 14-cycle and six 1-cycles, etc.). To count this, we think as follows. Arrange the 20 elements of [20] in sequence without repetition, and build a permutation from it represented in cycle notation by put the first two elements in a 2-cycle, then the 3^{rd} and 4^{th} elements in a 2-cycle, as well as the 5^{th} and 6^{th}, and 7^{th} and 8^{th}. Next, put the the 9^{th} through 12^{th} elements in a 4-cycle, and then the 13^{th} through 16^{th}, and 17^{th} through 20^{th} into four cycles as well. For example, if we write out our 20 numbers as
$$5\ \ 19\ \ 3\ \ 11\ \ 16\ \ 17\ \ 1\ \ 9\ \ 7\ \ 12\ \ 2\ \ 4\ \ 10\ \ 14\ \ 18\ \ 8\ \ 13\ \ 6\ \ 15\ \ 20$$
then we view this as representing the permutation
$$(5\ \ 19)(3\ \ 11)(16\ \ 17)(1\ \ 9)(7\ \ 12\ \ 2\ \ 4)(10\ \ 14\ \ 18\ \ 8)(13\ \ 6\ \ 15\ \ 20)$$
Notice that every permutation of [20] with four 2-cycles and three 4-cycles can be written with the four 2-cycles in the front (because we can always reorder the cycles), so we do get every permutation we are looking for in this way. However, this is a lot of overcount in this method. Notice that in each of the four

2-cycles in front, we can swap the order of the two 2 entries without changing the permutation. Thus, we get an overcount of 2^4 with these swappings. Furthermore, for each of the three 4-cycles, we can cyclically shift them in 4 ways without changing the actual permutation, so we get an overcount of 4^3 here. Finally, notice that we can rearrange the four 2-cycles up front in 4! ways, and also rearrange the three 4-cycles in 3! ways, without affecting the underlying permutation. It follows that the number of permutations of [20] that consist of four 2-cycles and three 4-cycles equals.

$$\frac{20!}{2^4 \cdot 4! \cdot 4^3 \cdot 3!}.$$

Definition 3.49. *Let σ be a permutation of $[n]$. An* inversion *of σ is an ordered pair (i, j) with $i < j$ but $\sigma(i) > \sigma(j)$. We let $Inv(\sigma)$ be the set of all inversions of σ.*

For example, consider the following permutations in one-line notation:

$$\sigma = 312546 \qquad \tau = 315246 \qquad \pi = 342516$$

In two-line notation, these are:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 1 & 2 & 5 & 4 & 6 \end{pmatrix} \qquad \tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 1 & 5 & 2 & 4 & 6 \end{pmatrix} \qquad \pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 2 & 5 & 1 & 6 \end{pmatrix}$$

while in cycle notation, these are

$$\sigma = (1\ 3\ 2)(4\ 5)(6) \qquad \tau = (1\ 3\ 5\ 4\ 2)(6) \qquad \pi = (1\ 3\ 2\ 4\ 5)(6)$$

Notice that τ and π are obtained by swapping just two elements in the one-line notation, i.e. by swapping two elements in the bottom row of the two-line notation. In terms of functions, τ and π are obtained by composing σ with a permutation consisting of one 2-cycle and four 1-cycles: we have $\sigma = \pi \circ (3\ 4)(1)(2)(5)(6)$ and $\tau = \pi \circ (2\ 5)(1)(3)(4)(6)$.

We now examine the inversions in each of these permutations. Notice that it is typically easier to determine these in first representations rather than in cycle notation:

$$Inv(\sigma) = \{(1, 2), (1, 3), (4, 5)\}$$
$$Inv(\tau) = \{(1, 2), (1, 4), (3, 4), (3, 5)\}$$
$$Inv(\pi) = \{(1, 3), (1, 5), (2, 3), (2, 5), (3, 5), (4, 5)\}$$

From this example, it may seem puzzling to see how the inversions are related. However, there is something quite interesting that is happening. Let's examine the relationship between $Inv(\sigma)$ and $Inv(\tau)$. By swapping the third and fourth positions in the second row, the inversion $(1, 3)$ in σ became the inversion $(1, 4)$ in τ, and the inversion $(4, 5)$ in σ became the inversion $(3, 5)$ in τ, so those match up. However, we added a new inversion by this swap, because although originally we had $\sigma(3) < \sigma(4)$, but the swapping made $\tau(3) > \tau(4)$. This accounts for the one additional inversion in τ. If instead we had $\sigma(3) > \sigma(4)$, then this swap would have lost an inversion. However, in either case, this example illustrates that a swapping of two adjacent numbers either increases or decreases the number of inversions by 1.

Lemma 3.50. *Suppose that σ is a permutation of $[n]$, and suppose τ is obtained from σ by swapping two adjacent entries in the one-line notation of σ. In other words, suppose that there is a k with $1 \leq k < n$ such that*

$$\tau(i) = \begin{cases} \sigma(i) & \text{if } i \neq k \text{ and } i \neq k+1 \\ \sigma(k+1) & \text{if } i = k \\ \sigma(k) & \text{if } i = k+1 \end{cases}$$

We then have that $|Inv(\sigma)|$ and $|Inv(\tau)|$ differ by 1.

Proof. Suppose that τ is obtained from σ by swapping the entries k and $k+1$. Notice that if $i, j \notin \{k, k+1\}$, then

$$(i, j) \in Inv(\sigma) \Longleftrightarrow (i, j) \in Inv(\tau)$$

Now given any i with $i < k$, we have

$$(i, k) \in Inv(\sigma) \Longleftrightarrow (i, k+1) \in Inv(\tau)$$
$$(i, k+1) \in Inv(\sigma) \Longleftrightarrow (i, k) \in Inv(\tau)$$

Similarly, given any j with $j > k+1$, we have

$$(k, j) \in Inv(\sigma) \Longleftrightarrow (k+1, j) \in Inv(\tau)$$
$$(k+1, j) \in Inv(\sigma) \Longleftrightarrow (k, j) \in Inv(\tau)$$

The final thing to notice is that

$$(k, k+1) \in Inv(\sigma) \Longleftrightarrow (k, k+1) \notin Inv(\tau)$$

because if $\sigma(k) > \sigma(k+1)$ then $\tau(k) < \tau(k+1)$, while if $\sigma(k) < \tau(k+1)$ then $\tau(k) > \tau(k+1)$. Since we have a bijection between $Inv(\sigma)\backslash\{(k, k+1)\}$ and $Inv(\tau)\backslash\{(k, k+1)\}$, while $(k, k+1)$ is exactly one of the sets $Inv(\sigma)$ and $Inv(\tau)$, it follows that $|Inv(\sigma)|$ and $|Inv(\tau)|$ differ by 1. □

A similar analysis is more difficult to perform on π because the swapping involved two non-adjacent numbers. As a result, elements in the middle had slightly more complicated interactions, and the above example shows that a swap of this type can sizably increase the number of inversions. Although it is possible to handle it directly, the key idea is to realize we can perform this swap through a sequence of adjacent swaps. This leads to the following result.

Corollary 3.51. *Suppose that σ is a permutation of $[n]$, and suppose τ is obtained from σ by swapping two entries in the one-line notation of σ. We then have that $|Inv(\sigma)|$ and $|Inv(\tau)|$ have different parities, i.e. one is even while the other is odd.*

Proof. Suppose that τ is obtained from σ by swapping positions k and ℓ, where $k < \ell$. We can assume that $\ell \geq k+2$ because otherwise $|Inv(\sigma)|$ and $|Inv(\tau)|$ differ by 1 and we are done. The key fact is that we can obtain this swap by performing an odd number of adjacent swaps. To see this, start by swapping k and $k+1$, then $k+1$ and $k+2$, then $k+2$ and $k+3$, etc. until we end by swapping $\ell-1$ and ℓ. Notice that there are $\ell - k$ many swaps here, and we end by shifting the entries in positions $k+1$ through ℓ by one to the left, and moving the entry in position k to the entry in position ℓ. Now we swap positions $\ell-2$ and $\ell-1$, then $\ell-3$ and $\ell-2$, etc. until we end by swapping k and $k+1$. Notice that there are $\ell-k-1$ many swaps here, and in the final product we have swapped the entries in positions k and ℓ of σ and left the rest in place. We have a total of $2k + 2\ell - 1 = 2(k+\ell) - 1$ many swaps. Now Lemma 3.50 says each of these adjacent swaps changes the number of inversions by 1 (either increasing or decreasing by 1), so each of these inversions changes the parity of the number inversions. Since there are an odd number of such swaps, we conclude that $|Inv(\sigma)|$ and $|Inv(\tau)|$ have different parities. □

3.6 Polynomial Coefficients

Definition 3.52. *Given a polynomial $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ with $a_n \neq 0$, we define $\deg(p(x)) = n$. We leave $\deg(0)$ undefined.*

For example, we have $\deg(x^2 + 5x - 1) = 2$ and $\deg(5) = 0$. Notice that

$$\deg(p(x)q(x)) = \deg(p(x)) + \deg(q(x))$$

for all nonzero polynomial $p(x)$ and $q(x)$, and this is one of the reasons why we leave $\deg(0)$ undefined. A fundamental fact about polynomials is the following.

Fact 3.53. *A polynomial of degree n has at most n roots.*

You will see a proof of this result in Abstract Algebra. Essentially, the key idea is that if a is a root of a polynomial $p(x)$, then it is possible to factor out $x - a$ from $p(x)$. Notice also that this is another reason why we leave $\deg(0)$ undefined, because *every* element of $\mathbb{R}$ is a root of the zero polynomial. Furthermore, from this fact, we obtain the following result.

Proposition 3.54. *Let $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ and $q(x) = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + a_0$ be polynomials of degree at most n. Suppose that $p(c) = q(c)$ for at least $n+1$ many $c \in \mathbb{R}$. We then have that $a_i = b_i$ for all i, so $p(x)$ and $q(x)$ are equal polynomials (and hence agree on all possible inputs).*

Proof. Consider the polynomial

$$p(x) - q(x) = (a_n - b_n)x^n + (a_{n-1} - b_{n-1})x^{n-1} + \cdots + (a_1 - b_1)x + (a_0 - b_0)$$

Notice that this polynomial has degree at most n but has at least $n+1$ many roots (because if $c \in \mathbb{R}$ is such that $p(c) = q(c)$, then c is a root of $p(x) - q(x)$). Since a polynomial of degree n has at most n roots, this is only possible if $p(x) - q(x)$ is the zero polynomial, i.e. if $a_i - b_i = 0$ for all i. We conclude that $a_i = b_i$ for all i. □

Since $\deg(p(x)q(x)) = \deg(p(x)) + \deg(q(x))$ for all nonzero polynomial $p(x)$ and $q(x)$, it follows that $\deg((x+1)^n) = n$ for all $n \in \mathbb{N}$. The Binomial Theorem tells us what the coefficients of the resulting polynomial are:

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

In other words, if we expand out

$$(x+1)(x+1)(x+1)\cdots(x+1)$$

and collect terms to form a polynomial of degree n, then the coefficient of x^k in the result is $\binom{n}{k}$. We next work to determine the coefficients of slightly more complicated polynomials.

Definition 3.55. *Given $n \in \mathbb{N}^+$, we define the following two polynomials:*

- $x^{\overline{n}} = x(x+1)(x+2)\cdots(x+n-1)$
- $x^{\underline{n}} = x(x-1)(x-2)\cdots(x-n+1)$

We also define $x^{\overline{0}} = 1 = x^{\underline{0}}$. Notice that $\deg(x^{\overline{n}}) = n = \deg(x^{\underline{n}})$ for all $n \in \mathbb{N}$.

For example, we have the following:

- $x^{\overline{0}} = 1$
- $x^{\overline{1}} = x = 0 + x$
- $x^{\overline{2}} = x(x+1) = 0 + x + x^2$
- $x^{\overline{3}} = x(x+1)(x+2) = 0 + 2x + 3x^2 + x^3$
- $x^{\overline{4}} = x(x+1)(x+2)(x+3) = 0 + 6x + 11x^2 + 6x^3 + x^4$

and also:

- $x^{\underline{0}} = 1$
- $x^{\underline{1}} = x = 0 + x$

- $x^{\underline{2}} = x(x-1) = 0 - x + x^2$
- $x^{\underline{3}} = x(x-1)(x-2) = 0 + 2x - 3x^2 + x^3$
- $x^{\underline{4}} = x(x-1)(x-2)(x-3) = 0 - 6x + 11x^2 - 6x^3 + x^4$

Theorem 3.56. *For every $n \in \mathbb{N}$, we have*

$$x^{\overline{n}} = \sum_{k=0}^{n} c(n,k) \cdot x^k$$

Proof. We prove the result by induction on $n \in \mathbb{N}$.

- *Base Cases:* We prove the result for $n = 0$ and $n = 1$ (we will need two base cases because we assume $n \geq 1$ in Theorem 3.48).

  - When $n = 0$, we have

    $$\begin{aligned} x^{\overline{0}} &= 1 \\ &= c(0,0) \\ &= c(0,0) \cdot x^0 \\ &= \sum_{k=0}^{0} c(n,k) \cdot x^k \end{aligned}$$

  - When $n = 1$, we have

    $$\begin{aligned} x^{\overline{1}} &= 0 + 1x \\ &= c(1,0) \cdot x^0 + c(1,1) \cdot x^1 \\ &= \sum_{k=0}^{1} c(n,k) x^k \end{aligned}$$

    Thus, the statement is true when $n = 0$ and $n = 1$.

- *Inductive Step:* Let $n \in \mathbb{N}^+$ be arbitrary and assume that the statement is true for n, i.e. assume that

  $$x^{\overline{n}} = \sum_{k=0}^{n} c(n,k) \cdot x^k$$

  Using the fact that $c(n,0) = 0 = c(n+1,0)$ and $c(n,n) = 1 = c(n+1,n+1)$, along with Theorem

3.48, we then have

$$\begin{aligned}
x^{\overline{n+1}} &= x^{\overline{n}}(x+n) \\
&= \left(\sum_{k=0}^{n} c(n,k)\cdot x^k\right)\cdot(x+n) \\
&= x\cdot\left(\sum_{k=0}^{n} c(n,k)\cdot x^k\right) + n\cdot\left(\sum_{k=0}^{n} c(n,k)\cdot x^k\right) \\
&= \left(\sum_{k=0}^{n} c(n,k)\cdot x^{k+1}\right) + \left(\sum_{k=0}^{n} n\cdot c(n,k)\cdot x^k\right) \\
&= \left(\sum_{k=1}^{n+1} c(n,k-1)\cdot x^k\right) + \left(\sum_{k=1}^{n} n\cdot c(n,k)\cdot x^k\right) \\
&= \left(\sum_{k=1}^{n} c(n,k-1)\cdot x^k\right) + c(n,n)\cdot x^{n+1} + \left(\sum_{k=1}^{n} n\cdot c(n,k)\cdot x^k\right) \\
&= 0\cdot x^0 + \left(\sum_{k=1}^{n}[c(n,k-1)+n\cdot c(n,k)]\cdot x^k\right) + c(n,n)\cdot x^{n+1} \\
&= c(n+1,0)\cdot x^0 + \left(\sum_{k=1}^{n} c(n+1,k)\cdot x^k\right) + c(n+1,n+1)\cdot x^{n+1} \\
&= \sum_{k=0}^{n+1} c(n+1,k)\cdot x^k
\end{aligned}$$

Thus, the statement is true for $n+1$.

The result follows by induction. □

Definition 3.57. *Let $k, n \in \mathbb{N}$. We define*

$$s(n,k) = (-1)^{n+k}c(n,k) = (-1)^{n+k}\begin{bmatrix} n \\ k \end{bmatrix}.$$

and call $s(n,k)$ the (signed) Stirling numbers of the first kind*.*

Notice that $s(0,0)=1$, $s(n,0)=0$ if $n \geq 1$, and $s(n,k)=0$ if $k>n$ because the same are true of $c(n,k)$ by definition.

Corollary 3.58. *For every $n \in \mathbb{N}^+$, we have*

$$x^{\underline{n}} = \sum_{k=0}^{n} s(n,k)\cdot x^k$$

Proof. One can prove this by induction as in the previous theorem, but there is another more clever method. Let $n \in \mathbb{N}^+$ be arbitrary. We know from the previous theorem that

$$x(x+1)(x+2)\cdots(x+(n-1)) = \sum_{k=0}^{n} c(n,k)\cdot x^k.$$

Since this is a polynomial equality, we can plug in any real value to obtain an equality of real numbers. Thus, for any $a \in \mathbb{R}$, we can plug $-a$ into the above polynomials to conclude that

$$(-a)((-a)+1)((-a)+2)\cdots((-a)+(n-1)) = \sum_{k=0}^{n} c(n,k)\cdot(-a)^k,$$

which implies that

$$(-1)^n \cdot a(a-1)(a-2)\cdots(a-(n-1)) = \sum_{k=0}^{n}(-1)^k c(n,k)\cdot a^k$$

Multiplying both sides by $(-1)^n$ it follows that

$$a(a-1)(a-2)\cdots(a-(n-1)) = \sum_{k=0}^{n}(-1)^{n+k} c(n,k)\cdot a^k$$

and hence

$$a(a-1)(a-2)\cdots(a-(n-1)) = \sum_{k=0}^{n} s(n,k)\cdot a^k$$

is true for all $a \in \mathbb{R}$. Since the polynomials $x^{\underline{n}} = x(x-1)(x-2)\cdots(x-(n-1))$ and $\sum_{k=0}^{n} s(n,k)\cdot x^k$ agree for all real numbers, we may use Proposition 3.54 to conclude that

$$x^{\underline{n}} = \sum_{k=0}^{n} s(n,k)\cdot x^k.$$

This completes the proof. □

We can interpret our two polynomial equalities in the following way. Let $n \in \mathbb{N}^+$ and consider the vector space V of all polynomials of degree at most n (as well as the zero polynomial). We know that $\{x^0, x^1, x^2, \dots, x^n\}$ is a basis for V. For each $\ell \in \mathbb{N}$ with $0 \leq \ell \leq n$, we have

$$x^{\overline{\ell}} = \sum_{k=0}^{\ell} c(\ell,k)\cdot x^k$$

and

$$x^{\underline{\ell}} = \sum_{k=0}^{\ell} s(\ell,k)\cdot x^k$$

so the (unsigned/signed) Stirling numbers of the first kind show to express $x^{\overline{\ell}} \in V$ and $x^{\underline{\ell}} \in V$ as linear combinations of the standard basis vectors in $\{x^0, x^1, x^2, \dots, x^n\}$. Can we reverse this process? In other words, can we express $1, x, x^2, \dots, x^n$ in terms of the vectors $\{x^{\underline{0}}, x^{\underline{1}}, x^{\underline{2}}, \dots, x^{\underline{n}}\}$? If this latter set is a basis for V, then this is indeed possible. One can show directly that $\{x^{\underline{0}}, x^{\underline{1}}, x^{\underline{2}}, \dots, x^{\underline{n}}\}$ is a linearly independent set of size $n+1$, so it must be a basis. Hence, it is at least theoretically possible. However, we can prove directly that is possible along with determining the coefficients with little work at this point.

Theorem 3.59. *For every $n \in \mathbb{N}$, we have*

$$x^n = \sum_{k=0}^{n} S(n,k)\cdot x^{\underline{k}}$$

Proof. When $n = 0$, we have $x^0 = 1$ and

$$\begin{aligned}\sum_{k=0}^{0} S(n,k) \cdot x^{\underline{k}} &= S(0,0) \cdot x^{\underline{0}} \\ &= 1 \cdot x^0 \\ &= 1\end{aligned}$$

as well, so the statement is true in this case. Suppose now that $n \geq 1$. Recall that Theorem 3.27 tells us that

$$m^n = \sum_{k=1}^{n} k! \cdot S(n,k) \cdot \binom{m}{k}$$

for all $m \in \mathbb{N}^+$ (because both sides count the number of functions from $[n]$ to $[m]$). Therefore, for any $m \in \mathbb{N}^+$, we have

$$\begin{aligned}m^n &= \sum_{k=1}^{n} k! \cdot S(n,k) \cdot \binom{m}{k} \\ &= \sum_{k=1}^{n} k! \cdot S(n,k) \cdot \frac{m!}{k! \cdot (m-k)!} \\ &= \sum_{k=1}^{n} S(n,k) \cdot \frac{m!}{(m-k)!} \\ &= \sum_{k=1}^{n} S(n,k) \cdot m(m-1)(m-2)\cdots(m-k+1) \\ &= \sum_{k=0}^{n} S(n,k) \cdot m(m-1)(m-2)\cdots(m-k+1)\end{aligned}$$

where the last line follows from the fact that $S(n,0) = 0$. Thus, the polynomial x^n and the polynomial

$$\sum_{k=0}^{n} S(n,k) \cdot x^{\underline{k}}$$

agree at every natural number m. Since these two polynomials have degree at most n and agree at infinitely many points, we may use Proposition 3.54 to conclude that

$$x^n = \sum_{k=0}^{n} S(n,k) \cdot x^{\underline{k}}$$

This completes the proof. □

We now know that

$$x^{\underline{\ell}} = \sum_{k=0}^{\ell} s(\ell,k) \cdot x^k$$

and

$$x^{\ell} = \sum_{k=0}^{\ell} S(\ell,k) \cdot x^{\underline{k}}$$

for all $\ell \in \mathbb{N}$. Let's return to the above setting, i.e. let $n \in \mathbb{N}^+$ and consider the vector space V of all polynomials of degree at most n (as well as the zero polynomial). Since $\{x^0, x^1, x^2, \ldots, x^n\}$ is a basis for V, and we've just seen that each x^ℓ is in the span of $\{x^{\underline{0}}, x^{\underline{1}}, x^{\underline{2}}, \ldots, x^{\underline{n}}\}$, it follows that $\{x^{\underline{0}}, x^{\underline{1}}, x^{\underline{2}}, \ldots, x^{\underline{n}}\}$ spans V. Since this is a spanning set of $n+1$ many vectors, it follows that this set is also basis of V. Furthermore, the above equalities show that the Stirling numbers give the change of basis matrices between these two bases. Thus, if we cut off the Stirling matrices $S(n,k)$ and $s(n,k)$ at some finite point, then the matrices must be inverses of each other.

$S(n,k)$	0	1	2	3	4	5	6	7
0	1	0	0	0	0	0	0	0
1	0	1	0	0	0	0	0	0
2	0	1	1	0	0	0	0	0
3	0	1	3	1	0	0	0	0
4	0	1	7	6	1	0	0	0
5	0	1	15	25	10	1	0	0
6	0	1	31	90	65	15	1	0
7	0	1	63	301	350	140	21	1

$s(n,k)$	0	1	2	3	4	5	6	7
0	1	0	0	0	0	0	0	0
1	0	1	0	0	0	0	0	0
2	0	-1	1	0	0	0	0	0
3	0	2	-3	1	0	0	0	0
4	0	-6	11	-6	1	0	0	0
5	0	24	-50	35	-10	1	0	0
6	0	-120	274	-225	85	-15	1	0
7	0	720	-1764	1624	-735	175	-21	1

Alternatively, instead of appealing to linear algebra, we can also determine this inverse relationship directly. For any $n \in \mathbb{N}$, we have

$$\begin{aligned}
x^n &= \sum_{\ell=0}^{n} S(n,\ell) \cdot x^{\underline{\ell}} \\
&= \sum_{\ell=0}^{n} [S(n,\ell) \cdot (\sum_{k=0}^{\ell} s(\ell,k) \cdot x^k)] \\
&= \sum_{\ell=0}^{n} \sum_{k=0}^{\ell} [S(n,\ell) \cdot s(\ell,k) \cdot x^k] \\
&= \sum_{\ell=0}^{n} \sum_{k=0}^{n} [S(n,\ell) \cdot s(\ell,k) \cdot x^k] \\
&= \sum_{k=0}^{n} \sum_{\ell=0}^{n} [S(n,\ell) \cdot s(\ell,k) \cdot x^k] \\
&= \sum_{k=0}^{n} [\sum_{\ell=0}^{n} S(n,\ell) \cdot s(\ell,k)] \cdot x^k
\end{aligned}$$

Therefore

$$\sum_{\ell=0}^{n} S(n,\ell) \cdot s(\ell,k) = \begin{cases} 1 & \text{if } n = k \\ 0 & \text{if } n \neq k \end{cases}$$

Similarly, we have

$$\sum_{\ell=0}^{n} s(n,\ell) \cdot S(\ell,k) = \begin{cases} 1 & \text{if } n = k \\ 0 & \text{if } n \neq k \end{cases}$$

In other words, the Stirling matrices $S(n,k)$ and $s(n,k)$ are inverses of each other.

Notice that we do not even need to cut off the matrices at some point to be $n \times n$ matrices (the fact that every row is eventually zero means the the matrix products make sense even for infinite matrices). In linear algebra terminology, the sets $\{x^0, x^1, x^2, \dots\}$ and $\{x^{\underline{0}}, x^{\underline{1}}, x^{\underline{2}}, \dots\}$ are both bases for the infinite-dimensional vector space of *all* polynomials (without degree restrictions), and the matrices $S(n,k)$ and $s(n,k)$ form the change of basis matrices for these two bases.

3.7 Countability and Uncountability

Recall that given finite sets A and B, if there exists a bijection $f \colon A \to B$, then $|A| = |B|$. Moreover, it's not hard to see that the converse of this is true is well, i.e. if A and B are finite sets with $|A| = |B|$, then there exists a bijection $f \colon A \to B$ (to see this, assume that $|A| = n = |B|$, list $A = \{a_1, a_2, \dots, a_n\}$ and $B = \{b_1, b_2, \dots, b_n\}$, and then define a bijection by letting $f(a_i) = b_i$ for all i). Now if A and B are infinite sets, then we have no obvious way to define the cardinality of A and B like we do for finite sets. However, it still makes sense to talk about bijections, and so one can simply *define* two (possibly infinite) sets A and B to have the same size if there is a bijection $f \colon A \to B$.

With this in mind, think about $\mathbb{N} = \{0, 1, 2, 3, \dots\}$ and the subset $\mathbb{N}^+ = \{1, 2, 3, 4, \dots\}$. Although $\mathbb{N}^+$ is a proper subset of $\mathbb{N}$ and "obviously" has one fewer element, the function $f \colon \mathbb{N} \to \mathbb{N}^+$ given by $f(n) = n + 1$ is a bijection, and so $\mathbb{N}$ and $\mathbb{N}^+$ have the same "size". For another even more surprising example, let $A = \{2n : n \in \mathbb{N}\} = \{0, 2, 4, 6, \dots\}$ be the set even natural numbers, and notice that the function $f \colon \mathbb{N} \to A$ given by $f(n) = 2n$ is a bijection from $\mathbb{N}$ to A. Hence, even though A intuitively seems to only have "half" of the elements of $\mathbb{N}$, there is still a bijection between $\mathbb{N}$ and A.

The next proposition shows that $\mathbb{N}$ is the "smallest" infinite set.

Proposition 3.60. *If A is an infinite set, then there is an injective function $f \colon \mathbb{N} \to A$.*

Proof. We define $f \colon \mathbb{N} \to A$ recursively. Pick an arbitrary $a_0 \in A$, and define $f(0) = a_0$. Suppose that $n \in \mathbb{N}$ and we have defined the values $f(0), f(1), \dots, f(n)$, all of which are elements of A. Since A is infinite, we have that $\{f(0), f(1), \dots, f(n)\} \neq A$. Thus, we can pick an arbitrary $a_{n+1} \in A$, and define $f(n+1) = a_{n+1}$. With this recursive definition, we have defined a function $f \colon \mathbb{N} \to A$. Notice that if $m < n$, then $f(n)$ was chosen to be distinct from $f(m)$ by definition, so $f(m) \neq f(n)$. Therefore, f is injective. □

With this in mind, we introduce a name for those infinite sets for which we can find a bijection with $\mathbb{N}$, and think of them as the "smallest" types of infinite sets.

Definition 3.61. *Let A be a set.*

- *We say that A is* countably infinite *if there exists a bijection $f \colon \mathbb{N} \to A$.*
- *We say that A is* countable *if it is either finite or countably infinite.*
- *If A is not countable, we say that A is* uncountable.

Suppose that A is countably infinite. We then have a bijection $f \colon \mathbb{N} \to A$, so we can arrange its elements in a list without repetitions by listing out $f(0), f(1), f(2), f(3), \dots$ to get:

$$a_0 \quad a_1 \quad a_2 \quad a_3 \quad \cdots$$

Conversely, writing out such a list without repetitions shows how to build a bijection $f \colon \mathbb{N} \to A$. Since working with such lists is more intuitively natural (although perhaps a little less rigorous), we'll work with countable sets in this way. What about lists that allow repetitions?

Proposition 3.62. *Let A be a set. The following are equivalent.*

1. *It is possible to list A, possibly with repetitions, as $a_0, a_1, a_2, a_3, \dots$.*
2. *There is a surjection $g\colon \mathbb{N} \to A$.*
3. *A is countable, i.e. either finite or countably infinite.*

Proof. $1 \leftrightarrow 2$: This is essentially the same as the argument just given. If we can list A, possibly with repetitions, as $a_0, a_1, a_2, a_3, \dots$, then the function $g\colon \mathbb{N} \to A$ given by $g(n) = a_n$ is a surjection. Conversely, if there is a surjection $g\colon \mathbb{N} \to A$, then $g(0), g(1), g(2), g(3), \dots$ is a listing of A.

$1 \to 3$: Suppose that there is a surjection $g\colon \mathbb{N} \to A$. If A is finite, then A is countable by definition, so we may assume that A is infinite. We define a new list as follows. Let $b_0 = a_0$. If we have defined $b_0, b_1, \dots, b_n$, let $b_{n+1} = a_k$, where k is chosen as the least value such that $a_k \notin \{b_0, b_1, \dots, b_n\}$ (such a k exists because A is infinite). Then

$$b_0 \quad b_1 \quad b_2 \quad b_3 \quad \cdots$$

is a listing of A without repetitions. Therefore, A is countably infinite.

$3 \to 1$: Suppose that A is countable. If A is countably infinite, then there is a bijection $f\colon \mathbb{N} \to A$, in which case

$$f(0) \quad f(1) \quad f(2) \quad f(3) \quad \cdots$$

is a listing of A (even without repetition). On the other hand, if A is finite, say $A = \{a_0, a_1, a_2, \dots, a_n\}$, then

$$a_0 \quad a_1 \quad a_2 \quad \cdots \quad a_n \quad a_n \quad a_n \quad \cdots$$

is a listing of A with repetitions. □

Our first really interesting result is that $\mathbb{Z}$, the set of integers, is countable. Of course, some insight is required because if we simply start to list the integers as

$$0 \quad 1 \quad 2 \quad 3 \quad 4 \quad \cdots$$

we won't ever get to the negative numbers. We thus use the sneaky strategy of bouncing back-and-forth between positive and negative integers.

Proposition 3.63. *$\mathbb{Z}$ is countable.*

Proof. We can list $\mathbb{Z}$ as

$$0 \quad 1 \quad -1 \quad 2 \quad -2 \quad \cdots$$

More formally, we could define $f\colon \mathbb{N} \to \mathbb{Z}$ by

$$f(n) = \begin{cases} -\frac{n}{2} & \text{if } n \text{ is even} \\ \frac{n+1}{2} & \text{if } n \text{ is odd} \end{cases}$$

and check that f is a bijection. □

The key idea used in previous proof can be abstracted into the following result.

Proposition 3.64. *If A and B are countable, then $A \cup B$ is countable.*

Proof. Since A is countable, we may list it as $a_0, a_1, a_2, a_3, \dots$. Since B is countable, we may list it as $b_0, b_1, b_2, b_3, \dots$. We therefore have the following two lists:

$$\begin{array}{ccccc} a_0 & a_1 & a_2 & a_3 & \cdots \\ b_0 & b_1 & b_2 & b_3 & \cdots \end{array}$$

We can list $A \cup B$ by going back-and-forth between the above lists as

$$a_0 \quad b_0 \quad a_1 \quad b_1 \quad a_2 \quad b_2 \quad \cdots$$

□

A slightly stronger result is now immediate.

Corollary 3.65. *If $A_0, A_1, \dots, A_n$ are countable, then $A_0 \cup A_1 \cup \cdots \cup A_n$ is countable.*

Proof. This follows from Proposition 3.64 by induction. Alternatively, we can argue as follows. For each fixed k with $0 \leq k \leq n$, we know that A_k is countable, so we may list it as $a_{k,0}, a_{k,1}, a_{k,2}, \dots$. We can visualize the situation with the following table.

$$\begin{array}{ccccc} a_{0,0} & a_{0,1} & a_{0,2} & a_{0,3} & \cdots \\ a_{1,0} & a_{1,1} & a_{1,2} & a_{1,3} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \\ a_{n,0} & a_{n,1} & a_{n,2} & a_{n,3} & \cdots \end{array}$$

We now list $A_0 \cup A_1 \cup \cdots \cup A_n$ by moving down each column in order, to obtain:

$$a_{0,0} \quad a_{1,0} \quad \cdots \quad a_{n,0} \quad a_{0,1} \quad a_{1,1} \quad \cdots \quad a_{n,1} \quad \cdots \quad \cdots$$

□

In fact, we can prove quite a significant extension of the above results. The next proposition is usually referred to by saying that "the countable union of countable sets is countable".

Proposition 3.66. *If $A_0, A_1, A_2, \dots$ are all countable, then $\bigcup\limits_{k=0}^{\infty} A_k = A_0 \cup A_1 \cup A_2 \cup \cdots$ is countable.*

Proof. For each $n \in \mathbb{N}$, we know that A_n is countable, so we may list it as $a_{k,0}, a_{k,1}, a_{k,2}, a_{k,3}, \dots$. We now have the following table.

$$\begin{array}{ccccc} a_{0,0} & a_{0,1} & a_{0,2} & a_{0,3} & \cdots \\ a_{1,0} & a_{1,1} & a_{1,2} & a_{1,3} & \cdots \\ a_{2,0} & a_{2,1} & a_{2,2} & a_{2,3} & \cdots \\ a_{3,0} & a_{3,1} & a_{3,2} & a_{3,3} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{array}$$

Now we can't list this by blindly walking down the rows or columns. We thus need a new, much more clever, strategy. The idea is to list the elements of the table by moving between rows and columns. One nice approach which works is to step along certain diagonals and obtain the following listing of $\bigcup\limits_{n=0}^{\infty} A_n$:

$$a_{0,0} \quad a_{0,1} \quad a_{1,0} \quad a_{0,2} \quad a_{1,1} \quad a_{2,0} \quad \cdots$$

The pattern here is that we are walking along the diagonals in turn, each of which is finite. Alternatively, we can describe this list as follows. For each $m \in \mathbb{N}$, there are only finitely many pairs $(i, j) \in \mathbb{N} \times \mathbb{N}$ with $i + j = m$. We first list the finitely many $a_{i,j}$ with $i + j = 0$, followed by those finitely many $a_{i,j}$ with $i + j = 1$, then those finitely many $a_{i,j}$ with $i + j = 2$, etc. This gives a listing of $\bigcup\limits_{k=0}^{\infty} A_k$. □

Theorem 3.67. $\mathbb{Q}$ *is countable.*

Proof. For each $k \in \mathbb{N}^+$, let $A_k = \{\frac{a}{k} : a \in \mathbb{Z}\}$. Notice that each A_k is countable because we can list it as

$$\frac{0}{k} \quad \frac{1}{k} \quad \frac{-1}{k} \quad \frac{2}{k} \quad \frac{-2}{k} \quad \cdots$$

Since

$$\mathbb{Q} = \bigcup_{k=1}^{\infty} A_k = A_1 \cup A_2 \cup A_3 \cup \cdots$$

we can use Proposition 3.66 to conclude that $\mathbb{Q}$ is countable. □

With all of this in hand, it is natural to ask whether uncountable sets exist.

Theorem 3.68. $\mathbb{R}$ *is uncountable.*

Proof. We need to show that there is no list of real numbers that includes every element of $\mathbb{R}$. Suppose then that $r_1, r_2, r_3, \dots$ is an arbitrary list of real numbers. We show that there exists $x \in \mathbb{R}$ with $x \neq r_n$ for every $n \in \mathbb{N}$. For each $n \in \mathbb{N}$, we write out the (nonterminating) decimal expansion of r_n as

$$a_n \quad . \quad d_{n,1} \quad d_{n,2} \quad d_{n,3} \quad d_{n,4} \quad \cdots$$

where $a_n \in \mathbb{Z}$ and each $d_{n,i} \in \mathbb{Z}$ satisfies $0 \leq d_{n,i} \leq 9$. We arrange our list of reals $r_1, r_2, r_3, \dots$ as a table

$$\begin{array}{ccccccc}
a_1 & . & d_{1,1} & d_{1,2} & d_{1,3} & d_{1,4} & \cdots \\
a_2 & . & d_{2,1} & d_{2,2} & d_{2,3} & d_{2,4} & \cdots \\
a_3 & . & d_{3,1} & d_{3,2} & d_{3,3} & d_{3,4} & \cdots \\
a_4 & . & d_{4,1} & d_{4,2} & d_{4,3} & d_{4,4} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}$$

For each $n \in \mathbb{N}$, let

$$e_n = \begin{cases} 3 & \text{if } d_{n,n} \neq 3 \\ 7 & \text{if } d_{n,n} = 3 \end{cases}$$

Let x be the real number with decimal expansion

$$. \quad e_1 \quad e_2 \quad e_3 \quad e_4 \quad \cdots$$

We claim that $x \neq r_n$ for every $n \in \mathbb{N}$. Let $n \in \mathbb{N}$ be arbitrary. Since $e_n \neq d_{n,n}$ by construction, it follows x and r_n disagree in the n^{th} decimal position. Therefore, since the (nonterminating) decimal expansions of x and r_n are different, it follows that $x \neq r_n$. □

4 Graph Theory

4.1 Graphs, Multigraphs, Representations, and Subgraphs

Definition 4.1. *A* graph *G is a pair (V, E) of sets such that:*

- *V is a nonempty set.*
- *E is a (possibly empty) set such that each element is subset of V of cardinality 2. In other words, $E \subseteq \mathcal{P}_2(V)$.*

Elements of V are called vertices*, and elements of E are called* edges*. We say that G is finite if V is finite (in which case E must be finite as well).*

For example, if we let

$$\begin{aligned} V &= \{1, 2, 3, 4, 5\} \\ E &= \{\{1, 2\}, \{1, 3\}, \{1, 5\}, \{3, 5\}\}, \end{aligned}$$

then $G = (V, E)$ is a graph.

Definition 4.2. *Let $G = (V, E)$ be a graph.*

- *Given an edge $e \in E$, we call the elements of e the* endpoints *of e, and we say that e is* incident *to these vertices.*
- *If $u, w \in V$ and $\{u, w\} \in E$, then we say that u and w are* adjacent *or* linked*.*

We can also view graphs as certain types of relations. Recall that a relation on a set V is a subset of V^2, i.e. a set of ordered pairs, while edges in a graph are *sets* with 2 elements. However, if we consider symmetric relations, then we kind of "ignore" the fact that edges are unordered because whenever we have the ordered pair (u, w) we also have the ordered pair (w, u). We need one other condition as well.

Definition 4.3. *A relation R on a set A is* irreflexive *if $(a, a) \notin R$ for all $a \in A$.*

Notice that irreflexive does not mean "not reflexive" (if $(a, a) \in R$ some $a \in A$ and $(a, a) \notin R$ for some other $a \in A$, then the relation is neither reflexive nor irreflexive). From this point of view, a graph can be described as a nonempty set V together with a relation on V that is symmetric and irreflexive. For example, we can interpret the above graph as follows:

$$\begin{aligned} V &= \{1, 2, 3, 4, 5\} \\ R &= \{(1, 2), (2, 1), (1, 3), (3, 1), (1, 5), (5, 1), (3, 5), (5, 3)\} \end{aligned}$$

Definition 4.4. *Let G be a finite graph with n vertices $v_1, v_2, \ldots, v_n$ listed in some order. We define an $n \times n$ matrix A called the* adjacency matrix *of G by letting*

$$a_{i,j} = \begin{cases} 1 & \text{if } v_i \text{ is adjacent to } v_j \\ 0 & \text{otherwise} \end{cases}$$

For the above example with vertices listed in order $1, 2, 3, 4, 5$, we have the adjacency matrix

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \end{pmatrix}$$

If we change the order of the vertices to be $4, 3, 5, 2, 1$, the we have the adjacency matrix

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

Notice that the adjacency matrix A of any finite graph is symmetric (i.e. $a_{i,j} = a_{j,i}$ for all i and j, or alternatively A is equal to its transpose), has each entry equal to either 0 or 1, and has all diagonal entries equal to 0. Furthermore, it's not hard to see that every matrix with these properties arises as the adjacency matrix of some finite graph.

Definition 4.5. *Let G be a finite graph with n vertices $v_1, v_2, \dots, v_n$ and m edges $e_1, e_2, \dots, e_m$ each listed in some order. We define an $n \times m$ matrix B called the* incidence matrix *of G by letting*

$$b_{i,j} = \begin{cases} 1 & \textit{if } v_i \textit{ is an endpoint of } e_j \\ 0 & \textit{otherwise} \end{cases}$$

For the above example with vertices listed in order $1, 2, 3, 4, 5$ and edges as $\{1, 2\}, \{1, 3\}, \{1, 5\}, \{3, 5\}$, we have the incidence matrix

$$B = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

Notice that incidence matrix has each entry equal to either 0 or 1, and has exactly two 1's in each column. Furthermore, it's not hard to see that every matrix with these properties arises as the incidence matrix of some graph.

Definition 4.6. *We define the following graphs.*

- *For each $n \in \mathbb{N}^+$, let K_n be the graph with vertex set $V = [n]$ and edge set*

  $$E = \{\{i, j\} : 1 \leq i \leq n, 1 \leq j \leq n, \text{ and } i \neq j\}$$

  equal to the set of all subsets of V of cardinality 2, i.e. every pair of distinct vertices are linked. We call K_n the complete graph *on n vertices.*

- *For each $n \in \mathbb{N}^+$, let P_n be the graph with vertex set $V = [n]$ and edge set*

  $$E = \{\{i, i + 1\} : 1 \leq i \leq n - 1\}.$$

  We call P_n the path graph *on n vertices.*

- *For each $n \in \mathbb{N}^+$ with $n \geq 3$, let C_n be the graph with vertex set $V = [n]$ and edge set*

  $$E = \{\{i, i + 1\} : 1 \leq i \leq n - 1\} \cup \{\{1, n\}\}.$$

  We call C_n the cycle graph *on n vertices.*

- *For each $m, n \in \mathbb{N}^+$, let $K_{m,n}$ be the graph with vertex set $V = [m + n]$ and edge set*

  $$E = \{\{i, j\} : 1 \leq i \leq m \text{ and } m + 1 \leq j \leq n\}.$$

  We call $K_{m,n}$ a complete bipartite graph.

In our definition of a graph, given any two vertices, either they are linked by one edge or they are not. Also, in a graph, edges always have two distinct endpoints, so there are no “loops”. By relaxing these conditions, we define a broader class of objects that we'll call multigraphs. Since two distinct vertices can have many different edges linking them in this context, the definition is more involved.

Definition 4.7. *A* multigraph G *is a triple* (V, E, f) *of sets such that:*

- V *is a nonempty set.*
- E *is a set disjoint from* V*.*
- f *is function with domain* E *such that* $f(e)$ *is a subset of* V *or cardinality either* 1 *or* 2 *for each* $e \in E$*. In other words,* $f\colon E \to \mathcal{P}_1(V) \cup \mathcal{P}_2(V)$*.*

Elements of V *are called* vertices*, and elements of* E *are called* edges*. We say that* G *is finite if both* V *and* E *are finite (notice that it possible that* V *is finite but* E *is infinite).*

Definition 4.8. *Let* $G = (V, E, f)$ *be a multigraph.*

- *Given an edge* $e \in E$*, we call the elements of* $f(e)$ *the* endpoints *of* e*, and we say that* e *is* incident *to these vertices.*
- *If* $u, w \in V$ *and there is an edge* $e \in E$ *with* $f(e) = \{u, w\}$*, then we say that* u *and* w *are* adjacent *or* linked.
- *We call an edge* $e \in E$ *a* loop *if* $f(e)$ *has only* 1 *element (i.e. if* e *has only* 1 *endpoint).*

For example, let

$$V = \{1, 2, 3\}$$
$$E = \{a, b, c, d, e\}$$

and define f by letting:

- $f(a) = \{1, 2\}$.
- $f(b) = \{1, 3\}$
- $f(c) = \{3\}$
- $f(d) = \{1, 2\}$
- $f(e) = \{1, 2\}$

We then have that $G = (V, E, f)$ is a graph. Intuitively, we can think of this graph as follows. There are three vertices labeled 1, 2, and 3. We have one edge linking vertices 1 and 3, three edges linking vertices 1 and 2, and one edge that is a loop at vertex 3 (so both of its endpoints are vertex 3).

Although the definitions of graphs and multigraphs are fundamentally different, we can interpret every graph as a multigraph. For example, recall our graph

$$V = \{1, 2, 3, 4, 5\}$$
$$E = \{\{1, 2\}, \{1, 3\}, \{1, 5\}, \{3, 5\}\},$$

We can interpret this as a mutligraph by letting keeping V, letting $E' = \{e_1, e_2, e_3, e_4\}$, and define f by

- $f(e_1) = \{1, 2\}$

- $f(e_2) = \{1, 3\}$
- $f(e_3) = \{1, 5\}$
- $f(e_4) = \{3, 5\}$

Alternatively, and much more simply, we can keep both V and E, and just let $f\colon E \to \mathcal{P}_2(V)$ be the function where $f(e) = e$ (since, after all, in a graph an element of E is a subset of V of cardinality 2). This always works as long as V and E are disjoint.

One can define analogues of the adjacency and incidence matrices for multigraphs, but there there is not a standard way to deal with loops.

- If G is a finite multigraph, and we've listed the vertices in order as $v_1, v_2, \dots, v_n$, then we define the $n \times n$ adjacency matrix A as follows. If $i \neq j$, let $a_{i,j}$ be the number of edges with endpoints v_i and v_j. For the diagonal entries, it is natural to let $a_{i,i}$ be the number of loops at v_i, but there is also a strong argument for letting $a_{i,i}$ be twice the number of loops at v_i (so we're counting the endpoint as having "multiplicity" 2).
- If G is a finite multigraph, and we've listed the vertices and edges in order as $v_1, v_2, \dots, v_n$ and $e_1, e_2, \dots, e_m$, then we define the $n \times m$ incidence matrix B as follows. If e_j is not a loop, let

  $$b_{i,j} = \begin{cases} 1 & \text{if } v_i \text{ is an endpoint of } e_j \\ 0 & \text{otherwise} \end{cases}$$

  as in the case for graphs. If e_j is a loop with single endpoint v_i, then we define $b_{k,j} = 0$ for all $k \neq i$, and we let $b_{i,j}$ equal either 1 or 2, depending on our preference (as in the adjacency matrix case).

Since we won't be dealing with adjacency and incidence matrices of multigraphs very often, we'll just stipulate which version we are using

Definition 4.9. *Let G be a multigraph and let $v \in V$. The* degree *of v, denoted by $d(v)$, is the number of edges incident to v, where each loop incident to v is counted twice.*

Proposition 4.10. *If G is a finite multigraph with m edges, then*

$$\sum_{v \in V} d(v) = 2m$$

Proof. Every edge has two endpoints (if we count the unique endpoint of any loop twice), so contributes 2 to the sum on the left hand-side. The result follows. □

Corollary 4.11. *A finite multigraph has an even number of vertices of odd degree.*

Proof. Let G be a finite multigraph with m edges. Proposition 4.10 tells us that

$$\sum_{v \in V} d(v) = 2m,$$

so

$$\sum_{v \in V} d(v)$$

is even. If there were odd number of vertices of odd degree, then this sum would be odd, which is a contradiction. □

Definition 4.12. *Let $G = (V_G, E_G)$ and $H = (V_H, E_H)$ be graphs. We say that H is a* subgraph *of G if $V_H \subseteq V_G$ and $E_H \subseteq E_G$.*

Definition 4.13. *Let $G = (V, E)$ be a graph.*

- *For any subset $F \subseteq E$, we let $G - F$ be the subgraph of G with vertex set equal to V and edge set equal to $E \backslash F$. If $F = \{e\}$, we write $G - e$ instead of $G - \{e\}$.*
- *For any subset $U \subseteq V$ with $U \neq V$, we let $G - U$ be the subgraph of G with vertex set $V \backslash U$ and edge set equal to $\{e \in E :$ Both endpoints of e are elements of $V \backslash U\}$. If $U = \{u\}$, we write $G - u$ instead of $G - \{u\}$.*
- *For any subset $U \subseteq V$ with $U \neq \emptyset$, we let $G[U]$ be the subgraph of G with vertex set U and edge set equal to $\{e \in E :$ Both endpoints of e are elements of $U\}$. Notice that $G[U] = G - (V/U)$. We call $G[U]$ the subgraph of G* induced *by U.*

Thus, $G - F$ is obtained by deleting all of the edges in F, while $G - U$ is obtained by deleting all of the vertices in U *as well as* all edges incident to some vertex in U. Intuitively, an induced subgraph of G is one obtained by only deleting vertices (and all of their associated edges), whereas a general subgraph of G is one obtained by deleting vertices (and all of their associated edges) along with possibly deleting additional edges whose endpoints are still alive. There can exist subgraphs of a graph G that are not induced subgraphs, such as the result of deleting one edge.

Definition 4.14. *Let $G = (V_G, E_G, f_G)$ and $H = (V_H, E_H, f_H)$ be multigraphs. We say that H is a* submultigraph *of G if $V_H \subseteq V_G$, $E_H \subseteq E_G$, and f_H is the restriction of f_G to the set E_H.*

One can similarly define $G - F$, $G - U$, and $G[U]$ for multigraphs.

Definition 4.15. *Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be graphs. An* isomorphism *from G_1 to G_2 is a bijection $g \colon V_1 \to V_2$ such that for all $u, w \in V_1$, we have that $\{u, w\} \in E_1$ if and only if $\{u, w\} \in E_2$.*

For example, consider the following two graphs. Let G_1 be the graph where

$$\begin{aligned} V_1 &= \{1, 2, 3, 4, 5\} \\ E_1 &= \{\{1, 2\}, \{1, 3\}, \{3, 4\}, \{3, 5\}, \{4, 5\}\} \end{aligned}$$

and let G_2 be the graph where

$$\begin{aligned} V_2 &= \{a, b, c, d, e\} \\ E_2 &= \{\{a, c\}, \{a, d\}, \{a, e\}, \{b, e\}, \{c, d\}\} \end{aligned}$$

We then have that the function $g \colon V_1 \to V_2$ defined by

$$\begin{aligned} g(1) &= e \\ g(2) &= b \\ g(3) &= a \\ g(4) &= c \\ g(5) &= d \end{aligned}$$

is an isomorphism.

Definition 4.16. *Given two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, we say that G_1 is* isomorphic *to G_2, and write $G_1 \cong G_2$, if there exists an isomorphism from G_1 to G_2.*

Intuitively, two graphs G_1 and G_2 are isomorphic when only can relabel the names of the vertices so that the graphs look the same, and the function g is precisely the "translation" between the names of the vertices on each side. In terms of pictures, saying that $G_1 \cong G_2$ is the same as saying that it is possible to draw the graphs in identical fashions, provided we can place the vertices anywhere that we would like.

Definition 4.17. *Let G_1 and G_2 be multigraphs. An* isomorphism *is a pair of bijections $g\colon V_1 \to V_2$ and $h\colon E_1 \to E_2$ such that for all $e \in E_1$, if $f_1(e) = \{u, w\}$, then $f_2(h(e)) = \{g(u), g(w)\}$, and if $f_1(e) = \{v\}$, then $f_2(h(e)) = \{g(v)\}$*

4.2 Walks, Paths, Cycles, and Connected Components

Definition 4.18. *Let $G = (V, E, f)$ be a multigraph.*

- *A* walk *in G is a sequence $v_0, e_1, v_1, e_2, v_2, \dots, v_{k-1}, e_k, v_k$ where each $v_i \in V$, each $e_i \in E$, and where $f(e_i) = \{v_{i-1}, v_i\}$ for all i (i.e. the endpoints of e_i are v_{i-1} and v_i). We allow walks to consists of a single vertex v_0 and no edges.*
- *A* trail *in G is a walk with no repeated edges, i.e. where $e_i \neq e_j$ whenever $i \neq j$.*
- *A* path *in G is a walk with no repeated vertices, i.e. where $v_i \neq v_j$ whenever $i \neq j$.*
- *A* closed walk *in G is a walk is a walk where $v_0 = v_k$. Similarly, a* closed trail *is a trail where $v_0 = v_k$.*
- *A u, w-walk in G is a walk with $v_0 = u$ and $v_k = w$ (similarly for a u, w-trail and a u, w-path).*

Definition 4.19. *Let G be a multigraph. Given a walk in G, we define the length of the walk to be the number of edges it contains, counting repetition. In other words, the length of the walk*

$$v_0, e_1, v_1, e_2, v_2, \dots, v_{k-1}, e_k, v_k$$

is k (which is one less than the number of vertices).

Proposition 4.20. *Let G be a multigraph.*

- *Every path in G is a trail.*
- *Every trail in G is a walk.*

Proof. Clearly every trail in G is a walk because a trail is by definition a walk. We need to prove that every path in G is a trail. Suppose then that

$$v_0, e_1, v_1, e_2, v_2, \dots, v_{k-1}, e_k, v_k$$

is a path in G. We need to show that this is a trail, so suppose for the sake of obtaining a contradiction that $e_i = e_j$ where $i < j$. Since $e_i = e_j$, we have that e_i and e_j have the same endpoints, which is to say that $\{v_{i-1}, v_i\} = \{v_{j-1}, v_j\}$. It follows that either $v_{i-1} = v_{j-1}$ or $v_{i-1} = v_j$. Since $i - 1 < i \leq j - 1$, in either case we have violated the definition of a path because we have found a repeated vertex. This is a contradiction, so our path must be a trail. □

Proposition 4.21. *Let G be a multigraph and let $u, w \in G$. The following are equivalent.*

1. *There is a u, w-walk in G.*
2. *There is a u, w-trail in G.*
3. *There is a u, w-path in G.*

Proof. $3 \to 2 \to 1$ are immediate from the previous proposition. We now prove that $1 \to 3$. Suppose that there is a u, w-walk in G. Fix a u, w-walk in G of shortest possible length, say it is:

$$v_0, e_1, v_1, \dots, v_{k-1}, e_k, v_k$$

where $v_0 = u$ and $v_k = w$. We argue that this walk is a u, w-path. Suppose for the sake of obtaining a contradiction that some vertex is repeated, say $v_i = v_j$ where $i < j$. We then have that

$$v_0, e_1, v_1, \dots, v_{i-1}, e_i, v_i, e_{j+1}, v_{j+1}, \dots v_{k-1}, e_k, v_k$$

is a u, w-walk because $v_i = v_j$ (so the set of endpoints of e_{j+1} equals $\{v_j, v_{j+1}\} = \{v_i, v_{j+1}\}$). Furthermore, this walk has length

$$i + (k - j) = k - (j - i) < k$$

Thus, we have produced a u, w-walk in G of shorter length, which is a contradiction. It follows that our above u, w-walk is in fact a u, w-path in G. □

Proposition 4.22. *Let G be a multigraph. Define a relation $\sim$ on V by letting $u \sim w$ mean that there is a u, w-walk in G. We then have that $\sim$ is an equivalence relation.*

Proof. We check the properties.

- Reflexive: For any $u \in V$, we have that the single vertex u is a u, u-walk in G, so $u \sim u$.
- Symmetric: Suppose that $u \sim w$. Fix a u, w-walk, say it is:

  $$u = v_0, e_1, v_1, e_2, v_2, \dots, v_{k-1}, e_k, v_k = w$$

  We then have that

  $$w = v_k, e_k, v_{k-1}, \dots, v_2, e_2, v_1, e_1, v_0 = u$$

  is a w, u-walk in G, so $w \sim u$.
- Suppose that $u \sim w$ and $w \sim y$. Fix a u, w-walk

  $$u = v_0, e_1, v_1, e_2, v_2, \dots, v_{k-1}, e_k, v_k = w$$

  and a w, y-walk

  $$w = x_0, f_1, x_1, f_2, x_2, \dots, x_{\ell-1}, f_\ell, x_\ell = y$$

  We then have that

  $$u = v_0, e_1, v_1, \dots, v_{k-1}, e_k, v_k = w = x_0, f_1, x_1, \dots, x_{\ell-1}, f_\ell, x_\ell = y$$

  is a u, y-walk in G, so $u \sim y$.

□

Definition 4.23. *Let G be a multigraph and let $\sim$ be the above relation on V. We know from our general theory of equivalence relations that the equivalence classes of $\sim$ are subsets of V that partition V. A* connecting component *of G is a subgraph of G of the form $G[U]$ for some equivalence class U of $\sim$.*

We know that each vertex of a multigraph G appears in a unique connected component of G because the equivalence classes of $\sim$ partition V. We now show that the same is true for edges.

Proposition 4.24. *Let G be a multigraph. Every edge of G appears in a unique connected component of G.*

Proof. Let $e \in E$ be arbitrary. Let u and w be the endpoints of E. Since u, e, w is a u, w-walk in G, we have $u \sim w$. Thus, if we let $U = \overline{u}$ be the equivalence class of U, then $u, w \in U$, and hence $e \in G[U]$. Furthermore, since the equivalence classes partition V, the vertices u and w are not in any other equivalence class, and hence e is not an element of any other connected component. □

Definition 4.25. *A multigraph G is* connected *if it has one connected component. In other words, G is a connected if there exists a u, w-walk in G for all $u, w \in V$.*

Proposition 4.26. *If G is a multigraph, then every connected component of G is a connected graph.*

Proof. Let G be a multigraph, and let $U \subseteq V$ be an equivalence class of $\sim$. Let $u, w \in U$ be arbitrary. Since u and w are elements of the same equivalence class, we have $u \sim w$, and hence we can fix a u, w-walk

$$u = v_0, e_1, v_1, e_2, v_2, \dots, v_{k-1}, e_k, v_k = w$$

in G. Notice that for each i with $1 \leq i \leq k$, we have that

$$v_0, e_1, v_1, e_2, v_2, \dots, v_{i-1}, e_i, v_i$$

is a walk in G, so $u \sim v_i$. It follows that $v_i \in U$ for all i with $1 \leq i \leq k$. Now given any i with $2 \leq i \leq k$, we have that both $v_{i-i} \in U$ and $v_i \in U$, so e_i is an edge of $G[U]$. It follows that the walk

$$u = v_0, e_1, v_1, e_2, v_2, \dots, v_{k-1}, e_k, v_k = w$$

is also a u, w-walk in $G[U]$. We have shown that for any two vertices $u, w \in U$, there is a u, w-walk in $G[U]$, so $G[U]$ is connected. □

Proposition 4.27. *Let G be a graph with vertices $v_1, v_2, \dots, v_n$ and let A be the adjacency matrix. For all $k \in \mathbb{N}^+$, the (i, j) entry of the matrix A^k equals the number of v_i, v_j-walks in G.*

Proof. We prove the result by induction on k.

- *Base Case:* Since
$$a_{i,j} = \begin{cases} 1 & \text{if } v_i \text{ is adjacent to } v_j \\ 0 & \text{otherwise} \end{cases}$$
and a walk of length 1 consists of a single edge, it follows that $a_{i,j}$ is the number of v_i, v_j-walks of length 1 in G.

- *Inductive Step:* Suppose then that the result is true for k. Letting $B = A^k$, we then have that $b_{i,j}$ is the number of v_i, v_j-walks of length k in G. Let $C = A^{k+1} = A^k A = BA$. Fix i and j, and let
$$L = \{\ell \in [n] : v_\ell \text{ is adjacent to } v_j\}$$
We have
$$\begin{aligned} c_{i,j} &= \sum_{\ell=1}^{n} b_{i,\ell} a_{\ell,j} \\ &= \sum_{\ell \in L} b_{i,\ell} a_{\ell,j} && (\text{since } a_{\ell,j} = 0 \text{ if } \ell \notin L) \\ &= \sum_{\ell \in L} b_{i,\ell} && (\text{since } a_{\ell,j} = 1 \text{ if } \ell \in L) \end{aligned}$$

Now given any v_i, v_j-walk of length $k+1$, the second to last vertex in the sequence must be a vertex adjacent to v_j, hence must equal v_ℓ for some $\ell \in L$. Thus, since G is a graph (i.e. not a multigraph), a v_i, v_j-walk of length $k+1$ is completely and uniquely determined by choice of v_ℓ for some $\ell \in L$ as the second to last vertex, together with a v_i, v_ℓ walk of length k. By induction, adding up the number of such walks amounts to calculating the last sum above. Therefore, $c_{i,j}$ is the is the number of v_i, v_j-walks of length $k+1$ in G. The result follows by induction.

□

Above, we defined the cycle graphs C_n for each $n \in \mathbb{N}^+$ with $n \geq 3$ as follows. Given $n \in \mathbb{N}^+$ with $n \geq 3$, we let C_n be the graph with vertex set $V = [n]$ and edge set

$$E = \{\{i, i+1\} : 1 \leq i \leq n-1\} \cup \{\{1, n\}\}.$$

For $n = 1$ and $n = 2$, we can also define *multigraphs* C_1 and C_2 as follows.

- C_1 is the multigraph with vertex set $[1] = \{1\}$ and one edge that is a loop at 1.
- C_2 is the multigraph with vertex set $[2] = \{1, 2\}$ and two edges, each of which have endpoints 1 and 2 (so there is one double edge).

Together together, the C_n for $n \in \mathbb{N}^+$ form the cycle (multi)graphs. We now define cycles *within* graphs.

Definition 4.28. *Let G be a multigraph. A* cycle *in G is a submultigraph of G that is isomorphic to C_n for some $n \in \mathbb{N}^+$.*

Although cycles are certain subgraphs of G, we often find them by finding closed walks without repeated vertices or edges.

Proposition 4.29. *Let G be a multigraph.*

1. *Suppose that $k \geq 1$ and that*
$$v_0, e_1, v_1, \dots, v_{k-1}, e_k, v_k$$
is a closed walk in G without repeated edges and without repeated vertices other than $v_0 = v_k$ (i.e. where $e_i \neq e_j$ whenever $1 \leq i < j < k$ and $v_i \neq v_j$ whenever $0 \leq i < j < k$). If we let $U = \{v_0, v_1, \dots, v_{k-1}\}$ and $F = \{e_1, e_2, \dots, e_k\}$, then $H = (U, F)$ is a submultigraph of G that is isomorphic to C_k, so $H = (U, F)$ is a cycle.

2. *Conversely, suppose that $H = (U, F)$ is a cycle of G. It is them possible to list the vertices of U as $v_0, v_1, \dots, v_{k-1}$ and list the edges of F as $e_1, e_2, \dots, e_k$ in such a way that*
$$v_0, e_1, v_1, \dots, v_{k-1}, e_k, v_k$$
is a closed walk without repeated edges and without repeated vertices other than $v_0 = v_k$ (i.e. where $e_i \neq e_j$ whenever $1 \leq i < j < k$ and $v_i \neq v_j$ whenever $0 \leq i < j < k$).

Proof. 1. We prove the result when $k \geq 3$ (the cases for $k = 1$ and $k = 2$ are similar, but there we need to treat C_k as a multigraph). Let
$$v_0, e_1, v_1, \dots, v_{k-1}, e_k, v_k$$
be a closed walk in G without repeated edges and without repeated vertices other than $v_0 = v_k$. Define a function $g \colon V_{C_k} \to U$ by letting $g(i) = v_i$ for all i. Also, define a function $h \colon E_{C_k} \to F$ by letting $h(\{i, i+1\}) = e_i$ for all i with $1 \leq i \leq k-1$, and letting $h(\{1, k\}) = e_k$. We then have that g and h are bijjective because the v_i and e_i are distinct. These functions give an isomorphism of $H = (U, F)$ with C_k, so H is a cycle.

2. Suppose that $H = (U, F)$ is a cycle of G. Fix bijections $g \colon V_{C_k} \to U$ and $h \colon E_{C_k} \to F$ that form an isomorphism. Let $v_i = g(i)$ for all i with $1 \leq i \leq k$, and let $v_0 = g(k)$. Also, let $e_i = h(\{i, i+1\})$ for all i with $1 \leq i < k$, and let $e_k = h(\{1, k\})$. Since g and h are bijections, it follows that the e_i are distinct and the v_i are distinct, other than $v_0 = v_k$. Furthermore, since g and h form an isomorphism, we have that
$$v_0, e_1, v_1, \dots, v_{k-1}, e_k, v_k$$
is a closed walk. This completes the proof.

□

Since walks are easier to understand and work with than isomorphisms, one may ask why we do not define cycles as these closed walks. The answer is that certain distinct closed walks give the "same" cycle. For example, consider the graph $G = (V, E)$ where:

$$V = \{1, 2, 3, 4\}$$
$$E = \{\{1, 2\}, \{1, 3\}, \{2, 3\}, \{3, 4\}\}$$

Although this graph only has one cycle, the following three closed walks are all distinct even though the "trace" the same cycle:

- $1, \{1, 2\}, 2, \{2, 3\}, 3, \{1, 3\}, 1$
- $2, \{2, 3\}, 3, \{1, 3\}, 1, \{1, 2\}, 2$
- $3, \{2, 3\}, 2, \{1, 2\}, 1, \{1, 3\}, 3$

There are 3 other possible such closed walks as well! Thus, if we want to *count* cycles, then our definition is superior.

The previous proposition lets us view cycles as arising from closed walks without repeated edges or vertices. One may ask whether we need this restriction. We certainly do not want to allow edges to repeat, because if we retrace our steps then that should not be a cycle (and the result will not be isomorphic to C_n). Simply saying that the edges do not repeat is also not enough. For example, consider the graph $G = (V, E)$ where:

$$V = \{1, 2, 3, 4, 5\}$$
$$E = \{\{1, 2\}, \{1, 3\}, \{2, 3\}, \{3, 4\}, \{3, 5\}, \{4, 5\}\}$$

This graph looks like a bow tie. If we follow the natural walk around this bow tie shape, then we do not repeat edges, but we do repeat the vertex 3 in the middle (in addition to the starting/ending vertex 1). This graph is not isomorphic to C_5, so we do not count it as a cycle.

How about if we only enforce that there are no repeated vertices? In trivial cases, this is not enough. For example, if G is a graph, and $e \in E$ is an edge with distinct endpoints u and w, then u, e, w, e, u is a closed walk without repeated vertices (other than the beginning/end), but it is not a cycle. However, for closed walks of length $k \geq 3$, having no repeated vertices automatically gives that there are no repeated edges.

Proposition 4.30. *Let G be a multigraph. Suppose that $k \geq 3$ and that*
$$v_0, e_1, v_1, \dots, v_{k-1}, e_k, v_k$$
is a closed walk without any repeated vertices other than $v_0 = v_k$. We then have that $e_i \neq e_j$ whenever $1 \leq i < j \leq k$. Thus, if we let $U = \{v_0, v_1, \dots, v_{k-1}\}$ and $F = \{e_1, e_2, \dots, e_k\}$, then $H = (U, F)$ is a submultigraph of G that is isomorphic to C_k, so $H = (U, F)$ is a cycle.

Proof. Notice that

$$v_0, e_1, v_1, \dots, e_{k-1}, v_{k-1}$$

is a path in G, so it is a trail in G by Proposition 4.20 is a trail. Hence $e_i \neq e_j$ whenever $1 \leq i < j \leq k-1$. We also have that

$$v_1, e_2, \dots, v_{k-1}, e_k, v_k$$

is a path in G, hence a trail, and so $e_i \neq e_j$ whenever $2 \leq i < j \leq k$. Finally, notice that $e_1 \neq e_k$ because v_1 is an endpoint of e_1, v_{k-1} is an endpoint of e_k, and $v_1 \neq v_{k-1}$ because $k - 1 \geq 2$ (as $k \geq 3$). The last statement now follows from Proposition 4.29. □

Despite the fact that a closed walk of length at least 1, i.e. a closed trail of length at least 1, need not be a cycle (as in our bow tie example), it turns out that if G contains such a closed trail, then G also contains a cycle.

Proposition 4.31. *If G contains a closed trail of length at least 1, then G contains a cycle.*

Proof. Fix a shortest possible closed trail of length at least 1, say it is:

$$v_0, e_1, v_1, \dots, v_{k-1}, e_k, v_k$$

Notice that if $k = 1$, then we have the closed trail v_0, e_1, v_0, which is a cycle (it is a loop and is isomorphic to C_1), so we are done. Suppose then that $k \geq 2$. We claim that the vertices in the list $v_0, v_1, \dots, v_{k-1}$ are all distinct. Suppose, for the sake of obtaining a contradiction, that this is not true. Fix i and j with $0 \leq i < j \leq k-1$ such that $v_i = v_j$. We then have that

$$v_i, e_{i+1}, v_{i+1}, \dots, v_{j-1}, e_j, v_j$$

is a closed trail of length $j - i$. Since $1 \leq j - i \leq k - 1$, this would give an example of a nontrivial closed trail of length strictly less than k, which is a contradiction. Thus, the vertices in the list $v_0, v_1, \dots, v_{k-1}$ are all distinct. Furthermore, since $v_0 = v_k$, we have that $v_i \neq v_k$ whenever $1 \leq i \leq k-1$. Therefore, the vertices in our closed trail

$$v_0, e_1, v_1, \dots, v_{k-1}, e_k, v_k$$

are all distinct (except for $v_0 = v_k$). Using Proposition 4.29, we conclude that G contains a cycle. □

Proposition 4.32. *Let G be a multigraph with the following properties.*

1. *V is finite.*
2. *$E \neq \emptyset$.*
3. *$d(v) \neq 1$ for all $v \in V$*

We then have that G contains a cycle.

Proof. If G has any loops or multiple edges, then it trivially has a cycle. Suppose then that G is a graph. Since V is finite, any path must have length at most $|V|$, and hence we may fix a longest possible path in G:

$$v_0, e_1, v_1, \dots, v_{k-1}, e_k, v_k$$

Since this is a path, no vertex is repeated, and so no edge is repeated either (because paths are trails). We also have $k \geq 1$ since G has at least one edge, and this edge is not a loop. Since e_1 has v_0 as an endpoint, we conclude that $d(v_0) \geq 1$. Now e_1 is not a loop, so there must be an edge $f \neq e_1$ such that f is incident to v_0. As f is not a loop, we know that f is incident to a vertex other than v_0. Let w be the other endpoint of f, so $w \neq v_0$. Now we must have that $w = v_i$ for some i with $0 \leq i \leq k$, because if $w \neq v_i$ for all i with $0 \leq i \leq k$, then

$$w, f, v_0, e_1, v_1, \dots, v_{k-1}, e_k, v_k$$

would be a longer path in G, contradicting our choice of a longest possible path in G. Thus, we can fix ℓ with $0 \leq \ell \leq k$ such that $w = v_\ell$. Notice that $\ell \neq 0$ because $w \neq v_0$. Also, $w \neq v_1$ because $f \neq e_1$ and we are assuming that G does not have multiple edges. Thus, we must have that $\ell \geq 2$. Now since the v_i are distinct, we know that that $w \neq v_i$ whenever $0 \leq i < \ell$. Therefore,

$$w, f, v_0, e_1, v_1, \dots, v_{\ell-1}, e_\ell, v_\ell$$

is a closed walk in G without repeated vertices (other than $w = v_\ell$). Furthermore, since $\ell \geq 2$, this closed walk has length at least 3, we may use Proposition 4.30 to conclude that G contains a cycle. □

Proposition 4.33. *Let G be a connected multigraph and let e be an edge of G. The following are equivalent.*

1. *$G - e$ is connected.*
2. *G has a cycle containing e.*

Proof. 1 → 2: Suppose that $G - e$ is connected. First notice that if e is a loop, then G certainly has a cycle containing e. Suppose then that e is not a loop. Let the endpoints of e be u and w. Since $G - e$ is connected, we can fix a u, w-path

$$u = v_0, e_1, v_1, e_2, v_2, \dots, v_{k-1}, e_k, v_k = w$$

in the graph $G - e$. We then have that

$$v_0, e_1, v_1, e_2, v_2, \dots, v_{k-1}, e_k, v_k, e, v_0$$

is a closed walk in G without repeated edges of vertices. Using Proposition 4.29, it follows that G has a cycle containing e.

2 → 1: Suppose that C is a cycle of G containing e. Let x and y be the endpoints of e (it is possible that $x = y$ if e is a loop). Let $\sim$ denote the connectivity relation in $G - e$. Now let $u, w \in V$ be arbitrary. We will show that $u \sim w$. Since G is connected, we may fix a u, w-path in G, say

$$u = v_0, e_1, v_1, \dots, v_{k-1}, e_k, v_k = w$$

Now if $e \neq e_i$ for all i, then this u, w-path exists in $G - e$, and we have $u \sim w$. Suppose instead that $e = e_i$ for some i. Since this is a path, it is also a trail by Proposition 4.20, and hence there is a unique ℓ with $1 \leq \ell \leq k$ such that $e = e_\ell$. Now the endpoints of e are x and y, so either $x = v_{\ell-1}$ and $y = v_\ell$, or $x = v_\ell$ and $y = v_{\ell-1}$.

- Suppose first that $x = v_{\ell-1}$ and $y = v_\ell$. Since $e \neq e_i$ whenever $1 \leq i \leq \ell - 1$, we have that

  $$u = v_0, e_1, v_1, \dots, e_{\ell-1}, v_{\ell-1} = x$$

  is a u, x-walk in $G - e$, so $u \sim x$. Similarly, since $e \neq e_i$ whenever $\ell + 1 \leq i \leq k$, we have that

  $$y = v_{\ell+1}, e_{\ell+2}, v_{\ell+2}, \dots, e_k, v_k = w$$

  is a y, w-walk in $G - e$, so $y \sim x$. Finally, since e lies in a cycle of G containing e, we see that $x \sim y$ by following such a cycle around in the other direction. Combining $u \sim x$, $y \sim w$, and $x \sim y$ with the fact that $\sim$ is an equivalence relation, we conclude that $u \sim w$. Thus, there is a u, w-walk in $G - e$.

- Suppose now that $x = v_\ell$ and $y = v_{\ell-1}$. Arguing as in the previous case, we have that $u \sim y$, $x \sim w$, and $y \sim x$, so $u \sim w$. Thus, there is a u, w-walk in $G - e$.

We have show that there is a u, w-walk in $G - e$ for all vertices u and w, so $G - e$ is connected. □

4.3 Trees and Forests

Definition 4.34. *A* tree *is a connected acyclic graph.*

Definition 4.35. *A* forest *is an acyclic graph.*

Although we've defined a trees and forests as certain types of *graphs*, notice that we can also define it as a connected acyclic multigraph because the lack of cycles rules out loops and multiple edges.

Proposition 4.36. *If G is a forest, then every connected component of G is a tree.*

Proof. Let H be a connected component of G. We know that H is connected by Proposition 4.26. Furthermore, H is acyclic, because a cycle in H would be a cycle in G. Therefore, H is a tree. □

Proposition 4.37. *If T is a tree with at least 2 vertices, then no vertex of T is isolated, i.e. no vertex of T has degree 0.*

Proof. Let $v \in V$ be arbitrary. Since T has at least 2 vertices, we can fix $w \in V$ with $w \neq v$. Now T is connected, so there exists a v, w-path in T. The first edge of this path must be incident to v, so $d(v) \geq 1$. □

Definition 4.38. *Let T be a tree. A* leaf *of T is a vertex of degree 1.*

Proposition 4.39. *If T is a finite tree with $n \geq 2$ vertices, then T has a leaf.*

Proof. Let T be a finite tree with $n \geq 2$ many vertices. Since T is connected and $n \geq 2$, we must have that $E \neq \emptyset$ by Proposition 4.37. Now if $d(v) \neq 1$ for all $v \in V$, then Proposition 4.32 would imply that T contains a cycle, which is a contradiction. Therefore, there must exist $v \in V$ with $d(v) = 1$, and such a v is a leaf. □

Proposition 4.40. *If v is a leaf of a tree T, then $T - v$ is a tree.*

Proof. Let T be a tree, and let v be a leaf of T. Let e be the unique edge incident to v. Since $T - v$ is a subgraph of T, it follows that $T - v$ is acyclic (a cycle in $T - v$ would be a cycle in T). Let u and w be arbitrary vertices of $T - v$. Since T is connected, we can fix a u, w-path in T. Notice that e and v can not occur on this path (because e is the the only edge incident to v, and so such a purported path would need to use e). Therefore, there is a u, w-path in $T - v$. Thus, $T - v$ is connected. □

The above two results allow one to prove results about finite trees by induction on the number of vertices. This is an extremely powerful tool.

Theorem 4.41. *If T is a tree with n vertices, then T has exactly $n - 1$ edges.*

Proof. By induction on n, i.e. we prove the statement that "Every tree on n vertices has exactly n edges" by induction. If $n = 1$, this is trivial. Suppose that $n \in \mathbb{N}^+$ and we know the result for n. Let T be a tree on $n + 1$ vertices. Fix a leaf v and the unique edge e incident to it. We then have that $T - v$ is a tree on n vertices, so has $n - 1$ edges by induction. Since T has one more edge that $T - v$, we conclude that T has $(n - 1) + 1 = n = (n + 1) - 1$ many edges. The result follows. □

Corollary 4.42. *If T is a finite tree with $n \geq 2$ vertices, then T has at least two leaves.*

Proof. Let T be a finite tree with $n \geq 2$ vertices. By Theorem 4.41, we know that T has exactly $n - 1$ edges. Thus

$$\sum_{v \in V} d(v) = 2(n - 1) = 2n - 2$$

Now if there is only one leaf, then $d(v) \geq 2$ for all other vertices v by Proposition 4.37, so

$$\begin{aligned}\sum_{v \in V} d(v) &\geq 1 + 2(n-1) \\ &= 2n - 1 \\ &> 2n - 2,\end{aligned}$$

a contradiction. Thus, T must have at least two leaves. □

Proposition 4.43. *If G is a finite forest with n vertices and k connected components, then G has $n - k$ edges.*

Proof. Let the components of G be $H_1, H_2, \dots, H_k$. Suppose that H_i has m_i many vertices for each i. Since each vertex lies in a unique connected component (recall that the vertices of a connected component are equivalence classes of $\sim$), we have

$$\sum_{i=1}^{k} m_i = n$$

Now each H_i is a tree by Proposition 4.36, so we know that H_i has $m_i - 1$ many edges for each i by Theorem 4.41. Since every edge is in a unique H_i by Proposition 4.24, it follows that the number of edges in G equals

$$\begin{aligned}\sum_{i=1}^{k} (m_i - 1) &= \left(\sum_{i=1}^{k} m_i\right) - k \\ &= n - k.\end{aligned}$$

This completes the proof. □

Definition 4.44. *Let G be a connected graph. A* spanning tree *of G is a subgraph T of G such that:*

- *$V_T = V_G$ (i.e. T is obtained from G by only deleting edges).*
- *T is a tree.*

Proposition 4.45. *Every finite connected graph has a spanning tree.*

Intuitively, we can argue this as follows. Suppose that G is a finite connected graph. If G has no cycles, then G itself is a spanning tree of G, and we are done. If G contains a cycle, then we pick an arbitrary edge e in a cycle, and notice that $G - e$ is a connected subgraph (by Proposition 4.33) with one fewer edge. If this subgraph of G has no cycles, then it is a spanning tree of G. Otherwise, if we remove another edge from a cycle in $G - e$, then the result is a connected subgraph with two fewer edges than G. Continue this process, and notice that we must stop because G has only finitely many edges. More formally, we can bypass this "continue" argument in the following way.

Proof. Let G be a finite connected graph. Notice that there is at least one connected subgraph H of G with $V_H = V_G$, namely G itself. Amongst all such connected subgraph H of G with $V_H = V_G$, choose one with the least possible number of edges, and call the resulting subgraph T. We then have that $V_T = V_G$ and that T is connected by definition. If T contained a cycle, then we could fix an arbitrary edge e in such a cycle, and notice that $T - e$ is a connected subgraph of G (by Proposition 4.33) with one fewer edge, a contradiction. Therefore, T is acyclic as well. It follows that T is a tree, and hence a spanning tree of G. □

Corollary 4.46. *If G is a finite connected graph with n vertices, then G contains at least $n - 1$ edges.*

Proof. Fix a spanning tree T of G. We then have that T contains $n-1$ edges, so G contains at least $n-1$ edges. □

Theorem 4.47. *Let G be a finite graph with n vertices. The following are equivalent.*

1. *G is a tree.*
2. *G is a connected graph with $n-1$ edges.*
3. *G is an acyclic graph with $n-1$ edges.*
4. *G is connected, but $G-e$ is disconnected for every edge e.*
5. *G is acyclic, but $G+e$ has a cycle for any new edge e having both endpoints in V_G.*

Proof. $1 \to 2$: Immediate from the definition of a tree and Theorem 4.41.

$1 \to 3$: Immediate from the definition of a tree and Theorem 4.41.

$1 \to 4$: Immediate from the definition of a tree and Proposition 4.33.

$1 \to 5$: Suppose that G is a tree, and that e is a new edge. Since $(G+e)-e = G$ is connected, e must be an element of some cycle of $G+e$ by Proposition 4.33. In particular, $G+e$ has a cycle.

$2 \to 4$: Immediate from Corollary 4.46.

$4 \to 1$: Since $G-e$ is disconnected for every edge e, Proposition 4.33 implies that no edge of G is contained in a cycle. Thus, G is acyclic.

$3 \to 1$: Since G is acyclic, it is a forest. Let k be the number of connected components of G. By Proposition 4.43, we know that G has $n-k$ edges. It follows that $n-k = n-1$, hence $k=1$. Therefore, G is connected, and hence a tree.

$5 \to 1$: We need to argue that G is connected. Let $u, w \in V$ and let e be a new edge with endpoints u and w. By assumption, we know that $G+e$ has a cycle, and since G is acyclic it must have a cycle containing e. Fix such a cycle, and walk around it the other way to obtain a u,w-path in G.

□

We now embark on a quest to count the number of trees of a given size. In other words, given $n \in \mathbb{N}^+$, how many trees are there with vertex set $[n]$? Let's consider the case when $n = 4$. To determine what we want to count, we first need to deal with a potential source of ambiguity in the question. For example, consider the trees T_1 and T_2 each with vertex set $[4]$, where the edge set of T_1 is

$$\{\{1,2\},\{1,3\},\{1,4\}\}$$

and the edge set of T_2 is

$$\{\{1,2\},\{2,3\},\{2,4\}\}$$

It is straightforward to check that these are each trees. At first sight, they clearly look different because the edge sets are distinct. However, it's not hard to see that the trees T_1 and T_2 are isomorphic (intuitively, we can draw T_1 by putting 1 in the middle with 3 edges out to to other vertices, and we can draw T_2 by putting 2 in the middle with 3 edges out to the other vertices). Thus, we need to ask whether we are actual counting the number of trees with vertex set $[n]$, or the number is isomorphism types of trees with vertex set $[n]$. Although both questions are interesting, we are going to do the former here. That is, we will consider the above two trees as different.

Suppose now that we want to count the number of trees with vertex set $[4]$. There are 16 such trees, and we give two ways to count this.

1. There are two isomorphism types for a tree with vertex set [4]. First, we can have a tree with a vertex of degree 3 that is adjacent to all other vertices (so there are 3 leaves). Second, we can have a tree with exactly 2 leaves and two vertices of degree 2 (recall that a tree on at least 2 vertices has at least 2 leaves, and that the sum of the degrees will be $2 \cdot (4-1) = 6$), and it's straightforward to check that such a tree is isomorphic to P_4. We now count the number of trees of each isomorphism type.
   - There are exactly 4 trees of the first type, because they are completely determined by the choice of the vertex of degree 3.
   - We now argue that there are exactly 12 trees of the second type. We can choose the two leaves in $\binom{4}{2} = 6$ ways. This then determines the vertices of degree 2 (which will be adjacent to each other). To determine the tree, we now need only pick which of these two is adjacent to the smaller leaf, and we have 2 choices. Thus, the number of such trees is $4 \cdot 2 = 12$.
2. Here is another argument. A tree with vertex set [4] will have exactly $4 - 1 = 3$ edges. Since there are $\binom{4}{2} = 6$ many possible edges, there are $\binom{6}{3} = 20$ many graphs with vertex set set [4] having exactly 3 edges. However, some of these will fail to be trees because they are not connected. This happens exactly when the graph consists of a cycle of length 3 and an isolated vertex, and there are 4 such graphs (because there are 4 choices for the isolated vertex). It follows that there are $20 = 4 = 16$ many trees with vertex set [4].

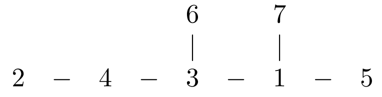
Counting the number of trees with vertex sets $[n]$ for other small values of n is also reasonably straightforward:

1. There is 1 tree with vertex set [1].
2. There is 1 tree with vertex set [2].
3. There are 3 trees with vertex set [3] (we need only the pick the unique vertex of degree 2).
4. There are 16 trees with vertex set [4] (as seen above).
5. There are 125 trees with vertex set [5], which can be argued through a slightly more complicated analysis than the one for [4].

However, as we move toward larger values of n, the above method become unwieldy. However, there is a natural pattern in the above numbers, and we will indeed prove the following result using a more sophisticated approach

Theorem 4.48 (Cayley's Formula)**.** *For each $n \in \mathbb{N}^+$, the number of trees with vertex set $[n]$ is n^{n-2}.*

In order to prove this, we will "code" each tree by a sequence of numbers of length $n - 2$, where each element of the sequence is an integer between 1 and n (inclusive). Given a tree T with vertex set $[n]$, we know that it has $n - 1$ edges. Consider the following tree:



The edge set of this tree is

$$\{\{1,3\},\{1,5\},\{1,7\},\{2,4\},\{3,4\},\{3,6\}\}$$

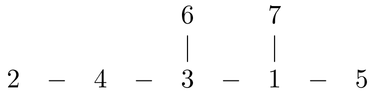
Of course, the edge set is a *set*, so we can reorder it any way we like without affecting the edges, and we can also reorder the two endpoints of an edge. Our first task will be to give an ordering to the edge set in a way that reflects the "structure" of the tree. To do this, given $n \geq 2$ and a tree T on $[n]$, we define a two sequences $(a_1, a_2, \ldots, a_{n-1})$ and $(p_1, p_2, \ldots, p_{n-1})$ as follows. Since T is a tree with at least two vertices, we know that T has a leaf by Proposition 4.39. Let a_1 be the smallest label of a leaf, and let p_1 be its unique neighbor. Now if we delete a_1 from T, then we know from Proposition 4.40 that $T - a_1$ is also a tree. If this tree has at least two vertices, then let a_2 be the smallest label of a leaf, and let p_2 be its unique neighbor. Continue until we end with a unique vertex. Once we have completed this process, list the sequences on top of each other as follows:

$$\begin{array}{cccccc} a_1 & a_2 & a_3 & \cdots & a_{n-2} & a_{n-1} \\ p_1 & p_2 & p_3 & \cdots & p_{n-2} & p_{n-1} \end{array}$$

Notice that the $n-1$ edges of T will be:

$$\{a_1, p_1\}, \{a_2, p_2\}, \dots, \{a_{n-1}, p_{n-1}\}$$

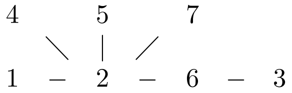
For example, given our tree



we obtain the following two sequences:

$$\begin{array}{cccccc} 2 & 4 & 5 & 6 & 3 & 1 \\ 4 & 3 & 1 & 3 & 1 & 7 \end{array}$$

For another example, given the tree



we obtain the following two sequences:

$$\begin{array}{cccccc} 1 & 3 & 4 & 5 & 6 & 2 \\ 2 & 6 & 2 & 2 & 2 & 7 \end{array}$$

We have the following properties:

Proposition 4.49. *Let T be a tree with vertex set $[n]$.*

1. *$a_1, a_2, \dots, a_{n-1}$ is a permutation of $[n-1]$.*

2. *$p_{n-1} = n$.*

3. *For all $k \in [n]$, $d(k)$ is the number of times that k occurs in $a_1, a_2, \dots, a_{n-1}, p_1, p_2, \dots, p_{n-1}$.*

4. *For all $k \in [n]$, $d(k)$ equals one plus the number of times that k occurs in $p_1, p_2, \dots, p_{n-2}$.*

Proof. We have the following:

1. Certainly every number appears at most once as an a_i because it gets deleted after appearing there. Furthermore, at each stage we have a tree on at least 2 vertices, so it has at least 2 leaves, and thus we never pick n.

2. Since we never have $a_i = n$, it follows that after $n-2$ stages we have two vertices, one of which is n. We pick the other as a_{n-1}, and thus $p_{n-1} = n$.

3. Since the edge set of T is
$$\{\{a_1, p_1\}, \{a_2, p_2\}, \dots, \{a_{n-1}, p_{n-1}\}\}$$
we see that $d(k)$ equals the number of times that k occurs in the two lists.

4. This follows immediately from 3 because every $k \in [n]$ occurs exactly once in $a_1, a_2, \dots, a_{n-1}, p_{n-1}$ by 1 and 2.

□

There is some unnecessary information in the two sequences. As we've seen, we don't need p_{n-1} because we know what it will be. In fact, we also don't need the a_i at all, because we can recover them from the sequence $(p_1, p_2, \ldots, p_{n-2})$.

Definition 4.50. *Given a tree T with vertex set $[n]$, we call the sequence $(p_1, p_2, \ldots, p_{n-2})$ the* Prüfer code *of T.*

To see this that we recover the a_i from the Prüfer code, let's first consider an example. Suppose that T is a tree with and we know the following values:

$$\begin{array}{cccccc} ? & ? & ? & ? & ? & ? \\ 5 & 1 & 7 & 5 & 2 & ? \end{array}$$

Since there are 6 entries in each row, we know that T is a tree with vertex set $[7]$. We know that $p_6 = 7$, so we can fill that in:

$$\begin{array}{cccccc} ? & ? & ? & ? & ? & ? \\ 5 & 1 & 7 & 5 & 2 & 7 \end{array}$$

Next, a_1 will be the smallest leaf. Using Proposition 4.49, we know that the degree of any vertex equals one plus the number of times that k occurs in $5, 1, 7, 5, 2$. Thus, we are looking for the least number that does not occur in this list. It follows that $a_1 = 3$, and we have:

$$\begin{array}{cccccc} 3 & ? & ? & ? & ? & ? \\ 5 & 1 & 7 & 5 & 2 & 7 \end{array}$$

To carry this forward, consider the following table:

vert	0	1	2	3	4	5
1	2	2	1	x	x	x
2	2	2	2	2	2	1
3	1	x	x	x	x	x
4	1	1	x	x	x	x
5	3	2	2	2	1	x
6	1	1	1	1	x	x
7	2	2	2	1	1	1

In the column labeled 0, we have put the degree of each vertex by adding 1 to the number of times that it occurs in $5, 1, 7, 5, 2$. Since 3 is the smallest leaf of T (as discussed above), and we know that it's unique neighbor is 5, then when we delete vertex 3 from the tree, we also decrease the degree of 5 by one. The column labeled 1 gives the degrees after this deletion. The smallest leaf remaining can now be seen to be 4, which gives the value of a_2. Since the unique neighbor of 4 in that tree is 1, we decrease the degree of 1 by one after deletion to form the next column. Continuing this process, we arrive at the following sequence:

$$\begin{array}{cccccc} 3 & 4 & 1 & 6 & 5 & 2 \\ 5 & 1 & 7 & 5 & 2 & 7 \end{array}$$

The above procedure illustrates the following result.

Proposition 4.51. *Let T be a tree with vertex set $[n]$. From $p_1, p_2, \ldots, p_{n-2}$, we can completely determine p_{n-1} along $a_1, a_2, \ldots, a_{n-1}$. Thus, if T and S are two trees with vertex set $[n]$ and the same Prüfee code $(p_1, p_2, \ldots, p_{n-2})$, then they have same values for $p_{n-1}, a_1, a_2, \ldots, a_{n-1}$ as well, so they have the same edge set and hence are the same tree. It follows that the function that takes a tree with vertex set $[n]$ and produces the Prüfer code $(p_1, p_2, \ldots, p_{n-2})$ of T is injective.*

Proof. Since T is a tree with vertex set $[n]$, know that $p_{n-1} = n$. Using Proposition 4.49, we can compute $d(k)$ for each $k \in [n]$. Since a_1 is the smallest leaf, we may let a_1 be the smallest element of $[n]$ that is not in the set $\{p_1, p_2, \dots, p_{n-2}\}$. In constructing these sequences, after we write down a_1 and p_1, we delete the leaf a_1 together with the unique edge incident to it. In the resulting tree, we no longer have a_1 as a vertex, and the degree of p_1 will be reduced by 1. Thus, the leaves of the resulting tree $T - a_1$ are then the numbers other than a_1 which do not occur in the list $p_2, p_3, \dots, p_{n-2}$. In other words, a_2 will be the smallest element of $[n]$ that is not in the set $\{a_1\} \cup \{p_2, p_3, \dots, p_{n-2}\}$. Now we delete a_2 and hence reduce the degree of p_2 by 1, so a_3 will be the smallest element of $[n]$ that is not in the set $\{a_1, a_2\} \cup \{p_3, \dots, p_{n-2}\}$. In general, we can recursively reconstruct the sequence $a_1, a_2, \dots, a_{n-1}$ because a_i is the smallest element of $[n]$ not in the set

$$\{a_1, a_2, \dots, a_{i-1}\} \cup \{p_i, p_{i+1}, \dots, p_{n-2}\},$$

which must exist since there are at most $n - 2$ many numbers. Therefore, we can reconstruct p_{n-1} along $a_1, a_2, \dots, a_{n-1}$. The remaining statements follow. □

We've shown that every sequence $(p_1, p_2, \dots, p_{n-2})$ occurs as the Prüfer code of at most one tree with vertex set $[n]$. We now show that every such code arises. Given any sequence $(p_1, p_2, \dots, p_{n-2})$, the idea is to define the number p_{n-1} and $a_1, a_2, \dots, a_{n-1}$ as in the proof of the previous result, and check that the resulting graph is a tree that produces the given code.

Proposition 4.52. *Let $n \in \mathbb{N}^+$, and suppose that $(q_1, q_2, \dots, q_{n-2})$ is sequence of integers with $1 \leq q_i \leq n$ for all i. Let $q_{n-1} = n$ and define $b_1, b_2, \dots, b_{n-1}$ recursively by letting b_i be the smallest element of $[n]$ not in the set*

$$\{b_1, b_2, \dots, b_{i-1}\} \cup \{q_i, q_{i+1}, \dots, q_{n-2}\}$$

(which must exist since there are at most $n - 2$ many numbers). If we let T be the graph with vertex set $[n]$ and edge set

$$\{\{b_1, q_1\}, \{b_2, q_2\}, \dots, \{b_{n-1}, q_{n-1}\}\},$$

then T is a tree with Prüfer code $(q_1, q_2, \dots, q_{n-2})$.

Proof. Notice that if $i < j$, then $b_j \neq b_i$ by the recursive definition of the sequence $b_1, b_2, \dots, b_{n-1}$ (because b_i will be an element of the set $\{b_1, b_2, \dots, b_{j-1}\}$. Thus, there are no repeated elements in the sequence $b_1, b_2, \dots, b_{n-1}$. Furthermore, since for each i, the set

$$\{b_1, b_2, \dots, b_{i-1}\} \cup \{q_i, q_{i+1}, \dots, q_{n-2}\}$$

has at most $n-2$ many elements, and we choose b_i to be the smallest element of $[n]$ not in this set, it follows that $b_i \in [n-1]$ for all i. Putting these facts together, we conclude that $b_1, b_2, \dots, b_{n-1}$ is a permutation of $[n-1]$. Now $b_i \neq q_i$ for all i by the recursive definition. Also, if $i < j$, then $\{b_i, q_i\} \neq \{b_j, q_j\}$ because $b_i \neq b_j$ from above and $b_i \neq q_j$ by the recursive definition. Thus, we can let T be the graph with vertex set $[n]$ and edge set

$$\{\{b_1, q_1\}, \{b_2, q_2\}, \dots, \{b_{n-1}, q_{n-1}\}\}.$$

Since $\{b_i, q_i\} \neq \{b_j, q_j\}$ whenever $i < j$, it follows that T has $n - 1$ many edges. We now show that T is acyclic. To see this, think about adding the edges of T in reverse order, i.e. add

$$\{b_{n-1}, q_{n-1}\}, \{b_{n-2}, q_{n-2}\}, \dots, \{b_2, q_2\}, \{b_1, q_1\}$$

one at a time. At each stage, we claim that the resulting graph is acyclic. This is certainly true for the graph with just the single edge $\{b_{n-1}, q_{n-1}\}$. Suppose that the we know that the graph with edge set

$$\{\{b_{n-1}, q_{n-1}\}, \{b_{n-2}, q_{n-2}\}, \dots, \{b_{i+1}, q_{i+1}\}\}$$

is acyclic. Now $b_i \notin \{b_{i+1}, \dots, b_{n-2}, b_{n-1}\}$ from above, and $b_i \notin \{q_{i+1}, \dots, q_{n-2}, q_{n-1}\}$ by definition. Thus, in the graph with edge set

$$\{\{b_{n-1}, q_{n-1}\}, \{b_{n-2}, q_{n-2}\}, \dots, \{b_{i+1}, q_{i+1}\}, \{b_i, q_i\}\}$$

we see that b_i has degree 1. Now any cycle in this graph would have to include the new edge, but such a cycle would have to include the vertex b_i, which is impossible because the degree of b_i is 1 in this graph. Therefore, the graph with the new edge is acyclic. By induction, it follows that T is acyclic. Since T is an acyclic graph with n vertices and $n-1$ edges, Theorem 4.47 implies that T is a tree.

Now that we know that T is a tree, we just need to check that the process of constructing the sequences to obtain the Prüfer code (i.e. finding the smallest leaf, deleting it, etc.) to produce

$$\begin{array}{cccccc} a_1 & a_2 & a_3 & \cdots & a_{n-2} & a_{n-1} \\ p_1 & p_2 & p_3 & \cdots & p_{n-2} & p_{n-1} \end{array}$$

results in our sequences

$$\begin{array}{cccccc} b_1 & b_2 & b_3 & \cdots & b_{n-2} & b_{n-1} \\ q_1 & q_2 & q_3 & \cdots & q_{n-2} & q_{n-1} \end{array}$$

Since the edge set of T equal $\{\{b_i, q_i\} : 1 \leq i \leq n-1\}$, we know that the degree of a vertex is just the number of times that it occurs in $b_1, b_2, \dots, b_{n-1}, q_1, q_2, \dots, q_{n-1}$. Furthermore, since we know that $q_{n-1} = n$ and that $b_1, b_2, \dots, b_{n-1}$ is a permutation of $[n-1]$, it follows that the degree of any vertex is 1 plus the number of times that it occurs in $q_1, q_2, \dots, q_{n-2}$. Now by definition of our recursive sequence, we have that b_1 is the smallest element of $[n]$ not in the set $\{q_1, q_2, \dots, q_{n-2}\}$, which means that b_1 is the smallest vertex of degree 1, and hence the smallest leaf. It follows that $a_1 = b_1$ and hence $p_1 = q_1$ (because $\{b_1, q_1\}$ is an edge of T). Since b_1 is a leaf of T, we have that $T - b_1$ is a tree, and we know that it has edge set

$$\{\{b_2, q_2\}, \{b_3, q_3\}, \dots, \{b_{n-1}, q_{n-1}\}\}.$$

In the resulting tree, we no longer have b_1 as a vertex, and the degree of q_1 will be reduced by 1. Thus, the leaves of the resulting tree $T - b_1$ are then the numbers other than b_1 which do not occur in the list $q_2, q_3, \dots, q_{n-2}$. By definition of b_2, we conclude then that b_2 will be the smallest leaf in $T - b_1$. Thus, $a_2 = b_2$ and $p_2 = q_2$. Now delete b_2 and hence reduce the degree of q_2 by 1, and a similar argument shows that b_3 will be the smallest leaf in the resulting tree. In this way (or by induction), it follows that $a_i = b_i$ and $p_i = q_i$ for all i. □

We can now prove Cayley's Formula.

Proof of Theorem 4.48. Let $n \in \mathbb{N}^+$. If $n = 1$, then there is trivially only 1 tree with vertex set $[1]$, and we have $1^{1-2} = 1$. Suppose then that $n \geq 2$. Define a function from trees with vertex set $[n]$ to $\{1, 2, \dots, n\}^{n-2}$ (i.e. the set of sequences of integers between 1 and n of length $n-2$) by assigning to each tree its Prüfer code. Notice that this function is injective by Proposition 4.51 and is surjective by Proposition 4.52. Therefore, the function is a bijection. Since $|\{1, 2, \dots, n\}^{n-2}| = n^{n-2}$, there are n^{n-2} many trees with vertex set $[n]$. □

4.4 Minimum Weight Spanning Trees and Kruskal's Algorithm

Let $\mathbb{R}^{\geq 0} = \{r \in \mathbb{R} : r \geq 0\}$. Suppose that G is a graph, and $w \colon E \to \mathbb{R}^{\geq 0}$ is a function. Given $e \in E$, we think of $w(e)$ as the *weight* of the edge e, and we think about it as the "cost" of including edge e in our graph G. Based on these costs, we can choose to either include an edge or not. A natural question is how to include enough edges so that that we still have a connected graph, but so that we minimize the resulting cost. Since we will only think about deleting edges, we consider subgraphs H of G with $V_H = V_G$. Now given such a subgraph $H = (V_G, E_H)$ of G, we define

$$w(H) = \sum_{e \in E_H} w(e)$$

to be the sum of the weights of the edges that are in H. Notice that if our connected subgraph H includes a cycle, then we know that we can remove an edge of that cycle and still have a connected subgraph by Proposition 4.33. Since the weight of every edge is nonnegative, deleting such an edge does not increase the cost. In other words, we want what is called a *minimum weight spanning tree* of G. This leads to the following problem.

Question 4.53. *Given a connected graph G and a function $w \colon E \to \mathbb{R}^{\geq 0}$, how do we build a spanning tree T of G such that $w(T)$ is as small as possible?*

There are (at least) two natural attempts to build such a spanning tree by making a series of choices that seem reasonable.

1. *Idea 1:* Start with no edges, and include edges from G one at a time. We then have to ask ourselves which edge to include next. Since we start with no edges, the idea is to include one edge without ever introducing a cycle. Since we are tying to minimize cost, the idea is to pick the cheapest edge that does not introduce a cycle at each stage.

2. *Idea 2:* Start with all of the edges of G, and delete edges one at a time. We then have to ask ourselves which edge to include next. We should only delete edges in cycles because we want a connected graph at the end. Since we are trying to minimize cost, the idea is to pick the most expensive edge that is included in a cycle at each stage.

Notice that each of these are *greedy* algorithms. In other words, at each step we are picking a choice that looks best *locally* at that moment, without any assurance that it will produce us a *globally* optimal solution in the end. In general, greedy algorithms do *not* produce globally optimal solutions. For example, if you want to climb a mountain, and you do it by always taking the one step that will increase your elevation most, then you may end up at the top of a tiny hill close by instead (because climbing the mountain may involve going down at some points). For another example, looking only one step ahead in chess may result in a move that looks excellent (say you kill a queen with your pawn), but it may be a globally bad move in that your opponent can checkmate you in the next move. We will see more precise examples of the failure of greedy algorithms in later sections. However, for the minimum weight spanning tree problem, it turns out that both of the above procedures do indeed produce minimum weight spanning trees!

We will study Idea 1, which as known as *Kruskal's algorithm*, because it is faster in practice (as we discuss below. However), let's first formalize this procedure more carefully. Suppose that we have a connected graph G and a function $w \colon E \to \mathbb{R}^{\geq 0}$. We build a sequence $H_0, H_1, H_2, \dots, H_{n-1}$ of subgraphs of G with the following properties:

- $V_{H_i} = V_G$ for each i.
- H_i has i edges for each i.
- H_i is acyclic for each i.

We start by letting H_0 be the subgraph of G consisting of all of the vertices of G, but no edges. Suppose that we have constructed H_i with the above properties. Let

$$S_i = \{e \in E_G : e \notin E_{H_i} \text{ and } H_i + e \text{ is still acyclic}\}$$

We pick an element of S_i such that $w(e)$ is as small as possible (if there are multiple such edges in S_i, we pick an arbitrary one), and we let $H_{i+1} = H_i + e$. Once we've proceeding through each of the stages, we then take H_{n-1} as our answer. Although this ends the description of Kruskal's algorithm, there are several extremely important questions.

1. Why does Kruskal's algorithm never get stuck? In other words, why is each S_i nonempty?

2. Why is H_{n-1} a spanning tree of G?

3. Assuming that H_{n-1} is a spanning tree of G, why is a minimum weight spanning tree? In other words, why is it the case that $w(H_{n-1}) \leq w(T)$ for all spanning trees T of G?

4. How do we implement it efficiently? After all, determining the elements of S_i by going through each edge e and checking if there is a cycle in $H_i + e$ seems to be costly.

In order to prove 1 (and eventually to given efficient methods to answer 4), we use the following result.

Proposition 4.54. *Let G be an graph. Let $u, w \in V$ be distinct vertices that are not adjacent. Let e be a new edge with endpoints u and w.*

1. *If u and w are in the same connected component of G, then e is an element of a cycle of $G + e$.*

2. *If G is acyclic and u and w are in distinct connected components of G, then $G + e$ is acyclic.*

Proof. 1. Since u and w are in the same connected component of G, we can fix a u, w-path

$$u, f_1, v_1, f_2, v_2, \dots, f_k, w$$

in G. Since this is a path, it is also a trail, and hence there are no repeated vertices or edges. We then have that

$$u, f_1, v_1, f_2, v_2, \dots, f_k, w, e, u$$

is a closed walk with repeated vertices or edges (because $e \notin E_G$ and hence $e \neq f_i$ for all i), so Proposition 4.29 implies that e is an element of a cycle of $G + e$.

2. Suppose that G is acyclic and that u and w are in distinct connected components of G. Suppose instead that $G + e$ contains a cycle C. Since G is acyclic, the cycle C must include the edge e. Using Proposition 4.29, there is closed walk without repeated vertices and edges corresponding to this cycle. By shifting the walk appropriately, we can write this closed walk as

$$w, e, u, f_1, v_1, f_2, v_2, \dots, f_k, w$$

Notice that $f_i \neq e$ for all i because there are no repeated edges. If follows that

$$u, f_1, v_1, f_2, v_2, \dots, f_k, w$$

is a u, w-walk in G, so u is in the same connected component of G, which is a contradiction. Therefore, $G + e$ must be acyclic.

□

Corollary 4.55. *Suppose that in the above algorithm we are at a stage i with $0 \leq i \leq n-2$ where we have that H_i is acyclic, has i edges, and satisfies $V_{H_i} = V_G$. We have the following:*

1. *$S_i = \{e \in E_G : e \notin E_{H_i}$ and the endpoints of e are in distinct connected components of $H_i\}$*

2. *$S_i \neq \emptyset$.*

3. *H_{i+1} is an acyclic graph with $i+1$ edges.*

Proof. 1. This follows immediately from the previous proposition and the fact that H_i is acyclic.

2. Notice that $|V_{H_i}| = |V_G| = n$, and $|E_{H_i}| = i \leq n-2$, so H_i is not connected by Corollary 4.46. Fix vertices u and w of H_i that are in distinct connected components of H_i. Since G is connected, there is a u,w-path in G, say
$$u = v_0, f_1, v_1, f_2, v_2, \dots, f_k, v_k = w$$
Since w is not in the same connected component as u in H_i, there is a smallest $i \geq 1$ such that v_i is not in the same connected component as u in H_i. We then have that v_{i-1} is in the same connected component as u in H_i. Thus, v_{i-1} and v_i are in distinct connected components in H_i, and so $f_i \in S_i$. Therefore, $S_i \neq \emptyset$.

3. Since H_{i+1} is obtained from H_i by adding a single edge from S_i, this follows immediately from part 1 above and part 2 of the previous proposition.

□

We've now seen that each $S_i \neq \emptyset$, so the above algorithm never gets stuck and results in acyclic subgraph H_{n-1} of G with $n-1$ edges and such that $V_{H_{n-1}} = V_G$. Since H_{n-1} is an acyclic graph with $n-1$ edges, Theorem 4.47 implies that H_{n-1} is a tree. Therefore, H_{n-1} is a spanning tree of G. We now answer the third question about why it is a minimum weight spanning tree.

Theorem 4.56. *H_{n-1} is a minimum weight spanning tree of G.*

Proof. We first prove by induction on i that H_i is contained in some minimal weight spanning tree of G (i.e. we argue that at each stage there is still the possibility of extending to a spanning tree of minimal weight). For the base case of $i = 0$, we have that H_0 has no edges, so certainly H_0 is contained in a minimum weight spanning tree.

For the inductive step, suppose that $0 \leq i \leq n-2$ and the statement is true for i, i.e. that H_i is contained in some minimum weight spanning tree of G. Fix a minimum weight spanning tree T of G such that H_i is contained in T. Suppose that the algorithm picks edge e, so $e \in S_i$ and $w(e) \leq w(f)$ for all $f \in S_i$. We need to argue that $H_{i+1} = H_i + e$ is contained in some minimum weight spanning tree of G. We have two cases.

- If e is an edge of T, then $H_i + e$ is contained in T, which is a minimum weight spanning tree of G.

- Suppose that e is not an edge of T. By Theorem 4.47, the graph $T + e$ has a cycle. Fix such a cycle C of $T + e$, and notice that e must be an edge of C because T is acyclic. Now $H_i + e$ is also acyclic because $e \in S_i$, so C must contain an edge f that is not an edge of $H_i + e$. Since $f \neq e$, it follows that f must be an edge of T. Since f is an element of a cycle of the connected graph $T + e$, we may use Proposition 4.33 to conclude that $T + e - f$ is connected. Combining this with the fact that $T + e - f$ has the same number of edges as T, which is $n - 1$, we conclude that $T + e - f$ is a spanning tree of G by Theorem 4.47. We also have that $H_i + f$ is a subgraph of T, so $H_i + f$ is acyclic, and hence $f \in S_i$ as well. Therefore, we must have
$$w(e) \leq w(f).$$
Thus $w(e) - w(f) \leq 0$, and hence
$$w(T + e - f) = w(T) + w(e) - w(f) \leq w(T).$$
Now T is minimum weight spanning tree, so we must have $w(T + e - f) = w(T)$, and hence that $T + e - f$ is also a minimum weight spanning tree. Since $H_{i+1} = H_i + e$ is contained in $T + e - f$, this completes the inductive step.

Since H_i is contained in a minimum weight spanning tree of G for all i with $0 \leq i \leq n-1$, it follows that H_{n-1} is contained in some minimum weight spanning tree T of G. We already know from above that H_{n-1} is a spanning tree of G. Thus, by Theorem 4.47, $H_i + e$ is not a tree for any new edge e. It follows that $H_{n-1} = T$, and hence H_{n-1} is a minimum weight spanning tree of G. □

We've finally answered our first three questions about Kruskal's Algorithm, so we know that it does indeed produce a minimum weight spanning tree of G. How do we implement it efficiently? As mentioned above, the difficult part is computing the sets S_i. We know from Corollary 4.55 that

$$S_i = \{e \in E_G : e \notin E_{H_i} \text{ and the endpoints of } e \text{ are in distinct connected components of } H_i\}$$

so instead of checking if an edge introduces a cycle, we can instead check if the endpoints are in distinct connected components of H_i. Of course, it is not immediately obvious how to do that. The essential idea is that we can keep track of the vertices of each of the connected components of the H_i throughout the algorithm by using the next result.

Proposition 4.57. *Let G be a graph, and let $u, w \in V$ be two vertices of G such that u and w are in distinct connected components of G. Consider the graph $G' = G + e$ where e is a new edge with endpoints u and w. Let $x \sim_G y$ mean that there is an x, y-walk in G and let $x \sim_{G'} y$ mean that there is an x, y-walk in G'. Moreover, for each $v \in V$, let C_v be the equivalence class of v under $\sim_G$, and let C'_v be the equivalence class of v under $\sim_{G'}$. Thus, C_v consists of the vertices of the connected component of v in G, while C'_v consists of the vertices of the connected component of v in G'. We have the following.*

1. *If $y \notin C_u$ and $y \notin C_w$, then $C'_y = C_y$.*
2. *If either $y \in C_u$ or $y \in C_w$, then $C'_y = C_u \cup C_w$.*

Proof. First notice that for any $a, b \in V$, if $a \sim_G b$, then $a \sim_{G'} b$ because an a, b-walk in G is an a, b-walk in G'.

1. Suppose that $y \notin C_u$ and $y \notin C_w$, so $y \not\sim_G u$ and $y \not\sim_G w$.
   - We first show that $C_y \subseteq C'_y$. Let $z \in C_y$ be arbitrary. We then have that $y \sim_G z$, so $y \sim_{G'} z$ from above, and hence $z \in C'_y$.
   - We now show that $C'_y \subseteq C_y$. Let $z \in C'_y$ be arbitrary. We then have that $y \sim_{G'} z$ and so we can fix a y, z-path P in G'. We claim that P does not contain the edge e. Suppose instead that P does include the edge e. Notice that since P is a path, it is a trail, so it must include e only once. Thus, either u or w occurs just before e on the path P, and if we cut off the path at this point, then we would have either an y, u-path in G or a y, w-path in G. Thus, either $y \sim_G u$ or $y \sim_G w$, each of which contradict our assumption. Thus, e does not occur in P. It follows that P is an y, z-path in G, so $y \sim_G z$ and hence $z \in C_y$.

   Combining these, we conclude that $C'_y = C_y$.

2. Suppose that $y \in C_u$, so $u \sim_G y$.
   - We first show that $C_u \cup C_w \subseteq C'_y$. Let $z \in C_u \cup C_w$ be arbitrary.
     - *Case 1:* Suppose that $z \in C_u$. We then have $u \sim_G z$. Since $u \sim_G y$, we can use symmetry and transitivity of $\sim_G$ to conclude that $y \sim_G z$. Therefore, $y \sim_{G'} z$ from above, and hence $z \in C'_y$.

- *Case 2:* Suppose that $z \in C_w$. We then have $w \sim_G z$, and so $w \sim_{G'} z$ from above. Since $u \sim_G y$, we also have $u \sim_{G'} y$ from above. Finally, notice that $u \sim_{G'} w$ via the new edge e. Using symmetry of $\sim_{G'}$, we have

$$y \sim_{G'} u \sim_{G'} w \sim_{G'} z$$

By transitivity of $\sim_{G'}$, if follows that $y \sim_{G'} z$, so $z \in C'_y$.

Therefore, in either case we have $z \in C'_y$. It follows that $C_u \cup C_w \subseteq C'_y$.

- We now show that $C'_y \subseteq C_u \cup C_w$. Let $z \in C'_y$ be arbitrary. We then have that $y \sim_{G'} z$, so we can fix a y,z-path P in G'.
  - *Case 1:* Suppose that P does not include the edge e. We then have that P is a y,z-path in G, so $y \sim_G z$. Since we also know that $u \sim_G y$, we can use transitivity of $\sim_G$ to conclude that $u \sim_G z$. In follows that $z \in C_u$, and hence $z \in C_u \cup C_w$.
  - *Case 2:* Suppose that P does include the edge e. Since P is path, we know that it is trail, and hence e occurs exactly once. On this path, we then have that u and w occur immediately before and after e. If u occurs immediately after e on P, then the portion of the path that starts with u and goes to the end is a u,z-path in G, so $u \sim_G z$ and hence $z \in C_u$. Similarly, if w occurs immediately after e on P, then we obtain a w,z-path in G, so $w \sim_G z$ and hence $z \in C_w$. Thus, we have $z \in C_u \cup C_w$.

  Therefore, in either case, we have $z \in C_u \cup C_w$. It follows that $C'_y \subseteq C_u \cup C_w$.

3. The case where $y \in C_w$, then the argument that $C'_y = C_u \cup C_w$ is completely analogous to the argument in 2.

□

We can use this result to efficiently implement Kruskal's Algorithm as follows. Suppose that G is a finite connected graph and $w \colon E \to \mathbb{R}^{\geq 0}$ is a weight function. First, sort the edges by weight in increasing order, and label all of the vertices with distinct numbers. Now go through the sorted list of edges in order once and do the following for each edge. Suppose that we are examining edge e with endpoints u and w. If u and w have the same label, do nothing but move on to the next edge. Suppose that u and w have distinct labels. We then have that u and w are in two distinct connected components of the current H_i, so $e \in S_i$, and its straightforward to check that it will be an element of S_i of minimal weight (because the edges are sorted and we're going through them in order). Thus, we add e to H_i to form H_{i+1}, and update the labels so that we "merge" all vertices that have a label equal to one of the labels of u and w (i.e. change all of the labels of vertices that have the same label as w to now all have the same label as u).

4.5 Vertex Colorings and Bipartite Graphs

Definition 4.58. *Let G be a graph and let $k \in \mathbb{N}^+$.*

- *A function $c \colon V \to [k]$ is a* k-coloring *of (the vertices of) G.*
- *We say that a coloring $c \colon V \to [k]$ is* proper *if $c(u) \neq c(w)$ whenever $u, w \in V$ are distinct adjacent vertices.*
- *We define $\chi(G)$, the* chromatic number *of G, to be the smallest $k \in \mathbb{N}^+$ such that G has a proper k-coloring.*

If G is a graph and $k \in \mathbb{N}^+$, then saying that $\chi(G) \leq k$ is equivalent to saying that there is a proper k-coloring of G. Notice also that if $k \leq \ell$, then any proper k-coloring of a graph G is automatically a proper ℓ-coloring of G. Thus, to prove that $\chi(G) = k$, we need to do two things:

- Show that there does indeed exist a proper k-coloring of G (this shows that $\chi(G) \leq k$).
- Show that there does *not* exist a proper $(k-1)$-coloring G (this shows that $\chi(G) \nleq k-1$).

Notice that $\chi(G) = 1$ if and only if G has no edges. For a more interesting example, the chromatic number of a cycle graph on 5 vertices is 3, i.e. $\chi(C_5) = 3$. To see that $\chi(C_5) \leq 3$, consider the following coloring $c\colon [5] \to [3]$ is a proper coloring of C_5:

- $c(1) = 1$
- $c(2) = 2$
- $c(3) = 1$
- $c(4) = 2$
- $c(5) = 3$

To show that $\chi(C_5) \nleq 2$, we need to show that there is no proper 2-coloring of C_5. Consider a supposed proper coloring $c\colon [5] \to [2]$ of C_5. We must have $c(1) \neq c(2)$ and $c(2) \neq c(3)$, so since the codomain of c has 2 elements, we must have $c(1) = c(3)$. Similarly, we have $c(3) \neq c(4)$ and $c(4) \neq c(5)$, so $c(3) = c(5)$. Therefore, we would need to have that $c(1) = c(5)$, contradicting the fact that 1 and 5 are adjacent. Thus, $\chi(C_5) \nleq 2$, and hence $\chi(C_5) = 3$.

How can we attempt to give a proper coloring of a finite graph G. One idea is do a *greedy* coloring. That is, fix an ordering $v_1, v_2, \dots, v_n$ of the vertices of G, say $v_1, v_2, \dots, v_n$. Now color the vertices in order by giving vertex v_i the least color that is possible. By formalizing this, we obtain the following result.

Proposition 4.59. *$\chi(G) \leq \Delta(G) + 1$ for all graphs finite G (where $\Delta(G)$ is the largest degree of a vertex of G).*

Proof. Let G be a finite graph. Fix an ordering $v_1, v_2, \dots, v_n$ of the vertices. We now define a coloring $c\colon V \to \mathbb{N}^+$ of the vertices in order recursively as follows. Let $c(v_1) = 1$. At stage $k+1$, once we've colored $v_1, v_2, \dots, v_k$, let

$$S_{k+1} = \{i \in [k] : v_i \text{ and } v_{k+1} \text{ are adjacent}\}$$

and notice that $|S_{k+1}| \leq \Delta(G)$. Define $c(v_{k+1})$ to be the least element of

$$\mathbb{N}^+ \backslash \{c(v_i) : i \in S_{k+1}\}$$

By doing this, we obtain a proper coloring c of G such that $c(v_i) \leq \Delta(G) + 1$ for all i because at each stage, the set $\{c(v_i) : i \in S_{k+1}\}$ that we've removed has at most $\Delta(G)$ many elements. Since c is a proper coloring of G using at most $\Delta(G) + 1$ many colors, we conclude that $\chi(G) \leq \Delta(G) + 1$. □

Notice that it is certainly possible that $\chi(G) < \Delta(G) + 1$. For example, if G is a graph with vertex set $[n]$, where n is adjacent to all other vertices but there are no other edges, then $\Delta(G) = n - 1$ but $\chi(G) = 2$. Thus, the above upper bound can be a bad upper bound in some cases. One may object that the greedy coloring described in the above proof does not necessarily use $\Delta(G) + 1$ many colors. For example, in our example G, a greedy coloring using any ordering of the vertices only actually uses 2 colors. This is true, but there is an even deeper problem. Although the greedy coloring does indeed give a proper coloring of G (and leads to above inequality), using the greedy coloring on a particular ordering of the vertices may *not* produce a coloring using exactly $\chi(G)$ many colors. For example, consider the graph P_4 having vertex set $[4]$ and edge set $\{\{1,2\}, \{2,3\}, \{3,4\}\}$. If one uses the ordering $1, 2, 3, 4$ of the vertices, then indeed the greedy coloring uses 2 colors. However, if one uses the ordering $1, 4, 2, 3$ of the vertices, then the greedy coloring uses 3 colors.

Definition 4.60. *A graph G is* bipartite *if it has a proper 2-coloring, i.e. if $\chi(G) \leq 2$. Equivalently, G is bipartite exactly when it is possible to partition V into two disjoint sets A and B such that every edge of G has one endpoint in A and one endpoint in B (so no edge of G has endpoints in the same set).*

Theorem 4.61. *Let G be a graph. The following are equivalent.*

1. *G is bipartite.*
2. *G has no cycles of odd length.*
3. *G has no closed walks of odd length.*

Proof.

- $1 \to 2$: We instead prove the contrapositive that $\neg 2 \to \neg 1$. Suppose then that G has a cycle of odd length. We show that G is not bipartite. Fix a closed walk

$$v_0, e_1, v_1, e_2, v_2, \ldots, v_{k-1}, e_k, v_k$$

without repeated edges or vertices (other than $v_0 = v_k$) and such that k is odd and at least 3 (since there are no loops in G). Since $k \geq 3$ is odd, we can fix $m \in \mathbb{N}^+$ with $k = 2m + 1$. Suppose now that $c \colon V \to [2]$ is proper coloring of G. We then must have $c(v_0) \neq c(v_1)$ and $c(v_1) \neq c(v_2)$, so since the codomain of c has 2 elements, it follows that $c(v_0) = c(v_2)$. A similar argument shows that $c(v_2) = c(v_4)$, and hence we conclude that $c(v_0) = c(v_4)$. In general, a simple induction shows that $c(v_0) = c(v_{2\ell})$ whenever $0 \leq \ell \leq m$. In particular, since $2m = k - 1$, we conclude that $c(v_0) = c(v_{k-1})$. Since $v_k = v_0$, this implies that $c(v_k) = c(v_{k-1})$, contradicting the fact that c is proper coloring of G. Therefore, G is not bipartite.

- $2 \to 3$: We instead prove the contrapositive that $\neg 3 \to \neg 2$. Suppose then G has a closed walk of odd length. Fix an odd length closed walk of smallest possible length, say it is

$$v_0, e_1, v_1, e_2, v_2, \ldots, v_{k-1}, e_k, v_k$$

with $v_0 = v_k$. Notice that $k \geq 3$ because G has no loops. We claim that there are no repeated vertices.

  - Suppose that there exists i with $0 < i < k$ such that $v_0 = v_i$. We then have $v_i = v_k$ as well, so

$$v_0, e_1, v_1, \ldots, v_{i-1}, e_i, v_i$$

and

$$v_i, e_{i+1}, v_{i+1}, \ldots, v_{k-1}, e_k, v_k$$

are both closed walks. Since the sum of the lengths of these walks is k, which is odd, either i is odd or $k - i$ is odd. In either case, we have produced a closed walk in G of shorter odd length, which is a contradiction.

  - Suppose that there exists i, j with $0 < i < j < k$ and such that $v_i = v_j$. We then have that

$$v_i, e_{i+1}, v_{i+1}, \ldots, v_{j-1}, e_j, v_j$$

and

$$v_0, e_1, v_1, \ldots, v_{i-1}, e_i, v_i, e_{j+1}, v_{j+1}, \ldots, v_{k-1}, e_k, v_k$$

are both closed walks. Since the sum of the lengths of these walks is k, which is odd, either $j - i$ is odd or $k - (j - i)$ is odd. In either case, we have produced a closed walk in G of shorter odd length, which is a contradiction.

Both cases lead to a contradiction, so our shortest closed walk of odd length has no repeated vertices and length at least 3. Therefore, G has a cycle of odd length by Proposition 4.30.

- $3 \to 1$: Suppose that G has no closed walks of odd length. We first that each connected component of G is bipartite. Consider an arbitrary connected component H of G. Fix an arbitrary $z \in V_H$. Define a coloring $c \colon V_H \to [2]$ as follows. Given $v \in V_H$, fix a shortest possible z, v-path in G, and define $c(v) = 1$ if this path has even length, and $c(v) = 2$ if this path has odd length. We claim that c is a proper coloring of H. To see this, suppose that $u, w \in V_H$ are adjacent and $c(u) = c(w)$. We then have a shortest possible z, u-path in H and a shortest possible z, w-path in H have the same parity (either both even or both odd), so the sum of their lengths is even. Thus, if we following a shortest z, u-path, then take then edge $\{u, w\}$ and then follow a shortest z, w-path backwards, we obtain a closed walk in G of odd length, which is a contradiction. Thus, H is bipartite.

  Since each of the connected components of G is bipartite, we can fix proper two coloring $c_H \colon V_H \to [2]$ of each connected component H of G. If we define $c \colon V_G \to [2]$ by letting $c(v) = c_H(v)$ for the unique connected component H containing v, then c is a proper coloring of G because two vertices in distinct connected components are not adjacent. Therefore, G is bipartite.

□

We know that a graph with n vertices has at most $\binom{n}{2} = \frac{n(n-1)}{2} = \frac{n^2-n}{2}$ many edges. How many edges can a bipartite graph with n vertices have? Recall that if $m, n \in \mathbb{N}^+$, then we defined $K_{m,n}$ to be the graph with vertex set $V = [m+n]$ and edge set

$$E = \{\{i, j\} : 1 \leq i \leq m \text{ and } m+1 \leq j \leq n\}.$$

Notice that $d(u) = n - m$ for all $u \in \{1, 2, \dots, m\}$, that $d(w) = m$ for all $w \in \{m+1, m+2, \dots, m+n\}$, and that $K_{m,n}$ has exactly mn many edges. Each $K_{m,n}$ is bipartite (by coloring the vertices $\{1, 2, \dots, m\}$ one color and the vertices $\{m+1, m+2, \dots, n\}$ the other), and has all possible edges between the two sets corresponding to the color classes. This is why we called $K_{m,n}$ a *complete bipartite graph*.

With this in mind, suppose that G is a bipartite graph with n vertices. Suppose that we fix a proper 2-coloring of G. If one of the color classes has m many vertices, then then other must have $n - m$ many vertices. Thus, the maximum number of edges that G can have is the number of edges in $K_{m,n-m}$, which we know is $m(n-m)$. In order to determine the maximum number of edges that a bipartite graph with n vertices can have, we need only figure out that largest possible value of $m(n-m)$ as let $m \in \mathbb{N}$ vary between $0 \leq m \leq n$.

Suppose that $n \in \mathbb{N}^+$. If n is even and we write $n = 2m$, then $K_{m,m}$ is a bipartite graph with $m^2 = (\frac{n}{2})^2 = \frac{n^2}{4}$ many edges. If n is odd and we write $n = 2m+1$, then $K_{m+1,m}$ is a bipartite graph with

$$m(m+1) = \frac{n-1}{2} \cdot \frac{n+1}{2} = \frac{n^2-1}{4}$$

many edges. Thus, in either case, we get

$$\left\lfloor \frac{n^2}{4} \right\rfloor$$

many edges. We now argue that this is best possible

Theorem 4.62. *Let $n \in \mathbb{N}^+$.*

- *If n is even, say $n = 2k$ with $k \in \mathbb{N}^+$, then the maximum number of edges that a bipartite graph with n vertices can have is*
$$\frac{n^2}{4} = k^2.$$
- *If n is odd, say $n = 2k+1$ with $m \in \mathbb{N}$, then the maximum number of edges that a bipartite graph with n vertices can have is*
$$\frac{n^2-1}{4} = k^2 + k.$$

Thus, in either case, the maximum number of edges that a bipartite graph with n vertices can have is $\lfloor \frac{n^2}{4} \rfloor$.

Proof. Let G be a bipartite graph with n vertices. Since G is bipartite, we can fix a proper 2-coloring of the vertices of G. Let A be the set of vertices given one color, and and let B be the set of vertices given the other color. Let $m = |A|$, so $|B| = n - m$. Each vertex in A is then adjacent to at most $|B|$ many vertices, so the number of edges in G is at most $|A| \cdot |B| = m(n - m)$. Thus, if we determine the maximum values of $m(n - m)$ as let m vary in the set $\{0, 1, 2, \dots, n\}$, then we will have a bound on the number of edges in any bipartite graph. In order to maximize this discrete function of m, we examine the continuous (indeed differentiable) function $f \colon \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = x(n - x) = nx - x^2$$

We want to maximize $f(x)$ on the closed interval $[0, n]$. We have

$$f'(x) = n - 2x$$

so $f'(x) > 0$ on $[0, \frac{n}{2}]$ and $f'(x) < 0$ on $[\frac{n}{2}, n]$. Thus, $f(x)$ is increasing on $[0, \frac{n}{2}]$ and decreasing on on $[\frac{n}{2}, n]$. Now if n is even, then $\frac{n}{2} \in \mathbb{N}$ with $0 \leq \frac{n}{2} \leq n$, and so the maximum occurs at $\frac{n}{2}$ with

$$f\left(\frac{n}{2}\right) = \frac{n}{2} \cdot \frac{n}{2} = \frac{n^2}{4}$$

Suppose that n is odd. Although the maximum of $f(x)$ occurs at $\frac{n}{2}$, this is not a natural number. However, $\frac{n-1}{2}$ is that largest natural number in the closed interval $[0, \frac{n}{2}]$, so as $f(x)$ is increasing on $[0, \frac{n}{2}]$, we conclude that the largest value of f at a natural number in the interval $[0, \frac{n}{2}]$ is

$$\begin{aligned} f\left(\frac{n-1}{2}\right) &= \frac{n-1}{2} \cdot \frac{n+1}{2} \\ &= \frac{n^2-1}{4}. \end{aligned}$$

Similarly, $\frac{n+1}{2}$ is that smallest natural number in the closed interval $[\frac{n}{2}, n]$, so as $f(x)$ is decreasing on $[\frac{n}{2}, n]$, we conclude that the largest value of f at a natural number in the interval $[\frac{n}{2}, n]$ is

$$\begin{aligned} f\left(\frac{n+1}{2}\right) &= \frac{n+1}{2} \cdot \frac{n-1}{2} \\ &= \frac{n^2-1}{4}. \end{aligned}$$

Thus, the maximum value of $m(n - m)$ across all m in the set $\{0, 1, 2, \dots, n\}$ equals $\frac{n^2-1}{4}$. □

Since the previous theorem gives an upper bound on the number of the number of edges in a bipartite graph, it also gives an upper bound on the number of edges in a graph that does not contain an odd cycle (so a 3-cycle, a 5-cycle, etc.). Somewhat surprisingly, if want to determine the maximum number of edges in a graph that does not contain a triangle (i.e. a 3-cycle), we obtain the exact same upper bound. To prove this, we use a different inductive approach.

Theorem 4.63. *We have the following.*

1. *If $k \geq 2$ and G is a graph with $n = 2k$ vertices and at least $k^2 + 1$ many edges, then G contains a triangle.*

2. *If $k \geq 1$ and G is a graph with $n = 2k + 1$ vertices and at least $k^2 + k + 1$ many edges, then G contains a triangle.*

Proof. 1. We prove this by induction on k.

- *Base Case:* Suppose that $k = 2$. Let G be a graph with $4 = 2 \cdot 2$ vertices and at least $2^2 + 1 = 5$ edges. We know that G is not bipartite by the previous theorem, so it has an odd cycle. Such a cycle must have length 3 (because there are only 4 vertices), so G contains a triangle.
- *Inductive Step:* Suppose that the statement is true for a fixed $k \geq 2$. Let G be a graph with $2(k+1) = 2k+2$ many vertices and at least $(k+1)^2 + 1$ many edges. Fix an edge in G, and call its endpoints u and w. If u and w have a common neighbor in G, then we have a triangle and we are done. Suppose then that u and w do not have a common height in G. For each of the other $2k$ vertices, at most one of u and w is adjacent to it, so there are at most $2k$ many other edges incident to at least one of u or w, not counting the edge $\{u, w\}$ itself. Thus, if we delete the two vertices u and w to form $G - \{u, v\}$, then the resulting graph has at most $2k+1$ many fewer edges than G. It follows that $G - \{u, v\}$ has at least

$$\begin{aligned}(k+1)^2 + 1 - (2k+1) &= k^2 + 2k + 1 + 1 - 2k - 1 \\ &= k^2 + 1\end{aligned}$$

many edges. By induction, $G - \{u, v\}$ has a triangle, so G does.

The result follows by induction.

2. We also prove this by induction on k.

- *Base Case:* Suppose that $k = 1$. Let G be a graph with $3 = 2 \cdot 1 + 1$ vertices and at least $1 \cdot 2 + 1 = 3$ edges. We then have that G is a triangle, so we are done.
- *Inductive Step:* Suppose that the statement is true for a fixed $k \geq 2$. Let G be a graph with $2(k+1) + 1 = 2k + 3$ many vertices and at least $(k+1)(k+2) + 1$ many edges. Fix an edge in G, and call its endpoints u and w. If u and w have a common neighbor in G, then we have a triangle and we are done. Suppose then that u and w do not have a common height in G. For each of the other $2k+1$ vertices, at most one of u and w is adjacent to it, so there are at most $2k+1$ many other edges incident to one of u or w, not counting $\{u, w\}$ itself. Thus, if we delete the two vertices u and w to form $G - \{u, v\}$, then the resulting graph has at most $2k+2$ many fewer edges than G. It follows that $G - \{u, v\}$ has at least

$$\begin{aligned}(k+1)(k+2) + 1 - (2k+2) &= k^2 + 3k + 2 + 1 - 2k - 2 \\ &= k^2 + k + 1 \\ &= k(k+1) + 1\end{aligned}$$

many edges. By induction, $G - \{u, v\}$ has a triangle, so G does.

The result follows by induction.

□

4.6 Matchings

Definition 4.64. *A* matching *in a graph is a set of edges such that no two distinct edges have a common endpoint.*

We will often be interested in matchings in bipartite graphs. The idea is that one side represents people and the other jobs/tasks. Historically, this was also viewed as men and women.

Definition 4.65. *Let G be a graph and let M be a matching in G.*

- *A set $S \subseteq V_G$. We say that M* saturates *S if every element of S is an endpoint of some edge in M.*

- *We say that M is a* perfect matching *if M saturates V, i.e. every vertex in G appears as an endpoint of some edge in M.*

- *We say that M is a* maximal matching *in G if $M \cup \{e\}$ is not a matching for every edge $e \notin E \backslash M$.*

- *We say that M is a* maximum matching *if it has at least as many edges as any other matching, i.e. if $|N| \leq |M|$ for all matchings N of G.*

Notice that any maximum matching in a graph G is a maximal matching. However, the converse is not true. For example, consider the graph P_4, so the vertex set is $[4] = \{1, 2, 3, 4\}$ and the edge set is $\{\{1, 2\}, \{2, 3\}, \{3, 4\}\}$. Notice that $M = \{\{2, 3\}\}$ is a maximal matching, but it is not a maximum matching because $M' = \{\{1, 2\}, \{3, 4\}\}$ is a matching with strictly more elements. Although it is relatively easy to determine if a given matching is a maximal matching (simply look through all of the edges in turn and check if each of their endpoints are saturated by M), it seems harder to determine if a matching is a maximum matching. The following definition will be essential to help us efficiently determine if a matching is a maximum matching.

Definition 4.66. *Let M be a matching in a graph G.*

- *An M-*alternating path *in G is a path*

$$v_0, e_1, v_1, e_2, v_2, \dots, v_{k-1}, e_k, v_k$$

in G where the e_i alternate between elements of M and element of $E \backslash M$ (i.e. either $e_1, e_3, e_5, \dots$ are all elements of M and $e_2, e_4, e_6, \dots$ are all elements of $E \backslash M$, or $e_1, e_3, e_5, \dots$ are all elements of $E \backslash M$ and $e_2, e_4, e_6, \dots$ are all elements of M).

- *An M-*augmenting path *in G is an M-alternating path*

$$v_0, e_1, v_1, e_2, v_2, \dots, v_{k-1}, e_k, v_k$$

in G where $k \geq 1$ and both v_0 and v_k are M-unsaturated (i.e. not the endpoint of any *edge in M).*

Notice that in an M-augmenting path, the first and last edges e_1 and e_k must both be elements of $E \backslash M$. However, an M-augmenting path is more than just an M-alternating path with this property because we require that *no* edge of M is incident to either v_0 or v_k, not just e_1 and e_k themselves. Before establishing our major theorem about M-augmenting paths, we first prove a useful lemma.

Lemma 4.67. *Let G be a finite graph with $d(v) \leq 2$ for all $v \in G$. We then have that every connected component of G is either a path or a cycle.*

Proof. Let H be an arbitrary connected component of G. Since G is finite, we know that H is finite, so we can fix a longest possible path

$$v_0, e_1, v_1, e_2, v_2, \dots, v_{k-1}, e_k, v_k$$

in G (note that possibly $k = 0$ and we have a trivial path). Since paths are trails, we know that the edges in this path are all distinct. For each i with $1 \leq i \leq n - 1$, we have that v_i is incident to the two distinct edges e_i and e_{i+1}. Since $d(v_i) \leq 2$ for each i with $1 \leq i \leq k - 1$, it follows that v_i is incident to no edges of G other than e_i or e_{i+1}. We now have a few cases.

- *Case 1:* Suppose that v_0 and v_k are incident to no edges besides e_1 and e_k respectively. We then have H consists of the vertices and edges on this path, and no other vertices/edges, so H is a path.

- *Case 2:* Suppose then v_0 is incident to another edge f. Notice that the other endpoint of f must be some v_i because otherwise we could extend our path to a longer one in H. Now this other endpoint can not be a v_i with $1 \leq i \leq k-1$ from above, so the other endpoint must be v_k. We then have that

$$v_0, e_1, v_1, e_2, v_2, \dots, v_{k-1}, e_k, v_k, f, v_0$$

  is a closed walk without repeated vertices/edges. Furthermore, no other edge can be incident to either v_0 or v_k because $d(v_0) \leq 2$ and $d(v_k) \leq 2$. We now have a cycle in H, and furthermore no other vertices or edges can be in H. Therefore, H is a cycle.

- *Case 3:* Suppose then v_k is incident to another edge f. Then following the argument in Case 2, the other endpoint of f must be v_0, and we conclude that H is a cycle.

□

We are now ready to prove our fundamental theorem about augmenting paths.

Theorem 4.68. *Let M be a matching in a finite graph G. The following are equivalent.*

1. *M is a maximum matching.*
2. *There is no M-augmenting path in G.*

Proof. We prove the contrapositive of each direction, so we show that M is *not* a maximum matching if and only if there exists an M-augmenting path in G.

Suppose first that there does exist an M-augmenting path in G, say it is

$$v_0, e_1, v_1, e_2, v_2, \dots, v_{k-1}, e_k, v_k$$

with $k \geq 1$. Since v_0 and v_k are both M-unsaturated, we know that $e_1 \in E\backslash M$ and $e_k \notin E\backslash M$. Since this path is M-alternating (because it is M-augmenting), we must have that $e_i \in E\backslash M$ for all odd i and $e_i \in M$ for all even i. In particular, it must be the case that k is odd. Let

$$M' = (M\backslash\{e_i : i \text{ is even}\}) \cup \{e_i : i \text{ is odd}\}$$

Notice that if $1 \leq i \leq k-1$, then v_i is incident to a unique edge in M' because it is incident to a unique edge of M. Also, since v_0 and v_k are both M-unsaturated, they are each incident to a unique edge in M'. Finally, for any vertex u not equal to any v_i, we have that u is incident to at most one edge in M' because it is incident to at most one edge in M. It follows that M' is matching in G. Now since k is odd, we have that $|M'| = |M| + 1$, so M' is matching in G with more elements than M. Therefore, M is not a maximum matching.

Suppose conversely that M is not a maximum matching. Fix a maximum matching M' in G, and notice that $|M| < |M'|$. Consider the subgraph H of G with vertex set $V_H = V_G$ and edge set the symmetric difference $E_H = M \Delta M'$, i.e. $E_H = (M\backslash M') \cup (M'\backslash M)$. Thus, an edge $e \in E_G$ appears in E_H when it is in exactly one of M or M'. Notice that since $|M'| > |M|$, we have

$$|M'\backslash M| > |M\backslash M'|.$$

For any $v \in V_H$, we know that v is incident to at most one edge in M and at most one edge in M', so we have that $d_H(v) \leq 2$ for all $v \in V_H$. Thus, by Lemma 4.67, each connected component of H is either a path or a cycle. Now any cycle in H must alternate edges between elements of $M\backslash M'$ and elements of $M'\backslash M$ because no vertex is incident two edges in M and no vertex is incident to edges in M'. Hence, each connected component in H that is a cycle has even length and an equal number of edges in both $M\backslash M'$

and $M'\backslash M$. Therefore, since $|M'\backslash M| > |M\backslash M'|$, there must exist a connected component in H that is path with strictly more edges in $M'\backslash M$ than in $M\backslash M'$. Fix such a path

$$v_0, e_1, v_1, e_2, v_2, \dots, v_{k-1}, e_k, v_k$$

Since each vertex is incident to at most edge of M and at most one edge of M', it follows that the e_i alternate between edges of $M'\backslash M$ and edges of $M\backslash M'$. Furthermore, since this path has strictly more edges in $M'\backslash M$ than in $M\backslash M'$, we must have $e_1, e_k \in M'\backslash M$. Thus, our path is an M-alternating path with $e_1, e_k \notin M$. Moreover, both v_0 and v_k must be M-unsaturated, because if either was incident to an edge in M, that edge would be in $M\backslash M'$ (since it is not e_1 and not e_k), which implies that it would appear in H, contradicting the fact that our path is a component in H. Therefore, we have shown the existence of an M-augmenting path in G. □

In many cases of applied interest, we are looking for matchings in bipartite graphs. Suppose then that G is bipartite graph. Fix a proper 2-coloring of G, and let X and Y be the corresponding color classes. We want to a simple necessary and sufficient condition for when there is a matching G that saturates X.

Definition 4.69. *Let G be graph. Given a set $T \subseteq V$, we define*

$$N(T) = \{v \in V : v \textit{ is adjacent to some element of } T\}$$

In other words, $N(T)$ is the set of neighbors *of elements of T.*

Theorem 4.70 (Hall's Marriage Theorem)**.** *Let G be a finite bipartite graph. Fix a proper 2-coloring of G, and let X and T be the corresponding color classes. There exists a matching of G that saturates X if and only if $|T| \leq |N(T)|$ for all $T \subseteq X$.*

Proof. Suppose first that there exists a matching of G that saturates X, and fix such a matching M. Let $T \subseteq X$ be arbitrary. Since M saturates X and $T \subseteq X$, we can define a function $f\colon T \to N(T)$ by letting $f(x)$ be the unique vertex that x is matched to in M. Since M is a matching, this function is injective, so $|T| \leq |N(T)|$.

Suppose conversely that there does *not* exist a matching of G that saturates X. We build a set $T \subseteq X$ with $|T| > |N(T)|$. First, fix a maximum matching M of G. Now M does not saturate X by assumption, so we can fix an $a \in X$ that is M-unsaturated. Consider the set of all M-alternating paths whose first vertex is a. Let B be the set of all endpoints of such paths that are elements of X, i.e.

$$B = \{x \in X : \text{There exists an } M\text{-alternating } a, x\text{-path in } G\}.$$

Similarly, let C be the set of all endpoints of such paths that are elements of Y, so

$$C = \{y \in Y : \text{There exists an } M\text{-alternating } a, y\text{-path in } G\}.$$

Thus, $B \subseteq X$, $C \subseteq Y$, and $a \in B$ (because the trivial path of just a is an M-alternating path). We have the following:

1. Every element of $B\backslash\{a\}$ is M-saturated and its matched partner is in C: Let $b \in B\backslash\{a\}$ be arbitrary. By definition of B, we can fix an M-alternating a, b-path P in G. Notice that P has even length of at least 2 because G is bipartite, $a, b \in X$, and $b \neq a$. Since the edges of P alternate between elements of $E\backslash M$ and M, and since P starts with an edge in $E\backslash M$ (because a is M-unsaturated), it follows that the last edge of P is an element of M. Thus, b is M-saturated. Furthermore, the penultimate vertex of P is matched to b, is an element of Y, and is the endpoint of the M-alternating path starting with a that is obtained by deleting the last vertex and edge of P. Therefore, the matched partner of b is an element of C.

2. Every element of C is M-saturated and its matched partner is in $B\backslash\{a\}$: Let $c \in C$ be arbitrary. By definition of C, we can fix an M-alternating a, c-path P in G. Notice that P has odd length because G is bipartite, $a \in X$, and $c \in Y$. Since the edges of P alternate between elements of $E\backslash M$ and M, and since P starts with an edge in $E\backslash M$ (because a is M-unsaturated), it follows that the last edge of this path is an element of $E\backslash M$. Furthermore, since M is a maximum matching, we know that P can not be an M-augmenting path by Theorem 4.68. Now P is M-alternating, and a is M-unsaturated, so it must be the case that c is M-saturated. Let $b \in X$ be the matched partner of c. Notice that b does not occur on P because c does not occur before the last vertex, and the last edge is an element of $E\backslash M$. Thus, if we add on the edge $\{c, b\}$ and the vertex b to the end of P, we obtain an M-alternating a, b-path in G, so $b \in B$. Also, notice that $b \neq a$ because a is M-unsaturated. Therefore, the matched parter of c is an element of $B\backslash\{a\}$.

3. $|C| = |B\backslash\{a\}|$: By 1, we can define a function from $B\backslash\{a\} \to C$ sending an element to its matched partner. This function is injective because M is a matching, so $|B\backslash\{a\}| \leq |C|$. Similarly, $|C| \leq |B\backslash\{a\}|$ by 2. It follows that $|C| = |B\backslash\{a\}|$

4. $N(B) \subseteq C$: Let $y \in N(B)$ be arbitrary. Since $y \in N(B)$, we can fix $b \in B$ such that b is adjacent to y. Since $b \in B$, we can fix an M-alternating a, b-path P in G. If y is a vertex on P, then by cutting off P at y we obtain an M-alternating a, y-path in G, so $y \in C$. Suppose then that y is not a vertex on P. As in 1, notice that P has even length and the last edge of P is an element of M. Since y is not a vertex on P and we know that the penultimate vertex of P is the matched partner of b, it follows that $\{b, y\} \notin M$. Thus, if we add on the edge $\{b, y\}$ and the vertex y to the end of P, we obtain an M-alternating a, y-path in G, so $y \in C$. It follows that $N(B) \subseteq C$.

Combining 3 and 4, we have

$$|N(B)| \leq |C| = |B\backslash\{a\}| = |B| - 1$$

Therefore, we may let $T = B$. □

Instead of thinking about finding a large matching, we now move on to consider finding a "good" matching in a certain sense. Suppose that we have two groups of n vertices and each one side has a ordering of the other in terms of preference. Thus, we have the complete bipartite graph $K_{n,n}$ where the vertex set is partitioned into two sets U (think of uppercase letters) and L (think of lowercase letters) of size n such that every vertex of U is adjacent to every vertex of L. Suppose furthermore that for each $A \in U$, we have a permutation of L which we can think of as an ordering $<_A$ of L. Similarly, for each $x \in L$, we have a permutation of Y which we can think of as an ordering $<_x$ of U. These orderings codify the preferences of the vertices. For example, if $U = \{A, B, C, D\}$ and $L = \{w, x, y, z\}$, we may have the following lists of preferences:

A	B	C	D
x	w	z	z
y	x	w	w
z	y	y	x
w	z	x	y

w	x	y	z
C	C	C	B
B	A	B	C
A	D	A	A
D	B	D	D

In this setting, there are of course many perfect matchings (in fact there are $n!$ many of them). However, we want to find a "good" matching. There are several ways one could define a notion of "good" (as many vertices as possible are paired with their first choice, as few vertices as possible are paired with their last choice, minimizing the sum of the ranks of the matched pairs, etc.), but we opt for a notion that tries to avoid "rogue" pairs.

Definition 4.71. *A matching is* stable *if there do not exist matched pairs $\{x, A\}, \{y, B\} \in M$ (where $A, B \in U$ and $x, y \in L$) such that $B <_y A$ and $x <_A y$, i.e. such that y prefers A to B and also A prefers y to x.*

If we think of our matching as provided marriages between uppercase people U and lowercase people L, then a stable matching is one where there does not exist two people who would prefer each other to their current spouses, thus warding off infidelity. Does every list of preferences have a stable matching? If so, is it unique? Also, how could we find one? There is a clever algorithm to form a matching that is useful in answering all of these questions. The idea is to pick one side and have it do a sequence of proposals to the other side. Say that we have the L vertices do the proposals to the U vertices. At the first stage, each L vertex approaches their first choice of a vertex U and proposes. Each U vertex that now has a proposal tells their favorite current suitor to come back the next round, and rejects the others telling them to never come back. We call the resulting pair engaged. In the next round, each engaged element of L returns to their engaged partner, and element of L that is not currently engaged proposes to their highest choice of a vertex in U that has not yet rejected them and proposes. At this point, each U vertex (even if they are currently engaged) that now has a proposal tells their favorite current suitor/engaged partner to come back the next round, and also rejects the others (possibly including the current engaged partner) telling them to never come back. As before, we call the resulting pairs engaged. We continue this process until we reach a round where everybody is engaged, and we take this matching. For example, consider the above lists of preferences:

A	B	C	D
x	w	z	z
y	x	w	w
z	y	y	x
w	z	x	y

w	x	y	z
C	C	C	B
B	A	B	C
A	D	A	A
D	B	D	D

With the elements of L proposing, we get the following run of the algorithm (where bold indicates the engaged element of L):

Round	A	B	C	D
1		$\mathbf{z}$	$\mathbf{w}, x, y$	
2	$\mathbf{x}$	$\mathbf{y}, z$	$\mathbf{w}$	
3	$\mathbf{x}$	$\mathbf{y}$	$w, \mathbf{z}$	
4	$\mathbf{x}$	$\mathbf{w}, y$	$\mathbf{z}$	
5	$\mathbf{x}, y$	$\mathbf{w}$	$\mathbf{z}$	
6	$\mathbf{x}$	$\mathbf{w}$	$\mathbf{z}$	$\mathbf{y}$

Thus, we get the matching

$$\{w, B\} \quad \{x, A\} \quad \{y, D\} \quad \{z, C\}$$

We can also switch and have elements of U proposing, giving the following run of the algorithm:

Round	w	x	y	z
1	$\mathbf{B}$	$\mathbf{A}$		$\mathbf{C}, D$
2	$\mathbf{B}, D$	$\mathbf{A}$		$\mathbf{C}$
3	$\mathbf{B}$	$\mathbf{A}, D$		$\mathbf{C}$
4	$\mathbf{B}$	$\mathbf{A}$	$\mathbf{D}$	$\mathbf{C}$

Notice that this results in the same matching. Of course, we now have a couple of questions. First, does this procedure always terminate in a matching? For example, is it possible that an element of the proposing set is rejected by everyone? If we know that the process does terminate in a matching, then the big question is the resulting matching stable? We start with the termination question. Suppose that the elements of L propose to the elements of U. By definition of the algorithm, we have the following properties:

1. If $x \in L$ is rejected by $A \in U$, then x never proposes to A again.
2. If $A \in U$ is engaged to an element of L at stage k, then A is engaged to some (possibly different) element of L at all later stages.

Furthermore, building on these facts, a closer look at the algorithm reveals that we also have the following properties.

3. If we fix $x \in L$, then sequence of elements of U that x proposes to each day is decreasing (not necessarily strictly) through its preference list and never skips anybody on their list.
4. If we fix $A \in U$, the sequence of elements of L that A is engaged to might start out empty, but then it is increasing (not necessarily strictly) though its preference list, and it might skip people on the list.

Now if we ever reach a stage where each element of U has a current proposal by some element of L, then we stop the algorithm and take the corresponding matching. With this in mind, we first argue that no element of L can be rejected by every element of U. To see this, suppose that we are at a stage where a given $x \in L$ gets rejected by the $(n-1)^{st}$ person on their list. By property 3 above, it follows that x has now been rejected by each of the first $n-1$ elements of their list. Now using property 2, it follows that on the next day, those $n-1$ people will each have a suitor, and then x will propose to the n^{th} person on their list. Thus, each of the n elements of U have a suitor, and since L also has n elements, it follows that each element of U has a unique suitor. Thus, the algorithm must terminate in a matching. Furthermore, since at least one rejection happens at each stage that does not produce the final matching, the above argument shows that algorithm terminates in at most $n(n-2)+2$ many steps (although this can be improved a bit). We now prove the algorithm produces a stable matching.

Theorem 4.72. *Suppose that the elements of L propose to the elements of U. The matching produced by the algorithm is stable.*

Proof. Suppose that the matching produces pairs $\{x, A\}$ and $\{y, B\}$. Suppose that y prefers A to B. Then at some stage, y must have proposed to A by property 3 above. Since y is not paired with A, it follows that A must have rejected y at some (possibly later) stage in favor of somebody that A preferred. Since the sequence of engagements of A only increases by property 4, it follows that A prefers x to y. □

A	B	C	D
x	w	z	z
y	x	w	w
z	z	y	x
w	y	x	y

w	x	y	z
C	D	C	B
B	A	B	C
A	C	A	A
D	B	D	D

With the lowercase letters proposing, we get the following run:

Round	A	B	C	D
1		**z**	**w**, y	**x**
2		y, **z**	**w**	**x**
3	**y**	**z**	**w**	**x**

Thus, we get the matching

$$\{w, C\} \quad \{x, D\} \quad \{y, A\} \quad \{z, B\}$$

If we run it in the other order, we get

Round	w	x	y	z
1	$\mathbf{B}$	$\mathbf{A}$		$\mathbf{C}, D$
2	$\mathbf{B}, D$	$\mathbf{A}$		$\mathbf{C}$
3	$\mathbf{B}$	$A, \mathbf{D}$		$\mathbf{C}$
4	$\mathbf{B}$	$\mathbf{D}$	$\mathbf{A}$	$\mathbf{C}$

Thus, we get the matching

$$\{w, B\} \quad \{x, D\} \quad \{y, A\} \quad \{z, C\}$$

Both of these matchings are stable by the above argument. In particular, there can be more than stable matching. This gives the questions of whether there is a "best" stable matching.

Definition 4.73. *Let $x \in L$ and let $A \in U$.*

- *We say that A is* feasible *for x if $\{x, A\}$ occurs in some stable matching. Similarly, we say that x is* feasible *for A if $\{x, A\}$ occurs in some stable matching.*
- *We say that A is* optimal *for x if A is the highest element of x's preference list that is feasible for x. Similarly, we say that x is* optimal *for A if x is the highest element of A's preference list that is feasible for A.*

Theorem 4.74. *If L does the proposing, then for all $x \in L$, the partner produces by the algorithm is optimal for x.*

Proof. We show that if x is rejected by A at some stage, then A is not feasible for x. We do this by induction on the stage of the construction.

- Suppose that A rejects x at the first stage. We need to show that A is not feasible for x. Let M be an arbitrary matching containing $\{x, A\}$. Since A rejects x at the first stage, we know that A is engaged to some y at the end of the first stage and that A prefers y to x. Fix B such that $\{y, B\} \in M$. We then have that A prefers y to x (from above) and y prefers A to B (since A was y's first choice as this is the first round), so M is not stable. Thus, any matching containing $\{x, A\}$ is not stable, so A is not feasible for x.
- Suppose now that we are at stage k, and we know that any $z \in L$ that has been rejected by some $C \in U$ at a stage before k, then C is not feasible for that z. Suppose now that A rejects x at stage k. We need to show that A is not feasible for x. Let M be an arbitrary matching containing $\{x, A\}$. Since A rejects x at the the stage k, we know that A is engaged to some y at the end of stage k and that A prefers y to x. Fix B such that $\{y, B\} \in M$. If B is not feasible for y, then M is not stable definition. Suppose then that B is feasible for y. Since y is engaged to A at stage k, we know that y was rejected by all elements of U above A in earlier rounds, and hence no element above A is feasible for y by induction. Since B is feasible for y, we must have that y prefers A to B. Combining this with the fact that A prefers y to x (from above), we conclude that M is not stable. Thus, any matching containing $\{x, A\}$ is not stable, so A is not feasible for x.

We have shown that if x is rejected by A at some stage, then A is not feasible for x. Since x is matched with the highest ranked element of U that does not reject x, it follows that the partner produced for x is optimal for x. □

Thus far, we've been working in the graph $K_{n,n}$ where each vertex on one side ranks the elements of the others. Suppose instead that we work in K_{2n} where each vertex ranks *all* other vertices. Think about this as taking a group of $2n$ people and trying to pair off roommates. Although this problem superficially seems completely analogue, the lack of two "sides" changes the situation dramatically. In fact, there may not be a stable matching! For example, suppose that $n = 2$, so we have 4 people who rank the other 3 as follows:

A	B	C	D
B	C	A	A
C	A	B	B
D	D	D	C

To see that there is no stable matching, simply look at who is matched with D (who is the lowest ranked vertex for all others).

- If we match $\{A, D\}$, then we must match $\{B, C\}$, and then A and C form an unstable pair.
- If we match $\{B, D\}$, then we must match $\{A, C\}$, and then A and B form an unstable pair.
- If we match $\{C, D\}$, then we must match $\{A, B\}$, and then B and C form an unstable pair.

Therefore, this is no stable matching.

4.7 Planar Graphs

Definition 4.75. *Let G be a graph. A* planar embedding *of G is way to draw G in the plane such that all vertices are represented by points, all edges are represented by continuous paths, and no two distinct edges cross (except at the endpoints). G is planar if there exists a planar embedding of G.*

One can make this definition more formal be representing edges by continuous functions $g\colon [0,1] \to \mathbb{R}^2$ such that $g(0)$ is one endpoint and $g(1)$ is the other endpoint. A careful treatment of this material relies on some important properties of continuous functions, and hence relies on some analysis and topology. We (obviously) don't have those tools, so we will proceed a bit more intuitively. However, rest assured that all of these results can be made very precise with the appropriate tools.

Definition 4.76. *Given a planar embedding of a graph G, we divide the plane (minus the image of the embedding) into regions that we call* faces*.*

Theorem 4.77 (Euler's Formula)**.** *Let G be a planar embedding of a connected planar multigraph. If this embedding has n vertices, m edges, and c faces, then $n - m + c = 2$.*

Proof. The proof is by induction on m.

- *Base Case:* Suppose that $m = 0$. Since G is connected, we must have $n = 1$ and hence $c = 1$ as well. Notice that $2 - 1 + 1 = 2$.
- *Inductive Step:* Suppose the result is true for all connected planar multigraphs with m edges. Let G be a connected planar multigraph with n vertices, $m + 1$ many edges, and c faces. We have two cases.
  - *Case 1:* G has no cycles, so G is a tree. We then have $m = n - 1$ by Theorem 4.41 and $c = 1$ (because there are no cycles), so

    $$\begin{aligned} n - m + c &= n - (n-1) + 1 \\ &= 2. \end{aligned}$$

  - *Case 2:* G has a cycle. Fix an edge e in a cycle, and let $G' = G - e$. Notice that G' is a connected planar graph (using Proposition 4.33) with n vertices and m edges. Furthermore, G' has $c - 1$ many faces because the faces on other side of that edge become one. By induction, we have

    $$n - m + (c - 1) = 2.$$

    Therefore

    $$n - (m + 1) + c = 2$$

    and hence the statement is true for G.

The result follows by induction.

□

Definition 4.78. *Suppose we have a planar embedding of a connected multigraph G. Given a face $f \in F$, we define the length of f, denoted $\ell(F)$ to be the length of the walk which traverses the boundary of the face. Notice that if F is on both "sides" of an edge, then that edge is counted twice.*

Proposition 4.79. *If G is a connected planar multigraph with m edges, then $\sum_{f\in F} \ell(f) = 2m$.*

Proof. Every edge has 2 "sides" so is counted twice on the left. □

Proposition 4.80. *Let G be a connected planar graph with n vertices and m edges. If $n \geq 3$, then $m \leq 3n-6$.*

Proof. Suppose that $n \geq 3$. For each face $f \in F$, we have $\ell(f) \geq 3$ because G is a connected graph with at least 3 vertices. Thus, the sum on the left above is at least $3c$ and so

$$2m = \sum_{f\in F} \ell(f) \geq 3c$$

It follows that $c \leq \frac{2}{3} \cdot m$. Now using Euler's Theorem, we have

$$\begin{aligned} m + 2 &= n + c \\ &\leq n + \frac{2}{3} \cdot m \end{aligned}$$

It follows that

$$\frac{1}{3} \cdot m \leq n - 2$$

and hence

$$m \leq 3n - 6.$$

□

Corollary 4.81. *K_5 is not planar.*

Proof. Notice that K_5 has 5 vertices and $\binom{5}{2} = 10$ edges. Since $3 \cdot 5 - 6 = 9$, it follows from Proposition 4.80 that K_5 is not planar. □

Proposition 4.82. *Let G be a connected planar graph with n vertices, m edges, and no triangles (i.e. no 3-cycles). If $n \geq 3$, then $m \leq 2n - 4$.*

Proof. For each face, we have at least 4 edges on its boundary. Thus, the sum on the left above is at least $4c$ and so $2m \geq 4c$. It follows that $c \leq \frac{1}{2} \cdot m$. Now we have

$$\begin{aligned} m + 2 &= n + c \\ &\leq n + \frac{1}{2} \cdot m \end{aligned}$$

It follows that

$$\frac{1}{2} \cdot m \leq n - 2$$

and hence

$$m \leq 2n - 4.$$

□

Corollary 4.83. *$K_{3,3}$ is not planar.*

Proof. Notice that $K_{3,3}$ is bipartite so has no triangles. Now $K_{3,3}$ has 6 vertices and $3 \cdot 3 = 9$ edges. Since $2n - 4 = 2 \cdot 6 - 4 = 8$, it follows from Proposition 4.82 that $K_{3,3}$ is not planar. □

Proposition 4.84. *If G is a finite planar graph, then there exists $v \in V$ with $d(v) \leq 5$.*

Proof. Suppose that G is finite planar graph with n vertices and m edges. If $n \leq 2$, then this result is trivial. Suppose then that $n \geq 3$. If $d(v) \geq 6$ for all $v \in V$, then

$$2m = \sum_{v \in V} d(v) \geq 6n$$

so $m \geq 3n$. However, this contradicts Proposition 4.80 (which holds even if G is not connected by Homework 16). Therefore, there must exist $v \in V$ with $d(v) \leq 5$. □

Proposition 4.85. *$\chi(G) \leq 6$ for every finite planar graph G.*

Proof. The proof is by induction on the number of vertices of G. Notice that if G has one vertex, then the result is trivial. Suppose then that $\chi(G) \leq 6$ for all finite planar graphs G with n vertices. Consider an arbitrary finite planar graph with $n + 1$ vertices. By Proposition 4.84, we can fix a vertex v with $d(v) \leq 5$. Now the graph $G - v$ is planar and has n vertices, so by induction we know that $\chi(G - v) \leq 6$. Fix a proper 6-coloring $c \colon V \backslash \{v\} \to [6]$ of $G - v$. Now v has at most 5 neighbors in G, so we there exists $i \in [6]$ such that $c(w) \neq i$ for all w adjacent to v. Thus, if we extend c by letting $c(v) = i$ for some such i, then we have a proper coloring of G using at most 6 colors. The result follows by induction. □

In fact, with a little more work, we can prove the following.

Theorem 4.86. *$\chi(G) \leq 5$ for every planar graph G.*

Proof. The proof is by induction on the number of vertices of G. Notice that if G has one vertex, then the result is trivial. Suppose then that $\chi(G) \leq 6$ for all finite planar graphs G with n vertices. Consider an arbitrary finite planar graph with $n + 1$ vertices. By Proposition 4.84, we can fix a vertex v with $d(v) \leq 5$. Now the graph $G - v$ is planar and has n vertices, so by induction we know that $\chi(G - v) \leq 6$. Fix a proper 6-coloring $c \colon V \backslash \{v\} \to [5]$ of $G - v$. Now if there is an $i \in [5]$ such that $c(w) \neq i$ for all w adjacent to v, then we obtain a proper 5-coloring of G as in the previous proposition.

Suppose then that all 5 colors occur on the neighbors of v. In some planar embedding of G, call the neighbors w_1, w_2, w_3, w_4, w_5 in a clockwise circle around v. Consider the subgraph $H_{1,3}$ of $G - v$ induced by the vertices currently colored 1 and 3. If w_1 and w_3 are in different connected components of $H_{1,3}$, then we can switch the colors 1 and 3 in the connected component of w_1 (so in particular w_1 is now colored 3), which then allows us to obtain a proper coloring of G with 5 colors by coloring w with 1.

Suppose then that w_1 and w_3 are in the same connected component of $H_{1,3}$. We can then fix a w_1, w_3-path P in $H_{1,3}$ of vertices alternating in color between 1 and 3. We now try the same strategy with w_2 and w_4 by considering the subgraph $H_{2,4}$ of $G - v$ induced by the vertices currently colored 2 and 4. Notice that we w_2 and w_4 can not be in the same connected component of $H_{2,4}$, because otherwise we would have w_2, w_4-path in $H_{2,4}$ of vertices alternating in color between 2 and 4, which would have to cross P, contradicting planarity (notice that the paths can't cross at a vertex because the vertices on the paths have different colors). Since w_2 and w_4 are not in the same connected component of $H_{2,4}$, we can switch the colors 2 and 4 in the connected component of w_2 (so in particular w_2 is now colored 4), which then allows us to obtain a proper coloring of G with 5 colors by coloring v with 2. The result follows by induction. □

In fact, the following much (much) harder result is true.

Theorem 4.87 (Four Color Theorem, Appel-Haken)**.** *$\chi(G) \leq 4$ for every planar graph G.*

We now examine convex regular polyhedra, which are 3-dimensional shapes each of whose faces is a convex polygon with the same number of edges, and such that the number of faces that meet at any point is equal throughout. These convex regular polyhedra can be viewed as graphs embedded on a sphere, but notice that a graph is can be embedded on the plane exactly when it can be embedded on the sphere. One can see this by picking a point on the sphere (not hit by any vertex/edge) and doing a stereographic projection. Thus, we can study these polyhedra by studying planar graphs such that $d(v)$ is constant for all $v \in V$ and $\ell(f)$ is constant for all $f \in F$. Let d be the common degree of vertices, and let ℓ be the common length of faces. Notice that $d \geq 3$ and $\ell \geq 3$. We also have

$$dn = 2m = \ell c$$

Now using Euler's Theorem, we conclude that

$$\begin{aligned} 2 &= n - m + c \\ &= \frac{2m}{d} - m + \frac{2m}{\ell} \end{aligned}$$

and hence

$$\frac{1}{d} + \frac{1}{\ell} = \frac{1}{2} + \frac{1}{m}$$

Since $\frac{1}{m} > 0$, it follows that

$$\frac{1}{d} + \frac{1}{\ell} > \frac{1}{2}$$

Now we know that $d \geq 3$, so if $\ell \geq 6$, then we would have

$$\frac{1}{d} + \frac{1}{\ell} \leq \frac{1}{3} + \frac{1}{6} = \frac{1}{2}$$

which is a contradiction. Similarly, we know that $\ell \geq 3$, so if $d \geq 6$, then we would have

$$\frac{1}{d} + \frac{1}{\ell} \leq \frac{1}{6} + \frac{1}{3} = \frac{1}{2}$$

which is a contradiction. Therefore, we must have $3 \leq d \leq 5$ and $3 \leq \ell \leq 4$. Furthermore, we can not have *both* $d \geq 4$ and $\ell \geq 4$ because this would imply that

$$\frac{1}{d} + \frac{1}{\ell} \leq \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

which is a contradiction. This gives the following possible pairs (d, ℓ):

$$(3,3) \quad (3,4) \quad (3,5) \quad (4,3) \quad (5,3)$$

Now from d and ℓ, we can compute m using the equation

$$\frac{1}{d} + \frac{1}{\ell} = \frac{1}{2} + \frac{1}{m}$$

Furthermore, once we have m as well, then we can determine n and c from the equation

$$dn = 2m = \ell c.$$

Doing all the calcuations, we conclude the following:

- $d = 3$ and $\ell = 3$: We then have $m = 6$, so $n = 4$ and $c = 4$.

- $d = 3$ and $\ell = 4$: We then have $m = 12$, so $n = 8$ and $c = 6$.
- $d = 3$ and $\ell = 5$: We then have $m = 30$, so $n = 20$ and $c = 12$.
- $d = 4$ and $\ell = 3$: We then have $m = 12$, so $n = 6$ and $c = 8$.
- $d = 5$ and $\ell = 3$: We then have $m = 30$, so $n = 12$ and $c = 20$.

d	ℓ	m	n	c	poly
3	3	6	4	4	tetrahedron
3	4	12	8	6	cube
3	5	30	20	12	dodecahedron
4	3	12	6	8	octahedron
5	3	30	12	20	icosahedron

4.8 Ramsey Theory

We begin with the following result.

Proposition 4.88. *Given an arbitrary coloring of the edges of K_6 with two colors, there always exists a triangle such that all edges have the same color.*

Proof. Consider an arbitrary coloring of the edges of K_6 with two colors. Call the colors red and blue. Pick an arbitrary vertex u. Since u is incident to 5 total edges, either u is incident to at least 3 red edges or u is incident to at least 3 blue edges. We now have two cases.

- *Case 1:* Suppose that u is incident to at least 3 red edges. Fix 3 such edges, and call the other endpoints of these edges w_1, w_2, and w_3. Now if the 3 edges with endpoints amongst the w_i are all blue, then w_1, w_2, and w_3 form a blue triangle. On the other hand, if any edge with endpoints w_i and w_j where $i \neq j$ is red, then we obtain a red triangle by looking at the vertices u, w_i, and w_j.
- *Case 2:* Suppose that u is incident to at least 3 blue edges. Fix 3 such edges, and call the other endpoints of these edges w_1, w_2, and w_3. Now if the 3 edges with endpoints amongst the w_i are all red, then w_1, w_2, and w_3 form a red triangle. On the other hand, if any edge with endpoints w_i and w_j where $i \neq j$ is blue, then we obtain a blue triangle by looking at the vertices u, w_i, and w_j.

Thus, we always obtain either a red triangle or a blue triangle (or possibly both). □

Is 6 best possible? In other words, is it true that every coloring of the edges of K_6 with two colors always has a monochromatic (i.e. either all red or all blue) triangle? Or is there a coloring of the edges of K_5 with two colors such that there are no monochromatic triangles? It turns out that the latter is true, so 6 is indeed best possible. To see this, consider K_5 with vertex set $[5] = \{1, 2, 3, 4, 5\}$. Think about these vertices as forming an outer 5-cycle in order. Color the edges of this cycle RED and color all edges BLUE. More formally, let

$$\begin{aligned} \text{RED} &= \{\{1,2\},\{2,3\},\{3,4\},\{4,5\},\{1,5\}\} \\ \text{BLUE} &= \{\{1,3\},\{1,4\},\{2,4\},\{2,5\},\{3,5\}\} \end{aligned}$$

Since any 3-element subset of the five vertices will contain at least one "consecutive" pair of numbers (where we consider 1 and 5 to be "consecutive"), and at least one that is not "consecutive", it follows that this coloring has no monochromatic triangle.

Can we generalize these ideas? If we color the edges of K_{10}, can you always find four vertices such that all edges between them have the same color, i.e. can you always find a monochromatic (or homogeneous) subgraph isomorphic to K_4? What about K_{20}?

Definition 4.89. *Let $k, \ell \in \mathbb{N}^+$. Define $R(k, \ell)$ to be the least $n \in \mathbb{N}^+$ (if it exists) such that whenever all of the edges of K_n are colored with either red or blue, either there exists a subset $A \subseteq V$ with $|A| = k$ such that all edges having both endpoints in A are red, or there exists a subset $B \subseteq V$ with $|B| = \ell$ such that all edges with both endpoints in B are blue.*

In other words, $R(k, \ell)$ is the least $n \in \mathbb{N}^+$ such that whenever all of the edges of K_n are colored with either red or blue, either there is a subgraph of K_n isomorphic to K_k in which all edges are red, or there is a subgraph of K_n isomorphic to K_ℓ in which all edges are blue. For example, we have the following simple facts:

- $R(k, 1) = 1$ and $R(1, \ell) = 1$ for all $k, \ell \in \mathbb{N}^+$. To see why $R(k, 1) = 1$, notice that for the trivial graph with vertex set $[1]$ and no edges, we may take $B = \{1\}$ is satisfy the definition. The other is symmetric.
- $R(k, 2) = k$ and $R(2, \ell) = \ell$ for all $k, \ell \in \mathbb{N}^+$. To see why $R(k, 2) \leq k$, notice that for the given any coloring of the edges of K_k, either there is at least blue edge whose endpoints we can take for B, or all edges are red and we can let $A = [n]$. Also, we have $R(k, 2) > k - 1$ because if we color all edges of K_{k-1} red, then no such A or B exist. Thus, $R(k, 2) = k$, and similarly $R(2, \ell) = \ell$.
- $R(3, 3) = 6$ from above.
- In general, we have $R(\ell, k) = R(k, \ell)$ because we can switch the role of read and blue.

It's not at all obvious that $R(k, \ell)$ exists for each $k, \ell \in \mathbb{N}^+$, but we will come up with a recursive upper bound for these values shortly. To prepare for this result, we first prove the following.

Proposition 4.90. *$R(4, 3)$ exists and in fact $R(4, 3) \leq 10$.*

Proof. Consider an arbitrary coloring of the edges of K_{10} with red and blue. Pick an arbitrary vertex u. Since u is incident to 9 total edges, either u is incident to at least 6 red edges or u is incident to at least 4 blue edges (otherwise, u would be incident to at most $5 + 3 = 8$ many edges). We now have two cases.

- *Case 1:* Suppose that u is incident to at least 6 red edges. Fix 6 such edges, and call the other endpoints of these edges $w_1, w_2, \dots, w_6$. Now look at the subgraph of K_{10} induced by these 6 vertices. Since $R(3, 3) = 6$ from above, either there is a red triangle amongst these vertices, or there is a blue triangle amongst these vertices. If there is a blue triangle, then we are done by letting B be the corresponding three vertices. Otherwise, there is a red triangle amongst the w_i, and then we can include u with the 3 vertices that make up the red triangle to obtain a subset $A \subseteq [10]$ with $|A| = 4$ such that all edges having both endpoints in A are red.
- *Case 2:* Suppose that u is incident to at least 4 blue edges. Fix 4 such edges, and call the other endpoints of these edges w_1, w_2, w_3, w_4. Now if the 6 edges with endpoints amongst the w_i are all red, then we can take $A = \{w_1, w_2, w_3, w_4\}$. Otherwise, there is a blue edge amongst the w_i, and we can include u with 2 endpoints of this edge to obtain a set $B \subseteq [10]$ with $|B| = 3$ such that all edges having both endpoints in B are blue.

This completes the proof. □

This idea generalizes to the next inductive argument.

Theorem 4.91. *For all $k, \ell \in \mathbb{N}^+$, we have that $R(k, \ell)$ exists, and in fact $R(k, \ell) \leq R(k-1, \ell) + R(k, \ell-1)$.*

Proof. We prove this by induction on the value of $k + \ell$. We know from above that $R(k, 1)$ and $R(k, 2)$ exist for all $k \in \mathbb{N}^+$, and also that $R(1, \ell)$ and $R(2, \ell)$ exist for all $\ell \in \mathbb{N}^+$. Now let $k, \ell \in \mathbb{N}$ with $k, \ell \geq 2$ and assume that both $R(k-1, \ell)$ and $R(k, \ell-1)$ exist. Let

- $c = R(k-1,\ell)$.
- $d = R(k,\ell-1)$.
- $n = R(k-1,\ell) + R(k,\ell-1) = c + d$.

We show that if we color all of the edges of K_n with red/blue, then either there exists a subset $A \subseteq V$ with $|A| = k$ such that all edges having both endpoints in A are red, or there exists a subset $B \subseteq V$ with $|B| = \ell$ such that all edges with both endpoints in B are blue.

Consider then an arbitrary coloring of the edges of K_n with red and blue. Pick an arbitrary vertex u. Since u is incident to $n-1 = c+d-1$ total edges, either u is incident to at least c red edges or u is incident to at least d blue edges (otherwise, u would be incident to at most $c-1+d-1 = c+d-2 = n-2$ many edges). We now have two cases.

- *Case 1:* Suppose that u is incident to at least c red edges. Fix c such edges, and call the other endpoints of these edges $w_1, w_2, \dots, w_c$. Now look at the subgraph of K_n induced by these c vertices. Since $c = R(k-1,\ell)$ from above, either there is a subset A of these vertices with $|A| = k-1$ such that all edges having both endpoints in A are red, or there exists a subset B of these vertices with $|B| = \ell$ such that all edges with both endpoints in B are blue. In the latter case, we are done by taking B. In the former case, we can let $A' = A \cup \{u\}$ and notice that $|A'| = k$ and A' has the required properties.
- *Case 2:* Suppose that u is incident to at least d blue edges. Fix d such edges, and call the other endpoints of these edges $w_1, w_2, \dots, w_d$. Now look at the subgraph of K_n induced by these d vertices. Since $d = R(k,\ell-1)$ from above, either there is a subset A of these vertices with $|A| = k$ such that all edges having both endpoints in A are red, or there exists a subset B of these vertices with $|B| = \ell-1$ such that all edges with both endpoints in B are blue. In the former case, we are done by taking A. In the latter case, we can let $B' = B \cup \{u\}$ and notice that $|B'| = \ell$ and B' has the required properties.

This completes the proof. □

Notice that using this result, we can immediately conclude that

$$R(4,3) \leq R(3,3) + R(4,2) = 6 + 4 = 10$$

as we showed in the special case above. Working in the other direction, we have the following.

Proposition 4.92. $R(4,3) \geq 9$.

Proof. We exhibit a coloring of the edges of K_8 with red and blue such that there is no red K_4 and no blue K_3. Take K_8 an label the vertices clockwise with the numbers from [8]. Color the edge $\{i,j\}$ with $i < j$ blue if $j - i \in \{1,4,7\}$, and color it red otherwise. Thus, we color $\{i,j\}$ with $i < j$ red if $j - i \in \{2,3,5,6\}$.

We first argue that there is no blue triangle. Suppose one exists, and let the smallest vertex be i. We would then need to choose two of the three potential vertices $\{i+1, i+4, i+7\}$ (where we "wrap around" beyond 8 if necessary) to add to it. This is impossible, because $4-1=3$, $7-4=3$, and $7-1=6$.

We now argue that there is no red K_4. Suppose one exists, and let the smallest vertex be i. We would then need to choose three of the four potential vertices $\{i+2, i+3, i+5, i+6\}$ to add to it. This is impossible, because we can't choose both $i+2$ and $i+3$, and we also can't choose both $i+5$ and $i+6$. □

In fact, we can improve the upper bound $R(4,3) \leq 10$ with a little work.

Proposition 4.93. $R(4,3) \leq 9$.

Proof. In the proof of Proposition 4.90, we argued that in an arbitrary red/blue coloring of the edges of K_{10}, if we take an arbitrary vertex u, then either u is incident to at least 6 red edges or u is incident to at least 4 blue edges. We now argue that in an arbitrary red/blue coloring of the edges of K_9, there exists a vertex u that is incident to either at least 6 red edges of at least 4 blue edges. From here, we can follow the proof of Proposition 4.90.

Suppose then that we have an arbitrary red/blue coloring of the edges of K_9. Since every vertex is incident to 8 edges, if there is no vertex u that is incident to either at least 6 red edges of at least 4 blue edges, then every vertex must be incident to exactly 5 red edges and exactly 3 blue edges. Thus, if we look at the subgraph of K_9 containing all of the vertices but only the red edges, then each of the 9 vertices has degree 5. Thus, the resulting graph would have an odd number of vertices of odd degree, which is a contradiction. □

Combining the two previous results, we conclude that $R(4,3) = 9$. Using Theorem 4.91, it follows that

$$\begin{aligned} R(4,4) &\leq R(3,4) + R(4,3) \\ &= 9 + 9 \\ &= 18 \end{aligned}$$

In fact, it can be shown that $R(4,4) = 18$. To conclude this, we need to show that there is a coloring of the edges K_{17} with red/blue such that there is no monochromatic K_4. This is possible as follows. Given $i, j \in \{0, 1, 2, \dots, 16\}$ with $i < j$, color the edge $\{i, j\}$ with $i < j$ red if $j - i$ is a quadratic residue modulo 17 (i.e. if some perfect square has $j - i$ as a remainder upon division by 17). In other words, we color $\{i, j\}$ red if $j - i \in \{1, 2, 4, 8, 9, 13, 15, 16\}$, and blue otherwise. Using some number theory, one can show that this coloring has the required properties.

Here is a table with many of the known values of $R(k, \ell)$, along with bounds on the numbers that we still do not now.

$k \backslash \ell$	2	3	4	5	6	7
2	2	3	4	5	6	7
3	3	6	9	14	18	23
4	4	9	18	25	35-41	49-61
5	5	14	25	43-49	58-87	80-143
6	6	18	35-41	58-87	102-165	113-298
7	7	23	49-61	80-143	113-298	205-540

In addition to the results of this table, it is also known that $R(3,8) = 28$ and $R(3,9) = 36$. However, the best bounds for $R(3,10)$ are that $40 \leq R(3,10) \leq 43$. Asymptotically, it is know that $R(3,k) \approx \frac{k^2}{\log k}$.

Theorem 4.94. *For any $k, \ell \in \mathbb{N}$ with $k, \ell \geq 2$, we have*

$$R(k,\ell) \leq \binom{k+\ell-2}{k-1}$$

Proof. We prove the result by induction on the value of $k + \ell$.

- *Base Case:* Suppose that $k + \ell = 4$. We then have that $k = 2$ and $\ell = 2$. Now $R(2,2) = 2$, and we have

  $$\binom{2+2-2}{2-1} = \binom{2}{1} = 2$$

  Thus, the statement is true when $k + \ell = 4$.

- Assume that $m \geq 4$ and that we know that the statement is true whenever $k + \ell = m$. Suppose now that we have values of $k, \ell \in \mathbb{N}$ with $k \geq 2$, $\ell \geq 2$, and $k + \ell = m + 1$. Notice that if $k = 2$, then $R(2, \ell) = \ell$ and
$$\binom{2+\ell-2}{2-1} = \binom{\ell}{1} = \ell$$
so the statement is true. Also, if $\ell = 2$, then $R(k, 2) = k$ and
$$\binom{2+k-2}{2-1} = \binom{k}{1} = k$$
so the statement is true. Suppose now that $k \geq 3$ and $\ell \geq 3$. We then have that $k - 1 \geq 2$, that $\ell - 1 \geq 2$, that $(k-1)+\ell = m$, and that $k+(\ell-1) = m$. Therefore, using Theorem 4.91 and induction, we have
$$\begin{aligned} R(k,\ell) &\leq R(k-1,\ell) + R(k,\ell-1) \\ &\leq \binom{(k-1)+\ell-2}{(k-1)-1} + \binom{k+(\ell-1)-2}{k-1} \\ &= \binom{k+\ell-3}{k-2} + \binom{k+\ell-3}{k-1} \\ &= \binom{k+\ell-2}{k-1} \end{aligned}$$
where the last line follow from Proposition 3.9.

The result follows by induction. □

Corollary 4.95. *For any $k \in \mathbb{N}$ with $k \geq 2$, we have*
$$R(k,k) \leq \binom{2k-2}{k-1}$$

Proof. Immediate from the previous result. □

Corollary 4.96. *For any $k \in \mathbb{N}$ with $k \geq 2$, we have $R(k,k) \leq 4^{k-1}$.*

Proof. Let $k \in \mathbb{N}$ with $k \geq 2$. We know from Corollary 3.11 that
$$\binom{2k-2}{0} + \binom{2k-2}{1} + \cdots + \binom{2k-2}{k-1} + \cdots + \binom{2k-2}{2k-2} = 2^{2k-2}.$$
Since every term in the above sum is nonnegative, it follows that
$$\begin{aligned} \binom{2k-2}{k-1} &\leq 2^{2k-2} \\ &= (2^2)^{k-1} \\ &= 4^{k-1}. \end{aligned}$$
The result now follows from the previous corollary. □

Theorem 4.97. *For all $k \geq 3$, we have $R(k,k) > 2^{k/2}$.*

Proof. Let n, and suppose that you color that edges of K_n randomly by flipping a coin. Given a subset S of $[n]$ with $|S| = k$, what is the probability that S is monochromatic? There are $2^{\binom{k}{2}}$ many ways to color the edges, and only two of them are monochromatic. Thus, the probability that S is monochromatic is

$$\frac{2}{2^{\binom{k}{2}}}$$

But this is just for this one particular set S. What is the probability that *some* such S is monochromatic? There are $\binom{n}{k}$ many subsets S, so the probability of the union is bounded by the sum of the probabilities, which gives

$$\binom{n}{k} \cdot \frac{2}{2^{\binom{k}{2}}}$$

Now if this number is less than 1, then we know that there is some positive probability that no monochromatic set of size k exists. So is this less than 1? If $n \leq 2^{k/2}$, we have

$$\begin{aligned} \binom{n}{k} \cdot \frac{2}{2^{\binom{k}{2}}} &< \frac{n^k}{k!} \cdot \frac{2}{2^{\binom{k}{2}}} \\ &\leq \frac{2n^k}{k! \cdot 2^{\binom{k}{2}}} \\ &\leq 2 \cdot \frac{n^k}{k! \cdot 2^{k(k-1)/2}} \\ &\leq 2 \cdot \frac{2^{k^2/2}}{k! \cdot 2^{k(k-1)/2}} \end{aligned}$$

Since

$$\frac{k^2}{2} - \frac{k(k-1)}{2} = \frac{k}{2}(k - (k-1)) = \frac{k}{2}$$

we get

$$\begin{aligned} \binom{n}{k} \cdot \frac{2}{2^{\binom{k}{2}}} &< 2 \cdot \frac{2^{k^2/2}}{k! \cdot 2^{k(k-1)/2}} \\ &= \frac{2 \cdot 2^{k/2}}{k!} \end{aligned}$$

which is certainly less than 1 if $k \geq 3$ (by induction). □