# MATH 61 LECTURE NOTES

SUDESH KALYANSWAMY

## 1. Logic

In this section, we give a brief introduction to mathematical logic, which is essential to the study of any branch of mathematics. If we think of mathematics as a language (and we should do this), then propositional logic is analogous to grammar.

### 1.1. Propositions and Operations.

Let's start with a definition.

**Definition.** A **proposition** is a statement (or declaration) that is either true or false.

**Example.**

1. The statement $p =$ "It is raining" is a proposition. It is either raining, or it is not raining.
2. If $q =$ "I take my umbrella", then $q$ is also a proposition. Either I take the umbrella, or I don't.
3. $p =$ "I have 5 dollars in my pocket."
4. $p(x) =$ "$x$ is a whole number" is a statement which depends on the value of $x$ which we plug in. Once we plug in a specific value of $x$, then $p(x)$ becomes a proposition. For example, $p(5)$ is the statement "5 is a whole number," which is true. On the other hand, $p(\pi) =$ "$\pi$ is a whole number," which is false.

Assigned to any proposition is a truth value, namely true or false. Next, we want to see how we can put propositions together to get more complex statements.

1.1.1. *Conjunction.* The conjunction of two propositions $p$ and $q$ is the statement $p$ *and* $q$, and is denoted $p \wedge q$.

**Example.**

(1) If $p =$ "2 is an even number" and $q =$ "3 is an odd number", then

$$p \wedge q = \text{"2 is an even number and 3 is an odd number."}$$

(2) If $p(x) =$ "$x < 4$" and $q(x) =$ "$x > 0$", then

$$p(x) \wedge q(x) = \text{"}x < 4 \text{ and } x > 0\text{"}.$$

Given the truth values of $p$ and $q$, how do we determine the truth value of $p \wedge q$? For example, the statement "My name is Sudesh and I am teaching math" be true since both propositions in the conjunction are true. However, the statement "I have a unicorn in my pocket and the sky is blue" should be false because I do not have a unicorn in my pocket. We can make this precise by saying that $p \wedge q$ is assigned the value of true if both $p$ and $q$ are true. Otherwise, $p \wedge q$ is false. We can summarize this cleanly and clearly in a *truth table* as follows:

| $p$ | $q$ | $p \wedge q$ |
|---|---|---|
| T | T | T |
| T | F | F |
| F | T | F |
| F | F | F |

This table records every possible combination of true and false for the statements $p$ and $q$ and gives the resulting truth value of $p \wedge q$ in the third column. Notice that the only time $p \wedge q$ is assigned a "T" is when both $p$ and $q$ are true.

**Example.**

1. Example (1) above is true because "2 is an even number" is true and "3 is an odd number" is true.
2. In example (2) above, if $x = 5$, then $p(5) \wedge q(5)$ is false because $p(5)$ is false (namely, "$5 < 4$" is false). However, $p(2) \wedge q(2)$ is true.

1.1.2. *Disjunction.* The disjunction of two propositions $p$ and $q$ is the statement $p$ *or* $q$, and is denoted $p \vee q$.

**Example.** (1) If $p =$ "A century is 10 years" and $q =$ "A century is 100 years", then

$$p \vee q = \text{"A century is 10 years or a century is 100 years."}$$

(2) If $p(x) =$ "$x$ is an integer" and $q(x) =$ "$x$ is a real number", then

$$p(x) \vee q(x) = \text{"}x\text{ is an integer or }x\text{ is a real number."}$$

The truth table for the disjunction is the following:

| $p$ | $q$ | $p \vee q$ |
|---|---|---|
| T | T | T |
| T | F | T |
| F | T | T |
| F | F | F |

Notice that the only time the disjunction is false is when both $p$ and $q$ are false. Also observe that the disjunction is not an "exclusive or." Both propositions can be true, it is not the same as an exclusive or which is "either $p$ is true or $q$ is true, but not both."

**Example.** (1) In example (1) above, the proposition $p \vee q$ is true since $q$ is true. Notice it did not matter that $p$ was false.

(2) For example (2) above, $p(5) \vee q(5)$ is true since 5 is both an integer and a real number. Observe that $p(\pi) \vee q(\pi)$ is true since $q(\pi)$ is true, but $p(1+i) \vee q(1+i)$ is false, since $1+i$ is neither an integer nor a real number.

1.1.3. *Negation.* The negation of a proposition $p$ is the proposition "not $p$." It simply reverses the truth value of $p$. It is denoted $\neg p$. The truth table is:

| $p$ | $\neg p$ |
|---|---|
| T | F |
| F | T |

**Example.** If $p =$ "Paris is the capital of England", then $\neg p =$ "Paris is not the capital of England." The truth value of $\neg p$ would be true since $p$ is false.

1.1.4. *Order of Operations.* There is an order of operations just as there was with addition, subtraction, multiplication, etc. The order of operations between $\neg, \vee$, and $\wedge$ is:

$$\neg, \quad \wedge, \quad \vee.$$

If you are looking for an acronym, I suppose it would be "NAO" (not, and, or), but I'm not quite sure how helpful this is. However, what this means is that if we have three statements $p, q$, and $r$, then $\neg p \vee q \wedge r$ should be read as

$$\neg p \vee q \wedge r = \neg p \vee (q \wedge r).$$

So, for example, if $p =$ "Albany is the capital of New York", $q =$ "My name is Sudesh" and $r =$ "we are at UCLA", then

$$\neg p \vee q \wedge r = \text{"Albany is not the capital of New York or my name is Sudesh and we are at UCLA"}$$

would be true, since $q \wedge r$ is true (as both $q$ and $r$ are true), meaning the disjunction would be true as well. We could make a truth table for the statement $\neg p \vee q \wedge r$ as follows:

| $p$ | $q$ | $r$ | $\neg p$ | $q \wedge r$ | $\neg p \vee (q \wedge r)$ |
|---|---|---|---|---|---|
| T | T | T | F | T | T |
| T | T | F | F | F | F |
| T | F | T | F | F | F |
| T | F | F | F | F | F |
| F | T | T | T | T | T |
| F | T | F | T | F | T |
| F | F | T | T | F | T |
| F | F | F | T | F | T |

1.1.5. *Problems.* Let's work through a couple of problems:

**Problem.** If $p$ is the statement "You run 10 laps daily" and $q$ is the statement "You are healthy", write the following in symbols: "You run 10 laps daily and you are not healthy."

*Solution.* This should be relatively straightforward: $p \wedge \neg q$.

**Problem.** Two propositions are called *logically equivalent* if they have the same truth content, i.e. if they have the same truth tables. Show that $\neg(p \wedge q)$ and $\neg p \vee \neg q$ are logically equivalent.

*Solution.* Let's construct the truth tables for both. For the first statement:

| $p$ | $q$ | $p \wedge q$ | $\neg(p \wedge q)$ |
|---|---|---|---|
| T | T | T | F |
| T | F | F | T |
| F | T | F | T |
| F | F | F | T |

For the second statement:

| $p$ | $q$ | $\neg p$ | $\neg q$ | $\neg p \vee \neg q$ |
|---|---|---|---|---|
| T | T | F | F | F |
| T | F | F | T | T |
| F | T | T | F | T |
| F | F | T | T | T |

Therefore, the two statements are logically equivalent.

## 1.2. Conditionals.

1.2.1. *Definition.* If you think back to the theorems you learned in calculus, they were all of the form "If (insert statement here), then (insert statement here)." This is precisely what a conditional is.

**Definition.** A *conditional proposition* is one of the form "If $p$, then $q$." Here, $p$ and $q$ are propositions. The statement $p$ is the *hypothesis*, and $q$ is the *conclusion*. This is typically denoted $p \implies q$.

The truth table for $p \implies q$ is:

| $p$ | $q$ | $p \implies q$ |
|---|---|---|
| T | T | T |
| T | F | F |
| F | T | T |
| F | F | T |

**Example.** The following are examples of conditional propositions:

(1) If you study hard, then you will do well.
(2) If I have a unicorn in my pocket, then it is blue. This is an example of a "vacuously true" statement, since the hypothesis is false.
(3) If $1 + 1 \leq 2$, then 5 is an odd number. This is a true statement.
(4) If 2 is an even number, then there are 1000 students in this class. This is a false statement, since the hypothesis is true but the conclusion is false.

**Problem.** Show that $\neg(p \implies q)$ is logically equivalent to $p \wedge \neg q$.

*Solution.* Again, we can construct truth tables:

| $p$ | $q$ | $p \implies q$ | $\neg(p \implies q)$ |
|---|---|---|---|
| T | T | T | F |
| T | F | F | T |
| F | T | T | F |
| F | F | T | F |

The truth table for $p \wedge \neg q$ is:

| $p$ | $q$ | $\neg q$ | $p \wedge \neg q$ |
|---|---|---|---|
| T | T | F | F |
| T | F | T | T |
| F | T | F | F |
| F | F | T | F |

Thus, the two are logically equivalent.

1.2.2. *Biconditionals.*

**Definition.** A *biconditional* is a proposition of the form $p \iff q$, read as "$p$ if and only if $q$." It is defined to be

$$p \iff q = (p \implies q) \wedge (q \implies p).$$

The truth table is as follows:

| $p$ | $q$ | $p \iff q$ |
|---|---|---|
| T | T | T |
| T | F | F |
| F | T | F |
| F | F | T |

In other words, $p \iff q$ is true if $p$ and $q$ have the same truth value.

**Example.** (1) $1 < 5$ if and only if $2 < 8$: This is true, since both statements are true (and thus share the same truth value).
(2) If $p(x) =$ "$x > 0$ if and only if $x^2 > 0$," then $p(5)$ is true, but $p(-5)$ is false (as the first statement is false and the second is true).
(3) Saying "It is raining if and only if I take my umbrella" means if it is raining, then I will have my umbrella, and if it is not raining, then I will not have my umbrella.

1.2.3. *Converse, Inverse, Contrapositive.* Given a conditional $p \implies q$, we have three new statements:

(1) Converse: $q \implies p$
(2) Inverse: $\neg p \implies \neg q$
(3) Contrapositive: $\neg q \implies \neg p$

For example, if we have the conditional proposition: "If it is Monday, then I am tired," then

(1) Converse: If I am tired, then it is Monday.
(2) Inverse: If it is not Monday, then I am not tired.
(3) Contrapositive: If I am not tired, then it is not Monday.

If the original statement is true, then the contrapositive should also be true. In the example above: the contrapositive would be false if I was not tired and it was Monday. But from the original statement, we know this cannot happen, since if it were Monday, then I should be tired. Thus, the contrapositive is also true.

The converse should not necessarily be true: it could happen that I am tired and it is not Monday. The original statement made no mention of my energy levels on Tuesdays, for example. Similarly, the inverse is not necessarily true.

In fact, you'll see on your homework that the inverse and converse are logically equivalent, and the original statement and its contrapositive are logically equivalent.

1.3. **Quantifiers.** The last topic of the section is quantifiers. Let's return to statements of the form $p(x)$. We remarked that, to make this a proposition, we have to plug in a specific value of $x$. For example: if $p(x) =$ "$x > 4$,", then $p(1)$ is false and $p(5)$ is true. However, we could also turn this into a proposition by saying: "For all real numbers $x$, $x > 4$." This statement would be false, since 1 is a real number but 1 is not less than 4. But suddenly, the statement depends on all $x$ in a given "set." We could also change it to "There exists a real number $x$ such that $x > 4$." This is now a proposition, and it is true, since 5 is a real number and $5 > 4$. Both of these illustrate *quanitifers*.

1.3.1. *Universal Quantifiers.* We use $\forall$ to denote "for all." The statement $\forall x, p(x)$ is read "for all $x$, the statement $p(x)$."

**Example.**

(1) $\forall$ integers $x, x^2 \geq 0$. This is true, since no matter what integer $x$ we plug in, $x^2$ is always $\geq 0$.

(2) $\forall x \geq 0, x - 5 \geq 0$ is false since there exists a value $x \geq 0$ for which $x - 5 \not\geq 0$ (take $x = 1$, for example).

The examples illustrate the connection between $\forall$ and conjunctions. If we can somehow list all the elements $x$ in the allowable universe, then we need $p(x)$ to be true for the first element AND the second element AND the third element, and so on.

1.3.2. *Existential Quanitifers.* The other type of quantifier is the existential one, denoted $\exists$. The proposition $\exists x, p(x)$ is read: "There exists an $x$ such that $p(x)$."

**Example.** (1) There exists a state in the USA beginning with the letter 'C': This is true, since California is such a state.

(2) $\exists$ integer $0 \leq x \leq 2$ such that $x^2 = 9$: This is false, since there is no integer value for $x$ between 0 and 2 for which $x^2 = 9$.

Just as $\forall$ was connected with "and," the existential $\exists$ is connected with disjunctions. In the first example above, for instance, the statement is true because if we list out all states, $p(x)$ is true for at least one of them. Similarly, example (2) is false since $p(0) \vee p(1) \vee p(2)$ is false (all propositions $p(i)$ being false).

1.3.3. *Negations and Order.* Negations with quantifiers are very easy:

(1) $\neg(\forall x, p(x)) = \exists x, \neg p(x)$

(2) $\neg(\exists x, p(x)) = \forall x, \neg p(x)$.

**Example.**

(1) $\forall$ real $x, x^2 > 0$. This statement is false. The negation would be $\exists$ real $x$ such that $x^2 \leq 0$, and this statement is true (take $x = 0$).

(2) There exists a sheep in the world which isn't white. This statement is true. The negation is: all sheep in the world are white. This statement is false.

The order of quantifiers is also very easy: read left to right.

**Example.** The statement:

$$\forall \text{ integers } x, \exists \text{ integer } y \text{ such that } x < y,$$

is read as:

$$\forall \text{ integers } x (\exists \text{ integer } y \text{ such that } x < y).$$

This statement is true. To "prove" this, we need to show that no matter what $x$ we plug in, there is some $y$ which is bigger. One could take $y = x + 1$.

The negation of the statement would be

$$\exists \text{ integer } x \text{ such that } \forall \text{ integers } y, x \geq y.$$

This statement is clearly false, since this is saying there is a "biggest integer," which is clearly false.

**Problem.** If $f(x) : \mathbb{R} \to \mathbb{R}$ is a function, then $f$ is continuous at $x = a$ if:

$$\forall \epsilon > 0, \exists \delta > 0 \text{ such that if } |x - a| < \delta, \text{ then } |f(x) - f(a)| < \epsilon.$$

How do you show $f(x)$ is NOT continuous at $a$, in symbols?

*Solution.* We just need the negation of the above statement. It would be:

$$\exists \epsilon > 0 \text{ such that } \forall \delta > 0, \neg(|x - a| < \delta \implies |f(x) - f(a)| < \epsilon).$$

The negation of the conditional is $|x - a| < \delta$ and $|f(x) - f(a)| \geq \epsilon$. Therefore, $f$ is not continuous at $x = a$ if:

$$\exists \epsilon > 0 \text{ such that } \forall \delta > 0, |x - a| < \delta \text{ and } |f(x) - f(a)| \geq \epsilon.$$

## 2. Proof Techniques

2.1. **Direct Proof, Proofs by Contradiction, and Proofs by Contrapositive.** Now that we have gone over the grammar of math, we can now practice writing sentences, albeit with a slightly limited vocabulary. We will highlight some of the basic proof techniques and look at some examples. A typical proposition/lemma/theorem in math looks like the following:

**Lemma.** Suppose $x$ is an integer. If $7x+9$ is even, then $x$ is odd.

Notice that it starts by defining the variables in the statement, telling us which set the variable belongs to. Then it follows by giving a statement, in this case a conditional. The goal in a proof is to show that this statement is always true. We could rephrase the lemma as follows:

$$\forall \text{ integers } x, \quad (7x+9 \text{ even} \implies x \text{ odd}).$$

Is this a reasonable statement? We should check a few cases. If $x=1$, then the conditional proposition:" If 16 is even, then 1 is odd" is true, since the hypothesis and conclusion are both true. If $x=2$, then the statement becomes "If 23 is even, then 2 is odd" which is vacuously true. To do this for every integer, we write a formal proof. For the proof to follow, we need to know that even integers are ones that can be written as $2k$ for some integer $k$, and odd integers are those which can be written as $2k+1$ for some integer $k$. Here's an example of a proof of the above lemma.

*Proof.* Pick integer $x$. If $7x+9$ is odd, then the conditional is vacuously true. If $7x+9$ is even, then $7x+9=2k$ for some integer $k$. Subtracting $6x+9$ from both sides yields $x=2k-6x-9$. We can write this as $x=2(k-3x-5)+1$. Therefore, $x$ is odd. □

This is what is known as a *direct proof.* Namely, we proved the statement by assuming the hypothesis and showing the conclusion was true. Observe that a lot of the proof was words, explaining each step along the way.

Recall that a conditional is logically equivalent to its contrapositive. So we could also prove the statement by trying to prove the contrapositive of the conditional. In this case, the statement would be the following:

**Lemma** (Contrapositive)**.** Let $x$ be an integer. If $x$ is even, then $7x+9$ is odd.

*Proof.* Choose an integer $x$. If $x$ is odd, then the statement is vacuously true. If $x$ is even, then $x=2k$ for some integer $k$. But then $7x+9=7(2k)+9=14k+9$, and this can be written as $2(7k+4)+1$. Since $7k+4$ is an integer (as $k$ is), this means $7x+9$ is odd, which is what we wanted to show. □

Both proofs are valid in proving the original lemma, they are just different techniques. This means when looking at a problem, you have to decide which technique you want to use. Most of the time, there is an easier path and a harder path.

**Problem.** Prove: Let $x$ be an integer. If $x^2-6x+5$ is even, then $x$ is odd.

*Proof.* In this case, the contrapositive is easier. We will show: if $x$ is even, then $x^2-6x+5$ is odd. Let $x$ be an integer. If $x$ is odd, the statement is vacuously true. If $x$ is even, then $x=2k$ for some integer $k$. Then

$$\begin{aligned} x^2-6x+5 &= (2k)^2-6(2k)+5 \\ &= 4k^2-12k+5 \\ &= 2(2k^2-6k+2)+1, \end{aligned}$$

which means $x^2-6x+5$ is odd, as desired. □

Here's an example when both techniques come in handy.

**Lemma.** Let $n$ be an integer. Then $n^2$ is odd if and only if $n$ is odd.

*Proof.* Remember if and only if is equivalent to: "If $n^2$ is odd, then $n$ is odd" AND "If $n$ is odd, then $n^2$ is odd." We have to prove both statements are true. Let's do the second one. Let $n$ be an integer. If $n$ is even, the statement is vacuously true. If $n$ is odd, then $n=2k+1$ for some integer $k$. Squaring yields $n^2=(2k+1)^2=4k^2+4k+1=2(2k^2+2k)+1$. Thus, $n^2$ is odd.

For the first statement, we will prove it by contrapositive. Namely, we will prove that if $n$ is even, then $n^2$ is even. Again, let $n$ be an integer. If $n$ is odd, the statement is vacuously true. If $n$ is even, then $n = 2k$ for some integer $k$. But then $n^2 = 4k^2 = 2(2k^2)$, meaning $n^2$ is even, as desired. □

There is another technique which is called proof by contradiction. An example would be the following:

**Lemma.** The number $\sqrt{2}$ is irrational.

*Proof.* We will prove this statement by assuming it is rational and arriving at a contradiction. Suppose $\sqrt{2}$ is rational. Then $\sqrt{2} = a/b$ for some integers $a$ and $b$. We can assume that this is written in lowest terms. Squaring both sides yields $2 = a^2/b^2$, so $a^2 = 2b^2$. This means $a^2$ is even, which by the previous lemma implies $a$ is even. Thus, $a = 2k$ for some integer $k$. Substituting this into the equality above yields $4k^2 = 2b^2$, which means $b^2 = 2k^2$. Again, this means $b$ is even. But this contradicts the fact that we were in lowest terms. Therefore, $\sqrt{2}$ is irrational. □

Why is this logically valid? Look at problem 3(b) of homework 1 and see if you can see why this is a valid form of proof. Let's look at one more to see if we can determine which proof technique to use. Before we do so, we need a definition.

**Definition.** Let $x$ and $y$ be integers. We say $x$ *divides* $y$ if there exists an integer $k$ such that $xk = y$. It is denoted $x|y$.

**Problem.** Let $a$ and $b$ be integers. Prove that if $5 \not| xy$, then $5 \not| x$ and $5 \not| y$.

*Proof.* The easiest method here is contrapositive. That is: if $5|x$ or $5|y$, then $5|xy$. So let $x$ and $y$ be integers. We know that either $5|x$ or $5|y$ (or both). Assume $5|x$. Then $x = 5k$ for some integer $k$. But then $xy = 5(ky)$, and $ky$ is also an integer, so this implies $5|xy$, as desired.

If, on the other hand, $5|y$, then a similar argument will yield $5|xy$ as well. □

Before we move on to induction, let us demonstrate two examples which show how NOT to write proofs.

**Problem.** Let $x > 0$. Then $x + \frac{4}{x} \geq 4$.

*Bad proof.* Let $x > 0$. Then multiplying both sides of the inequality by $x$ gives $x^2 + 4 \geq 4x$ (the inequality is the same because $x > 0$). Subtracting $4x$ from both sides gives

$$x^2 - 4x + 4 \geq 0.$$

We can factor the left side as $(x-2)^2 \geq 0$. This is true since squaring always yields a nonnegative number. □

This is a bad proof because we were assuming the statement we wanted to show was true. The work in this proof is good side work, but it should not be used in the proof. A valid proof would be to trace through this argument in reverse:

*Good proof.* Let $x > 0$. Then $(x-2)^2 \geq 0$ as squaring always yields nonnegative numbers. Expanding and some algebra yields $x^2 + 4 \geq 4x$. Since $x > 0$, we can divide and keep the inequality the same way to get

$$x + \frac{4}{x} \geq 4,$$

which is what we wanted to show. □

The same idea goes for the following:

**Problem.** Prove there exists an integer $m$ such that $\frac{m-7}{2m+4} = 5$.

*Solution.* You might be tempted to just solve the equality for $m$ and see if there is an integer solution. This is the right thing to do...on the side. Doing this gives $m = -3$. So a valid proof is:

*Proof.* Take $m = -3$. Then $\frac{m-7}{2m+4} = \frac{-10}{-2} = 5$, as desired. □

## 2.2. **Induction.**

2.2.1. *Weak Induction.* Suppose we have a proposition of the form

$$\forall n \geq 1, p(n),$$

where $p(n)$ is some statement which depends on $n$. We want to show that this proposition is true, meaning $p(n)$ is true for any integer $n \geq 1$ which we plug in. One strategy is to use induction. An inductive proof has two steps:

(1) **Base Case**: Show $p(1)$ is true.
(2) **Inductive Step**: If $p(k)$ is true, then $p(k+1)$ is true.

Once this is done, then we can conclude our proposition is true. Indeed, step (1) gives that $p(1)$ is true. Then step (2) says that since $p(1)$ is true, we know $p(2)$ is true as well. Then applying step (2) again gives that since $p(2)$ is true, so is $p(3)$, and so on. In this way, we can get $p(n)$ to be true for all $n \geq 1$.

*Remark.* (1) It is unimportant that the proposition was for all $n \geq 1$. If we replace 1 by some other integer, say $n_0$, and we want to show $\forall n \geq n_0$, $p(n)$, then just replace the base case with showing $p(n_0)$ is true and doing the inductive step as described above.
(2) In symbols, what this is saying is that "$\forall n \geq 1, p(n)$" is equivalent to

$$p(1) \wedge (\forall k, p(k) \implies p(k+1)).$$

Let's look at some examples of this.

**Problem.** Prove that $1 + 3 + 5 + \ldots + (2n-1) = n^2$ for $n \geq 1$.

*Solution.* We use the steps outlined above:

(1) Base case: $n = 1$. If $n = 1$, then the formula $p(1)$ reduces to $1 = 1^2$, which is certainly true. Thus, the base case holds.
(2) Inductive step: Assume $p(k)$ is true. For this problem, this means

$$1 + 3 + 5 + \ldots + (2k-1) = k^2$$

holds. We want to show this implies $p(k+1)$, which means we want to somehow show that

$$1 + 3 + 5 + \ldots + (2(k+1)-1) = (k+1)^2,$$

or, simplifying, that $1+3+\ldots+(2k+1) = (k+1)^2$. How to do this? Observe the difference between the left hand side of $p(k)$ and the left hand side of $p(k+1)$ is the term $2k+1$. So let's add $2k+1$ to both sides of the equation for $p(k)$:

$$1 + 3 + 5 + \ldots + (2k-1) + (2k+1) = k^2 + (2k+1).$$

But the right hand side of this new equation is precisely $(k+1)^2$, which is what we wanted. Therefore, the inductive step holds.

Before we do the next example, recall that the factorial operator is defined as:

$$n! = \begin{cases} 1 & \text{if } n = 0 \\ 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n & \text{if } n > 0 \end{cases} .$$

**Problem.** Prove $n! \geq 2^{n-1}$ for $n \geq 1$.

*Solution.* Again, we do the two steps:

(1) Base case: $n = 1$. Plugging in $n = 1$ to both sides gives $1! \geq 2^0$, which is true as both sides are equal to 1.
(2) Inductive step: Assume $p(k)$ is true, meaning $k! \geq 2^{k-1}$. We want to show $(k+1)! \geq 2^{k+1-1} = 2^k$. Well, we know $(k+1)! = (k+1)k!$. By the inductive hypothesis, we know

$$(k+1)! = (k+1)k! \geq (k+1)2^{k-1}.$$

But since $k \geq 1$, we know $k+1 \geq 2$, so

$$(k+1)2^{k-1} \geq 2 \cdot 2^{k-1} = 2^k,$$

which is what we wanted. Therefore, the inductive step holds.

Before looking at two more examples, we recall one more concept.

**Definition.** Let $m$ and $n$ be integers. We say $m$ *divides* $n$ (or $n$ is divisible by $m$) if there exists an integer $k$ such that $n = mk$. We write $m|n$ if $m$ divides $n$.

**Example.** We have $2|6$ since $2 \cdot 3 = 6$, but $7 \not| 17$ as 17 is not divisible by 7.

**Problem.** Prove $5|6^n - 1$ for all $n \geq 0$.

*Solution.* We use induction again.

(1) Base case: $n = 0$ (notice it is not $n = 1$ in this example). Plugging in $n = 0$ gives $5|6^0 - 1$, or $5|0$. This is true since $5 \cdot 0 = 0$.

(2) Inductive step: assume $5|6^k - 1$. We want to show $5|6^{k+1} - 1$. Well, since $5|6^k - 1$, there exists an $m$ such that $5m = 6^k - 1$. We need to find an $m'$ such that $5m' = 6^{k+1} - 1$. We need to somehow break down the $6^{k+1}$ to use the inductive hypothesis. Observe that

$$6^{k+1} - 1 = 6^k \cdot 6 - 1 = 6^k + 5 \cdot 6^k - 1 = 5m + 5 \cdot 6^k = 5(m + 6^k).$$

Thus, $5|6^{k+1} - 1$ (the $m'$ we were looking for was $m + 6^k$), and the inductive step holds.

The next example is equally interesting and looks similar, but there is something strange that needs to be resolved first.

**Problem.** Prove $8|7^n - 1$ if $n \geq 0$ is even.

*Proof.* The difficulty of this problem is that the formula is only true for even positive integers, not odd positive integers. Now, we could modify the framework of induction to deal with this (indeed, changing the inductive step to: if $p(k)$ is true, then $p(k + 2)$ is true, would do). But in an effort to keep these two steps the same, we change the problem. Showing the divisibility for even $n$ is the same as showing $8|7^{2m} - 1$ for *all integers* $m \geq 0$. This works because all even numbers have the form $2m$ for some $m$. Now we don't need to worry about even or odd, and we can do the induction the way we have been doing it:

(1) Base case: $m = 0$. We get $8|0$, which is true. Thus, the base case holds.

(2) Inductive step: If $8|7^{2k} - 1$, then $8|7^{2(k+1)} - 1$. Since we are assuming the former, we get $8l = 7^{2k} - 1$ for some $l$. We want to find $l'$ so that $8l' = 7^{2k+2} - 1$. Operating as in the previous problem, observe that

$$7^{2k+2} - 1 = 7^{2k} \cdot 49 - 1 = 7^{2k} - 1 + 48 \cdot 7^{2k} = 8l + 48 \cdot 7^{2k} = 8(l + 6 \cdot 7^{2k}).$$

Thus $8|7^{2k+2} - 1$ (with $l' = l + 6 \cdot 7^{2k}$), and the inductive step holds.

□

2.2.2. *Strong Induction.* Notice that in all the induction examples in the previous subsection, to prove $p(k+1)$ we only needed $p(k)$, never any cases before that. There will be times, however, when we need to assume all cases up to $p(k)$, not just $p(k)$ itself. This is where strong induction comes in. Strong induction works as follows:

(1) **Base Case(s)**: Show $p(1)$ is true.

(2) **Inductive Step**: If $p(1), p(2), \dots, p(k)$ are all true, then $p(k + 1)$ is true.

*Remark.* (1) The only difference between the two forms of induction is the hypothesis of the inductive step. Strong induction assumes more. However, strong induction and weak induction are equivalent when working with the statements we will be dealing with in this course.

(2) In symbols, strong induction is the equivalence of the statement "$\forall n \geq 1, p(n)$" with

$$p(1) \wedge (\forall k(\forall m \leq k, p(m)) \implies p(k)).$$

Let's look at some example of this.

**Problem.** Define a sequence $T_n$ as follows: $T_1 = T_2 = T_3 = 1$, and $T_n = T_{n-1} + T_{n-2} + T_{n-3}$ for $n \geq 4$. Prove $T_n < 2^n$ for all $n \geq 1$.

*Solution.* (1) Base case: In this problem, every term depends on the three previous terms. So, for example, $T_4$ will depend on $T_1$, $T_2$, and $T_3$. It may be useful (and indeed it is necessary) to prove three bases cases instead of just one. So, for this problem, if $n = 1$, then we get $T_1 = 1 < 2^1$, which is true. If $n = 2$, then $T_2 = 1 < 2^2$, which is also true. And lastly, $T_3 = 1 < 2^3$ which is again true.

(2) Inductive step: Assume $T_m < 2^m$ for all $m \leq k$. We want to show $T_{k+1} < 2^{k+1}$. Well, we know $T_{k+1} = T_k + T_{k-1} + T_{k-2}$, and so by the inductive hypothesis,

$$T_{k+1} = T_k + T_{k-1} + T_{k-2} < 2^k + 2^{k-1} + 2^{k-2} = 2^{k+1}\left(\frac{1}{2} + \frac{1}{4} + \frac{1}{8}\right) = 2^{k+1} \cdot \frac{7}{8} < 2^{k+1},$$

which is what we wanted to prove. Thus, the inductive step holds.

*Remark.* Notice that proving three base cases was essential because the inductive step required the three previous cases. If we tried to replicate the inductive step for $n = 4$, for example, we would get

$$T_4 = T_3 + T_2 + T_1,$$

and it was essential that we knew that the inductive hypothesis was true for each of these three cases. This is why having three base cases was necessary.

**Problem.** Prove that every amount of postage which is at least 12 cents can be made up from 4 and 5 cent stamps.

*Solution.* If you don't know where to start, listing out some cases might help. So $12 = 4 + 4 + 4$, which is good. Next, $13 = 4 + 4 + 5$ and $14 = 4 + 5 + 5$. Then $15 = 5 + 5 + 5$, and by the time we get to 16 we notice that we can just take the partition of 12 above and add a 4. This suggests that we want to somehow relate the problem for $n$ to the problem for $n - 4$, because we can just tack on a 4 at the end. Formally, we have the base cases of 12, 13, 14, and 15 above. For the inductive step, assume the statement is true for all $m \leq k$. We want to show it is true for $k + 1$. Well $k + 1 = (k - 3) + 4$. By the inductive hypothesis, $k - 3 = 4j + 5l$ for some $j$ and $l$. But then $k + 1 = 4(j + 1) + 5l$, which is what we wanted to show.

**Problem.** Define the Fibonacci sequence $F_n$ as $F_1 = F_2 = 1$ and $F_n = F_{n-1} + F_{n-2}$ for $n \geq 3$. Prove $F_1^2 + F_2^2 + \ldots + F_n^2 = F_n F_{n+1}$ for $n \geq 1$.

*Solution.* (1) Base case: $n = 1$. We get $F_1^2 = F_1 F_2$, and this is true as $1 = 1$.

(2) Inductive Step: Assume $F_1^2 + \ldots + F_k^2 = F_k F_{k+1}$. We want to show $F_1^2 + \ldots + F_k^2 + F_{k+1}^2 = F_{k+1} F_{k+2}$. Well, we know

$$F_1^2 + \ldots + F_k^2 + F_{k+1}^2 = F_k F_{k+1} + F_{k+1}^2 = F_{k+1}(F_k + F_{k+1}) = F_{k+1} F_{k+2},$$

which is what we wanted. Thus, the inductive step holds.

*Remark.* How did we know that we only needed one base case and not more as in previous problems? The easiest way is to look at the inductive step to see how many of the previous steps you need. Since in this problem we only needed $p(k)$ and not $p(m)$ for any $m < k$, we only needed one base case.

## 3. Sets

We have, to this point, covered logic and induction. Logic is analogous to grammar rules of any language, and induction is an application of those rules to form sentences. Now, we need to describe some of the vocabulary. In this section, we'll examine sets, and in the next section, functions.

### 3.1. Sets and Subsets: Definitions.

**Definition.** A *set* is a collection of distinct objects, either finite or infinite. A set is denoted with braces {} with the collection of objects listed inside.

**Example.** The following are examples of sets:

(1) $A = \{1, 2, 3, 4\}$: This is a set (as indicated by the braces) with the objects being the numbers $1, 2, 3$, and 4.

(2) $B = \{\{1, 2\}, 3, x\}$: This is a set whose objects are the set containing 1 and 2, the number 3, and the letter $x$.

(3) We have the following useful sets: $\mathbb{N}$ (natural numbers), $\mathbb{N}_0$ (natural numbers along with 0), $\mathbb{Z}$ (integers), $\mathbb{Q}$ (rational numbers), $\mathbb{R}$ (real numbers), and $\mathbb{C}$ (complex numbers).

(4) $C = \{x \in \mathbb{R} : x^2 - 3x + 2 = 0\}$. The colon in the middle of the set is read as "such that." Therefore, $C$ is the set of real $x$'s *such that* $x^2 - 3x + 2 = 0$. We could also have written $C = \{1, 2\}$, because both sets are the same collection of objects.

(5) $D = \emptyset$: this is the symbol for the empty set, the set containing no objects.
(6) $E = \{\emptyset\}$: this is a set containing the emptyset. The set $E$ itself is NOT empty.

**Definition.** If an object $a$ is in the set $A$, we say $a$ is *an element of* $A$, and we write $a \in A$.

In the above examples, $2 \in A$, $\{1, 2\} \in B$, $\emptyset \in E$, but there is no element of $D$. Next we want to define what it means for one set to contain another.

**Definition.** Let $A$ and $B$ be sets. We say $A$ is a *subset* of $B$, denoted $A \subseteq B$, if every element of $A$ is also an element of $B$. In symbols, $A \subseteq B$ if

$$\forall x, x \in A \implies x \in B.$$

**Example.** Consider the sets from the first example above.
(1) $\{2, 3\} \subseteq A$, since $2 \in A$ and $3 \in A$.
(2) $\{1\} \not\subseteq B$, since $B$ does not contain the element 1.
(3) $\mathbb{N} \subseteq \mathbb{N}_0 \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$
(4) If $A = \{x \in R : \ln(x) > 1\}$ and $B = \{x \in \mathbb{R} : x \geq 3\}$, then $B \subseteq A \subseteq \mathbb{R}$, since every element of $B$ is also an element of $A$, and every element of $A$ is also an element of $\mathbb{R}$.
(5) If $G = \{x \in \mathbb{Z} : 2|x\}$ and $H = \{x \in \mathbb{Z} : 4|x\}$, then $H \subseteq G$, but $G \not\subseteq H$. Indeed, if an integer is divisible by 4, it is certainly divisible by 2. However, $2 \in G$ but $2 \notin H$.

If we have two sets $A$ and $B$ such that $A \subseteq B$ and $B \subseteq A$, then $A = B$. This is the best way to show the equality of sets. We will examine this later.

*Remark.* The empty set $\emptyset$ is a subset of all sets, since the definition of subsets vacuously holds.

**Problem.** List all subsets of the set $\{1, 2, 3\}$.

*Solution.* There is one subset with zero elements: $\emptyset$. There are three subsets with 1 element: $\{1\}$, $\{2\}$, $\{3\}$. There are three subsets with two elements: $\{1, 2\}$, $\{1, 3\}$, and $\{2, 3\}$. Finally, there is one subset with three elements, namely $\{1, 2, 3\}$. This gives eight total subsets.

**Definition.** The *cardinality* of a set $A$ is the number of elements of $A$ (if finite). It is denoted $|A|$, or sometimes $\#A$.

**Example.** Again, consider the sets from the first example.
(1) $|A| = 4$
(2) $|B| = 3$
(3) $|D| = 0$
(4) $|E| = 1$.
(5) If $A = \{\mathbb{R}, \mathbb{Z}\}$, then $|A| = 2$, since $A$ is a set containing two elements: the set $\mathbb{R}$ and the set $\mathbb{Z}$.

3.2. **Operations.** There are tons of things you can do with sets, and we will go through many of the important operations here.

3.2.1. *Union.*

**Definition.** Given two sets $A$ and $B$, their *union*, denoted $A \cup B$, is the set

$$A \cup B = \{x : x \in A \vee x \in B\}.$$

So the union is the set of elements that belong to either $A$ or $B$.

**Example.** If $A = \{1, 2, 3\}$ and $B = \{3, 4, 7\}$, then $A \cup B = \{1, 2, 3, 4, 7\}$.

To show an element is in the union, you must show it belongs to either $A$ or $B$.

**Problem.** If $A = \{x \in \mathbb{Z} : 2|x\}$ and $B = \{x \in \mathbb{Z} : 3|x\}$, prove $1 \notin A \cup B$.

*Solution.* If $1 \in A \cup B$, then $1 \in A$ or $1 \in B$. But $1 \notin A$ as $2 \not| 1$, and $1 \notin B$ as $3 \not| 1$. Therefore, $1 \notin A \cup B$.

We can illustrate this operation with a *venn diagram* as in Figure 1. In Figure 2, you'll see the two sets $A$ and $B$, represented by the circles, and the shaded area represents $A \cup B$.

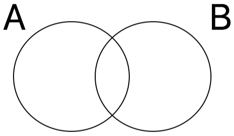


Figure 1. A Venn Diagram

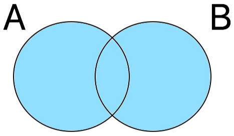


Figure 2. The shaded region represents $A \cup B$

3.2.2. *Intersection.*

**Definition.** Given two sets $A$ and $B$, their *intersection*, denoted $A \cap B$, is the set

$$A \cap B = \{x : x \in A \wedge x \in B\}.$$

From this definition, we see the intersection is the set of elements which belong to both $A$ and $B$. Thus, to show an element belongs to the intersection, you must show that it belongs to both sets. The venn diagram for the intersection is shown in Figure 3.

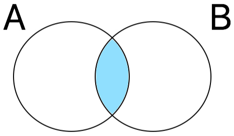


Figure 3. The shaded region represents $A \cap B$

**Problem.** Prove $A \cap B \subseteq A \cup B$.

*Solution.* If $x \in A \cap B$, then $x \in A$ and $x \in B$. But then by definition of the union, $x \in A \cup B$ (as $x \in A$, for example). As $x$ was arbitrary, $A \cap B \subseteq A \cup B$.

**Problem.** If $A = \{x \in \mathbb{Z} : 2|x\}$ and $B = \{x \in \mathbb{Z} : 3|x\}$, prove $A \cap B = C := \{x \in \mathbb{Z} : 6|x\}$.

*Solution.* Here we practice a proof technique with sets. To show to sets are equal, we show each is a subset of the other. In this case, we need to show $A \cap B \subseteq C$ and $C \subseteq A \cap B$. To show the former, let $x \in A \cap B$. Thus, $2|x$ and $3|x$, so there exist integers $m$ and $n$ such that $x = 2m$ and $x = 3n$. The first equation says $x$ is even. From the second equation, we see this implies $n$ is even, thus $n = 2n'$ for some $n'$. Substituting this back into the second equation, we get $x = 3 \cdot 2n' = 6n'$, which means $6|x$. Therefore, $x \in C$ and $A \cap B \subseteq C$.

For the reverse inclusion $C \subseteq A \cap B$, let $x \in C$. Thus, $x = 6m$ for some integer $m$. But this means $x = 2(3m) = 3(2m)$. The first equality means $x \in A$, and the second equality means $x \in B$. Therefore, $x \in A \cap B$, and we have the reverse inclusion as well. Putting both inclusions together gives $A \cap B = C$.

**Definition.** Two sets $A$ and $B$ are said to be *disjoint* if $A \cap B = \emptyset$.

Disjoint sets can be represented as in Figure 4.

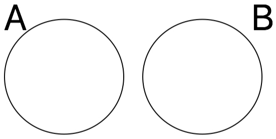


FIGURE 4. Disjoint sets

3.2.3. *Complement.*

**Definition.** If $A$ and $B$ are sets, then the *relative complement* of $B$ in $A$, denoted $A - B$ or $A \backslash B$, is defined as

$$A - B = \{x : x \in A \wedge x \notin B\}.$$

The relative complement is the set of elements in $A$ which are not in $B$. An illustration of this set is shown in Figure 5.

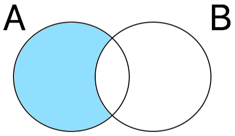


FIGURE 5. The relative complement $A - B$

**Example.** If $A = \{1, 2, 3\}$ and $B = \{1, 3, 4\}$, then $A - B = \{2\}$, since 2 is the only element of $A$ which is not in $B$.

**Problem.** Prove $A - B = A - (A \cap B)$ for sets $A$ and $B$.

*Solution.* Again, we show both inclusions. First, let's show $A - B \subseteq A - (A \cap B)$. Take $x \in (A - B)$. Then $x \in A$ and $x \notin B$. But since $x \notin B$, we know $x \notin A \cap B$. Thus, $x \in A - (A \cap B)$ and we have the desired inclusion.

For the reverse $A - (A \cap B) \subseteq A - B$, take $x \in A - (A \cap B)$. Thus $x \in A$ and $x \notin A \cap B$. We want to say this implies $x \notin B$. Assume the opposite. If $x \in B$, then as $x \in A$ as well, we know $x \in A \cap B$, a contradition. Thus, $x \notin B$, meaning $x \in A - B$. Thus, $A - (A \cap B) \subseteq A - B$, and we have the equality of sets.

*Remark.* If our set $A$ is understood to be a subset of a larger set $U$, then we define the *complement* of $A$, denoted $A^c$, to be the relative complement of $A$ in $U$. But the universe $U$ should be defined at some point. For example, if $A = \{x \in \mathbb{Z} : 2|x\}$, and $U = \mathbb{Z}$, then $A^c = \{\text{odd integers}\}$. However, if $U = \mathbb{Q}$, we get a different complement (exercise for the reader: what is it?). Therefore, it is extremely important that $U$ be understood from the context.

### 3.2.4. *Power Set.*

**Definition.** If $A$ is a set, then the *power set of* $A$, denoted $\mathcal{P}(A)$, is the set of all subsets of $A$.

**Example.** If $A = \{1, 2, 3\}$, then $\mathcal{P}(A)$ was determined in the first problem of section 3.1 above. Namely,

$$\mathcal{P}(A) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}.$$

Notice that $\mathcal{P}(A)$ is a set of sets.

*Remark.* Observe that in the above example, $|A| = 3$, and $|\mathcal{P}(A)| = 8 = 2^3$. This is no coincidence. We will encounter several ways to prove that the number of elements of the power set is $2^{|A|}$, provided $A$ is finite.

**Problem.** Prove that $|\mathcal{P}(A)| = 2^{|A|}$, provided $A$ is finite.

*Solution.* We proceed by induction. If $|A| = 0$, then $A = \emptyset$ and $\mathcal{P}(A) = \{\emptyset\}$. Therefore, $|\mathcal{P}(A)| = 1$ and the base case holds.

For the inductive step, assume the statement is true if $|A| = k$. We want to show it is true if $|A| = k+1$. To do this, let $a \in A$ and $A' = A\backslash\{a\}$. Subsets of $A$ either contain $a$ or do not contain $a$. Subsets of $A$ which do not contain $a$ are precisely subsets of $A'$. By the inductive hypothesis, there are $2^k$ of these. On the other hand, subsets which do contain $a$ are of the form $\{a\} \cup S$, where $S$ is a subset of $A'$. Again by the inductive hypothesis, there are $2^k$ of these. Therefore, there are $2^k + 2^k = 2^{k+1}$ subsets of $A$, as desired.

**Problem.** Show $\mathcal{P}(A \cap B) = \mathcal{P}(A) \cap \mathcal{P}(B)$.

*Proof.* We, once again, show both inclusions. To prove $\mathcal{P}(A \cap B) \subseteq \mathcal{P}(A) \cap \mathcal{P}(B)$, let $S \in \mathcal{P}(A \cap B)$. By definition, this means $S \subseteq A \cap B$ is a subset. But by definition of the intersection, this implies $S \subseteq A$ and $S \subseteq B$, so $S \in \mathcal{P}(A)$ and $S \in \mathcal{P}(B)$. Thus, $S \in \mathcal{P}(A) \cap \mathcal{P}(B)$.

For the reverse inclusion $\mathcal{P}(A)\cap\mathcal{P}(B) \subseteq \mathcal{P}(A\cap B)$, let $S \in \mathcal{P}(A)\cap\mathcal{P}(B)$. Then $S \subseteq A$ and $S \subseteq B$ by definition of the power set. Thus, $S \subseteq A\cap B$, and thus $S \in \mathcal{P}(A\cap B)$, as desired. Therefore, $\mathcal{P}(A)\cap\mathcal{P}(B) \subseteq \mathcal{P}(A\cap B)$, and we have an equality of sets. □

### 3.2.5. *Cartesian Product.*

**Definition.** If $A$ and $B$ are sets, then the *cartesian product* of $A$ and $B$, denoted $A \times B$, is the set of ordered pairs:

$$A \times B = \{(x, y) : x \in A \wedge y \in B\}.$$

Here, ordered pair means $(a, b) \neq (b, a)$, and $(a, b) = (c, d)$ if and only if $a = c$ and $b = d$.

**Example.**

1. If $A = \{1, 2\}$ and $B = \{a, b\}$, then
$$A \times B = \{(1, a), (1, b), (2, a), (2, b)\}.$$
2. If $A = B = \mathbb{R}$, then $A \times B = \{(x, y) : x \in R, y \in R\}$. In multivariable calculus and linear algebra, this is what is denoted $\mathbb{R}^2$. We see now that this coincides with $\mathbb{R} \times \mathbb{R}$.
3. If $A = \emptyset$ and $B$ is any set, then $A \times B = \emptyset$. Conversely, if $A \times B = \emptyset$, then $A = \emptyset$ or $B = \emptyset$.

We won't deal with the following concept much, but it is good to have it introduced.

**Definition.** If $A_1, \ldots, A_n$ are sets, then $A_1 \times A_2 \times \cdots \times A_n$ is the set

$$A_1 \times A_2 \times \cdots \times A_n = \{(a_1, a_2, \ldots, a_n) : a_i \in A_i \quad \forall i\}.$$

We have the following fact:

**Proposition.** If $A_1, \ldots, A_n$ are sets, then

$$|A_1 \times A_2 \times \cdots \times A_n| = |A_1| \cdot |A_2| \cdot |A_3| \cdot \cdots \cdot |A_n|.$$

## 4. Functions

In this section, we examine functions on sets. You probably already have some intuition for functions, having studied them in depth since middle school or high school. But here, we will introduce the vocabulary of functions in a way you most likely have not seen before.

### 4.1. Definition and Examples.

**Definition.** Let $X$ and $Y$ be sets.

1. A *function* $X$ to $Y$ is a subset $f$ of $X \times Y$ such that for all $x \in X$, there exists a unique $y \in Y$ such that $(x, y) \in f$.
2. The set $X$ is called the *domain* of $f$. The set $Y$ is called the *codomain* of $f$.

*Notation.* A function $f$ from $X$ to $Y$ is typically denoted $f : X \to Y$, and if $(x, y) \in f$, then we write $f(x) = y$.

In words, a function assigns to each point in $X$ a unique value in $Y$.

**Example.** Let $X = \{1, 2, 3\}$ and $Y = \{a, b\}$.

1. The subset $f = \{(1, a), (2, b), (3, a)\}$ is a function $f : X \to Y$, since every $x$ has a unique $y$-value associated to it. In the notation described above, $f(1) = a$, $f(2) = b$, $f(3) = a$.
2. The subset $f = \{(1, a), (2, a), (3, a)\}$ is also a function $f : X \to Y$.
3. If $f = \{(1, a), (2, b)\}$, then $f$ is not a function $f : X \to Y$, since there is no point in $f$ corresponding to the value 3 in the domain.
4. If $f = \{(1, a), (1, b), (2, a), (3, b)\}$, then $f$ cannot describe a function $f$ from $X$ to $Y$ because the value 1 in the domain has multiple $y$-values associated to it, meaning the uniqueness part of the definition fails.
5. If $X = \mathbb{R}$ and $Y = \mathbb{R}$, then $f = \{(x, x^2) : x \in \mathbb{R}\}$ is a function $f : \mathbb{R} \to \mathbb{R}$, and corresponds to our usual function $f(x) = x^2$.
6. If $X = \mathbb{R}$ and $Y = \mathbb{R}$, then $f(x) = \frac{1}{x}$ is NOT a function since there is $y$-value corresponding to the value 0 in the domain. However, if we let $X' = \mathbb{R} - \{0\}$, then $f$ is a function $f : X' \to Y$.

**Definition.** The set $\{y : \exists x \in X \text{ such that } (x, y) \in f\}$ is called the *range* of $f$, and it will be denoted $R(f)$.

**Example.**

1. In example (1) above, the range is $\{a, b\}$.
2. In example (2) above, the range is $\{a\}$.

### 4.2. Injection, Surjection, Bijection.

We will now take a look at different types of functions.

**Definition.** Let $X$ and $Y$ be sets and $f : X \to Y$ a function. We say $f$ is *injective* (or *1-1*) if for all $y \in Y$, there exists at most one $x \in X$ such that $(x, y) \in f$. In this case $f$ is called an *injection.*

*Remark.* There are two equivalent formulations of injective:

1. $\forall y \in R(f), \exists!\, x \in X, (x, y) \in f$. ($\exists!$ means "there exists a unique").
2. $\forall x_1 \in X \forall x_2 \in X (f(x_1) = f(x_2) \implies x_1 = x_2)$.

Observe that $f$ is NOT 1-1 if $\exists x_1 \in X \exists x_2 \in X(f(x_1) = f(x_2) \wedge x_1 \neq x_2)$. That is $f$ is not injective if there exist distinct $x$-values which have the same $y$-value.

**Example.** (1) If $X = \{1, 2, 3\}$ and $Y = \{a, b, c\}$, then $f = \{(1, a), (2, b), (3, c)\}$ is injective.

2. With the same $X$ and $Y$, $f = \{(1, a), (2, b), (3, a)\}$ is not injective.
3. $f : \mathbb{N} \to \mathbb{R}$ given by $f(n) = 2^n - n^2$ is not injective since $f(2) = f(4)$ (but $2 \neq 4$).

Next, we define what a surjective function is.

**Definition.** Let $X$ and $Y$ be sets and $f : X \to Y$ a function. We say $f$ is *surjective* (or *onto*) if for all $y \in Y$, there exists $x \in X$ such that $(x, y) \in f$. In this case $f$ is called an *surjection.*

Thus, a surjective function is one where every value in the codomain comes from at least one value in the domain. Equivalently, a function is surjective if $R(f) = Y$. Using the definition, we see that a function is NOT surjective if $\exists y \in Y, \forall x \in X, (x, y) \notin f$.

**Example.** (1) If $f : \mathbb{R} \to \mathbb{R}$ is given by $f(x) = x^2$, then $f$ is not surjective.

2. If $Y = \{y \in \mathbb{R} : y \geq 0\}$, then $f : \mathbb{R} \to Y$ given by $f(x) = x^2$ is surjective.
3. $f : \mathbb{Z} \to \mathbb{Z}$, $f(n) = 2n - 1$ is not surjective, as the range is odd numbers.

**Definition.** A function which is both injective and surjective is called *bijective.* Thus, a function is bijective if for all $y \in Y$ there exists a unique $x \in X$ such that $(x, y) \in f$.

If $f : X \to Y$ is bijective, then an inverse function exists. Namely, if $f \subseteq X \times Y$ is bijective, then we can define the inverse $f^{-1} \subseteq Y \times X$ as $(y, x) \in f^{-1} \iff (x, y) \in f$. Using the other notation, $f(x) = y \iff f^{-1}(y) = x$.

**Example.** (1) If $X = \{1, 2, 3\}$ and $Y = \{a, b, c\}$, then $f : X \to Y$ given by $f = \{(1, b), (2, c), (3, a)\}$ is bijective. The inverse is $f^{-1} : Y \to X$ given by $f^{-1} = \{(b, 1), (c, 2), (a, 3)\}$.

(2) If $X = \mathbb{R}$ and $Y = \{y \in \mathbb{R} : y > 0\}$, then $f(x) = e^x$ is bijective. The inverse is $f^{-1}(y) = \ln(y)$.

### 4.3. Composition.

**Definition.** If $f : X \to Y$ and $g : Y \to Z$ are functions (where $X$, $Y$, and $Z$ are sets), then we have a composition $g \circ f : X \to Z$ given by

$$g \circ f = \{(x, z) \in X \times Z : \exists y \in Y \text{ such that } (x, y) \in f \wedge (y, z) \in g\}.$$

**Example.** If $X = \{a, b, c\}$, $Y = \{1, 2, 3\}$, $Z = \{1, 2, 3, 4\}$, and $f : X \to Y$ and $g : Y \to Z$ are the functions

$$f = \{(a, 1), (b, 2), (c, 2)\}, \quad g = \{(1, 2), (2, 3), (3, 4)\},$$

then

$$g \circ f = \{(a, 2), (b, 3), (c, 3)\}.$$

### 4.4. Problems.

With all these definitions in hand, let's look at some problems.

**Problem.** If $f : X \to Y$ and $g : Y \to Z$ are injective functions, prove $g \circ f$ is injective.

*Solution.* We want to prove if $(g \circ f)(x_1) = (g \circ f)(x_2)$, then $x_1 = x_2$. If $g(f(x_1)) = g(f(x_2))$, then as $g$ is injective, this implies $f(x_1) = f(x_2)$. Now as $f$ is injective, this implies $x_1 = x_2$, as desired.

*Remark.* Is the above true if $f$ or $g$ is not injective? The answer is no. For example, if $X = \{1, 2\}$ and $Y = \{a, b\}$, and $f : X \to Y$ and $g : Y \to X$ are given by

$$f = \{(1, a), (2, a)\}, \quad g = \{(a, 1), (b, 2)\},$$

then $f$ is not injective, but $g$ is. Notice that

$$g \circ f = \{(1, 1), (2, 1)\}$$

is not injective. It is similarly easy to construct functions $f$ and $g$ such that $f$ is injective and $g$ is not injective, and $g \circ f$ is not injective.

Before looking at our next problem, let's go over some notation. If $f : X \to Y$ and $A \subseteq X$ and $B \subseteq Y$ are subsets, then we define

$$f(A) = \{f(a) : a \in A\} \subseteq Y, \quad f^{-1}(B) = \{x \in X : f(x) \in B\} \subseteq X.$$

The first set is called the *image* of $A$ and the second set is called the *inverse image* of $B$.

**Example.** If $f : \{1, 2, 3\} \to \{a, b\}$ is given by $f = \{(1, a), (2, b), (3, a)\}$, and $A = \{1, 2\}$, then $f(A) = \{a, b\}$. If $B = \{a\}$, then $f^{-1}(B) = \{1, 3\}$.

**Problem.** Prove a function $f : X \to Y$ is injective if and only if $f(A \cap B) = f(A) \cap f(B)$ for all subsets $A, B \subseteq X$.

*Solution.* Assume first that $f(A \cap B) = f(A) \cap f(B)$ for all subsets $A, B \subseteq X$. We want to show $f$ is injective, i.e. if $f(x) = f(y)$, then $x = y$. Assume $f(x) = f(y)$ but $x \neq y$. Then let $A = \{x\}$ and $B = \{y\}$. Then $A \cap B = \emptyset$, meaning $f(A \cap B) = \emptyset$. However, $f(A) \cap f(B) = \{f(x)\} \neq \emptyset$, a contradiction. Therefore, $f$ is injective.

Assume now $f$ is injective, and let $A, B \subseteq X$. We want to prove $f(A \cap B) = f(A) \cap f(B)$. We will show both inclusions. To show $f(A \cap B) \subseteq f(A) \cap f(B)$, observe that $A \cap B \subseteq A$ and $A \cap B \subseteq B$. Therefore, $f(A \cap B) \subseteq f(A)$ and $f(A \cap B) \subseteq f(B)$, which implies $f(A \cap B) \subseteq f(A) \cap f(B)$.

For the reverse inclusion, let $y \in f(A) \cap f(B)$. Then $y = f(a) = f(b)$ for some $a \in A$ and $b \in B$. Since $f$ is injective, this $a = b$. But this means $a = b \in A \cap B$, which implies $y \in f(A \cap B)$, as desired.

4.5. **Countability.** Recall that when we defined the cardinality of a set, we said that $|A|$ was the number of elements of $A$ provided $A$ had finitely many elements. Why did we impose this restriction? If there are infinitely many elements, why do we not say that $|A| = \infty$? This is the question we seek to answer in this section.

Here are some facts which are probably intuitively clear. Let $X$ and $Y$ be finite sets.

(1) There is an injection $f : X \to Y$ if and only if $|X| \leq |Y|$.
(2) There is a surjection $f : X \to Y$ if and only if $|X| \geq |Y|$.
(3) There is a bijection $f : X \to Y$ if and only if $|X| = |Y|$.

We will not go into the proof here, but they should make sense. Now, let us examine infinite sets. We know a few examples: $\mathbb{N}$, $\mathbb{Q}$, $\mathbb{Z}$, $\mathbb{R}$.

**Definition.** An infinite set $X$ is called *countably infinite* or *countable* if there is a bijection $f : \mathbb{N} \to X$.

*Remark.* A lot of books include finite sets in the set of countable sets, but we will not do that here.

What are some examples of countable sets?

**Example.** (1) $\mathbb{N}$ is countable (take $f : \mathbb{N} \to \mathbb{N}$ to be the map $f(n) = n$).
(2) $\mathbb{N}_0$ is countable: let $f : \mathbb{N} \to \mathbb{N}_0$ be $f(n) = n - 1$ which is clearly bijective.
(3) $\mathbb{Z}$ is countable: to give $f : \mathbb{N} \to \mathbb{Z}$ it suffices to list the elements of $\mathbb{Z}$ in some order in such a way that every integer is accounted for. We can do this as:

$$0, 1, -1, 2, -2, 3, -3, \ldots.$$

(4) $\mathbb{Q}$ is countable: again, it suffices to give a list. We do this as follows. Construct a table of rational numbers where the $i$-th row and $j$-th column of this table is the rational number $i/j$. Then proceed to list them as shown in Figure 6. In the process, skip any rationals that have already been accounted for. For example, $2/2$ is unnecessary since it is the same as $1/1$. Then, to the list, add each of the negatives right after the corresponding positives, and throw 0 in the beginning. This provides our list.

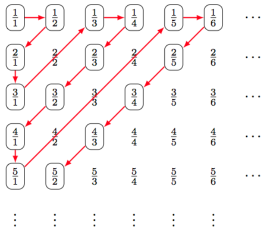


FIGURE 6. Rationals are Countable. Image from divisbyzero.com

Take a second to appreciate what example (4) shows. The rationals are everywhere. Between any two rational numbers there are infinitely many rational numbers. And yet there is somehow the "same number" of rationals as there are whole numbers.

The following proposition will allow us to build lots of countable sets.

**Proposition.** (1) If $A$ and $B$ are countable, then $A \cup B$ is countable.
(2) Let $\{A_i\}_{i\in\mathbb{N}}$ be a collection of countable sets. Then $\bigcup_{i\in\mathbb{N}} A_i$ is countable.
(3) If $X$ and $Y$ are countable sets, then so is $X \times Y$.

*Proof.* Number (1) is left as an exercise (we did it in class). For (2), we can assume that the $A_i$'s have no elements in common. Let $a_{ij}$ be the elements in $A_i$, where $j \in \mathbb{N}$. We can list them in a table like we did for $\mathbb{Q}$ above, and then we can perform the same trick as we did in that example to prove that the union is countable.

Statement (3) follows from (1). If $X = \{x_i\}_{i\in\mathbb{N}}$ and $Y = \{y_j\}_{j\in\mathbb{N}}$, then

$$X \times Y = \bigcup_{i\in\mathbb{N}} \{(x_i, y_j) : j \in \mathbb{N}\}.$$

Each set in the union is countable (in bijection with $Y$), so $X \times Y$ is countable by (1). □

Notice that this proposition tells us that $\mathbb{N} \times \mathbb{N}$ is countable, $\mathbb{Q} \times \mathbb{Q} \times \mathbb{Q} \times \mathbb{Q}$ is countable, and so on.

What is the punchline of all this? What is very interesting is that $\mathbb{R}$ is NOT countable, even though it is infinite.

**Proposition.** $\mathbb{R}$ is not countable.

*Proof.* Since there is a bijection $g : (0, 1) \to \mathbb{R}$, it suffices to prove $(0, 1)$ is not countable. If it were countable, there would be a bijection $f : \mathbb{N} \to (0, 1)$, and composing with $g$ would give a bijection $\mathbb{N} \to \mathbb{R}$. So we seek to show $(0, 1)$ is not countable. A number in $(0, 1)$ can be written in an infinite decimal expansion. It is not unique, however. For example, $.49999\overline{9} = .5$. To fix a standard expansion, we take the one ending in 0's, not 9's. Suppose $(0, 1)$ were countable, so that we could list them as

$$1 \mapsto 0.a_{11}a_{12}a_{13}\cdots$$
$$2 \mapsto 0.a_{21}a_{22}a_{23}\cdots$$
$$3 \mapsto 0.a_{31}a_{32}a_{33}\cdots$$
$$\vdots$$

We will create an element not on the list. Call our new number $s = 0.s_1s_2\cdots$. To construct it, let $s_1$ be a number different from $a_{11}$, $s_2$ a number different from $a_{22}$, $s_3$ a number different from $a_{33}$, and so on. We claim $s$ is not on our list. Indeed, it differs from the $i$-th element in the $i$-th decimal place, and so could not be the $i$-th number. This is a contradiction, and thus $\mathbb{R}$ is not countable. □

So with our definitions, $\mathbb{R}$ is "larger" than $\mathbb{N}$ (or $\mathbb{Z}$, or $\mathbb{Q}$), in that the latter sets are countable and $\mathbb{R}$ is not (it is called *uncountable*). We can actually show more:

**Proposition.** There is no bijection between a set $A$ and $\mathcal{P}(A)$.

*Proof.* Suppose there was a bijection $f : A \to \mathcal{P}(A)$. Then for each $a \in A$, $f(a)$ is a subset of $A$. We will construct a set $S \in \mathcal{P}(A)$ which is not in the image of $f$. To do this, we need to decide, for each $a \in A$, whether $a$ belongs to $S$ or not. We do this as follows: if $a \in f(a)$, then $a \notin S$, and if $a \notin f(a)$, then $a \in S$. Repeat this for all $a \in A$ to construct $S$.

We claim $S \notin R(f)$. Indeed, if it was, then $S = f(b)$ for some $b \in A$. Is the element $b \in S$? By construction, it is in $f(b) = S$ if and only if $b \notin f(b) = S$, a contradiction. Therefore, $S \notin R(f)$ and hence $f$ is not a bijection. □

**Corollary.** There is no "largest set."

## 5. Sequences and Strings

In this section, we will study sequences, which you typically encounter in a Calculus BC course, and strings, which is a special type of functions.

### 5.1. Sequences.

**Definition.** A *sequence* is a function $f : D \to S$, where $D \subset \mathbb{Z}$ is a set of consecutive numbers and $S$ is any set (typically $\mathbb{R}$).

**Example.**

(1) Let $D = \mathbb{N}$ and $S = \mathbb{R}$, and let $f : D \to S$ be the function $f(n) = \frac{1}{n}$. This is the sequence
$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots.$$
We sometimes write $f(n) = a_n = \frac{1}{n}$, and the terms of the sequence are $a_1, a_2, \ldots$ (since $n \in \mathbb{N}$).

(2) Let $D = \{1, 2, 3, 4\}$ and $S = \mathbb{N}$. Let $f : D \to S$ be $f(1) = 1$, $f(2) = 3$, $f(3) = 3$, and $f(4) = 2$. This is the sequence
$$1, 3, 3, 2.$$

**Definition.** A sequence is called *finite* if $D$ is finite.

Since you study sequences in depth in calculus, we will not go over them in much detail here. However, we will briefly study one interesting idea which you probably have not seen before.

5.1.1. *Finding Polynomial Expressions for Sequences.* Suppose we have the sequence

$$1, 4, 7, 10, 13, \ldots.$$

We want to find the formula for the function $f : \mathbb{N}_0 \to \mathbb{Z}$ which produces this sequence. In this case, the function is linear, and is given by the formula $f(n) = 3n + 1$.

How did we know the function was linear? Looking at consecutive terms in the sequence, we notice that the difference is always three. This tells us that the function grows linearly. From this, we can say the function is $f(n) = an + b$ for integers $a$ and $b$, and then we can use the fact that $f(0) = 1$ and $f(1) = 4$ to find what $a$ and $b$ need to be:

$$1 = a(0) + b$$
$$4 = a(1) + b,$$

and solving for $a$ and $b$ yields $a = 3$ and $b = 1$.

The moral of the this example is that linear things are easy. Given a more complicated sequence, how do we know what polynomial fits the sequence? Let's take a look at squares for a moment:

$$0, 1, 4, 9, 16, 25, \ldots.$$

Here, unlike with the linear expression, the differences between the terms are not constant:

$$1, 3, 5, 7, 9, \ldots.$$

However, observe that if we look at *second differences* (namely, the differences of the differences), then they are constant:

$$2, 2, 2, 2, \ldots.$$

This suggests that to determine whether a sequence is quadratic, we should see whether the second differences are constant. We can verify this as follows: consider the function $f(n) = an^2 + bn + c$. The first few terms are

$$c, a + b + c, 4a + 2b + c, 9a + 3b + c, 16a + 4b + c, 25a + 5b + c, \ldots.$$

The differences are:

$$a + b, 3a + b, 5a + b, 7a + b, 9a + b, \ldots,$$

and the second differences are

$$2a, 2a, 2a, 2a, \ldots.$$

So with quadratic expressions, the second differences are constant, and we can even determine the leading coefficient $a$ from this quantity (namely, it is the constant second difference divided by 2). That is pretty remarkable.

As one would expect, constant third differences come from cubics, constant fourth differences come from quartics, and so on.

**Problem.** Find a general formula for the sequence

$$2, 4, 8, 14, 22, 32, \ldots.$$

*Solution.* Let's look at differences:

$$2, 4, 6, 8, 10, \ldots.$$

These aren't constant, so we can look at second differences:

$$2, 2, 2, 2, \ldots.$$

Since these are constant, our expression is quadratic. (We even know the leading coefficient is 1, but even if we don't use that trick it is easy to do from here.) Let $f(n) = an^2 + bn + c$. We know $f(0) = 2$, $f(1) = 4$ and $f(2) = 8$. Plugging these in gives the three equations:

$$2 = c$$

$$4 = a + b + c$$
$$8 = 4a + 2b + c.$$

After solving, we find $a = 1$, $b = 1$, and $c = 2$. Thus, the formula is

$$f(n) = n^2 + n + 2.$$

5.2. **Strings.**

**Definition.** Let $X$ be a set. A *string over* $X$ is a finite sequence $f : D \to X$ (meaning $D$ is finite), where we take $D = \{1, 2, 3, \ldots, n\}$.

**Example.** (1) $X = \{a, b, c\}$, $D = \{1, 2, 3, 4, 5\}$, and $f : D \to X$ given by the sequence $a, a, b, c, b$.
(2) $X = \{x, y, z, w\}$, $D = \{1, 2, 3, 4, 5, 6\}$, $g : D \to X$ given by $x, y, y, z, w, x$.

We typically write strings as sequences of characters without any commas. So the first sequence above would be $aabcb = a^2bcb$ (here the $a^2$ means 2 consecutive $a$'s). The second string can be written as $xyyzwx = xy^2zwx$.

If $D = \emptyset$, then the unique sequence $D \to X$ is denoted $\lambda$ (the null string).

*Notation.* The set of all strings over $X$ is denoted $X^*$, and the set of all non-null strings over $X$ is denoted $X^+$.

One more piece of terminology:

**Definition.** A *bit string* is a string on the set $X = \{0, 1\}$.

5.3. **Operations on Strings.** Like all objects we have encountered so far, there are operations we can perform.

5.3.1. *Length.*

**Definition.** The *length* of a string $f : D \to X$ over $X$ is $|D|$, and it is denoted $|f|$.

This should be pretty self-explanatory. We list a few examples:

**Example.** (1) If $\alpha = aabab = a^2bab$ is a string over $X = \{a, b, c\}$, then $|\alpha| = 5$.
(2) If $\beta = a^3b^4a^{32}$ over the same set $X$, then $|\beta| = 39 = 3 + 4 + 32$.

5.3.2. *Concatenation.*

**Definition.** Let $\alpha$ and $\beta$ be two strings over $X$. The *concatenation* is denoted $\alpha\beta$, and it is the string $\alpha$ followed by the string $\beta$. Formally, if $\alpha : \{1, 2, \ldots, n\} \to X$ and $\beta = \{1, 2, 3, \ldots, m\} \to X$, then $\alpha\beta : \{1, 2, \ldots, m + n\} \to X$ is given by

$$(\alpha\beta)(i) = \alpha(i) \quad \text{for } i = 1, 2, \ldots, n$$

and

$$(\alpha\beta)(j) = \beta(j - n) \quad \text{for } j = n + 1 \ldots, n + m.$$

*Remark.* It is clear from the definition that the length of $\alpha\beta$ is $|\alpha\beta| = |\alpha| + |\beta|$.

**Example.** Let $X = \{a, b\}$. If $\alpha = aba^2$ and $\beta = babab^2$, then

$$\alpha\beta = aba^2babab^2.$$

Notice

$$\beta\alpha = babab^2aba^2 \neq \alpha\beta,$$

so concatenation is not commutative.

*Remark.* Observe that for any string $\alpha$ over $X$, we have

$$\lambda\alpha = \alpha\lambda = \alpha.$$

This means that $\lambda$ is the "identity element" of strings.

**Problem.** Let $X$ be any nonempty set. Let $f : X^* \times X^* \to X^*$ be the map $f(\alpha, \beta) = \alpha\beta$. Is $f$ injective? Surjective? Bijective?

*Solution.* The map $f$ being injective means $f(\alpha, \beta) = f(\alpha', \beta')$ implies $\alpha = \alpha'$ and $\beta = \beta'$. But this is not true. If $x \in X$, then the string $xxx = xx \cdot x = x \cdot xx$ can be decomposed in two distinct ways. Therefore $f$ is not injective.

However, $f$ is surjective. Take $\alpha \in X^*$. Then $\alpha = f(\alpha, \lambda)$, so $f$ is surjective. The map $f$ is not bijective as it is not injective.

5.3.3. *Reverse.*

**Definition.** If $X$ is a set and $\alpha$ is a string on $X$, then the *reverse* string, denoted $\alpha^R$ is just the string $\alpha$ backwards. Formally, if $n = |\alpha|$, then

$$\alpha^R(i) = \alpha(n+1-i) \quad \text{for } i = 1, 2, \ldots, n.$$

Again, it is clear from the definition that $|\alpha^R| = |\alpha|$ and that $(\alpha^R)^R = \alpha$.

**Example.** If $\alpha = aba^2$ is a string on $X = \{a, b\}$, then $\alpha^R = aaba = a^2ba$.

**Problem.** Let $X$ be a nonempty set, and let $f : X^* \to X^*$ be $f(\alpha) = a^R$. Prove $f$ is a bijection.

*Solution.* To prove $f$ is injective, let $f(\alpha) = f(\beta)$, so $\alpha^R = \beta^R$. Then applying the reverse operation to both sides gives $\alpha = \beta$, meaning $f$ is injective.

To prove $f$ is surjective, simply observe that $f(\alpha^R) = \alpha$. Thus, $f$ is a bijection.

5.3.4. *Substrings.*

**Definition.** If $\alpha$ is a string over a nonempty set $X$, then a string $\beta$ is a *substring* of $\alpha$ if there exist $\gamma, \delta \in X^*$ such that $\alpha = \gamma\beta\delta$.

Informally, a substring is just a subsequence of consecutive characters in the larger string.

**Example.** If $X = \{a, b, c\}$ and $\alpha = aabcbbac$, then $\beta = bcb$ is a substring with $\gamma = aa$ and $\delta = bac$.

**Problem.** Let $X$ be a nonempty set, and let $f : X^* \to \mathcal{P}(X^*)$ be the map with $f(\alpha)$ being the set of all substrings of $\alpha$. Is $f$ injective? Surjective? Bijective?

*Solution.* Suppose $f(\alpha) = f(\beta)$. We want to see if $\alpha = \beta$. Well $\alpha$ is a substring of itself, and since $f(\alpha) = f(\beta)$, it must be a substring of $\beta$. Therefore $\beta = \gamma\alpha\delta$ for some strings $\gamma$ and $\delta$. By the same reasoning $\beta$ is a substring of $\alpha$, so there exists $\gamma'$ and $\delta'$ such that $\alpha = \gamma'\beta\delta'$. Together, they say $|\alpha| = |\beta|$, and the only way this can happen is if $\gamma = \delta = \gamma' = \delta' = \lambda$. Thus, $\alpha = \beta$ and $f$ is injective.

The map $f$ is not surjective. Any set in $\mathcal{P}(X^*)$ which does not include $\lambda$ cannot be in the image of $f$. Therefore, $f$ is neither surjective nor bijective.

## 6. Relations

### 6.1. Definition and Examples.

**Definition.** A *binary relation* $R$ from a set $X$ to a set $Y$ is a subset $R \subseteq X \times Y$.

*Notation.* (1) If $(x, y) \in R$, then we write $xRy$ or $x \sim y$.
(2) If $X = Y$, then we say $R$ is a relation on $X$.

**Example.** Let $X = \{\text{Bill, Joe, Alice, Dave}\}$ and $Y = \{\text{Math, Physics, Chemistry, History}\}$. Then

$$R = \{(\text{Bill, Math}), (\text{Alice, Physics}), (\text{Bill, History}), (\text{Dave, Chemistry}), (\text{Joe, Physics})\} \subset X \times Y$$

is an example of a relation from $X$ to $Y$.

One could imagine that this relation encodes classes different students are taking. We could also define relations with rules:

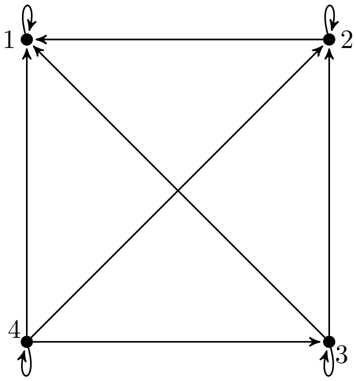
**Example.** (1) Let $X = \{1, 2, 3, 4\}$ and $Y = \{1, 2, 3, 4, 5\}$. Then we can define a relation $R$ by saying $x \sim y$ if $x|y$. We could list out $R$ by writing all the ordered pairs:

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

(2) $X = \{1, 2, 3, 4\}$. Define a relation on $X$ by writing $x \sim y$ if $x \geq y$. Thus,

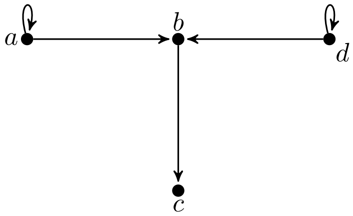
$$R = \{(1,1),(2,1),(3,1),(4,1),(2,2),(3,2),(4,2),(3,3),(4,3),(4,4)\}.$$

We could also encode this information in the following diagram:



The vertices in the graph are just the elements of $X$, and there is an arrow from $x$ to $y$ precisely when $x \sim y$.

(3) Using the previous idea, if $X = \{a, b, c, d\}$, then the following illustrates a relation on $X$:



This is the relation:

$$R = \{(a,a),(a,b),(b,c),(d,b),(d,d)\}.$$

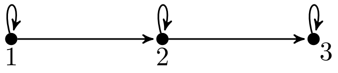
6.2. **Types of Relations.** There are four types of relations that we will focus on, and three of them will be very important in the next section.

6.2.1. *Reflexive.*

**Definition.** A relation $R$ on a set $X$ is called *reflexive* if $\forall x \in X$, $(x,x) \in R$. That is, every $x \in X$ is related to itself.

*Remark.* Using what we know about negation, we can say that a relation is NOT reflexive if $\exists x \in X$ such that $(x,x) \notin R$.

**Example.** (1) If $X = \{1, 2, 3\}$ and $R = \{(1,1),(1,2),(2,3),(2,2),(3,3)\}$, then $R$ is reflexive. To verify this, just go through each element $x$ of $X$ and verify that $(x,x) \in R$. We can also see this graphically. The graph of this relation is:



The relation $R$ being reflexive means every vertex has a loop to itself.

(2) If $X = \{a, b, c, d\}$ and $R = \{(a,a),(b,c),(b,d),(b,b),(c,a),(d,d)\}$, then $R$ is not reflexive as $(c,c) \notin R$.

**Problem.** Suppose we define a relation on $\mathbb{Z}$ by saying $x \sim y$ if $3|(x-y)$. Is this relation reflexive?

*Solution.* We need to see if $x \sim x$ for any $x \in \mathbb{Z}$. So let $x \in \mathbb{Z}$. The statement $x \sim x$ is equivalent to asking whether $3|(x-x)$, or $3|0$. This is true as $3 \cdot 0 = 0$. Therefore, $x \sim x$, and since $x \in \mathbb{Z}$ was arbitrary, we have proved this relation is reflexive.

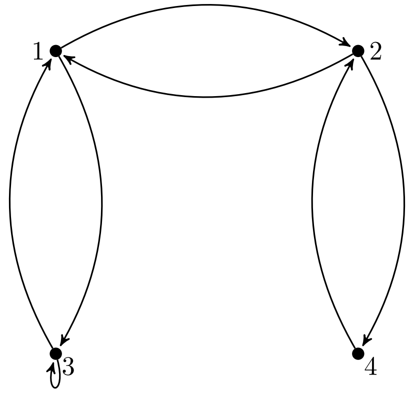
6.2.2. *Symmetric.*

**Definition.** A relation $R$ on a set $X$ is called *symmetric* if $\forall x \in X\ \forall y \in X\ (x, y) \in R \implies (y, x) \in R$.

*Remark.* Again, we can negate this statement to find that $R$ is not symmetric if $\exists x \in X\ \exists y \in X$ with $(x, y) \in R \wedge (y, x) \notin R$.

**Example.** (1) Let $X = \{1, 2, 3, 4\}$ and define a relation by $x \sim y$ if $x|y$. This is not symmetric, as $(2, 4) \in R$ and $(4, 2) \notin R$.

(2) If $X = \{1, 2, 3, 4\}$ and $R = \{(1, 2), (2, 1), (3, 1), (3, 3), (1, 3), (4, 2), (2, 4)\}$. This is symmetric. To verify this, one could go through and check each pair of elements and verify that we can switch the order and still remain in $R$. For example, $(1, 2) \in R$ and $(2, 1) \in R$. We also have $(3, 1) \in R$ and $(1, 3) \in R$, and so on. We could also, again, look at the graph.



One can tell symmetry from the graph if every pair of vertices either has zero edges or two edges (one in each direction) between them. So the fact that 3 and 4 have no edges between them is ok, but since 3 has an edge going to 1, we must verify that 1 also has an edge going to 3, and so on.

**Problem.** Consider the relation on $\mathbb{Z}$ defined previously: $x \sim y$ if $3|(x - y)$. Prove this relation is symmetric.

*Solution.* Suppose $x, y \in \mathbb{Z}$ with $x \sim y$. We want to show $y \sim x$. Since $x \sim y$, we know there exists $k \in \mathbb{Z}$ with $3k = x - y$. But then $y - x = 3(-k)$, meaning $3|(y - x)$. Therefore $y \sim x$, as desired. Therefore, the relation is symmetric.

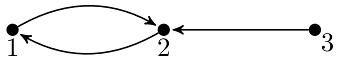
6.2.3. *Antisymmetric.*

**Definition.** A relation $R$ on a set $X$ is called *antisymmetric* if $\forall x \in X\ \forall y \in X\ ((x, y) \in R \wedge (y, x) \in R) \implies x = y$. That is, the only way two elements can each be related to the other is if they are the same element.

*Remark.* A relation is not antisymmetric if $\exists x \in X\ \exists y \in X$ with $(x, y) \in R \wedge (y, x) \in R$ but $x \neq y$.

**Example.** (1) The relation in example (1) of 6.2.2 is antisymmetric. Indeed, if $x \sim y$ and $y \sim x$, then $x|y$ and $y|x$. But the only way this can happen is if $x = y$.

(2) If $X = \{1, 2, 3\}$ and $R = \{(1, 2), (2, 1), (3, 2)\}$, then $R$ is not antisymmetric: take $x = 1$ and $y = 2$ in the remark above. We could also tell from the graph:



The problem with this graph is that there are two edges between 1 and 2. A relation is antisymmetric if there is at most one edge between any two distinct vertices.

It is important to note that symmetric and antisymmetric are not opposites of one another. Example (2) above is an example of a relation which is neither symmetric nor antisymmetric. The relation on $\mathbb{Z}$ given by $x \sim y$ if $x = y$ is both symmetric and antisymmetric.

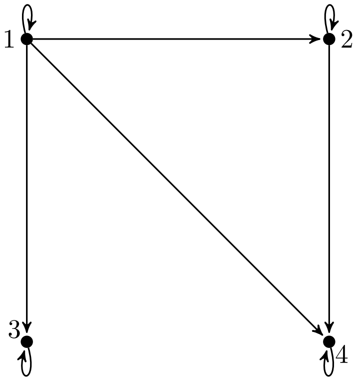
6.2.4. *Transitive.* This is the last property that we will need.

**Definition.** A relation $R$ on a set $X$ is *transitive* if $\forall x \in X\ \forall y \in X\ \forall z \in X\ ((x,y) \in R \wedge (y,z) \in R) \implies (x,z) \in R$.

**Example.** (1) Let $X = \{1,2,3,4\}$ and let $R$ be the relation $x \sim y$ if $x|y$ as before. Observe that

$$R = \{(1,1),(2,2),(3,3),(4,4),(1,2),(1,3),(1,4),(2,4)\}.$$

This is transitive. For example, $(1,2) \in R$ and $(2,4) \in R$, and we check that $(1,4) \in R$ as well. The graph of this relation is:



Transitivity translates to: if there is an edge from $x$ to $y$ and an edge from $y$ to $z$, then there is an edge from $x$ to $z$.

(2) Again look at the relation on $\mathbb{Z}$ defined by $x \sim y$ if $3|(x-y)$. This relation is transitive. If $x \sim y$ and $y \sim z$, then there exist $m, n$ with $3m = x - y$ and $3n = y - z$. But then

$$x - z = (x-y) + (y-z) = 3m + 3n = 3(m+n),$$

which means $3|(x-z)$, as desired. Therefore, the relation is transitive.

6.3. **Problems.** Let's look at a few problems.

**Problem.** Define a relation on $\mathbb{R}$ by saying $(x,y) \in R$ if $x \leq y$. Determine whether this relation is reflexive, symmetric, antisymmetric, transitive, or some combination of these.

*Solution.* This relation is reflexive as $x \leq x$ for any $x \in \mathbb{R}$. It is not symmetric though: $x \leq y$ does not mean $y \leq x$. For example, $4 \leq 5$ but $5 \not\leq 4$. It is antisymmetric: if $x \leq y$ and $y \leq x$, then $x = y$, which is the exact definition of antisymmetry. And finally, it is transitive. If $x \leq y$ and $y \leq z$, then $x \leq z$. Therefore, the relation is reflexive, antisymmetric, and transitive.

**Problem.** Same problem as above for the relation on $\mathbb{R}$ given by $x \sim y$ if $x - y \in \mathbb{Q}$.

*Solution.* This relation is reflexive as $x - x = 0 \in \mathbb{Q}$. It is symmetric as well: take $x, y \in \mathbb{R}$. If $x \sim y$, then $x - y = k \in \mathbb{Q}$. But then $y - x = -k \in \mathbb{Q}$, so $y \sim x$ as well. Therefore the relation is symmetric. It is not antisymmetric, $(0,1) \in R$ and $(1,0) \in R$ but $1 \neq 0$. Lastly, it is transitive. If $x, y, z \in \mathbb{R}$ with $x \sim y$ and $y \sim z$, then $x - y = k \in \mathbb{Q}$ and $y - z = k' \in \mathbb{Q}$. But then $x - z = k + k' \in \mathbb{Q}$. Therefore, $x \sim z$ and the relation is transitive.

**Problem.** If $R$ and $S$ are relations on $X$, then we can examine the relation $R \cap S$. If $R$ and $S$ are both transitive, prove $R \cap S$ is transitive.

*Solution.* Suppose $x, y, z \in X$ with $(x,y) \in R \cap S$ and $(y,z) \in R \cap S$. We want to show $(x,z) \in R \cap S$. Since $(x,y) \in R \cap S$, we know $(x,y) \in R$ and $(x,y) \in S$, and similarly for $(y,z)$. But as $R$ is transitive, this implies $(x,z) \in R$. Similarly, as $S$ is transitive, we know $(x,z) \in S$. Therefore $(x,z) \in R \cap S$, and so $R \cap S$ is transitive.

*Remark.* There is nothing special about transitivity in the previous example. If $R$ and $S$ both have property $P$ (where $P$ is one of the four properties examined above), then $R \cap S$ will also have property $P$.

As an illustration of this intersection, let $R$ be the relation on $\mathbb{Z}$ defined by $(x,y)\in R$ if $2|(x-y)$. Let $S$ be the relation on $\mathbb{Z}$ given by $(x,y)\in S$ if $3|(x-y)$. What is $R\cap S$? If $(x,y)\in R\cap S$, then $2|(x-y)$ and $3|(x-y)$, so $6|(x-y)$. The claim is that $R\cap S$ is precisely this set, namely,

$$R\cap S=\{(x,y)\in\mathbb{Z}\times\mathbb{Z}:6|(x-y)\}.$$

We have proved the inclusion $R\cap S\subseteq\{(x,y)\in\mathbb{Z}\times\mathbb{Z}:6|(x-y)\}$. For the reverse, assume $(a,b)\in\{(x,y)\in\mathbb{Z}\times\mathbb{Z}:6|(x-y)\}$. We want $(a,b)\in R\cap S$. Since $6|(b-a)$, we know $2|(b-a)$, so $(a,b)\in R$. We also know $3|(b-a)$, so $(a,b)\in S$. Therefore, $(a,b)\in R\cap S$, as desired.

## 7. Equivalence Relations

### 7.1. Definition and Examples.

We now turn our attention to a very specific type of relation, known as an equivalence relation. This concept comes up repeatedly in math classes in every area of math, so we devote an entire section to it.

**Definition.** A relation $R$ on a set $X$ is called an *equivalence relation* if $R$ is reflexive, symmetric, and transitive.

Since equivalence relations are important, we will look at several examples.

**Example.** Let $R$ be the relation on $\mathbb{R}$ defined by $x\sim y$ if $x-y\in\mathbb{Z}$. To check this is an equivalence relation, we need to check the three properties:

1. *Reflexive*: Let $x\in\mathbb{R}$. Then $x-x=0\in\mathbb{Z}$, meaning $x\sim x$, which is what we wanted. Therefore, $R$ is reflexive.
2. *Symmetric*: Let $x,y\in\mathbb{R}$ with $x\sim y$. We want to show $y\sim x$. Since $x\sim y$, we know $x-y=k\in\mathbb{Z}$. But then $y-x=-k\in\mathbb{Z}$. Therefore, $y\sim x$, and $R$ is symmetric.
3. *Transitive*: Let $x,y,z\in\mathbb{R}$ with $x\sim y$ and $y\sim z$. We want to show $x\sim z$. Since $x\sim y$, we know $x-y=k\in\mathbb{Z}$, and since $y\sim z$, we know $y-z=l\in\mathbb{Z}$. But then
$$x-z=(x-y)+(y-z)=k+l\in\mathbb{Z}.$$
Thus, $x\sim z$, and $R$ is transitive.

Therefore, $R$ is an equivalence relation.

**Example.** Fix a positive integer $n$. We define a relation $R$ on $\mathbb{Z}$ by saying $a\sim b$ if $n|(a-b)$. We saw in the previous section (with $n=3$) that this relation is reflexive, symmetric, and transitive. There was nothing special about 3 in those examples. Replace the 3 by an $n$ and the same work shows this is an equivalence relation for any $n\in\mathbb{Z}$.

**Example.** Let $S$ and $T$ be sets, and let $f:S\to T$ be a function. Define a relation $R$ on $S$ by saying $x_1\sim x_2$ if $f(x_1)=f(x_2)$. This relation is clearly reflexive and symmetric. If $x_1\sim x_2$ and $x_2\sim x_3$, then $f(x_1)=f(x_2)$, and $f(x_2)=f(x_3)$, and therefore $f(x_1)=f(x_3)$, meaning $x_1\sim x_3$. Thus, $R$ is transitive, making it an equivalence relation.

### 7.2. Equivalence Classes.

**Definition.** Let $R$ be an equivalence relation on a set $X$. For $a\in X$, define the *equivalence class* of $a$ to be the set of all elements equivalent to $a$. It is denoted $[a]$. In other words,

$$[a]=\{x\in X:x\sim a\}.$$

**Example.** In the first example above, we have

$$[0]=\{\ldots,-3,-2,-1,0,1,2,3,\ldots\}=\mathbb{Z},$$
$$[2]=\{\ldots,-3,-2,-1,0,1,2,3,\ldots\}=\mathbb{Z},$$
$$[1.1]=\{\ldots,-3.9,-2.9,-1.9,-.9,.1,1.1,2.1,\ldots\}.$$

How do we get these? So see what $[0]$ is, we need to see which elements are equivalent to zero. We know $x\sim 0$ if and only if $x-0\in\mathbb{Z}$, but this just means $x\in\mathbb{Z}$. This is why $[0]=\mathbb{Z}$.

For the second one, we have that $x\sim 2$ precisely when $x-2\in\mathbb{Z}$. But if $x-2\in\mathbb{Z}$, then $x\in\mathbb{Z}$ as well, which explains the second class.

Lastly, $x \sim 1.1$ if $x - 1.1 \in \mathbb{Z}$. This means $x = 1.1 + k$ for $k \in \mathbb{Z}$. This is how we get the last set.

**Example.** If $n = 3$ in the second example, how can we describe $[2]$? Well,

$$[2] = \{\dots, -1, 2, 5, 8, 11, \dots\}.$$

Taking a closer look, we see that $[2]$ is just the set of integers which have a remainder of 2 when divided by 3. More generally, $[a]$ will be the set of integers which have the same remainder as $a$ when divided by 3. If $n$ is just an arbitrary integer, then $[a]$ will be the set of integers which share the same remainder as $a$ when divided by $n$.

Taking an even closer look at the relation on $\mathbb{R}$ given by $x \sim y$ if $x - y \in \mathbb{Z}$, observe $[0] = [2]$. Why is that? It happened because $0 \sim 2$ under $R$, so that anything equivalent to 2 is equivalent to 0 (by transitivity), and vice versa. This is a more general fact.

**Lemma.** If $R$ is an equivalence relation on a set $S$, then for any $a, b \in S$, $[a] = [b]$ if and only if $a \sim b$.

*Proof.* Let $a, b \in S$. If $[a] = [b]$, then as $a \in [a]$, we know $a \in [b]$, and thus $a \sim b$.

Conversely, let $a \sim b$. Then $[a] \subseteq [b]$ because if $c \in [a]$, then $c \sim a$, and since $a \sim b$, we know by transitivity that $c \sim b$. Thus $c \in [b]$. The reverse inclusion is similar. Thus, $[a] = [b]$, as desired. □

Consider the relation on $\mathbb{Z}$ given by $a \sim b$ if $5|(b - a)$, an equivalence relation as above. Then we know $[2] = [7]$ since $2 \sim 7$, but $[2] \neq [1]$ as $2 \not\sim 1$.

**Definition.** Let $R$ be an equivalence relation on a set $S$. If we consider the equivalence class $[a]$, then $a$ is called a *representative* for the equivalence class.

*Remark.* All the work above shows that there could be many representatives for the same class. For example, if $R$ on $\mathbb{R}$ is the relation $x \sim y$ if $x - y \in \mathbb{Z}$, then every equivalence class has infinitely many representatives. Take the class $[0]$. We know $[0] = [1]$, so 1 is a representative for the same class. In fact, any integer $k$ can be a representative for the class, as $[0] = [k]$ for any $k \in \mathbb{Z}$. In this example, there are infinitely many possible representatives.

### 7.3. **Quotients.**

**Definition.** Let $X$ be a set and $R$ an equivalence relation on $X$. Then the set of equivalence classes under $R$ is denoted $X/R$ or $X/\sim$. We call this set $X$ *modulo* $R$ or the *quotient of* $X$ *by* $R$.

**Example.** If $R$ is the relation on $\mathbb{R}$ given by $x \sim y$ if $x - y \in \mathbb{Z}$, then what is $\mathbb{R}/\sim$? We know that two real numbers define the same equivalence class if they differ by an integer. Notice that every class has a representative which lies in the interval $[0, 1)$. For example, the class $[5.1234]$ is the same as the class $[.1234]$. Moreover, if $a, b \in [0, 1)$, then $[a] \neq [b]$, since there is no way $a$ and $b$ can differ by an integer (as they are less than distance 1 from one another). Thus,

$$\mathbb{R}/R = \{[c] : c \in [0, 1) \subset \mathbb{R}\}.$$

**Example.** Fix a positive integer $n$. Let $R$ be the relation on $\mathbb{Z}$ given by $x \sim y$ if $n|(x - y)$. We typically denote $\mathbb{Z}/R$ by $\mathbb{Z}/n$ to emphasize that it depends on $n$. What is $\mathbb{Z}/n$? The claim is that

$$\mathbb{Z}/n = \{[0], [1], [2], \dots, [n - 1]\}.$$

To see this, notice that every equivalence class has some representative between 0 and $n - 1$ (just take the remainder when divided by $n$). On the other hand, all these classes are inequivalent since the difference between any two of the representatives is less than $n$, and hence cannot be divisible by $n$.

### 7.4. **Equivalence Relations and Partitions.** We have the following proposition.

**Proposition.** Let $X$ be a set and $R$ an equivalence relation on $X$. Then every element of $X$ belongs to a unique equivalence class.

*Proof.* Let $x \in X$. Then $x \in [x]$, so $x$ belongs to some equivalence class. We want to show this class is unique. Suppose $x$ belongs to another class $[y]$, i.e. $x \in [y]$. Then by definition, $x \sim y$. But then by the lemma proved above, we get $[x] = [y]$. Thus, the class is unique. □

What this proposition tells us is that equivalence classes divide our original set $X$ into different disjoint pieces. We can make this precise with the following definition.

**Definition.** Let $X$ be a set. A family $\mathcal{F}$ of subsets of $X$ *partitions* $X$ if every element of $X$ belongs to exactly one member of $\mathcal{F}$.

**Example.** Let $X = \{1, 2, 3, 4\}$, and let $\mathcal{F} = \{\{1, 2\}, \{3\}, \{4\}\}$. Then $\mathcal{F}$ partitions $X$. On the other hand, if $\mathcal{G} = \{\{1, 2\}, \{1, 3, 4\}\}$, then $\mathcal{G}$ is not a partition of $X$ because 1 belongs to two classes, not just one.

Given a set $X$ and an equivalence relation $R$, then we get a partition of $X$ by just considering the equivalence classes. That is, $X/R$ partitions $X$. The converse is true as well.

**Proposition.** If $\mathcal{F}$ is a partition on $X$, then $\mathcal{F}$ determines an equivalence relation on $X$ by declaring that $x \sim y$ if and only if $\exists A \in \mathcal{F}$ such that $x \in A$ and $y \in A$.

*Proof.* We show the three properties. First, take $x \in X$. Then $x \in A$ for some $A \in \mathcal{F}$ by definition of the partition. Thus, $x \sim x$, and the relation is reflexive.

Next, suppose $x \sim y$. Thus, $x \in A$ and $y \in A$ for some $A \in \mathcal{F}$. But this clearly means $y \sim x$ (conjunctions are symmetric). Thus, the relation is symmetric.

Finally, suppose $x \sim y$ and $y \sim z$ for elements $x, y, z \in X$. Then $x \in A$ and $y \in A$ for some $A \in \mathcal{F}$, and $y \in B$ and $z \in B$ for some $B \in \mathcal{F}$. However, by definition, $y$ must belong to a unique member of $\mathcal{F}$, and thus $A = B$. Therefore, $x \sim z$, and the relation is transitive. □

7.5. **Functions and Quotient Spaces.** Defining functions when there are quotient spaces involved is not always the easiest thing to do. We do, however, have one natural function.

**Definition.** Let $X$ be a set, and $R$ an equivalence relation on $X$. Then there is a *natural projection* $p : X \to X/R$ given by $p(x) = [x]$.

One question would be whether we could go the other way. Namely, is there a function $X/R \to X$ mapping $[x] \mapsto x$? We could certainly try to define a function this way, but we run into trouble. The problem is that this may not be a function. For example, consider the map $f : \mathbb{Z}/5 \to \mathbb{Z}$ mapping $[x] \to x$. This says $f([2]) = 2$, and $f([3]) = 3$. But it also means $f([7]) = 7$, which is bad since $[2] = [7]$. So $f([2]) = 2$ and 7, which contradicts the definition of a function.

The moral of the story is that when defining functions with some quotient space as the domain, we need to check it is *well-defined*, which means that it does not depend on the choice of representative. Let us look at an example.

**Problem.** Show that $f : \mathbb{Z}/5 \to \mathbb{Z}/5$ given by $f([x]) = [2x]$ is a well-defined function.

*Solution.* What we want to show is that if $[x] = [y]$, then $f([x]) = f([y])$. Well, if $[x] = [y]$, then $x \sim y$, so $5|(x - y)$. We want to show $[2x] = [2y]$, i.e. that $5|(2x - 2y)$. But notice that as $5|(x - y)$, there is an integer $k$ such that $5k = x - y$. Then $5(2k) = 2x - 2y$, meaning $5|(2x - 2y)$, as desired. Thus, this function is well-defined.

*Remark.* The map $f : \mathbb{Z}/5 \to \mathbb{Z}$ given by $f([x]) = 2x$ is NOT well-defined. Notice $f([1]) = 2$, but if we change the representative to 6, then $f([6]) = 12 \neq 2$, which means this map depends on the choice of representative. Therefore, it is not a function.

7.6. **Problems.** Let us look at some problems.

**Problem.** Let $S$ and $T$ be sets, and let $f : S \to T$ be a function. Consider the relation $R$ on $S$ given by $s_1 \sim s_2$ if $f(s_1) = f(s_2)$. This is an equivalence relation. Define a map $g : S/\sim \to T$ given by $g([s]) = f(s)$.

(a) Show $g$ is a well-defined function.
(b) Show that $g$ is injective.

*Solution.* (a) Suppose $[s_1] = [s_2]$. We want to show $g([s_1]) = g([s_2])$. If $[s_1] = [s_2]$, then $s_1 \sim s_2$, meaning $f(s_1) = f(s_2)$. But then

$$g([s_1]) = f(s_1) = f(s_2) = g([s_2]),$$

meaning $g$ is well-defined.

(b) Suppose $g([s_1]) = g([s_2])$. We want to show $[s_1] = [s_2]$. By definition of $g$, we know $f(s_1) = f(s_2)$, so $s_1 \sim s_2$. This implies $[s_1] = [s_2]$, as desired.

**Problem.** Consider the map $g : \mathbb{Z}/n \times \mathbb{Z}/n \to \mathbb{Z}/n$ given by $g([x], [y]) = [xy]$. Show that $g$ is well-defined.

*Solution.* Suppose $[x] = [x']$ and $[y] = [y']$. We want to show $g([x], [y]) = g([x'], [y'])$. Since $x \equiv x' \mod n$, there exists a $k$ such that $x - x' = nk$. Similarly, there exists an $l$ such that $y - y' = nl$. We want to show $n|(xy - x'y')$. It is not immediately clear how to do this, but we observe we can write

$$xy = (x' + nk)(y' + nl) = x'y' + n(lx' + ky') + n^2kl.$$

Thus, $xy - x'y' = n(lx' + ky' + nkl)$, which is what we needed. Therefore $g$ is well-defined.

*Remark.* This gives us a way of multiplying modulo $n$: $[x] \cdot [y] = [xy]$, and this is well-defined by the above problem. In the HW, you will show addition is well-defined as well. Another way of writing what was shown above is that if $x \equiv x' \mod n$ and $y \equiv y' \mod n$, then $xy \equiv x'y' \mod n$.

To show you the power of the above remark, consider the following problem:

**Problem.** Compute the remainder when $8^7$ is divided by 3.

*Solution.* There are several approaches to this problem given what we've done. The first is to say that we are looking for $8^7 \mod 3$, i.e. $[8^7]$ in quotient space $\mathbb{Z}/3$. By the above, this is just $[8]^7$. We know $[8] = [2]$, so this is just $[2]^7$. We could just list out the powers of $[2]$ in $\mathbb{Z}/3$: $[2]^2 = [4] = [1]$, so $[2]^3 = [2]$, and $[2]^4 = [1]$, and so on. We end up with $[2]^7 = [2]$, and thus the remainder is 2.

We could also have noticed that $[8] = [-1]$, so $[8]^7 = [-1]^7 = [(-1)^7] = [-1] = [2]$, so the remainder is 2.

**Problem.** Is the well-defined function $f : \mathbb{Z}/6 \to \mathbb{Z}/6$ given by $f([x]) = [2x + 3]$ injective?

*Solution.* Since $\mathbb{Z}/6$ only has 6 elements, we should be able to just check it:

$$f([0]) = [3], \quad f([1]) = [5], \quad f([2]) = [1], \quad f([3]) = [3].$$

At this point we can stop since $f([0]) = f([3])$, which means $f$ is not injective.

## 8. Counting

This section is the start of our next major topic. There will be many subsections here, and since this is an incredibly important chapter, we will be looking at lots of examples.

### 8.1. Basic Counting Principles.

At the end of the day, most of counting boils down to "choices." Keeping track of what choices you have the make will make counting problems that much easier. Consider the following example:

**Problem.** Dinner at a restaurant consists of an appetizer, main course, and dessert. There are five choices of appetizer, six main courses, and four possible desserts. How many possible dinners are there?

You may have encountered this sort of problem in school already. If you recall, the answer boils down to multiplication: $5 \cdot 6 \cdot 4 = 120$ possible dinners. This example illustrates the **multiplication principle**.

**Principle** (Multiplication Principle)**.** Let $A_1, A_2, \ldots, A_k$ be ordered events. Suppose there are $n_1$ possible outcomes for event $A_1$, $n_2$ possible outcomes for event $A_2$, and so on. Then the total number of possible outcomes for: $A_1$ then $A_2$ then $A_3$, and so on, is $n_1n_2n_3 \cdots n_k$.

Intuitively, one can imagine that the number of outcomes is in bijection with $A_1 \times A_2 \times \cdots \times A_k$, which we know has size $|A_1| \cdot |A_2| \cdot \cdots \cdot |A_k|$. Regardless of how you want to think about it, this principle is arguably the most useful counting principle.

**Example.** Suppose we are in the situation in the earlier problem. Suppose dessert and appetizer are optional. How many dinners are possible? Well now, instead of five choices of appetizer, we actually have six: the five choices, plus the option of getting nothing, which is a sixth choice. Similarly, we now have five choices for dessert (the four plus the option of getting nothing). Thus, there are $6 \cdot 6 \cdot 5$ possible dinners.

Let us look at some examples you may not have seen before.

**Problem.** How many strings of length 4 can be made from the letters $A$, $B$, $C$, $D$, $E$, $F$ if:

(a) Repetition is allowed.
(b) Repetition is not allowed.
(c) Repetition is not allowed, and the first letter must be $B$.
(d) Repetition is not allowed, and the first letter cannot be a $B$.

*Solution.* (a) If repetition is allowed, then there are 6 choices for each letter in the string, thus there are $6^4$ total strings.

(b) We now have to be more careful. There are 6 choices for the first letter. After we choose this letter, there are only 5 choices remaining for the second one. Then there will be only 4 choices for the third, and then 3 for the last one. Thus, there are $6 \cdot 5 \cdot 4 \cdot 3 = 360$ total strings.

(c) If $B$ must be the first letter, then there is only one choice for the first letter. After this is chosen, there are 5 choices for the second letter as repetition is not allowed, then 4 for the third letter, and 3 for the last. Therefore, there are $1 \cdot 5 \cdot 4 \cdot 3 = 60$ total strings.

(d) Since the first letter cannot be $B$, there are only 5 options for that one. For the second letter, it cannot be that first letter, so there are 5 options for this one as well (since we are allowing $B$ now). Then there are 4 for the third letter and 3 for the last. Thus, there are $5 \cdot 5 \cdot 4 \cdot 3 = 300$ total options.

*Remark.* We could also have approached (d) above in the following way: to count the number of strings that don't start with $B$, we could take the total number of strings with repetition not allowed and subtract off the number that do start with $B$. This would give $360 - 60 = 300$, which is exactly the number we found in (d). The general principle is that if $A \subset B$, then $|B \backslash A| = |B| - |A|$.

**Problem.** Suppose $X$ and $Y$ are sets with $|X| = n$ and $|Y| = m$. How many function $f : X \to Y$ are there?

*Solution.* To describe a function, we must assign a unique $y \in X$ to each $x \in X$. Thus, we have $n$ choices to make (one for each $x \in X$), and for each choice, we have $m = |Y|$ options to choose from. Therefore, there are $n^m$ total functions.

**Problem.** Let $X$ be a set with $|X| = n$. Use the multiplication principle to prove $|\mathcal{P}(X)| = 2^n$.

*Solution.* We need to figure out how to reduce the problem of counting subsets to choices. If we think about it some, we notice that to determine a subset, we just need to decide whether a given element is in the subset or not. If $X = \{x_1, \ldots, x_n\}$, then to determine a subset $S$, we need to decide whether $x_i \in S$ or $x_i \notin S$, for all $i$. This means we have $n$ choices, each with two options. Thus, by the multiplication principle, $|\mathcal{P}(X)| = 2^n$.

**Problem.** Let $X$ be a set with $|X| = n$. How many ordered pairs of subsets $(A, B)$ are there which satisfy $A \subseteq B \subseteq X$?

*Solution.* This problem seems like it might be difficult. Let's look at an example first: if $X = \{1\}$, then $\mathcal{P}(A) = \{\emptyset, \{1\}\}$, and there are three such pairs: $(\emptyset, \emptyset)$, $(\emptyset, \{1\})$, $(\{1\}, \{1\})$. Let's try and reason as we did in the previous problem. For a good ordered pair $(A, B)$ and a given $x \in X$, there are three options: either $x \in A$, $x \in B \backslash A$, or $x \in X \backslash B$. So for each $x \in X$, we have three options. As there are $n$ elements of $X$, we find the total number of ordered pairs is $3^n$.

Looking at these problems, there is a strong relationship between the multiplication principle and conjunctions. Indeed, the principle says that if we have choice 1 AND choice 2, then the number of total options is the product of the number for each choice. But what is the corresponding principle for disjunctions? This is the addition principle.

**Principle.** Let $X_1, X_2, \ldots, X_k$ be sets with $|X_i| = n_i$ and $X_i \cap X_j = \emptyset$ for $i \neq j$. Then

$$|X_1 \cup X_2 \cup \cdots \cup X_k| = n_1 + n_2 + \ldots + n_k.$$

Again, this should be intuitively clear. If we have a bunch of disjoint sets (i.e. with no overlap), then the number of elements in the union should be just the sum of the number of elements of each individual set.

**Problem.** Suppose we have five computer science books, three math books, and two art books. How many ways are there of selecting two books, each from different subjects?

*Solution.* Since the books must come from different subjects, there are three possibilities: we choose a math and computer science book, a math and art book, or a computer science and art book. And each of these sets are disjoint. How many ways are there of selecting a math and computer science book. Here we see conjunction and choices, so we will need the multiplication principle. There are 3 choices of math book and 5 of computer science, so there are $3 \cdot 5 = 15$ ways of selecting a math book and a computer science book. Similarly, there are 6 ways of choosing a math and art book, and 10 ways of selecting a computer science and art book. By the addition principle, there are $15 + 10 + 6 = 31$ ways of selecting two books, each from different subjects.

**Problem.** How many strings of length 4 can be made from the letters $A$, $B$, $C$, $D$, $E$, $F$ if repetition is not allowed, and the first letter must be $B$ or a $C$.

*Solution.* Let $X_1$ be the set of strings which start with $B$ and $X_2$ the set of strings which start with $C$. They are disjoint (you can't start with both $B$ and $C$), so we can use the addition principle to find $|X_1 \cup X_2|$. We computed $|X_1|$ in the first problem of the multiplication principle, and we got 60. Similarly, there are 60 strings which start with $C$. Thus, by the addition principle, there are $60 + 60 = 120$ total such strings.

What happens if the sets are not disjoint? Then we have the method of inclusion/exclusion:

**Theorem 8.1** (Inclusion/Exclusion)**.** If $X$ and $Y$ are sets, then $|X \cup Y| = |X| + |Y| - |X \cap Y|$.

*Proof.* We get this from the addition principle. We are going to break up $X$, $Y$, and $X \cup Y$ into disjoint sets and use the addition principle for each.

*Claim.* $X = (X - Y) \cup (X \cap Y)$, and these sets are disjoint: The inclusion $(X - Y) \cup (X \cap Y) \subseteq X$ should be clear. For the reverse inclusion, take $x \in X$. If $x \in Y$, then $x \in X \cap Y$. If $x \notin Y$, then $x \in X \backslash Y$. In either case, $x \in (X\backslash Y) \cup (X \cap Y)$, and we have an equality of sets. They are disjoint because if $x \in (X\backslash Y) \cap (X \cap Y)$, then $x \in Y$ and $x \notin Y$, a contradiction.

By the addition principle, $|X| = |X\backslash Y| + |X \cap Y|$.

*Claim.* $Y = (Y - X) \cup (X \cap Y)$, and these sets are disjoint: It is the same proof as the first claim.

Thus, $|Y| = |Y\backslash X| + |X \cap Y|$. You showed on your homework that $X \cup Y = (X\backslash Y) \cup (Y\backslash X) \cup (X \cap Y)$. It is easy to prove these sets are disjoint as well. Thus,

$$|X \cup Y| = |X\backslash Y| + |Y\backslash X| + |X \cap Y|.$$

Using our first two equalities, we get

$$|X \cup Y| = |X| - |X \cap Y| + |Y| - |Y \cap X| + |X \cap Y| = |X| + |Y| - |X \cap Y|.$$

□

Let's look at how to apply this.

**Problem.** A committee of three people (president, vice president, and treasurer) is to be made up from Alice, Bob, Cindy, Dave, Ed, Fred. How many different committees are possible if either Alice or Dave (or both) needs to be on the committee.

*Solution.* Let $X_A$ be the set of committees with Alice as a member, and let $X_D$ be the set of committees with Dave as a member. We want $|X_A \cup X_D|$. Here, $X_A \cap X_D \neq \emptyset$, so we need to use inclusion/exclusion. For $|X_A|$, we can use the multiplication principle. There are actually three choices to make. First, we need to decide which position Alice will fill. There are three choices for this. Next, we have to choose people for the other two positions. There are 5 choices for the first and 4 for the last. Thus, $|X_A| = 3 \cdot 5 \cdot 4 = 60$. Similarly, $|X_D| = 60$.

For $X_A \cap X_D$, we have three choices. First, we choose the position Alice will fill. Next, we choose the position Dave will fill. Lastly, we choose the person for the final position. Thus, there are $3 \cdot 2 \cdot 4 = 24$ ways of doing this. Thus,

$$|X_A \cup X_D| = 60 + 60 - 24 = 96.$$

**Problem.** How many numbers between 1 and 100 are divisible by 2 or 5 (or both)?

**Problem.** Let $X_2$ be the set of numbers divisible by 2 and $X_5$ the set of numbers divisible by 5. We want $|X_2 \cup X_5|$. Well $|X_2| = 100/2 = 50$ and $|X_5| = 100/5 = 20$. Lastly, $X_2 \cap X_5$ is the set of numbers divisible by both 2 and 5, which are just multiples of 10. There are 10 such numbers. Thus,

$$|X_2 \cup X_5| = 50 + 20 - 10 = 60.$$

We'll end with one more problem.

**Problem.** How many 8 bit strings neither start with 11 nor end with 000?

*Solution.* Let $A$ be the set of 8 bit strings which start with 11, and $B$ the set of 8 bit strings which end with 000. We want the set of 8 bit strings $x$ such that $x \notin A$ and $x \notin B$. Letting $U$ denote all 8 bit strings, this is the same as saying we want

$$x \in U \backslash (A \cup B).$$

To count the number of such strings, we can use inclusion/exclusion:

$$|U \backslash (A \cup B)| = |U| - |A \cup B| = |U| - (|A| + |B| - |A \cap B|).$$

We just need to compute the terms. First, $|U| = 2^8$ since there are 8 choices, each with 2 options. For $|A|$, we only have 6 choices since the first two bits are determined for us already. Thus, $|A| = 2^6$. Similarly, in computing $|B|$, we notice we only have 5 choices since the last three bits must be 0. Thus, $|B| = 2^5$. Lastly, we need $|A \cap B|$. But to be in $A$ and $B$ means that we need to start with 11 and end with 000, so the only choices are the three digits between them. Thus, $|A \cap B| = 2^3$. Putting this all together gives:

$$|U \backslash (A \cup B)| = 2^8 - (2^6 + 2^5 - 2^3) = 256 - (64 + 32 - 8) = 168.$$

## 9. Permutations and Combinations

9.1. **Permutations.** We will now look at permutations and combinations. Permutations are a very specific application of the multiplication principle. To give the general setup, let $X$ be a set with $|X| = n$. Suppose we want to pick and order $r$ of the $n$ elements of $X$. For example, if $X = \{x_1, \ldots, x_n\}$, then if we want to pick and order 3 of the elements, we could have $\{x_1, x_2, x_3\}$, or $\{x_2, x_1, x_3\}$, or $\{x_3, x_2, x_1\}$, and so on. Let's make this a definition.

**Definition.** Given $n$ distinct elements $x_1, \ldots, x_n$, an *$r$-permutation* of these elements is an ordering of an $r$-element subset of $\{x_1, \ldots, x_n\}$. The number of $r$-permutations is denoted $P(n, r)$.

*Remark.* We only care about $r \leq n$, since it doesn't make sense to choose more than $n$ distinct elements from a set of $n$ elements.

What should $P(n, r)$ be? It is an application of the multiplication principle. We want to select and order $r$ things from the $n$ objects. Well, there are $n$ choices for the first element. Then there are $(n - 1)$ elements for the second, since we cannot choose the first element again. Then there are $(n - 2)$ choices for the third element, and so on. Thus, we have:

**Theorem 9.1.** $P(n, r) = n(n-1)(n-2)\cdots(n-(r-1)) = \frac{n!}{(n-r)!}$.

*Proof.* The first equality is the multiplication principle, and the second is straightforward from the definition of factorials. □

**Example.** How many 2-permutations are there of the elements $a, b, c$? This should be $P(3, 2) = \frac{3!}{(3-2)!} = 6$. We could get this from the multiplication principle by saying there are 3 choices for the first slot and 2 for the second slot, giving $3 \cdot 2 = 6$ total 2-permutations of the three elements. We can even list them:

$$ab, ba, ac, ca, bc, cb.$$

**Problem.** A committee of three people (president, vice president, treasurer) is to be chosen from a group of 10 people. How many committees are possible?

*Solution.* We have ten people, and we want an ordered set of three people (ordered because the positions are different). Thus, the total number of committees is $P(10, 3) = \frac{10!}{7!} = 720$. Notice, again, that this is just the multiplication principle at work.

**Definition.** A *permutation* of a set $X = \{x_1, \ldots, x_n\}$ is an $n$-permutation of the elements $x_1, \ldots, x_n$.

How many permutations are there? There $P(n,n)=\frac{n!}{(n-n)!}=n!$ such permutations. We will need this in a second.

9.2. **Combinations.**

**Definition.** Given a set $X=\{x_1,\dots,x_n\}$, an *r-combination* (where $r\leq n$) is an unordered selection of $r$ elements of $X$. The number of $r$-combinations is denoted $C(n,r)$.

The only difference between permutations and combinations is the importance of order. Permutations are ordered, and combinations are unordered. This helps us figure out what $C(n,r)$ should be.

**Example.** If $X=\{a,b,c\}$, then there are three 2-combinations: $ab,ac,bc$. Notice the difference with permutations. Here, $ab$ and $ba$ is the same unordered set of elements, but they would be different elements when doing permutations.

**Definition.** Let $n\geq 0$ and $0\leq r\leq n$. We let $\binom{n}{r}$ denote the quantity

$$\binom{n}{r}=\frac{n!}{r!(n-r)!}.$$

**Proposition.** $C(n,r)=\binom{n}{r}$.

*Proof.* Let $X=\{x_1,\dots,x_n\}$. Given an $r$-combination of $X$, there are $r!$ rearrangements of these $r$ elements. This produces all $r$ permutations involving these chosen elements. Thus,

$$P(n,r)=C(n,r)\cdot r!,$$

which means

$$C(n,r)=\frac{P(n,r)}{r!}=\binom{n}{r}.$$

□

**Corollary.** The quantity $\binom{n}{r}$ is an integer.

**Problem.** How many length 8 bit strings are there which contain exactly 5 zeroes?

*Solution.* Again, let's look at choices. Before, we went bit by bit and looked at choices (for example, when showing there are $2^8$ length 8 bit strings). But here, we need to decide where the 0's are going to be. There are 8 slots, and we have to choose 5 for the zeroes. It doesn't matter which order we choose the slots in. Thus, there are $\binom{8}{5}=\frac{8!}{3!5!}$ length 8 bit strings which contain exactly 5 zeroes.

**Problem.** How many strings can be formed using all the letters of BOOKKEEPER.

*Solution.* To use all the letters means it needs to be a string of length 10. There need to be 3 E's, 2 K's, 2 O's, and 1 each of B, P, and R. There are 10 total slots, and we have to choose 3 of them for the E's: there are $\binom{10}{3}$ such choices. Then, out of the 7 remaining slots, we must choose 2 for the K's: there are $\binom{7}{2}$ ways to do this. Then there are $\binom{5}{2}$ ways of choosing the slots for the O's, then $\binom{3}{1}$ for the B, $\binom{2}{1}$ for the P, and $\binom{1}{1}$ for the R. Thus, there are

$$\binom{10}{3}\binom{7}{2}\binom{5}{2}\binom{3}{1}\binom{2}{1}\binom{1}{1}=\frac{10!}{3!2!2!1!1!1!}$$

rearrangements of the letters at BOOKKEEPER.

*Remark.* The other way to look at this solution is the following: if all the letters were different, then there would be 10! rearrangements of the letters. But they aren't all different. If we had EEEKKOOBPR, then any rearrangement of the E's would produce the same string, and any rearrangement of the K's would produce the same string, and so on. So we have overcounted. The quantity in the denominator is the amount we've overcounted by: the number of rearrangements of each of the individual characters.

**Problem.** A grid path from $(0,0)$ to $(n,n)$ is a sequence of up moves and right moves to go from the origin to $(n,n)$. How many grid paths are there?

*Solution.* A grid path is a sequence of $U$ and $R$ with $n$ U's and $n$ R's. This is the same as length $2n$ bit strings with $U$ and $R$ instead of 1 and 0. There are $\binom{2n}{n}$ such strings, so there are this many grid paths.

**Problem.** How many non-negative solutions are there to

$$x_1 + x_2 + x_3 = 10?$$

(Non-negative means each $x_i \geq 0$.)

*Solution.* We can list a few: $(0,0,10)$, $(1,3,6)$, $(2,5,3)$, and so on. But we need a systematic way of counting them. We can use a technique called "stars and bars" or "balls and buckets". We can think of this problem as having 10 balls, and we need to distribute them among three distinguishable buckets (distinguishable as $(2,5,3)$ is different than $(5,2,3)$). We can line the 10 balls up in any order since they are the same. We can then take two bars and partition off the balls, forming three groups. The number of balls in each group gives $x_1$, $x_2$, and $x_3$. For example:

$$\cdot\cdot \,|\, \cdot\cdot\cdot\cdot\cdot \,|\, \cdot\cdot\cdot$$

represents $(2,5,3)$. So now we must consider how many arrangements of these 10 balls and 2 bars there are. We can think that there are $10 + 2 = 12$ slots, and we need to choose 2 for the bars. The remaining slots are for the balls. Thus, there are $\binom{12}{2}$ arrangements of the balls and the bars, giving $\binom{12}{2}$ partitions.

9.3. **Problems.** Now let's look at some where we have to decide which counting principle/method to use.

**Problem.** Alice is taking a multiple choice test. There are 10 questions each with 4 possible answers. In how many ways can Alice answer the questions if:

(a) Alice must answer all the questions?
(b) Alice has the option of leaving questions blank?
(c) Alice may leave no more than 1 questions blank?
(d) Alice must answer at least 8 questions correctly?

*Solution.* (a) If she must answer all the questions, then for each one, there are four choices, meaning there are $4^{10}$ possible ways of answering the questions, by the multiplication principle.

(b) Now there are five choices for each question, namely the four answers or leaving it blank. Thus, there are $5^{10}$ ways.

(c) Now we have to mix parts (a) and (b). There are cases here, either she leaves zero blank or one blank. To leave zero blank means we are in part (a), thus there are $4^{10}$ for this case. For the other case, first we have to pick which one she leaves blank, and there are $\binom{10}{1} = 10$ ways of doing this. After choosing the one she will leave blank, we just need her to decide how to answer the remaining questions, giving $4^9$ possibilities. Thus, by the addition principle, there are

$$4^{10} + 10 \cdot 4^9$$

ways of answering the questions.

(d) First, we must decide which ones she will answer correctly. She either answers 8, 9 or 10 correctly. In the first case, there are $\binom{10}{8}$ ways of choosing the ones she answers correctly. There are $3^2$ ways of choosing answers for the two she gets wrong. Thus, there are $\binom{10}{8}3^2$ ways of doing 8 correctly. Similarly, there are $\binom{10}{9}3^1$ ways of doing 9 correctly, and $\binom{10}{10}$ ways of doing them all correctly. Thus, there are

$$\binom{10}{8}3^2 + \binom{10}{9}3 + \binom{10}{10} = 45 \cdot 9 + 30 + 1 = 436$$

total ways of answering the questions.

**Problem.** Consider the set $X = \{1, 2, 3, \ldots, 8\}$.

(a) How many subsets contain the element 5?
(b) How many subsets of size four are there which contain the element 5?
(c) How many subsets of size four are there such that either the minimum of the elements is 5 or the maximum of the four elements is 5?
(d) How many subsets of size four are there such that the product of three of the elements in the subset is 6.

*Solution.* (a) We just have to decide whether each of the other six elements are going to be in our subset our not. There are two options for each element (in the subset or not in the subset), giving $2^7$ possibilities.

(b) We just have to choose the other three elements in the subset. There are $\binom{7}{3}$ ways of doing this.

(c) If the maximum is 5, then we have to fill the subset with elements less than 5. There are $\binom{4}{3} = 4$ ways of doing this. If the minimum is less than 5, then we have to fill the subset with elements greater than 5. There are $\binom{3}{3}$ ways of doing this. Finally, there is no way of having both the minimum and maximum be 5, so these two events are disjoint. By the addition principle, the total number is $\binom{4}{3} + \binom{3}{3} = 5$.

(d) For the product of three of the entries to be 6, we must have $1, 2, 3$ be in the subset. There is only a choice for the last element, and there are $\binom{5}{1} = 5$ ways of picking it.

**Problem.** Find the number of grid paths from $(0,0)$ to $(6,6)$ which:

(a) Go through $(2,2)$.
(b) Go through $(4,4)$.
(c) Go through $(2,2)$ or $(4,4)$.
(d) Do not go through $(2,2)$ and $(4,4)$.

*Solution.* (a) To go through $(2,2)$, we first count paths to $(2,2)$ and then paths from $(2,2)$ to $(6,6)$. In the first case, we have four moves to make (two up and two right). There are $\binom{4}{2}$ ways of picking when the rights come, giving $\binom{4}{2} = 6$ paths to $(2,2)$. Then, we have to pick 4 rights and 4 ups to get to $(6,6)$. There are $\binom{8}{4}$ ways of picking when the rights should be made. By the multiplication principle, there are $\binom{4}{2}\binom{8}{4}$ total paths through $(2,2)$.

(b) The work is similar to (a), and we get $\binom{8}{4}\binom{4}{2}$.

(c) Let $A$ be the set of paths which go through $(2,2)$, and $B$ the set which go through $(4,4)$. We want $|A \cup B|$. By inclusion/exclusion, it is $|A| + |B| - |A \cap B|$. We found the first two terms in the last two parts. To find $|A \cap B|$, we want the number of paths which go through $(2,2)$ and $(4,4)$. By the multiplication principle and similar logic as in (a), we get $\binom{4}{2}\binom{4}{2}\binom{4}{2}$ ways of doing this. Thus, there are

$$\binom{4}{2}\binom{8}{4} + \binom{8}{4}\binom{4}{2} - \binom{4}{2}\binom{4}{2}\binom{4}{2}$$

total paths through $(2,2)$ or $(4,4)$.

(d) This is the total number of paths minus the ones which go through either $(2,2)$ or $(4,4)$, which is

$$\binom{12}{6} - \left(\binom{4}{2}\binom{8}{4} + \binom{8}{4}\binom{4}{2} - \binom{4}{2}\binom{4}{2}\binom{4}{2}\right).$$

**Problem.** Let $X_n = \{1, 2, 3, \dots, n\}$.

(a) How many functions $f : X_n \to X_m$ are there?
(b) If $n \leq m$, how many injections $f : X_n \to X_m$ are there?
(c) If $n = m$, how many bijections $f : X_n \to X_m$ are there?
(d) If $n \geq 2$, how many surjections $f : X_n \to X_2$ are there?

*Solution.* (a) For each of the $n$ elements in $X_n$, there are $m$ options in $X_m$ for where to send them. This gives $m^n$ total functions.

(b) We have $m$ choices for where to send 1, $(m-1)$ choices for where to send 2, and so on. This gives

$$m(m-1)(m-2)\cdots(m-n+1) = P(m,n)$$

total injections.

(c) This is the same thing as a permutation of $X_n$, which gives $n!$ total bijections.

(d) Well, there are $2^n$ total functions to $X_2$. Let $A_1$ be the set of functions such that $R(f) = \{1\}$, and $A_2$ the set of functions such that $R(f) = \{2\}$. We don't want functions which are in $A_1$ or $A_2$, since they are not surjective. There is one function in $A_1$ and one function in $A_2$. There are zero functions in $A_1 \cap A_2$. Thus, by inclusion/exclusion, we have

$$2^n - 2$$

total surjections.

## 10. Binomial Theorem and Binomial Coefficients

As it turns out, the numbers $\binom{n}{r}$ defined in the previous section are pretty useful, and they also have several fascinating properties. We will examine one application of these numbers, namely the binomial theorem. To start out, let's look at $\binom{2}{r}$ for all the possible $r$'s:

$$\binom{2}{0} = 1, \quad \binom{2}{1} = 2, \quad \binom{2}{2} = 1.$$

These numbers probably don't look interesting like this, but observe that when we foil out $(x+y)^2$, we get

$$(x+y)^2 = x^2 + 2xy + y^2 = \binom{2}{0}x^2 + \binom{2}{1}xy + \binom{2}{2}y^2.$$

So the coefficients seem to be the numbers $\binom{2}{r}$ for all the possible $r$. It is also easy to check that

$$(x+y)^3 = x^3 + 3x^2y + 3xy^2 + y^3 = \binom{3}{0}x^3 + \binom{3}{1}x^2y + \binom{3}{2}xy^2 + \binom{3}{3}y^3.$$

This generalizes into the following theorem:

**Theorem** (Binomial Theorem)**.** If $n \geq 1$, then

$$(x+y)^n = x^n + \binom{n}{1}x^{n-1}y + \binom{n}{2}x^{n-2}y^2 + \ldots + \binom{n}{n-1}xy^{n-1} + y^n = \sum_{k=0}^{n}\binom{n}{k}x^{n-k}y^k.$$

We will give one proof now, and you will see the proof by induction on your HW.

*Proof.* Write

$$(x+y)^n = (x+y)(x+y)(x+y)\cdots(x+y).$$

This will be a sum of monomials of degree $n$. What is the coefficient of $x^{n-k}y^k$? One way to think about it is that to get a $x^{n-k}y^k$, we need to $k$ of the terms in the product to contribute a $y$ and the rest to contribute $x$'s. There are $\binom{n}{k}$ ways of picking the terms in the product to give $y$'s, and then the rest are forced to contribute $x$'s. Thus, the coefficient of $x^{n-k}y^k$ is $\binom{n}{k}$. Doing this for all $0 \leq k \leq n$ gives the theorem. □

The numbers $\binom{n}{k}$ are known as *binomial coefficients* because they are the coefficients which come up in the binomial theorem.

**Example.** Suppose we wanted to expand $(x+y)^5$. The theorem says it would be

$$(x+y)^5 = x^5 + \binom{5}{1}x^4y + \binom{5}{2}x^3y^2 + \binom{5}{3}x^2y^3 + \binom{5}{4}xy^4 + y^5 = x^5 + 5x^4y + 10x^3y^2 + 10x^2y^3 + 5xy^4 + y^5.$$

**Problem.** Expand $(a-2b)^4$.

*Solution.* Just put $x = a$ and $y = -2b$ in the binomial theorem:

$$\begin{aligned}(a-2b)^4 &= a^4 + \binom{4}{1}a^3(-2b) + \binom{4}{2}a^2(-2b)^2 + \binom{4}{3}a(-2b)^3 + (-2b)^4 \\ &= a^4 + 4a^3(-2b) + 6a^2(4b^2) + 4a(-8b^3) + 16b^4 \\ &= a^4 - 8a^3b + 24a^2b^2 - 32ab^3 + 16b^4.\end{aligned}$$

This idea can be generalized to expressions of the form $(x+y+z)^n$ or even $(x_1 + x_2 + \ldots + x_k)^n$. Let's look at a problem.

**Problem.** Consider $(x+y+z)^5$. What is the coefficient of $x^2y^2z$ in the expansion? How many terms are there in the expansion (after combining)?

*Solution.* As we had before, we know

$$(x+y+z)^5 = (x+y+z)(x+y+z)(x+y+z)(x+y+z)(x+y+z).$$

How do we get $x^2y^2z$? We need $x$'s from 2 of the terms, $y$ from 2 of the terms, and the remaining term will contribute a $z$. The number of ways of picking the terms contributing $x$'s is $\binom{5}{2}$. Once these are chosen, the number of ways of picking the terms to contribute $y$'s is $\binom{3}{2}$ (i.e. $\binom{5-2}{2}$). The remaining term will contribute the $z$ (or if you want, there is $\binom{1}{1} = 1$ way of picking the term giving the $z$). Thus, the coefficient of $x^2y^2z$ is

$$\binom{5}{2}\binom{3}{2}\binom{1}{1} = \frac{5!}{2!3!} \cdot \frac{3!}{2!1!} = \frac{5!}{2!2!1!}.$$

As for the number of terms, we know that the expansion is a sum of elements of the form $x^iy^jz^k$, where $i + j + k = 5$ and $i, j, k \geq 0$. We did this sort of thing in the counting section: using the balls in the buckets or stars and bars strategy, we get that it is $\binom{7}{2} = \binom{5+3-1}{3-1}$.

We can use the binomial theorem to give us some nice properties about the binomial coefficients, but before we do that, we will examine properties we can get without the theorem.

**Proposition.** For $1 \leq k \leq n$, we have

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

**Example.** This proposition is saying, for example, that

$$\binom{5}{2} + \binom{5}{3} = \binom{6}{3}.$$

This is true, as $10 + 10 = 20$.

*Proof.* The quantity $\binom{n+1}{k}$ counts the number of subsets of $\{1, 2, 3, \ldots, n+1\}$ of size $k$. We can count this another way: count the subsets $S$ such that $1 \in S$, and the subsets such that $1 \notin S$. If $1 \in S$, then we have to choose the remaining $k-1$ elements of the subset from a set of $n$ elements: this is $\binom{n}{k-1}$. If $1 \notin S$, then we have to fill the entire subset from $n-1$ elements: this is $\binom{n}{k}$. By the addition principle, the total number of subsets of size $k$ is

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

Thus,

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

□

*Remark.* One could also provide an algebraic proof that simply uses the definition of the binomial coefficients and combines the fractions.

We can use the proposition and the fact that $\binom{n}{0} = \binom{n}{n} = 1$ to get a nice visual representation of the binomial coefficients, known as Pascal's triangle. The triangle looks like:

$$\binom{0}{0}$$

$$\binom{1}{0} \quad \binom{1}{1}$$

$$\binom{2}{0} \quad \binom{2}{1} \quad \binom{2}{2}$$

$$\binom{3}{0} \quad \binom{3}{1} \quad \binom{3}{2} \quad \binom{3}{3}$$

$$\vdots$$

We know that the outer perimeter of the triangle will be all 1's since $\binom{n}{0} = \binom{n}{n} = 1$. The proposition says that we can get any of the middle terms by adding the two terms above it. For example, $\binom{3}{1} = \binom{2}{0} + \binom{2}{1}$. Filling in the numbers for each of the binomial coefficients gives

$$1$$

$$1 \quad 1$$

$$\begin{array}{ccccccccc} & & 1 & & 2 & & 1 & & \\ & 1 & & 3 & & 3 & & 1 & \\ 1 & & 4 & & 6 & & 4 & & 1 \\ & & & & \vdots & & & & \end{array}$$

and so on. What is the usefulness of the triangle? For one thing, the binomial theorem tells us that the $n$-th row (i.e. the row whose second entry is $n$) gives the coefficients of $(x+y)^n$. So, for example,

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

which is gotten by reading the fourth row of the triangle for the coefficients. (Note: The 1 on top is row 0.) It is also nice in that many of the properties we prove in a little while have nice representations on the triangle.

So what are some of the properties? Well, looking at the triangle, one thing that might stand out is the symmetry of the left half and the right half. For example, $\binom{3}{1} = \binom{3}{2}$, and $\binom{4}{1} = \binom{4}{3}$. We could conjecture that $\binom{n}{k} = \binom{n}{n-k}$. In fact, this is true.

**Proposition.** We have $\binom{n}{k} = \binom{n}{n-k}$ for any $0 \leq k \leq n$.

We will present two proofs: an algebraic proof, and a combinatorial proof.

*Algebraic Proof.* Just working out the definition, we get

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

□

*Combinatorial Proof.* The quantity $\binom{n}{k}$ counts the number of subsets of $\{1,2,3,4\ldots,n\}$ of size $k$. Instead building a subset by picking the $k$ elements to be in the subset, we can just as easily pick the $n-k$ elements to not be in the subset. There are $\binom{n}{n-k}$ such choices, which means the number of subsets of size $k$ is also $\binom{n}{n-k}$. Thus, $\binom{n}{k} = \binom{n}{n-k}$. □

In this instance, the algebraic proof seems quicker, so why not stick to the algebraic proofs all the time? We will see that as the identities get harder, so will the algebra. The ability to count things in two different ways is a nice tool to have at your disposal.

Let's look at another property:

**Proposition.** For any $n \geq 0$, we have

$$\sum_{k=0}^{n} \binom{n}{k} = \binom{n}{0} + \binom{n}{1} + \ldots + \binom{n}{n} = 2^n.$$

*Proof.* Take the binomial theorem and plug in $x = y = 1$. Then

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

or

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

which is exactly what we wanted to show. □

*Remark.* In terms of Pascal's triangle, this says that the sum of the entries in the $n$-th row is $2^n$.

The algebraic proof here is nice because it gives us another proof of a theorem we have seen twice now:

**Corollary.** If $X$ is a set with $|X| = n$, then $|\mathcal{P}(X)| = 2^n$.

*Proof.* To count the number of subsets, we can count the number of subsets of size $k$, for $0 \leq k \leq n$, and add them up by the addition principle. The number of subsets of size $k$ is $\binom{n}{k}$, so

$$|\mathcal{P}(X)| = \binom{n}{0} + \binom{n}{1} + \ldots + \binom{n}{n} = 2^n.$$

□

Let's look at two more properties.

**Proposition.** For any $m \geq 0$ and $n \geq 0$, we have

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

*Proof.* Fix $n \in \mathbb{N}_0$. We prove this by induction on $m$. The base case is $m = 0$, in which case the equation reads

$$\binom{n}{n} = \binom{n+1}{n+1},$$

which is true as both are 1.

For the inductive step, assume the statement is true for $m = k$. We want to show it is true for $m = k+1$. Well,

$$\begin{aligned} \sum_{j=0}^{k+1} \binom{n+j}{n} &= \sum_{j=0}^{k} \binom{n+j}{n} + \binom{n+k+1}{n} \\ &= \binom{n+k+1}{n+1} + \binom{n+k+1}{n} \\ &= \binom{n+k+2}{n+1}, \end{aligned}$$

which is what we wanted to prove. Therefore, the inductive step holds. □

**Proposition.** For $m, n, r \in \mathbb{N}_0$ and $r \leq m$ and $n$, we have

$$\sum_{k=0}^{r} \binom{n}{k}\binom{m}{r-k} = \binom{n+m}{r}.$$

**Example.** As an example of the proposition, let $m = 3$, $n = 5$, and $r = 3$. Then

$$\binom{8}{3} = \binom{5}{0}\binom{3}{3} + \binom{5}{1}\binom{3}{2} + \binom{5}{2}\binom{3}{1} + \binom{5}{3}\binom{3}{0}.$$

*Proof.* We give a combinatorial proof. Suppose we had a group of $m$ women and $n$ men, and we want a subset of $r$ people. Well, we could either have 0 men and $r$ women, 1 man and $r-1$ women, 2 men and $r-2$ women, and so on. If we have $k$ men, we have $r-k$ women. The number of ways of choosing this is $\binom{n}{k}\binom{m}{r-k}$ by the multiplication principle, and we add this up for every possible $k$. This is the left side of what we want to prove.

For the other side, notice that there are $\binom{n+m}{r}$ total subsets of $r$ people. Thus, as both sides count the same quantity, we have equality. □

**Corollary.** For any $n \in \mathbb{N}_0$, we have

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

*Proof.* Take $m = n = r$ in the proposition. □

*Remark.* You will see a different proof of this proposition on the HW.

## 11. Recurrence Relations

11.1. **Finding Recurrence Relations.** Our next section is recurrence relations. For those of you taking computer science, this will probably be a familiar topic as recursion is a programming technique typically taught in computer science classes. We have seen recurrence a few times already in the induction section in the beginning of the course (e.g. Fibonacci numbers). In this section, we'll define them properly and look at techniques to solve them.

**Definition.** (1) A *recurrence relation* is a sequence $a_1, a_2, \ldots$ such that $a_n$ is defined in terms of the previous terms $a_0, \ldots, a_{n-1}$. That is

$$a_n = f(a_0, a_1, \ldots, a_{n-1})$$

for some function $f$.

(2) A recurrence relation has *initial conditions* if the values of finitely many initial terms are given.

**Example.** Recall that the Fibonacci numbers were defined by $F_1 = F_2 = 1$ and $F_n = F_{n-1} + F_{n-2}$ for $n \geq 3$. This is a recurrence relation, and the initial conditions are the $F_1 = F_2 = 1$. This produces the sequence

$$1, 1, 2, 3, 5, 8, 13, 21, 34, \ldots.$$

A natural question would be: is there a formula for $F_n$ which depends only on $n$? We will examine this later.

Let's look at some some problems where we have to come up with our own recurrence relations.

**Problem.** How many ways are there of tiling an $2 \times n$ board with dominoes of size $2 \times 1$ and $1 \times 2$? Give the recurrence relation and the initial conditions.

*Solution.* Usually, the best way is to figure out how the solution must start. If $n = 1$, there is only one way to do it. If $n = 2$, there are two ways of doing it (namely, with two $1 \times 2$ pieces or two $2 \times 1$ pieces). So let $n \geq 3$, and let $a_n$ be the number of tilings of a $2 \times n$ board. If we were to work left to right, we could either start by placing a $2 \times 1$ piece down, in which case we have to fill the remaining $2 \times (n-1)$ board, or we could place two $1 \times 2$ pieces horizontally, in which case we have the fill the remaining $2 \times (n-2)$ board. In the first case, there are $a_{n-1}$ such tilings. In the second case, there are $a_{n-2}$ tilings. Thus, by the addition principle,

$$a_n = a_{n-1} + a_{n-2}, \quad n \geq 3.$$

From the beginning, we know $a_1 = 1$ and $a_2 = 2$.

*Remark.* Notice that this recurrence relation looks like the Fibonacci relation, and indeed it is. However, our initial conditions are different: it is the Fibonacci numbers starting with $F_2$ as opposed to starting with $F_1$. But notice that the initial conditions change the sequence.

**Problem.** Suppose you invest 1000 dollars with 7% interest compounded annually. Let $A_n$ be the amount of money after $n$ years. Find a recurrence relation and give the initial conditions.

*Solution.* We want $A_n$: the amount of money after $n$ years. It should be the amount of money after $n-1$ years, plus interest. Thus,

$$A_n = A_{n-1} + .07A_{n-1} = 1.07A_{n-1}.$$

We need only one initial condition since the relation only needs one previous term: $A_0 = 1000$. Thus, $A_n = 1.07A_{n-1}$ for $n \geq 1$ and $A_0 = 1000$.

**Problem.** Let $S_n$ be the number of subsets of $\{1, 2, 3, \ldots, n\}$. Give a recurrence relation and initial conditions for $S_n$.

*Solution.* We want to try to get $S_n$ in terms of previous terms. This means we want to somehow reduce the number of elements in our set. Well, a subset with either contain $n$ or not contain $n$. In the first case, our subset is just a subset of $\{1, 2, 3, \ldots, n-1\}$. In the second case, the subset is $\{n\} \cup S'$, where $S'$ is a subset of $\{1, 2, 3, \ldots, n-1\}$. Thus, by the addition principle, the number of subsets is $S_n = S_{n-1} + S_{n-1} = 2S_{n-1}$. We need one initial condition, and this relation will work for $n \geq 1$. Our initial condition is $S_0 = 1$.

These should be relatively straightforward examples. Now, let's look at some problems which require some more thought.

**Problem.** Let $S_n$ be the number of $n$-bit strings that do not contain the pattern 111. Give a recurrence relation and initial conditions.

*Solution.* Again, the best way is to examine how such a string must start. There are three ways it could start: with a 0, with a 10, or with a 110 (that is, it can start with either no 1's, one 1 or two 1's). These cases are disjoint, so we can use the addition principle for each. If the string starts with a 0, then we just have to give the remaining $(n-1)$-bits while ensuring that 111 does not come up. There are $S_{n-1}$ such possibilities. In the second case, there are $S_{n-2}$ ways of filling the rest of the string. Finally, in the last case, there are $S_{n-3}$ possibilities. Thus,

$$S_n = S_{n-1} + S_{n-2} + S_{n-3}.$$

What are the initial conditions? We need to consider three initial conditions since our relation needs three previous terms, and then our relation will work for $n \geq 4$. We find $S_1 = 2$ (0 or 1), $S_2 = 4$ $(00, 01, 10, 11)$ , and $S_3 = 7$.

**Problem.** The Tower of Hanoi is a puzzle which works like this: there are three pegs mounted on a board, and $n$ disks of different sizes with holes in their centers so that they can be placed around the peg. All the disks start out around peg 1 with the smallest disk on top and largest on the bottom (so ordered by increasing sizes). The goal is to move all the disks to peg 2 in such a way that once a disk is placed on a peg, only a disk of smaller diameter may be placed on top of it. Let $H_n$ be the minimum number of moves needed to do this for $n$ disks. Find a recurrence relation and give initial conditions.

*Solution.* Clearly, $H_1 = 1$. For $n$ disks, the following must happen: the first $n-1$ disks must be moved to peg 3, then the largest disk is moved from peg 1 to peg 2, and then all the smaller disks must be moved from peg 3 to peg 2. The minimum number of moves for the first step is $H_{n-1}$, then we have one move for step 2, and finally $H_{n-1}$ moves for step 3. Thus,

$$H_n = H_{n-1} + 1 + H_{n-1} = 2H_{n-1} + 1.$$

This works for $n \geq 2$, and we know the initial condition is $H_1 = 1$.

**Problem.** How many grid paths from $(0,0)$ to $(n,n)$ never go below the line $y = x$? Give a recurrence relation and initial conditions.

*Solution.* Let $C_n$ be the total number of such paths. If $n = 0$, there is only one path, so $C_0 = 1$, and $C_1 = 1$. One easily checks that $C_2 = 2$. Let $p$ be a path which does not go below the diagonal. Let $(i+1, i+1)$ be the first time $p$ touches the diagonal (other than the starting point). Therefore, $i \leq n-1$. The path $p$ must start with an up move and then meet the point $(i, i+1)$ (since it can't get to $(i+1, i+1)$ with an up move). Thus, $p$ consists of an move followed by the path $p_1$ to $(i, i+1)$ followed by a right move followed by a path $p_2$ from $(i+1, i+1)$ to $(n, n)$ which does not go below the diagonal. There are $C_i$ paths from $(0, 1)$ to $(i, i+1)$, and $C_{n-i-1}$ paths for $p_2$. By the multiplication principle, the number of such paths $p$ is $C_i C_{n-i-1}$. We can do this for any $0 \leq i \leq n-1$, so we have

$$C_n = C_0 C_{n-1} + C_1 C_{n-2} + \ldots + C_{n-1} C_0.$$

This technically works for any $n \geq 2$:

$$C_2 = C_0 C_1 + C_1 C_0 = 1 + 1 = 2,$$

$$C_3 = C_0 C_2 + C_1 C_1 + C_2 C_0 = 2 + 1 + 2 = 5.$$

Thus, this gives our recurrence relation and the initial conditions are $C_0 = 1$ and $C_1 = 1$.

*Remark.* The numbers $C_n$ in the above problem are called *Catalan numbers.*

11.2. **Solving Recurrence Relations.** Now that we have seen how to make recurrence relations, we want to be able to find formulas which satisfy them. We will examine two techniques here:

(1) Iteration to find a formula,
(2) Solving linear homogeneous equations (we will define these later).

We examine technique (1) by doing an example.

**Example.** If $a_n = a_{n-1} + 5$ and $a_1 = 2$, let's find a formula for $a_n$. Well, we know $a_n = a_{n-1} + 5$. This means $a_{n-1} = a_{n-2} + 5$, so substituting this into the first equation gives $a_n = (a_{n-2} + 5) + 5 = a_{n-2} + 2 \cdot 5$. Since $a_{n-2} = a_{n-3} + 5$, we can substitute this in now to get $a_n = a_{n-3} + 3 \cdot 5$. We can repeat this process until we are left with $a_1$ and other terms. What will that other term be? Well, looking above, we see $a_n = a_{n-k} + 5k$. If $k = n - 1$, we get $a_n = a_1 + 5(n - 1)$, so $a_n = 2 + 5(n - 1) = 5n - 3$. Saying $k = n - 1$ is just saying that we had to repeat this process $n - 1$ times to get to the beginning term.

We can prove that $a_n = 5n - 3$ is the right solution by checking the initial conditions and the recurrence relation. Notice if we plug in $n = 1$, we get $a_1 = 2$, which is correct. We also know that $a_{n-1} = 5(n-1) - 3 = 5n - 8$. We can just check the recurrence relation, i.e. does $a_n = a_{n-1} + 5$? Well, $a_n = 5n - 3$ and $a_{n-1} = 5n - 8$, so yes, they are equal. Thus, this $a_n$ does satisfy the correct recurrence relation.

Let's use this technique in other problems. Before we do, we present the following fact:

**Proposition.** $1 + r + r^2 + r^3 + \ldots + r^k = \frac{r^{k+1}-1}{r-1}$.

**Problem.** Find the general solution for the minimum number of moves in the Hanoi problem, $H_n$.

*Solution.* We know $H_1 = 1$ and $H_n = 2H_{n-1} + 1$. Thus, $H_{n-1} = 2H_{n-2} + 1$, so

$$H_n = 2(2H_{n-2} + 1) + 1 = 2^2 H_{n-2} + 2 + 1.$$

Since $H_{n-3} = 2H_{n-2} + 1$, we can substitute this in to get

$$H_n = 2^2(2H_{n-3} + 1) + 2 + 1 = 2^3 H_{n-3} + 2^2 + 2 + 1.$$

Repeating the process we get

$$H_n = 1 + 2 + 2^2 + \ldots + 2^{k-1} + 2^k H_{n-k}.$$

Plugging in $k = n - 1$ yields

$$H_n = 1 + 2 + 2^2 + \ldots + 2^{n-2} + 2^{n-1} H_1 = 1 + 2 + 2^2 + \ldots + 2^{n-1} = 2^n - 1.$$

This is our desired formula.

We can again prove it is correct: $H_1 = 2^1 - 1 = 1$, which is true. And

$$2H_{n-1} + 1 = 2(2^{n-1} - 1) + 1 = 2^n - 2 + 1 = 2^n - 1 = H_n,$$

which means our solution satisfies the correct recurrence relation.

**Problem.** Use the recurrence from the previous section to prove that if $|X| = n$, then $|\mathcal{P}(X)| = 2^n$.

*Solution.* We found the recurrence relation $S_0 = 1$ and $S_n = 2S_{n-1}$ ($n \geq 1$) for the number of subsets of $\{1, 2, 3, \ldots, n\}$ . So,

$$\begin{aligned} S_n &= 2S_{n-1} \\ &= 2^2 S_{n-2} \\ &= 2^3 S_{n-3} \\ &= \cdots, \end{aligned}$$

meaning $S_n = 2^k S_{n-k}$. Taking $k = n$, we get $S_n = 2^n S_0 = 2^n$, which shows that $\{1, 2, 3, \ldots, n\}$ has $2^n$ subsets, as desired.

So technique (1) boils down to repeated use of the recurrence relation and recognizing a pattern. Technique (2) will be easier in some sense, in that we don't need to recognize any patterns.

**Definition.** A *linear homogeneous recurrence relation of order k with constant coefficients* is a recurrence relation of the form

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \ldots + c_k a_{n-k}, \quad c_k \neq 0.$$

*Remark.* You should think of this order as giving you the how far back in the sequence we need to go in the recurrence relation.

**Example.** (1) The Fibonacci sequence given by $F_1 = F_2 = 1$ and $F_n = F_{n-1} + F_{n-2}$ is a linear homogeneous recurrence relation of order 2.

(2) The number of $n$-bit strings which do not contain the string 111 satisfies a linear homogenous recurrence relation of order 3.

Before looking at how to solve these, we will need the following proposition.

**Proposition.** Let $a_n$ be a sequence given by a linear homogeneous recurrence relation of order $k$ (for any $k \geq 1$). If $S_n$ and $T_n$ are two solutions of the recurrence relation, then so is $cS_n + dT_n$ for any constants $c, d$.

*Proof.* Suppose $a_n$ satisfies

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

Saying $S_n$ and $T_n$ are solutions of the recurrence relation means

$$S_n = c_1 S_{n-1} + \ldots + c_k S_{n-k},$$

$$T_n = c_1 T_{n-1} + \ldots + c_k T_{n-k}.$$

Let $U_n = cS_n + T_n$. We want to show $U_n$ satisfies the same recurrence relation. Well, multiplying the first equation by $c$ and the second by $d$ and adding gives

$$U_n = c_1(cS_{n-1} + dT_{n-1}) + \ldots + c_k(cS_{n-k} + dT_{n-k}),$$

so

$$U_n = c_1 U_1 + \ldots + c_k U_{n-k},$$

as desired. □

Let's look at a concrete example of this:

**Example.** Consider the recurrence relation $a_n = 5a_{n-1} - 6a_{n-2}$. We can show $S_n = 2^n$ and $T_n = 3^n$ both satisfy the recurrence relation. For example,

$$S_n - 5S_{n-1} + 6S_{n-2} = 2^n - 5 \cdot 2^{n-1} + 6 \cdot 2^{n-2} = 2^{n-2}(2^2 - 5 \cdot 2 + 6) = 0,$$

which is what we wanted. One can similarly check that $T_n$ satisfies the recurrence relation. Thus, for any constants, $c$ and $d$, $c \cdot 2^n + d \cdot 3^n$ satisfies the recurrence relation.

**Problem.** Suppose we have a linear homogeneous recurrence relation of order $k$ given by

$$a_n - c_1 a_{n-1} + \ldots - c_k a_{n-k} = 0.$$

If $r$ is a root of the polynomial

$$f(t) = t^k - c_1 t^{k-1} - \ldots - c_k = 0,$$

then $S_n = r^n$ is a solution of the recurrence relation.

*Solution.* We want to show

$$S_n - c_1 S_{n-1} - \ldots - c_k S_{n-k} = 0.$$

Well, the left side is precisely

$$r^n - c_1 r^{n-1} - \ldots - c_k r^{n-k} = r^{n-k}(r^k - c_1 r^{k-1} - \ldots - c_k) = r^{n-k} f(r) = 0,$$

which is what we wanted.

We can use this problem to give us a general method for solving order 2 linear homogeneous recurrence relations:

**Method.** Suppose we have a recurrence relation of the form $a_n = c_1 a_{n-1} + c_2 a_{n-2}$, where $c_1$ and $c_2$ are constants.

(1) Find the two roots of the polynomial $p(t) = t^2 - c_1 t - c_2$, call them $r_1$ and $r_2$.

(2) (a) If $r_1 \neq r_2$, then the sequence $S_n$ given by $S_n = cr_1^n + dr_2^n$ is a solution to the recurrence relation, where the constants $c$ and $d$ can be determined from initial conditions.

(b) if $r_1 = r_2$, then the sequence $S_n$ given by $S_n = cr_1^n + dnr_1^n$ is a solution to the recurrence relation, where the constants can be determined from the initial conditions.

We will not prove this method in detail, but it mostly follows from the previous problem. The only thing to really prove is 2(b), and you can simply check that this is a solution to the recurrence relation.

Let's look at some examples.

**Problem.** Find a general formula for the Fibonacci sequence $F_n = F_{n-1} + F_{n-2}$ and $F_0 = 0$, $F_1 = F_2 = 1$.

*Solution.* We want to examine roots of the polynomial $t^2 - t - 1 = 0$, which are $r_1 = \frac{1+\sqrt{5}}{2}$ and $r_2 = \frac{1-\sqrt{5}}{2}$. Here, $r_1 \neq r_2$, so a general solution is

$$F_n = c\left(\frac{1+\sqrt{5}}{2}\right)^n + d\left(\frac{1-\sqrt{5}}{2}\right)^n.$$

To get $c$ and $d$, we can use the initial conditions:

$$0 = F_0 = c + d,$$
$$1 = F_1 = cr_1 + dr_2.$$

So $d = -c$, and plugging this into the second equation gives

$$1 = cr_1 - cr_2 = c(r_1 - r_2) = \sqrt{5}c.$$

Thus, $c = \frac{1}{\sqrt{5}}$ and $d = -c = -\frac{1}{\sqrt{5}}$. Our solution is therefore

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

*Remark.* It is remarkable that $F_n$ is an integer and yet our formula involves the irrational number $\sqrt{5}$. Try to use the binomial theorem to show that this formula is an integer.

**Problem.** Find a solution to the recurrence relation $a_n = 6a_{n-1} - 9a_{n-2}$, where $a_0 = 5$ and $a_1 = 9$.

*Solution.* Now, we must look at roots of $f(t) = t^2 - 6t + 9 = (t-3)^2$. So there is a repeated root of $t = 3$. Thus, by 2(b) of the method, our general solution looks like

$$a_n = c \cdot 3^n + dn \cdot 3^n.$$

Since $a_0 = 5$, we get $5 = c + 0d$, so $c = 5$. Then, since $a_1 = 9$, we get

$$9 = 3c + 3d = 15 + 3d,$$

so $d = -2$. Thus, our solution is

$$a_n = 5 \cdot 3^n - 2n \cdot 3^n.$$

Even if our recurrence relation doesn't look like this, we may be able to perform a transformation to make it a linear homogeneous relation.

**Problem.** Suppose $a_n$ satisfies $a_n^3 = 3a_{n-1}^3 - 2a_{n-2}^3$, with $a_0 = 1$ and $a_1 = \sqrt[3]{3}$. Find a formula for $a_n$.

*Solution.* This is certainly not a linear homogeneous relation, but if we make the transformation $b_n = a_n^3$, then we get

$$b_n = 3b_{n-1} - 2b_{n-2},$$

which is a linear recurrence relation. To solve this for $b_n$, we need the roots of $t^2 - 3t + 2$ which are $t = 1, 2$. Thus, our solution for $b_n$ looks like

$$b_n = c + d \cdot 2^n.$$

Since $a_0 = 1$, we get $b_0 = a_0^3 = 1$, and since $a_1 = \sqrt[3]{2}$, we get $b_1 = 2$. Plugging these into our solution gives the two equations:

$$1 = c + d,$$
$$3 = c + 2d,$$

so $d = 2$ and $c = -1$. Thus, $b_n = -1 + 2 \cdot 2^n = -1 + 2^{n+1}$. However, we wanted $a_n$, so using the fact that $b_n = a_n^3$, we can say

$$a_n^3 = -1 + 2^{n+1},$$

which means $a_n = \sqrt[3]{-1 + 2^{n+1}}$.