A first course on linear algebra

Ricardo Souza

August 28, 2019

Contents

Chapter 1

Basic set theory

1.1 Naive set theory

1.1.1 Axioms and you

Most, if not all, concepts in mathematics are phrased in the language of set theory: Geometric figures are just collections of points, transformations between two different objects are the collections of all the transitional states inbetween etc.

Hence, it makes sense to give some more formal foothold when studying any area of maths by beginning with some basic set theory.

But then, why *naive*?

Well, formal mathematics (that is, all contemporary and modern mathematics for more than a hundred years) is based on what we like to call *axioms* - you can think of them as the "rules of the game", in some sense.

Let me give you all an example of a well-accepted axiom of Euclidean geometry:

AXIOM:

Given any two distinct points, there is one, and only one, line through them.

Some people say it's "something that you can't prove", but it's not exactly that - axioms are either things that you don't *want* to prove, and just want to assume as truth (maybe because it is, indeed, impossible to prove it) or things that are, in some vague sense, "natural" or "self-evident".

Either way, the correct mindset to approach axioms is to think of them as the building blocks with which you build maths - just like atoms are the building blocks of matter -: by combining different axioms in different ways you get different results - the so called "theorems".

That's what maths is all about: Working with axioms and already proven theorems to prove new theorems. It's kinda like a game of scrabble, where the axioms are not only the blocks you (and everyone else) has in their hands, but also the rules of the game and the game board, and the theorems are the words you can make - subject to the rules of the game, the pieces and the board.

Hence, *naive* set theory is called so not because it is a theory of naive sets, but because it's a theory that's not properly formalized, and relies heavily on intuition and common sense.

In proper, axiomatic, set theory you'd have to define what is, and what isn't, a set. In naive set theory, however, we can just hand-wave it and say

NAIVE AXIOM(?):

Any collection of things is a set.

Now, as to *why* this isn't formal, it's due to the fact that it leads to a logical contradiction - a paradox. We're gonna show this contradiction in what follows, but if it doesn't interest you (you filthy, you) you can just skip the next section. It's fine, I won't judge you (actually, I will).

1.1.2 Russell's Paradox

Imagine that every random collection of random things is a set. Then it is only natural to consider the collection of all sets. But, since it is a collection (duh) it is also a set. But since it is the collection of all sets, it is an element of itself.

That's weird, ok? Try thinking of any sets - I'll give you plenty of time, don't worry - that are like that: they contain themselves as elements. You can't, right?

While that's not a contradiction, per se, it *really is weird*.

So let us consider N the collection of all non-weird sets - that is, the collection of all sets that do not contain themselves as elements.

Now, one naturally asks the question: Is N itself a weird set? That is, $N \in N$?

Well, I don't know. But if it was, then, by definition, all elements of N are non-weird sets, so N, being an element of N, would have to be a non-weird set - that is, $N \notin N$. So... If we assume that $N \in N$ we can logically infer that $N \notin N$...

Okay, maybe we made a mistake all along by assuming that $N \in N$! Yeah, that must be the case! Clearly, N can't be a weird set!... But then, since N isn't weird, it must be an element of N (since N contains *all* non-weird sets)... That is, $N \in N$. So if we assume $N \notin N$ we can logically deduce that $N \in N$.

We have just proven that $N \in N$ and $N \notin N$ are *logically equivalent*. But by the **Principle of Non-Contradiction** (something can't be simultaneously both true and false) those two can't be equivalent!

So, by assuming that there is a set containing all sets we can logically derive a contradiction - that, my friends, is the definition of a paradox.

This is the famous **Russell's Paradox** and it applies in broader contexts - it basically means that, from a logical POV, self-references are *kinda weird and you shouldn't actually do that*.

For instance, if you put as an axiom that "anything that can be stated can be proven", then you could ask "can I prove that there is something that cannot be proven?" and the answer would have to be *yes*, since you said (by axiom) that everything had to be provable. But that's a contradiction - by forcing everything to have a proof you have proven that you cannot prove everything.

This was proposed by philosopher-mathematician Bertrand Russell to show that maths really does need a formal framework to work with - otherwise we might be working in a system where contradictions arise (as we have seen).

There is, however, a solution to this. We have a set of axioms for set theory called the Zermello-Frankel axioms, which are a list of axioms that do not generate that kind of contradiction. It is,

however, *impossible* to prove whether it does or doesn't generate *any* paradox (this is due to a bunch of hard maths/philosophy that is waaaaay out of the scope of this text).

Just know that if you ever see ZFC anywhere you can rest safe because you're working with a (relatively) safe set of axioms.

1.2 Basic results and properties of sets

1.2.1 Equality always

As we have previously stated, a *set* is a collection of objects. We will usually denote a set by a capital letter (not always), such as X, A or B.

Since we cannot (as seen in the previous section) consider "the set of all sets", fix any set X. Now, X might be any set - numbers, birds, colours, the numerous of ways you can insult someone's mum etc.

When we have an object that is in that set we say that it is an **element** of that set, and usually denote it by a non-capital letter (once again, not always). In symbols, if we want to say that a is an element of X we would write that as $a \in X$ - which should be read as "a is an element of X", "a is in X" or even "X contains a as an element".

EXAMPLE(S):

Let E be the collection of all even integers. So $2 \in E$ and $28 \in E$, but $5 \notin E$ ("5 isn't in E", or "5 isn't an even integer") and *dog* $\notin E$ (because *dog* is **not** an even integer). Actually, you can see this as a formal proof of the well known fact that all dogs are *odd*.

You can, however, take all the elements of a set and ask if they satisfy a certain condition.

EXAMPLE(S):

Following up on the previous example, let ϕ denote the proposition "*can be written in english with only three letters*". Now we can consider the *subset* of E formed by all elements of E that also satisfy ϕ (if $x \in E$ is such an element, we simply write $\phi(x)$ to denote "x satifies ϕ"). This is written as follows:

$$E_\phi := \{x \in E \mid \phi(x)\}.$$

Let us break this down bit-by-bit:

- The symbol E_ϕ is non-standard notation that we're introducing here to mean "*the set E subject to the condition ϕ*";
- The symbol $:=$ means "*equals, by definition*". This can be used in two distinct ways: During a logical regression, we can use this symbol to justify one step by saying "this thing that I'm claiming is true, is actually true by definition"; or we can use it to define new terms - we're basically saying "the LHS is a new symbol whose meaning I'm defining to be the RHS" - kinda like attributing a value to a variable.

In this text we're **always** going to use this symbol with the second meaning - so in the preceding expression the := means "I'm defining E_ϕ to mean $\{x \in E \mid \phi(x)\}$".

- The brackets, in mathematics, almost always denote a *set*, and always are presented with the following structure: $\{A \mid B\}$.

  The A part is *what kind of elements does this set have*. In the example above, $x \in E$ means that the elements we're working with are even integers.

  The B part is *which condition these elements are subject to*. In the example above, $\phi(x)$ means that the elements of this set must satisfy ϕ.

Now that that's out of the way, what is E_ϕ? What are *the even integers that can be written in english using only three letters*? There are only three such numbers: **two**, **six** and **ten**. So we write $E_\phi = \{2, 6, 10\}$.

Definition 1.2.1. *Two sets A and B are said to be **equal** if they have the same elements. This means that every element of A is an element of B, and every element of B is an element of A.*
In this case we write $A = B$.

Let us give some examples of equalities.

EXAMPLE(S):

- Let A be the set of all animals that are wooly, fluffy and go *baa*, and let B be the set of all sheep. Clearly $A = B$.
- Let A be the set of roots of the polynomial $x^2 - x$ and let $B = \{0, 1\}$. It is an easy exercise to see that these two sets are the same.
- However, $A = \mathbb{N}$ the set of all natural numbers, and $B = \mathbb{Z}^{\geq 0}$ the set of non-negative integers, are **not** equal sets. You can see this in any proper course of number/set theory, but the elements of $\mathbb{Z}$ are always signed: -2, $+6$, $+1$ etc. (aside from 0), whereas the elements of $\mathbb{N}$ are **not** signed: 1, 6 etc. So $1 \notin \mathbb{Z}$ and $+1 \notin \mathbb{N}$, and therefore $A \neq B$.

Remark 1.2.2

In mathematics, a *definition* is the term we use to "assign" a new value to a certain term. In the definition above, we assigned a meaning to the phrase "two sets are equal".
Please be aware that this text will be filled with definitions of this kind, so take your time to get accostumed to them.

Notice, however, that we can sort of "relax" the conditions of the preceding definition. For instance, consider the following case:

EXAMPLE(S):

Let $A = \mathbb{N}$ the set of all natural numbers and $B = E$ the set of all even natural numbers. Notice that $A \neq B$ - for instance, 3 is in A, but not in B - so they can't be equal.
On the other hand, notice that it is impossible to produce such a counterexample starting from B: No matter which element you choose in B it will always be a natural number, of course, and therefore it will also be an element of A.
So these two sets, although not-equal, are not *entirely* different.

Definition 1.2.3. *Let A and B be two sets such that every element of B is also an element of A. In this case, we say that A* ***contains*** *B* ***as a subset*** *- or more simply that B* ***is a subset of*** *A, which we'll denote in symbols by $B \subseteq A$.*

EXAMPLE(S):

- In the preceding example, we see that $B \subseteq A$.
- Take any set A, and let $B = A$. We then ask the question: Is B a subset of A? Well, by definition, $B \subseteq A$ if, and only if, every element of B is also an element of A... But this is trivially true - since $B = A$!

  This gives us some insight on our first result:

Proposition 1.2.4. *For any set A we have that $A \subseteq A$.*

Proof

We want to show that every element $a \in A$ is also an element of A. But that's trivial. The result follows. □

Remark 1.2.5

In mathematics, a *proof* of a proposition/lemma/theorem/corolary is nothing more than a logical reasoning explaining why what we said is true. Proofs are to mathematics as scientific experiments are to sciences. This is what mathematicians do and work with all their lives. One could argue that maths is the science of reasoning and arguing.

Now we have our first non-trivial result:

Proposition 1.2.6. *Let A and B be two sets. Then $A = B$ if, and only if, $A \subseteq B$ and $B \subseteq A$.*

Proof

Assume that $A = B$. We want to show that $A \subseteq B$ and $B \subseteq A$, but this is trivial in light of the preceding proposition.

Assume now that $A \subseteq B$ and $B \subseteq A$. We want to show that $A = B$ - that is, every element of A is an element of B, and every element of B is an element of A.
Notice, however, that the phrase "every element of A is an element of B" is the definition of the symbol $A \subseteq B$, and the phrase "every element of B is an element of A" is the definition of the symbol $B \subseteq A$ - both of which we are assuming to be true.
Therefore, we have just proven that $A = B$, as stated, which finishes the proof. □

Remark 1.2.7

In mathematics, an *if, and only if,* statement is the equivalent of a logical equivalence. Basically, whenever we say "*this* holds if, and only if, *that* holds" what that means is that *this* and *that* are equivalent: *this* is true precisely when *that* is true, and *this* is false precisely when *that* is also false.
Without going too much into propositional logic, we usually write "a if, and only if, b" in symbols as $a \iff b$, which is logically equivalent to saying that "a being true is sufficient for us to prove that b is also true" and "b being true is sufficient for us to prove that a is also true". In symbols we would write these, respectively, as $a \implies b$ and $b \implies a$ - which should be read as "a implies b" and "b implies a", respectively.
That's what we did in the preceding proposition: If a="$A = B$" and b="$A \subseteq B$ and $B \subseteq A$", we proved that assuming a we can conclude b, and that assuming b we can conclude a - that is, we proved that a implies b and b implies a - which is logically equivalent to proving that a and b are equivalent.

This proposition is the most common tool used by mathematicians to prove that two sets are equal: We simply prove that each one contains the other - therefore, they must be equal.

EXAMPLE(S):

Let A be the set of roots of the polynomial $x^2 - x$ - that is, the set of numbers r such that $r^2 - r = 0$ - and $B = \{0, 1\}$. We claim that $A = B$.
First, let us show that $B \subseteq A$ - that is, both 0 and 1 are roots of $x^2 - x$. This is done by a simple verification:

$$0^2 - 0 = 0 - 0 = 0 \qquad \text{and} \qquad 1^2 - 1 = 1 - 1 = 0$$

so they are, indeed, roots of $x^2 - x$ - and therefore, $B \subseteq A$.
Now, to prove that $A \subseteq B$ we need to show that those are the only two possible roots.
To do that, let r be any root of $x^2 - x$ - that is, $r^2 - r = 0$. But then, $r^2 = r$, by adding r on both sides, and we see that $r = 0$ is indeed a solution to this equation ($0^2 = 0$). So if we assume that $r \neq 0$ we can divide both sides by r and get $\frac{r^2}{r} = \frac{r}{r}$ which is the same as $r = 1$, which was a unique solution being $r = 1$.
Hence we have proven that any root r of $x^2 - x$ is either 0 or 1, and therefore $A \subseteq B$.
Finally, since $A \subseteq B$ and $B \subseteq A$ we can finally say that $A = B$, as we had previously stated.

Definition 1.2.8. *We say that A is a **proper subset** of B if A is a subset of B, but B isn't a subset of A. In this case we use the symbol $A \subset B$.*

EXAMPLE(S):

Consider $A = \mathbb{N}$ the set of natural numbers, and $B = E$ the set of even natural numbers. We clearly have $B \subseteq A$ and $A \nsubseteq B$, so we can see that B is a *proper* subset of A - that is, $B \subset A$.

Finally, we can use all that we've done so far to construct a very special set - the empty set.

EXAMPLE(S):

Let $\mathbb{N}$ be the set of natural numbers and let ϕ be the proposition "is not a natural number". For instance, $\phi(\text{car})$ is just "car is not a natural number", which is true.
Now we can do just as we did before and consider

$$\mathbb{N}_\phi := \{n \in \mathbb{N} \mid \phi(n)\}$$

that is, the set of all natural numbers which are not natural numbers.
What **is** this set? Is there any natural number that isn't a natural number? Of course not! So this is a set *which has no elements*.
Take now $\mathbb{Z}$ the set of all integers and let ψ be the proposition "is not an integer". We can then define, once more,

$$\mathbb{Z}_\psi := \{n \in \mathbb{Z} \mid \psi(n)\}$$

that is, the set of all integers which aren't integers.
This set is, once again, empty.
This begets the question: $\mathbb{N}_\phi = \mathbb{Z}_\psi$ - that is, are two empty sets always equal?

Definition 1.2.9. *Given any set X we call the **empty set defined by** X to be the set of all elements of X which aren't elements of X, denoted by $\varnothing_X$.*

Theorem 1.2.10. *Given any two sets A and B, then $\varnothing_A = \varnothing_B$.*

Proof

If they were different, then there would either be some element of $\varnothing_A$ which is not in $\varnothing_B$, or some element of $\varnothing_B$ which is not in $\varnothing_A$. But both of these are impossible, since both sets are empty.
So they can't be different, and, therefore, $\varnothing_A = \varnothing_B$ □

Corolary 1.2.11. *For any set A, its empty set $\varnothing_A$ is uniquely determined.*

Corolary 1.2.12. *There a unique empty set.*

Remark 1.2.13

In mathematics, a *corolary* is a result that follows immediately from something that came before it - sometimes even foregoing a proof because of how immediate this conclusion is.

Definition 1.2.14. *We're going to define the* ***unique empty set*** *to be the empty set of any set, which will be denoted in symbols by* $\varnothing$.

Proposition 1.2.15. *For any set* A *we have that* $\varnothing \subseteq A$. *Furthermore, we have that* $A \subseteq \varnothing$ *if, and only if,* $A = \varnothing$.

Proof

If $\varnothing \not\subseteq A$, then there'd be some element in $\varnothing$ that was not in A. But $\varnothing$ is empty, therefore $\varnothing \not\subseteq A$ is false, and hence $\varnothing \subseteq A$.

For the second statement, we clearly have $A \subseteq \varnothing$ if $A = \varnothing$, by definition of set equality. But if we assume that $A \subseteq \varnothing$, we can now use the first statement of this proof, which proves that $\varnothing \subseteq A$, to conclude, by definition of set equality, that $A = \varnothing$, and this finishes the proof. □

1.2.2 United we stand, intersected we... Fall?

Now that we have a basic understanding of sets and subsets, we're going to build new sets from existing ones.

Definition 1.2.16. *Let* A *and* B *be two sets. The* ***union*** *of* A *and* B *is another set - denoted by* $A \cup B$ *defined by the following properties:*

(a) $A \cup B$ *contains both* A *and* B *as subsets;*

(b) Any other set C *that contains both* A *and* B *as subsets also contains* $A \cup B$ *as a subset.*

First things first, let us show that this definition makes sense - that is, that given two sets, their union is a unique set:

Lemma 1.2.17. *Let* A *and* B *be two sets, and* C *and* D *be two sets satisfying the above definition. Then* $C = D$.

Proof

Since C and D are unions of A and B, they contain both of them as subsets (item (a)). Now, since C satisfies (a) and D satisfies (b), we get that $D \subseteq C$. Similarly, since D satisfies (a) and C satisfies (b), we get that $C \subseteq D$.
It follows that $C = D$, and so the union of two sets is indeed well-defined, □

With that out of the way, let us show some examples to build some intuition:

EXAMPLE(S):

Let A be the set of all dogs and B be the set of all cats. Let C be the set of all animals. We ask: $C = A \cup B$?
Certainly, C satisfies (a) (since all dogs and all cats are animals), but does it satisfy (b)?
Well, certainly not! Because the set D of all mammals also contains A and B, but it clearly doesn't contain C (because not every animal is a mammal - for instance, there are birds).
Now we ask: Ok, since C is not the union of A and B, maybe D is?
Well, no, because we can consider E - the set of all mammal quadrupeds - and see that it contains both A and B as susbets, but not D.
And so on, and so forth...

How can we make sure that we don't get an endless regression - that is, we're always inching closer to the result, but never truly getting there?

Well, in formal set theory, for instance ZFC, you can always use your axioms to guarantee the existence of such a set. Here, however, we're going to have to appeal to intuition:

Proposition 1.2.18. *Given two sets A and B and any set C containing both A and B, their union is precisely the subset of C given by the proposition $\phi =$ "is in any one of the sets A or B".*

Proof

First, we'll show that $A \subset C_\phi$ and $B \subseteq C_\phi$.
To do that we'll just use the definition: Take $a \in A$ (resp. $b \in B$). Since $A \subseteq C$ (resp. $B \subseteq C$) we have that $a \in C$ (resp. $b \in C$). We then ask: is $\phi(a)$ (resp. $\phi(b)$) true? Well, it trivially is - $\phi(x)$ is true if, and only if x is in A or B - and a (resp. b) certainly is. Therefore, for any $a \in A$ (resp. $b \in B$) we can conclude $\phi(a)$ (resp. $\phi(b)$) - and therefore, $a \in C_\phi$ (resp. $b \in C_\phi$). This shows that $A \subseteq C_\phi$ (resp. $B \subseteq C_\phi$) - and therefore, C_ϕ satifies item (a) of the definition of set union.
Now, take any set D such that D contains both A and B as subsets. If we show that D also contains C_ϕ as a subset, we'll have shown that C_ϕ satisfies the definition of union - and therefore it must be the union.
To do that, take any $x \in C_\phi$. Then, by definition, $\phi(x)$ is true - that is, $x \in A$ or $x \in B$. But since both $A \subseteq D$ and $B \subseteq D$ hold, it doesn't matter if $x \in A$ or $x \in B$ is true - as long as one of them is true, we can conclude that $x \in D$. And since this holds for any $x \in C_\phi$, we have just shown that $C_\phi \subseteq D$.
Since the D we chose was general, the result follows. □

This is important: We now have a way to construct the union of two sets - just take any set containing both of them and restrict it to be only the elements from the original sets.

That, however, requires the existence of some set containing both of them - and that's where ZFC comes in: There's an axiom that states that there always exists a set containing any amount of other sets.

Since we're foregoing axioms here, we're going to provide a "construction" that should be enough for most purposes:

EXAMPLE(S):

Following up on the previous example, we can now see that $A \cup B$ is any one of "the set of all animals which are cats or dogs", "the set of all mammals which are cats or dogs" or "the set of all mammal quadrupeds which are cats or dogs" - any one of those work, by what we've already proven.
We could, however, give a more explicit construction: $A \cup B$ is just the set of all cats and dogs.

EXAMPLE(S):

Another, even more constructive example: Let $A = \{a, b, c\}$ and $B = \{c, d, e, f\}$. Then $A \cup B = \{a, b, c, d, e, f\}$ (prove it using the definition if you're not convinced).

Finally, let's end our discussions on the union with the following alternative characterization of it:

Lemma 1.2.19. *Let A and B be sets. Then $x \in A \cup B$ if, and only if, $x \in A$ or $x \in B$.*

Proof

One side of this proof is trivial and follows from the definition of set union.
Let us prove then that $x \in A \cup B$ implies $x \in A$ or $x \in B$.
Define $N = \{x \in A \cup B \mid x \notin A \text{ and } x \notin B\}$ the collection of all elements of $A \cup B$ which are in neither A nor B - which is, by definition, a subset of $A \cup B$.
We can now define $U = \{x \in A \cup B \mid x \notin N\}$ the collection of all elements of $A \cup B$ which are not in N - which is, by definition, a subset of $A \cup B$.
We claim that U contains A and B as subsets. This is easy to see: Take y in either A or B (doesn't matter which). Since $A \cup B$ contains both of them, $y \in A \cup B$. But since y came from either A or B, it cannot be in N (by definition of N) - so it must be in U (by definition of U). It follows that both A and B are contained in U.
But this is a conundrum, because $A \cup B$ is *contained* in every set that contains A and B (by definition of set union) - in particular, since U contains A and B this means that U *also contains* $A \cup B$.
This shows that $U = A \cup B$, and therefore $N = \varnothing$.
Finally, to show that $x \in A \cup B$ implies $x \in A$ or $x \in B$, take any $x \in A \cup B$ and notice, by what we've done, that this is the same as saying that $x \in U$. But, then again, this is the same as saying that $x \notin N$ - that is x is in A or B, just as stated. This finishes the proof. □

Going in the opposite direction of unions, there is the concept of intersections. If unions take two sets to build a bigger one, interceptions take two sets to build a smaller one:

Definition 1.2.20. *Let A and B be two sets. The **intersection** of A and B is another set - denoted by $A \cap B$ defined by the following properties:*

(a) $A \cap B$ is contained in both A and B as a subset;

(b) Any other set C that is contained both A and B as a subset is also contained $A \cap B$ as a subset.

Remark 1.2.21

Notice that the two definitions are basically the same, just changing, in some sense, the "order" of the inclusions $\subseteq$.

Now, let us proceed to prove essentially the same results for intersections as we did for unions:

Lemma 1.2.22. *Let A and B be two sets, and C and D be two sets satisfying the above definition. Then $C = D$.*

Proof

Since C and D are intersections of A and B, they are contained in both of them as subsets (item (a)). Now, since C satisfies (a) and D satisfies (b), we get that $C \subseteq D$. Similarly, since D satisfies (a) and C satisfies (b), we get that $D \subseteq C$.
It follows that $C = D$, and so the intersection of two sets is indeed well-defined, □

Contrary to unions, however, we cannot refine intersections. We can, however, still give a construction of the intersection:

Lemma 1.2.23. *Let A and B be sets. Then $x \in A \cap B$ if, and only if, $x \in A$ and $x \in B$.*

Proof

One side of this proof is trivial and follows from the definition of set intersection.
Let us prove then that $x \in A$ and $x \in B$ implies $x \in A \cap B$.
Define $N = \{x \in A \text{ and } x \in B \mid x \notin A \cap B\}$ the collection of all elements which are, at once, in both A and B, but not in $A \cap B$ - which is, by definition, a subset of both A and B.
We can now define $I = \{x \in A \text{ and } x \in B \mid x \notin N\}$ the collection of all elements of both A and B which are not in N - which is, by definition, a subset of both A and B. This implies, by definition of set intersection, that $I \subseteq A \cap B$.
We claim now that I contains $A \cap B$ as a subset. This is easy to see: Take any $y \in A \cap B$. Since $A \cap B \subseteq A$ and $A \cap B \subseteq B$, we see that $y \in A$ and $y \in B$. So this y is an element of both A and B which is in $A \cap B$ - which is the definition of an element of I. This shows that $y \in I$.
But this is a conundrum, because $A \cap B$ *contains* every set that is contained in both A and B (by definition of set intersection).
This shows that $I = A \cap B$, and therefore $N = \varnothing$.
Finally, to show that $x \in A$ and $x \in B$ implies $x \in A \cap B$, take any $x \in A$ and $x \in B$ and notice, by what we've done, that this is the same as saying that $x \notin N$ (since it is empty). But, then again, this is the same as saying that $x \in I$ - that is x is in $A \cap B$, just as stated.
This finishes the proof. □

Finally, let's do some examples:

EXAMPLE(S):

- Let $A = \{1, 2, 3, 4\}$ and $B = \{1, 3, 5, 7, 9\}$. Then, $A \cap B = \{1, 3\}$.
- If A is the set of all even integers, and B is the set of all odd integers, then $A \cap B = \varnothing$.
- If A is the set of all cats and B is the set of all brown animals, then $A \cap B$ is the set of all brown cats.

As a final topic on this section, let us consider another construction.

Definition 1.2.24. *Given any set X we denote the **set of all subsets of** X by $\mathcal{P}(X)$ (or 2^X) and call it the **power set** of X.*

Remark 1.2.25

Note that, at this point, we have not defined products and sums of sets - even less exponents. So, for now, the symbol 2^X is just that - a symbol. It has no meaning resembling the powering of real numbers.
We will, however, as this text progresses, show two reasons why this notation makes sense, and we'll expand it to be able to take any set to the power of any other set.

Okay, before anything else, let us do some examples:

EXAMPLE(S):

Let $A = \{1, 2, 3\}$. What is $\mathcal{P}(A)$? Well, by definition it is the set of all subsets of A. Well then - what are the subsets of A?
We can list a few: $\varnothing$, A, $\{1\}$, $\{2\}$, $\{3\}$, $\{1, 2\}$, $\{1, 3\}$ and $\{2, 3\}$. But are there any others? Well, assume $B \subseteq A$. Then we can ask if B has any elements or not. If it doesn't, great!, because $B = \varnothing$, which we've already accounted for.
If it does, we can ask if it contains 1. And then, we can ask if it contains 2 and 3. And depending on those answers we can pinpoint B exactly, and see that it is, indeed, in the list above (e.g., if it contains 1 and 2, but not 3, then $B = \{1, 2\}$, which is on the list above).
At this point, it is easy to see that

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

The preceding reasoning, however, gives us our first insight into how to understand the symbol 2^A: Making a subset of A is the same as asking each element of A if it is, or not, in there.
Imagine the elements of A are cards in a deck and you want to make a hand. Making a hand is the same as going through the deck, card by card, and choosing which cards you want to keep, or not.
Since every card has two options (to be, or not to be), the amount of hands is precisely 2 to the number of cards.

Note that in this particular example, $2^A = \mathcal{P}(A)$ has precisely $2^3 = 8$ elements, while A has precisely 3 elements.

Now that we're talking about power sets, we can define one of the most important concepts of set theory:

Definition 1.2.26. *Let X be a set and 2^X its power set. Given any $A \in 2^X$, we define its* ***complement*** *to be the set denoted by $X \setminus A$, which is given by*

$$X \setminus A := \{x \in X \mid x \notin A\}.$$

That is, the complement of a set is the collection of all elements that *do not* belong to that set.

EXAMPLE(S):

Following up on the previous example, let $B = \{1\}$. Then what is $A \setminus B$? Well, by definition, it's the collection of all elements of A that are not in B - that is, 2 and 3, so $A \setminus B = \{2, 3\}$. Call $C = A \setminus B$. What is, then, $A \setminus C$? Once again, by definition, it's the set of all elements of A which are not in C - that is, 1, so $A \setminus C = B$.

And finally, just before wrapping up this section, let us give one final definition and example:

Definition 1.2.27. *Let A and B be any two sets. We define $A \setminus B$ to be equal to $(A \cup B) \setminus B$ - that is, $A \setminus B$ is the complement of B in $A \cup B$.*

EXAMPLE(S):

Let $A = \{a, b, c, d, e, f, g, h, i, j\}$ and $B = \{a, e, i, o, u\}$. Then $A \setminus B$ is, by definition, the set of all elements of $A \cup B$ which are not in B. So writing $A \cup B = \{a, b, c, d, e, f, g, h, i, j, o, u\}$ we see that $A \setminus B$ is just $\{b, c, d, f, g, h, j\}$.
Similarly, $B \setminus A$ is the set of all elements of $A \cup B$ which are not in A - that is, $B \setminus A = \{o, u\}$.

To really wrap up this section, then, we're gonna make a list of properties for the things we've just described. You're welcome to try to prove them, although most of them are really trivial (that is, they follow immediately from the definitions or a quick observation).

Proposition 1.2.28. *Let A, B, C be any three subsets of a given, fixed, set X. Then the following properties always hold:*

(1) $A \cup B = B \cup A$;

(2) $A \cup (B \cup C) = (A \cup B) \cup C$;

(3) $A \cup \varnothing = A$;

(4) $A \cup X = X$;

(5) $A \cup A = A$;

(6) $A \cap B = B \cap A$;

(7) $A \cap (B \cap C) = (A \cap B) \cap C$;

(8) $A \cap \varnothing = \varnothing$;

(9) $A \cap X = A$;

(10) $A \cap A = A$;

(11) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$;

(12) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$;

(13) $A \cup B = A$ *if, and only if,* $B \subseteq A$;

(14) $A \cap B = B$ *if, and only if,* $B \subseteq A$;

(15) $A \subseteq B$ *implies* $A \setminus B = \varnothing$;

(16) $A \cap B = \varnothing$ *implies* $A \setminus B = A$ *and* $B \setminus A = B$;

(17) $X = (X \setminus A) \cup A$;

(18) $(A \setminus B) \cup (B \setminus A) \cup (A \cap B) = A \cup B$;

(19) $(A \setminus B) \cap (B \setminus A) = \varnothing$;

(20) $X \setminus A \in 2^X$;

(21) $X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B)$;

(22) $X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B)$;

(23) $A \setminus (A \setminus B) = B$;

(24) $A \setminus (A \cap B) = A \setminus B$.

Remark 1.2.29

In the preceding proposition, as well as in maths as a whole, we usually save the parenthesis to mean "this should be done first". For instance, $A \cup (B \cup C)$ means "the union of A and the union of B and C", whereas $(A \cup B) \cup C$ means "the union of the union of A and B and C".

1.3 Time to do some actual set theory, none of this introductory bullshit

1.3.1 How does this function

Understanding functions is, basically, the most important thing in all of mathematics - and that's not an overstatement. Even if you forego set theory, the concept of a function still makes sense and it's still at the center of any mathematical discussion.

Since this is a naive introduction to set theory, we're not gonna bother with certain technicalities and simply define:

Definition 1.3.1. *Let A and B be two sets. A formula ϕ is said to be of **function type** (or a **function**) from A to B if for any $a \in A$ there's a unique $b \in B$ such that $\phi(a, b)$.*

*In that case, we will write that as $\phi(a) = b$ and say that b **is the image of** a **under** ϕ.*

EXAMPLE(S):

Let $A = B = \mathbb{N}$ the set of natural numbers, and let $\phi(x, y) =$ "y is the square of x". Then ϕ is clearly a function: for any $a \in A$, there is a unique $b \in B$ such that $\phi(a, b)$, and that b is precisely a^2. So we write this as $\phi(a) = a^2$.
Now, define $\psi(x, y) = \phi(y, x)$. Is ψ also a function? The answer is no: Indeed, for any $a \in A$, there exists, at most, one $b \in B$ such that $\psi(a, b)$. But the thing is - there are some a for which there is no b! For instance, for $a = 3$, there is no b such that $\psi(3, b)$. So ψ can't be a function.

Definition 1.3.2. *Let A, B be sets and ϕ be a function from A to B. We will call A the **domain** of the function and B its **codomain**, somtimes written as $A = \mathrm{Dom}(\phi)$ and $B = \mathrm{Cod}(\phi)$.*

In this case, we wil also use the notation $\phi : A \to B$ or $A \xrightarrow{\phi} B$ to say that "ϕ is a function whose domain is A and whose codomain is B".

Definition 1.3.3. *Two functions $f, g : A \to B$ between the same two sets are said to be **equal** if $f(a) = g(a)$ for all $a \in A$. That is, $f(a, g(a))$ and $g(a, f(a))$ hold for all $a \in A$.*

EXAMPLE(S):

Let $A = B = \mathbb{R}$ the set of real numbers, and let $f, g : A \to B$ be functions defined by $f(x) = \sqrt{x^2}$ and $g(x) = \begin{cases} x, \text{ if } x \geq 0 \\ -x, \text{ otherwise.} \end{cases}$
We claim that $f = g$.
To see that, take any real number, $x \in \mathbb{R}$. Now, if $x \geq 0$, then $g(x) = x$. Furthermore, $f(x) = \sqrt{x^2} = x$, so $f(x) = g(x)$. On the other hand, if $x < 0$, we have $g(x) = -x$ and $f(x) = \sqrt{x^2} = \sqrt{(-x)^2}$ and since $x < 0$, $-x$ must be greater than 0, so $f(x)$ is simply $-x$.
Therefore, $f(x) = g(x)$ for all $x \in \mathbb{R}$ (since any real number is either negative or non-

negative), and we see that $f = g$, as stated.

Definition 1.3.4. *If $f : A \to B$ is a function such that $f(a) = b$ for some $a \in A$ and $b \in B$, we then say that f **takes** a **to** b, which will be written as $a \mapsto b$.*

EXAMPLE(S):

The functions f, g of the previous example can be rewritten as

$$\begin{aligned} f : A &\to B \\ a &\mapsto \sqrt{a^2} \end{aligned}$$

and

$$\begin{aligned} g : A &\to B \\ a &\mapsto \begin{cases} a, \text{ if } a \geq 0 \\ -a, \text{otherwise.} \end{cases} \end{aligned}$$

Before we move forward, a couple of important definitions:

Definition 1.3.5. *Given any function $f : A \to B$, the set of all elements of B which are image of some element of A under f will be called the **image of** A **under** f (or just the image of f) and denoted by $f(A)$ (or $\mathrm{Im}(f)$).*

*Analogously, given any $X \subseteq A$, we denote the set of all elements of B which are image of some element of X under f by **image of** f **when restricted to** X, and denote it by $f(X)$.*

Proposition 1.3.6. *For any function $f : A \to B$, and any $X \subseteq A$, $f(X) \subseteq B$.*

Proof

Trivial, by the definition of image of a function. □

Definition 1.3.7. *Given any function $f : A \to B$ and any point $b \in f(A)$, we define the **inverse image of** b **under** f to be the set $f^{-1}(b) := \{a \in A \mid f(a) = b\}$ of all points in A whose image under f is precisely b.*

*Analogously, given any $Y \subseteq f(A)$, we define the **inverse image of** Y **under** f to be the set $f^{-1}(X) := \{a \in A \mid f(a) \in Y\}$ of all points in A whose image under f is in Y.*

Proposition 1.3.8. *For any function $f : A \to B$, and any $Y \in f(A)$, $f^{-1}(Y) \subseteq A$.*

Proof

Trivial, by the definition of inverse image of a function. □

Finally we can start working with some very important classes of functions: Injections, surjections and bijections.

Definition 1.3.9. *A function $f : A \to B$ is called an **injection** if $f(a) = f(a')$ implies $a = a'$.*

Remark 1.3.10

This is logically equivalent to saying that a function is an injection if diferent points of the domain have different images in the codomain.

EXAMPLE(S):

Let $f, g : \{1, 2, 3\} \to \{a, b, c, d\}$ be defined by: $f(1) = a, f(2) = b, f(3) = c$ and $g(1) = g(2) = g(3) = d$. Then f is injective and g is clearly not injective.

Let, now, $h : \mathbb{R} \to \mathbb{R}$ be defined by $h(x) = x^2$. Is h injective?
Well, suppose x and x' are such that $h(x) = h(x')$. This means that $x^2 = x'^2$. Taking square roots on both sides we get that $|x| = |x'|$ - which can be further simplified to mean $x = \pm x'$. In other words, we see that if two points have the same square, then they must differ only by a sign. That's good and all, but also shows us that two numbers that differ by a sign have the same image under h - and therefore h cannot be injective.
For instance, $2 \neq -2$, but $h(2) = h(-2) = 4$.

Definition 1.3.11. *A function $f : A \to B$ is called a **surjection** if for any $b \in B$ there is some $a \in A$ such that $f(a) = b$.*

EXAMPLE(S):

Following up on the previous example, neither f nor g are surjections: $d \in \{a, b, c, d\}$ isn't in the image of any point over both f and g.
Let us then define $f' : \{a, b, c, d\} \to \{1, 2, 3\}$ by putting $f'(a) = 1, f'(b) = 2, f'(c) = 3, f'(d) = 3$. Now, f' is indeed a surjection.

Notice that h too isn't a surjection: -1 isn't the image of any real number under h. However, if we define $h' : \mathbb{R} \to \mathbb{R}^{\geq 0}$, where $\mathbb{R}^{\geq 0}$ is the set of all non-negative real numbers, by putting $h'(x) := h(x)$, we see that h' is, now, a surjection.

Remark 1.3.12

Note that, in the example above, we defined h' by putting $h'(x) := h(x)$. Does that mean that $h' = h$?
The answer is **no**: By the definition of function equality, for two functions to be equal they must have the same domain and codomain.
This is a very important distinction, and one that most mathematicians and students rarely pay attenttion to.

Finally, we can define:

Definition 1.3.13. *A function $f : A \to B$ is called a **bijection** if it is both an injection and a surjection.*

EXAMPLE(S):

None of the previous examples are bijections, so we have to come up with new examples.
Let $f : \{a, b, c\} \to \{1, 2, 3\}$ be defined by $f(a) = 1, f(b) = 2, f(c) = 3$. Then f is both injective and surjective, and, therefore, a bijection by definition.

Let $g : \mathbb{R}^{\geq 0} \to \mathbb{R}^{\geq 0}$ be defined by $g(x) = x^2$. Then g is both injective (since there's only one sign on the domain) and surjective (since that are no negatives on the codomain), and, therefore, bijective by definition.

Let $h : \{a, b, c\} \to \{a, b, c\}$ be defined by $h(a) = a, h(b) = c, h(c) = b$. Is h a bijection? Well, it clearly is both injective and surjective, so it has to be by definition.
Notice that bijections don't have to abide by our expectations (such is life).

Lemma 1.3.14. *If $f : A \to B$ is injective, then there is a bijection $g : A \to f(A)$.*

Proof

Let $g : A \to f(A)$ be defined by $g(a) := f(a)$ for all $a \in A$.

- g is injective:

  To see that, take $a, a' \in A$ such that $g(a) = g(a')$. Then, by definition, this implies $f(a) = f(a')$, and since f is injective, this in turn implies $a = a'$, so g is injective.

- g is surjective:

  To see that, take any $b \in f(A)$. By definition of image, there exists some $a \in A$ such that $b = f(a)$. But now, by defintion of g, this means that $b = g(a)$.

  We have just shown that every point in the codomain of g is the image of some point in the domain of g under g - this means that g is surjective.

Since g is both injective and surjective, it is, by definition, a bijection, which ends the proof. □

We can use this lemma to easily determine whether a function is, or isn't, an injection.

EXAMPLE(S):

Let $f : \{a, b, c\} \to \{1, 2\}$ be defined by $f(a) = 1, f(b) = 2, f(c) = 1$. Is f injective?
Well, f takes two different points (a and c) to the same point (1), so it can't be injective.

Actually - is it possible for there to be an injective function from $\{a, b, c\}$ to $\{1, 2\}$?
Let's try making one: First, we choose an image for a - it can be either 1 or 2 - doesn't

matter which. Now, to choose an image for b we can't choose the same point as we chose for a - otherwise f won't be injective. So b's image is now uniquely determined: the only point left in $\{1, 2\}$ after we take out $f(a)$. Finally, when we try to choose an image for c, it can't be $f(a)$, nor can in be $f(b)$ (otherwise, f wouldn't be injective). But $\{1, 2\} = \{f(a), f(b)\}$ - that is, if $f(c)$ can't be $f(a)$ and it can't be $f(b)$, then there's **nothing** that it can be!
But, on the other hand, since f is a function, we **have to** take c somewhere. This means that we **have to** repeat either $f(a)$ or $f(b)$.
This shows that there are no injective functions from $\{a, b, c\}$ to $\{1, 2\}$.

Lemma 1.3.15. *If $f : A \to B$ is surjective, then there is a bijection $g : f(A) \to B$.*

Proof

Let $g : f(A) \to B$ be defined by $g(b) := b$ for all $b \in f(A)$.

- g is surjective:

  To see that, take any $b \in B$. Since f is surjective, for each point in B there is at least one point in A which is its inverse image under f - in particular, there is some $a \in A$ such that $f(a) = b$. But this means that $b \in f(A)$, by definition of image of f. Now, since $b \in f(A)$, we see that $g(b) = b$ and, therefore, g is surjective.

- g is injective:

  To see that, take any two points $b, b' \in f(A)$ such that $g(b) = g(b')$. But, by definition of g, this is the same as saying $g = g'$ - therefore g is injective.

Since g is both injective and surjective, it is, by definition, a bijection, which ends the proof. □

Analogously to injections, this lemma gives us a clear cut method for distinguishing surjections:

EXAMPLE(S):

Let $f : \{1, 2\} \to \{a, b, c\}$ be given by $f(1) = a$ and $f(2) = b$. Clearly, then, f isn't surjective, because there is one point in its codomain (c) which is not the image of any point of the domain under f.

And then we ask: Can there ever be a surjective function from $\{1, 2\}$ to $\{a, b, c\}$?
Once again, let's try building one: First, we choose $f(1)$. It can be anything, so choose anything. Now to choose $f(2)$, there's also no restrictions, but remember that we're trying to make a function that "covers" $\{a, b, c\}$ with guys from $\{1, 2\}$, so even though we could put $f(2) := f(1)$, it makes sense to choose $f(2)$ to be anything aside from $f(1)$... And we're done.
Notice, however, that no matter **how** we do that choice, there'll always be some point left in $\{a, b, c\}$. Therefore, there can be no surjections from $\{1, 2\}$ to $\{a, b, c\}$.

These last two examples give us a nice intuition of what injections and surjections measure: Injections measure how much "smaller" the domain is, when compared to the codomain, and surjections measure how much "bigger" the codomain is, when compared to the domain.

This allows us to consider one final example:

EXAMPLE(S):

Let $f : \{a, b, c\} \to \{1, 2, 3\}$ be a function. Can f be a bijection?
Let's try: First, we choose any of $\{1, 2, 3\}$ to be $f(a)$. Now, since we want f to be a bijection, it needs to be injective and surjective, so we can't choose $f(b) = f(a)$, so choose $f(b)$ to be any of $\{1, 2, 3\} \setminus \{f(a)\}$. Again, by the same reasoning, choose $f(c)$ to be any of $\{1, 2, 3\} \setminus \{f(a), f(b)\}$ - which isn't really a choice, since there's only one point left.
And we're done! By construction, $f(a) \neq f(b)$, $f(a) \neq f(c)$ and $f(b) \neq f(c)$ (so f is injective) and all of $\{1, 2, 3\}$ have inverse images.
This is a strong intuition that we want to build at this point:
Bijections between two sets tell us if they have the same amount of points. In many ways, then, bijections can be thought of as a relabeling of your set - or even, in some cases, as the *definitive and improved* notion of set equality.
And it makes sense - why should the sets $\{a, b, c\}$ and $\{1, 2, 3\}$ be treated as being different? You might argue that $1 + 2 = 3$, but $a + b$ doesn't even make sense - but the point here is that even $1 + 2$ doesn't make sense. There's no operations being taken into consideration, nothing. Just sets with elements. The only information we have is that "$\{a, b, c\}$ is a set with three distinct things inside it" and that "$\{1, 2, 3\}$ is a set with three distinct things inside it". *What* those things are doesn't really matter to us from a set-theoretical POV. What matters is that there are some things.

To expand in that idea - that bijections are the new equality - we're gonna start a more technical subsection.

The reader is encouraged to **not** skip this section, although I don't own you, so you do you. This subsection will have many proofs, so it's good for practicing your proofs, but not only that - the reasoning employed here is central to understanding what's behind many of the most intricate results in linear algebra.

1.3.2 Bijection is the new equality

Definition 1.3.16. *Given any two functions $f : A \to B$ and $g : B \to C$, we call the function $g \circ f : A \to C$ defined by $(g \circ f)(a) := g(f(a))$ the **composition** of f and g.*

Definition 1.3.17. *Given any set X, we call the function $\mathrm{id}_X : X \to X$ defined by $\mathrm{id}_X(x) := x$ the **identity function** of X.*

Definition 1.3.18. *Given a function $f : A \to B$, we say that f **is an isomorphism** if there is some function $g : B \to A$ such that $f \circ g = \mathrm{id}_B$ and $g \circ f = \mathrm{id}_A$. In this case, we say that g is an **inverse** for f.*

Proposition 1.3.19. *Function composition is associative - that is, if* $A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D$*, then* $h \circ (g \circ f) = (h \circ g) \circ f$*.*

Proof

Take $a \in A$. Then,

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

and therefore $(h \circ (g \circ f))(a) = ((h \circ g) \circ f)(a)$ for any $a \in A$ which, by the definition of function equality, implies that $h \circ (g \circ f) = (h \circ g) \circ f$. □

Proposition 1.3.20. *Given any function* $f : A \to B$*, we have that* $f = \mathrm{id}_B \circ f = f \circ \mathrm{id}_A$*.*

Proof

Take any $a \in A$. Then:

$$(\mathrm{id}_B \circ f)(a) = \mathrm{id}_B(f(a)) = f(a) = f(\mathrm{id}(a)) = (f \circ \mathrm{id}_A)(a)$$

which implies, by the definition of function equality, that $\mathrm{id}_B \circ f = f = f \circ \mathrm{id}_A$. □

Proposition 1.3.21. *Let* $f : A \to B$ *be an isomorphism, and let* $g, h : B \to A$ *be two inverses for* f*. Then* $g = h$*.*

Proof

This follows from the two preceding propositions and the definition of isomorphism:

$$g = g \circ \mathrm{id}_B = g \circ (f \circ h) = (g \circ f) \circ h = \mathrm{id}_A \circ h = h.$$

□

Definition 1.3.22. *Given an isomorphism* f*, we will denote its (unique!) inverse by* f^{-1}*.*

Definition 1.3.23. *A function* $f : A \to B$ *is called a* ***monomorphism*** *if given any other two functions* $g, h : C \to A$*, we have that* $f \circ g = f \circ h$ *implies* $g = h$*.*

EXAMPLE(S):

Let $f : \{1, 2, 3\} \to \{a, b, c, d\}$ be defined by $f(1) = a, f(2) = b, f(3) = c$. We claim that f is a monomorphism.

To see that, take any $g, h : C \to \{1, 2, 3\}$ such that $f \circ g = f \circ h$. In particular, for any $x \in C$ we have that $f(g(x)) = f(h(x))$. Well, this means that $f(g(x))$ is either a, b or c. In either case, we know precisely who $g(x)$ is (for instance, if $f(g(x))) = b$, then $g(x) = 2$, since 2 is the only point which is taken to b via f).
But since $f(g(x)) = f(h(x))$, there's a unique $u \in \{1, 2, 3\}$ such that $y = g(x) = h(x)$. In particular, $g(x) = h(x)$.
This shows that $g = h$, and, therefore, f is a monomorphism.

Theorem 1.3.24. *A function is a monomorphism if, and only if, it is an injection.*

Proof

Assume that $f : A \to B$ is injective. Then given $g, h : C \to A$ such that $f \circ g = f \circ h$ we want to show that $g = h$. Since f is injective, $f(g(c)) = f(h(c))$ implies $g(c) = h(c)$, for all $c \in C$. It follows, then, that $g = h$ and f is monic.

Conversely, if $f : A \to B$ is monic, define $g, h : \{c\} \to A$ by putting $g(c) = a$ and $h(c) = a'$ for two $a \neq a' \in A$ fixed. Now, since f is monic, by assumption, and since $g \neq h$, we have that $f \circ g \neq f \circ h$ (otherwise we would have f monic, $f \circ g = f \circ h$ and $g \neq h$ all being true, which is impossible). But then:

$$f(a') = f(h(c)) = f \circ h(c) \neq f \circ g(c) = f(g(c)) = f(a)$$

that is, $a \neq a'$ assures us that $f(a) \neq f(a')$, and so f is injective.
Notice that we used the fact that there are two distinct points in A: a and a'. If, however, A has only one point it is even simpler: Any function from a set with a single point has to be injective - in particular, monomorphisms whose domain are a single point are injective.
This finishes the proof. □

Lemma 1.3.25. *Every isomorphism is a monomorphism.*

Proof

Let $f : A \to B$ be an isomorphism and $g, h : C \to A$ any two functions such that $f \circ g = f \circ h$. Then, since f is an isomorphism, there is a unique inverse $f^{-1} : B \to A$ such that $\mathrm{id}_A = f^{-1} \circ f$ and $\mathrm{id}_B = f \circ f^{-1}$. Therefore:

$$g = \mathrm{id}_A \circ g = (f^{-1} \circ f) \circ g = f^{-1} \circ (f \circ g) = f^{-1} \circ (f \circ h) = (f^{-1} \circ f) \circ h = \mathrm{id}_A \circ h = h$$

and we see that f is monic. □

Definition 1.3.26. *A function $f : A \to B$ is called an **epimorphism** if given any other two functions $g, h : B \to C$, we have that $g \circ f = h \circ f$ implies $g = h$.*

EXAMPLE(S):

Let $f : \{a, b, c\} \to \{1, 2\}$ be defined by $f(a) = f(b) = 1$ and $f(c) = 2$. We claim that f is epic.
To see that, take any two functions $g, h : \{1, 2\} \to C$ such that $g \circ f = h \circ f$. We want to show that $g = h$ - that is, for any $x \in \{1, 2\}$, we have that $g(x) = h(x)$.
But since f is surjective (check!), there is some $y \in \{a, b, c\}$ such that, then $x = f(y)$, so

$$g(x) = g(f(y)) = (g \circ f)(y) = (h \circ f)(y) = h(f(y)) = h(x)$$

and therefore we see that $g = h$, which proves that f is indeed an epimorphism.

Theorem 1.3.27. *A function $f : A \to B$ is an epimorphism if, and only if, it is a surjection.*

Proof

First, let us assume that $f : A \to B$ is surjective. Then, given $g, h : B \to C$ such that$g \circ f = h \circ f$ we wish to show that $g = h$. We can just proceed as above: Proving that $g = h$ is the same as proving that for all $x \in B$, we have that $g(x) = h(x)$, but since f is surjective, by assumption, we have that there is some $y \in A$ such that $x = f(y)$. It follows then that

$$g(x) = g(f(y)) = (g \circ f)(y) = (h \circ f)(y) = h(f(y)) = h(x)$$

and therefore $h = g$, which shows that f is epic.
Conversely, assume that f is epic, and let $C = \{c, c'\}$. Now pick a point $b \in B$ and define $g, h : B \to C$ by putting $g(x) = c$ for all $x \in B$, $h(b) = c'$ and $h(x) = c$ if $x \neq b$.
Now, $g(b) \neq h(b)$, so $g \neq h$. Since f is epic, we must then have that $g \circ f \neq h \circ f$, by definition of epimorphism. This means that there is some $a \in A$ such that $g(f(ay)) \neq h(f(a))$.
Since g takes everyone to c, the only possible value of $h(f(a))$ that could be different from that is is $h(f(a)) = c'$, but the only element of B that is taken to c' by h is b - this means that $f(a) = b$.
We have just proven that given any $b \in B$ there is some $a \in A$ such that $f(a) = b$ - that is, f is surjective, which finishes the proof. □

Lemma 1.3.28. *Every isomorphism is an epimorphism.*

Proof

Let $f : A \to B$ be an isomorphism and $g, h : B \to C$ two functions such that $g \circ f = h \circ f$. Since f is an isomorphism, it has an inverse $f^{-1} : B \to A$. Therefore:

$$g = g \circ \mathrm{id}_B = g \circ (f \circ f^{-1}) = (g \circ f) \circ f^{-1} = (h \circ f) \circ f^{-1} = h \circ (f \circ f^{-1}) = h \circ \mathrm{id}_B = h$$

and we see that $g = h$, which proves that f is an epimorphism. □

Definition 1.3.29. *A function $f : A \to B$ is called a **bimorphism** if it is a mono-epimorphism.*

Clearly, by what we've already shown, bijections and bimorphisms are the same thing. However, we can do one better than that:

Lemma 1.3.30. *Every monomorphism $f : A \to B$ has a left-inverse, that is, a function $g : B \to A$ such that $\mathrm{id}_A = g \circ f$, and every function which has a left-inverse is a monomorphism.*

Proof

Let $f : A \to B$ be a monomorphism and consider $f(A)$. Since f is monic, by theorem 1.3.24 we see that f is injective, which means that for all $b \in f(A)$ we have that $f^{-1}(b)$ is a single point in A.
We then define $g : B \to A$ by

$$g(b) = \begin{cases} f^{-1}(b), & \text{if } b \in f(A) \\ a, & \text{otherwise,} \end{cases}$$

where $a \in A$ is any (literally any) element of A.
We claim that this g is a left-inverse for f. Indeed, for any $x \in A$ we have

$$(g \circ f)(x) = g(f(x)) = f^{-1}(f(x)) = x = \mathrm{id}_A(x)$$

since $f(x) \in f(A)$ for all $x \in A$. Therefore, we have shown that for all x we have $(g \circ f)(x) = x = \mathrm{id}_A(x)$ - which implies, by the definition of function equality, that $g \circ f = \mathrm{id}_A$.

Take now $f : A \to B$ a function that has a left-inverse $g : B \to A$ - that is, $\mathrm{id}_A = g \circ f$. Now take two functions $h, j : C \to A$ such that $f \circ h = f \circ j$. We want to show that $h = j$ (and therefore, f is monic).
Since $f \circ h = f \circ j$, we can compose g on the left on both sides of the equation to obtain $g \circ (f \circ h) = g \circ (f \circ j)$, which, by proposition 1.3.19, is the same as $(g \circ f) \circ h = (g \circ f) \circ j$, and since g is a left-inverse to f, we can further affirm that this is the same as $\mathrm{id}_A \circ h = \mathrm{id}_A \circ j$. Finally, by the definition of id_A, we see that this implies $h = j$ - and therefore f is monic, as stated.
This finishes the proof. □

Lemma 1.3.31. *Every epimorphism $f : A \to B$ has a right-inverse, that is, a function $g : B \to A$ such that $\mathrm{id}_B = f \circ g$, and every function which has a right-inverse is an epimorphism.*

Proof

Let $f : A \to B$ be an epimorphism, and consider $f(A)$ its image. By theorem 1.3.27, we know that g is a surjection. Since g is a surjection, then, for every $b \in B$ the set $f^{-1}(b)$ is well defined (by definition of surjection).
Now, choose $a_b \in f^{-1}(b)$ for each $b \in B$ (here the index is simply so we know where it came

from), and consider the function $g : B \to A$ taking each b to the a_b we chose above.
This is clearly a function (check!), and so we can do, for evrey $b \in B$:

$$(f \circ g)(b) = f(g(b)) = f(a_b) = b = \mathrm{id}_B(b)$$

and therefore $f \circ g$ and id_B are equal in every point - which means that they're equal, and g is a right-inverse for f, as stated.

Take now $f : A \to B$ a function with a right-inverse $g : B \to A$ - that is, $f \circ g = \mathrm{id}_B$. Now take two functions $h, j : B \to C$ such that $h \circ f = j \circ f$. We want to show that $h = j$ (and, therefore, f is epic).
Since $h \circ f = j \circ f$, we can compose g on the right on both sides of the equation to obtain $(h \circ f) \circ g = (j \circ f) \circ g$, which, by proposition 1.3.19, is the same as $h \circ (f \circ g) = j \circ (f \circ g)$, and since g is a right-inverse to f, we can further affirm that this is the same as $h \circ \mathrm{id}_B = j \circ \mathrm{id}_B$. Finally, by the definition of id_B, we see that this implies $h = j$ - and therefore f is epic, as stated.
This finishes the proof. □

Lemma 1.3.32. *If a function $f : A \to B$ is such that $g, h : B \to A$ are a left- and a right-inverse, respectively, then $g = h$, f is an isomorphism and g is its inverse.*

Proof

It follows trivially by the following computation:

$$g = g \circ \mathrm{id}_B = g \circ (f \circ h) = (g \circ f) \circ h = \mathrm{id}_A \circ h = h,$$

which shows at once that $g = h$. This means that $\mathrm{id}_A = g \circ f$ and $\mathrm{id}_B = f \circ g$ - and therefore g is an inverse to f, which shows that f is an isomorphism, as stated. □

Theorem 1.3.33. *A function f is an isomorphism if, and only if, it is a bimorphism.*

Proof

In light of lemmas 1.3.25 and 1.3.28, we see that every isomorphism is monic and epic and, therefore, a bimorphism.
Conversely, by lemmas 1.3.30 and 1.3.31 we see that any bimorphism has both a left- and a right-inverse. But now, lemma 1.3.32 shows us that since every bimorphism has a left- and a right-inverse, it must be an isomorphism, which ends the proof. □

Corolary 1.3.34. *Every bijection has a unique inverse.*

And now, finally, to end this section, some technical results that appear all the time in mathematics.

Lemma 1.3.35. *Let $A \xrightarrow{f} B \xrightarrow{g} C$ be two functions such that $g \circ f$ is monic (resp. epic). Then f is also monic (resp. g is also epic).*

Proof

If $g \circ f$ is monic, by lemma 1.3.30, we see that there is some function $h : C \to A$ that is a left-inverse to $g \circ f$ - that is, $\mathrm{id}_A = h \circ (g \circ f)$. But now, by proposition 1.3.19, we see that $h \circ (g \circ f) = (h \circ g) \circ f$ and, therefore, $\mathrm{id}_A = (h \circ g) \circ f$ and we see that $h \circ g$ is a left-inverse for f. Now the converse of lemma 1.3.30 tells us that since f has a left-inverse, it must be monic.

Anaolgously, if $g \circ f$ is epic, by lemma 1.3.31, we see that there is some function $h : B \to C$ that is a right-inverse to $g \circ f$ - that is, $\mathrm{id}_B = (g \circ f) \circ h$. But now, by proposition 1.3.19, we see that $(g \circ f) \circ h = g \circ (f \circ h)$ and, therefore, $\mathrm{id}_B = g \circ (f \circ h)$ and we see that $f \circ h$ is a right-inverse to g. Now, the converse of lemma 1.3.31 tells us that since g has a right-inverse, it must be epic.
The result then follows. □

Lemma 1.3.36. *$A \xrightarrow{f} B \xrightarrow{g} C$ be two functions. Then the following hold:*

(a) If both f and g are monic, then so is $g \circ f$;

(b) If both f and g are epic, then so is $g \circ f$;

(c) If both f and g are iso, then so is $g \circ f$;

(d) If $g \circ f$ and f are iso, then so is g;

(e) If $g \circ f$ and g are iso, then so is f.

Proof

(a) Assume both f and g are monic, and let $f' : B \to A$ and $g' : C \to B$ be their respective left-inverses. We claim that $f' \circ g'$ is a left-inverse to $g \circ f$. Indeed:

$$(f' \circ g') \circ (g \circ f) = f' \circ (g' \circ g) \circ f = f' \circ \mathrm{id}_B \circ f = f' \circ f = \mathrm{id}_A$$

so $g \circ f$ is monic.

(b) Assume both f and g are epic, and let $f' : B \to A$ and $g' : C \to B$ be their respective right-inverses. We claim that $f' \circ g'$ is a right-inverse to $g \circ f$. Indeed:

$$(g \circ f) \circ (f' \circ g') = g \circ (f \circ f') \circ g' = g \circ \mathrm{id}_B \circ g' = g \circ g' = \mathrm{id}_C$$

so $g \circ f$ is epic.

(c) Follows immediately from (a) and (b). In this case, since f and g are iso, we have inverses f^{-1} and g^{-1}, and we can readily see that $f^{-1} \circ g^{-1}$ is an inverse for $g \circ f$.

(d) Since f and $g \circ f$ are iso, there are inverses f^{-1} and $(g \circ f)^{-1}$. From this, we have that

$$\mathrm{id}_A = (g \circ f)^{-1} \circ (g \circ f),$$

and composing by f^{-1} on the right on both sides of this equation gives us

$$f^{-1} = ((g \circ f)^{-1} \circ (g \circ f)) \circ f^{-1}.$$

Now we see that

$$\begin{aligned} f^{-1} &= ((g \circ f)^{-1} \circ (g \circ f)) \circ f^{-1} \\ &= (g \circ f)^{-1} \circ ((g \circ f) \circ f^{-1}) \\ &= (g \circ f)^{-1} \circ (g \circ (f \circ f^{-1})) \\ &= (g \circ f)^{-1} \circ (g \circ \mathrm{id}_B) = (g \circ f)^{-1} \circ g. \end{aligned}$$

Finally, by composing f on the left on both sides of the equation we get

$$\mathrm{id}_B = (f \circ (g \circ f)^{-1}) \circ g,$$

and so $f \circ (g \circ f)^{-1}$ is a left-inverse for g.

But it is also clearly a right-inverse:

$$g \circ (f \circ (g \circ f)^{-1}) = (g \circ f) \circ (g \circ f)^{-1} = \mathrm{id}_C$$

so g has an inverse and is, therefore, an isomorphism.

(e) Since g and $g \circ f$ are iso, there are inverses g^{-1} and $(g \circ f)^{-1}$. From this, we have that

$$\mathrm{id}_C = (g \circ f) \circ (g \circ f)^{-1},$$

and composing by g^{-1} on the left on both sides of this equation gives us

$$g^{-1} = g^{-1} \circ ((g \circ f) \circ (g \circ f)^{-1}).$$

Now we see that

$$\begin{aligned} g^{-1} &= g^{-1} \circ ((g \circ f) \circ (g \circ f)^{-1}) \\ &= (g^{-1} \circ (g \circ f)) \circ (g \circ f)^{-1} \\ &= ((g^{-1} \circ g) \circ f) \circ (g \circ f)^{-1} \\ &= ((\mathrm{id}_B) \circ f) \circ (g \circ f)^{-1} = f \circ (g \circ f)^{-1} \end{aligned}$$

Finally, by composing g on the right on both sides of the equation we get

$$\mathrm{id}_B = f \circ ((g \circ f)^{-1} \circ g),$$

so f has an inverse and is, therefore, an isomorphism.

□

And finally we end this section with a definition.

Remark 1.3.37

Notice that for finite sets, $\#A$ is precisely the formalization of the intuitive notion of "number of elements of A".