Set Theory for High School Students: Rough Notes

Anguel Nikolov
highschoolsettheory@gmail.com

2026

These notes represent the rough structure of a high school course on set theory. They are not meant to be used as a primary learning resource and are not complete.

1 Introduction

- Language check: how comfortable are you with English?
- Relax and ask questions, please!

Practicing mathematics means studying the consequences achieved by *logical inference* from a collection of assumptions, called *axioms*. We use our intuition and curiosity to decide which assumptions to study, what consequences to seek and *how* to reach them, but not to *determine* if something is indeed a consequence of the assumptions. That may only be determined by demonstrating how it follows from the assumptions by the logical rules of inference.

The axioms we'll study in this course are those of set theory. Initially, set theory was developed as the study of different infinities: comparing them, ordering them and studying their relationship with other mathematical objects. It was later found that almost all of mathematics can be meaningfully developed within set theory itself. We will experience both of these aspects over time.

As interesting as set theory is in itself, studying it in this course will only be our secondary objective. We will use it as an occasion for demonstrating and practicing the mathematical method described above, which will be our main goal. In other words, we will focus on the mathematical rather than on the set-theoretical.

1.1 Sets: Intuition, Examples and Illustrations

A set is any *unordered* collection of *distinct* objects which *we can consider as an object itself*. There need not be any connection between these objects other than the fact that we are considering them together. An object can either be an element of a given set or not be an element of a given set and that is all that we are able to state about its relationship to that set. In other words, we can only distinguish sets by looking at which elements belong to them.

Here is how Cantor, the father of set theory, put it:

> „Unter einer ‚Menge' verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten m unserer Anschauung oder unseres Denkens (welche die ‚Elemente' von M genannt werden) zu einem Ganzen."

Translated to English,

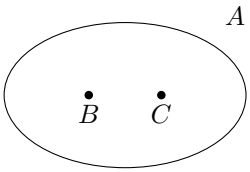
> "By a 'set' we understand every collection M into a whole of definite, distinct objects m of our intuition or our thought (which are called the 'elements' of M)."

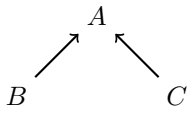
Some examples of sets are:

- the set of all odd natural numbers,
- the set whose elements are the numbers 1, 2, 17, and 823409,
- the empty set,
- the set of the sets of natural numbers, containing the number 5.

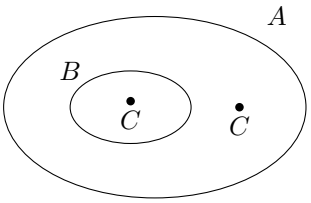
Here are a couple of ways to illustrate sets. Note that imagining sets as bags of objects can be misleading.

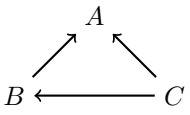
The set A with elements B and C.





The set A with elements B and C, but C is also an element of B.





All objects we study will be sets. The question if x is an element of y will be applicable to any two objects x and y of our universe. There will be no special proto-elements, which are not themselves sets.

This may seem too restrictive. For example, we will study sets of natural numbers, so shouldn't natural numbers also be objects in our universe? We will find that the natural numbers (and all other objects we may want to study) can in fact be implemented from sets. That is to say that we can find sets which, with some careful interpretation, behave just like the natural numbers.

The advantage of this is simplicity. We need not specify different axioms for sets and for proto-elements and can use the results we prove for all objects.

1.2 Language of Set Theory

To apply the mathematical method, we need to use a language to express our axioms and conclusions. An obvious choice may be to use a natural language like English or Bulgarian. However, natural languages suffer from significant disatvantages:

- it's cumbersome to define what a valid statement is;
- the meaning of valid statements is often ambiguous;
- there is a lot of self-reference;
- there is additional subtext, which can only complicate our work.

Therefore, we'll also introduce a special mathematical language for our purpose. We'll still use natural language alongside it, but we'll make sure to be confident that we can formulate all our axioms, proofs and conclusions in the mathematical language as well.

An expression in our mathematical language will be called a *formula*. For each formula, we'll care about three things:

- whether or not it's written correctly;
- what it means;
- whether or not we can prove that it's true.

1.2.1 Names for Objects

Variables are the names we ordinarily give to the objects we talk about in our formulas. For variables we'll usually use Latin letters, lowercase or uppercase: $a, b, c, \ldots, x, y, z, A, B, C, \ldots, X, YZ$. We may also combine these with additional symbols like subscripts, primes or seconds: x_0, x', x'', A_1, B_2. Later on, we'll also introduce Greek letters for special variables: α, β, γ and so on.

Later, we'll introduce another kind of name, one which refers to a specific object we defined somewhere, a sort of abbreviation. For this we'll use the symbol '$\leftrightharpoons$' and make sure that the objects we introduce names for are unique.

1.2.2 Primitive Properties

There are only two primitve properties of sets in our language: belonging ('$\in$') and equality ('$=$').

As we progress, we'll define additional properties, but they will only be abbreviations. For this we'll also use the symbol '$\leftrightharpoons$'. Any formula we write with abbreviations can be rewritten without them without changing it's meaning or provability.

To express that the set x is an element of the set y, we write the formula:

$$x \in y$$

When we say that two objects x and y are equal, what we mean is that whenever we mention x in that context, we can substitute it for y and maintain the truth value of the formula. We express this with

$$x = y$$

Primitive properties can be formulated for any two objects and indeed for an object with itself: $x = x$ and $x \in x$ are both correctly written formulas.

1.2.3 Logical Connectives

If φ and ψ are formulas, then we can also construct the compound formula whose meaning is:

- "φ is not true", called the *negation* of φ and written

$$\neg\varphi$$

- "both φ is true and ψ is true", called the *conjunction* of φ and ψ and written

$$\varphi \,\&\, \psi$$

- "at least one of φ or ψ is true", called the *disjunction* of φ and ψ and written

$$\varphi \vee \psi$$

- "whenever φ is true, ψ is also true" or "if φ, then ψ", called the *implication* of φ and ψ and written

$$\varphi \to \psi$$

- "φ is true exactly when ψ is true" or "φ, if and only if ψ", called the *equivalence* of φ and ψ and written

$$\varphi \leftrightarrow \psi$$

These connectives may be expressed in terms of one another. For example $\varphi \leftrightarrow \psi$ is the same as $(\varphi \rightarrow \psi)$ & $(\psi \rightarrow \varphi)$ and $\varphi \vee \psi$ is the same as $\neg(\neg\varphi \,\&\, \neg\psi)$.

We'll now also introduce our first abbreviations:

$$x \notin y \leftrightharpoons \neg(x \in y)$$

and

$$x \neq y \leftrightharpoons \neg(x = y)$$

1.2.4 Quantifiers

The final construction of our language are the two quantifiers. They allow us to say that something is true for certain values of some variable.

The *universal quantifier* ('$\forall$') tells us that some formula φ is true for all values of x. This is written as

$$\forall x\varphi$$

and can be read out as "for all x, φ".

The *existential quantifier* ('$\exists$') tells us that some formula φ is true for at least one value of x. This is written as

$$\exists x\varphi$$

and can be read out as "There exists x, such that ϕ".

The two quantifiers can also be expressed in terms of one another: $\forall x\varphi$ is the same as $\neg\exists x(\neg\varphi)$ and $\exists x\varphi$ is the same as $\neg\forall x(\neg\varphi)$

Some very useful abbreviations when working with quantifiers are the following:

- $(\forall x \in A)\varphi \leftrightharpoons \forall x(x \in A \rightarrow \varphi)$
- $(\exists x \in A)\varphi \leftrightharpoons \exists x(x \in A \,\&\, \varphi)$

These allow us to say that φ is true for all elements of a set A or for at least one of them, respectively.

1.2.5 Example Formulas

- $\forall x(x = x)$
- $\forall x(x \neq x \rightarrow x \in y)$
- $\exists y(x \neq y)$
- $\forall y \forall z((z \in y \,\&\, y \in x) \rightarrow z \in x)$
- $\forall x(x \notin y)$
- $\exists A \exists x(\forall y(y \notin x) \,\&\, x \in A \,\&\, \forall z(z \in A \rightarrow \exists z'(\forall u(u \in z' \leftrightarrow u = z \vee u \in z) \,\&\, z' \in A)))$

1.2.6 Set-Theoretical Properties

We'll call a formula, which leaves at least one variable unbound by quantifiers a *set-theoretical property* of that variable. Typically, this formula will be satisfied for some values of the variable and not for others. We say that the values for which the formula is satisfied are the objects possesing this property.

Here are some examples of set-theoretical properties:

- $x \in x$
- $x \neq x$
- $\exists y(y \in x)$
- $\forall y(y \in x \leftrightarrow y \in y)$

Usually we'll denote a formula φ, with a free variable x as $\varphi(x)$ when we want to look at it as the property φ for the object x.

We can also have properties that involve more than one variable. For example, if ψ is the formula $\forall z(z \in y \leftrightarrow (z = x \vee z \in x))$, we could refer to it as $\psi(x, y)$ when we think of it as a property of x and y.

1.3 Naive Set Theory

Suppose we naively give axioms for set theory based on the intuition given earlier (a set is any collection of objects which we consider as an object itself) and on the language we introduced.

Axiom of Schema of Comprehension (FALSE): For every set-theoretical property $\varphi(x)$, there is a set of all the objects which satisfy the property, i.e.

$$\exists A \forall x(x \in A \leftrightarrow \varphi(x))$$

This will unfortunately will be unsuccessful. The reason for this is known as *Russel's paradox*. Let's consider the axiom schema for the property

$$\varphi(x) \leftrightharpoons x \notin x$$

i.e. there is a set R, such that

$$\forall x(x \in R \leftrightarrow x \notin x)$$

Is it true that $R \in R$?

Case 1: If $R \in R$, then as an element of R, R satisfies the property of all elements of R, namely $R \notin R$, which is a contradiction.

Case 2: If $R \notin R$, then the condition $(R \notin R)$ for being an element of R is satisfied by R, so $R \in R$, which is also a contradiction.

The axiom of comprehension leads to a contradiction. Not every collection of objects, which satisfy a set-theoretical property is a set. Some are in a sense too big. We need better axioms and must be very careful about what we find to be a set.

2 Zermelo–Fraenkel Axioms

Ax0 (Existence): There exists a set: $\exists x(x = x)$

Ax1 (Extensionality): Two sets are equal if they have the same elements:

$$\forall x \forall y(x = y \leftrightarrow \forall z(z \in x \leftrightarrow z \in y))$$

Equality means having identical behaviour w.r.t. all properties and operations of the language. This is how we prove equality.

Ax2 (Pairing): For any two sets, there exists a set containing exactly those two elements:

$$\forall x \forall y \exists z \forall w(w \in z \leftrightarrow w = x \vee w = y)$$

We now have our first mechanism for proving the existence of new sets.

Prop: There is only one such set. That is to say: if for some sets x and y, z and z' are such that $\forall w(w \in z \leftrightarrow w = x \vee w = y)$ and $\forall w(w \in z' \leftrightarrow w = x \vee w = y)$, then $z = z'$ (z and z' are two names for the same set).

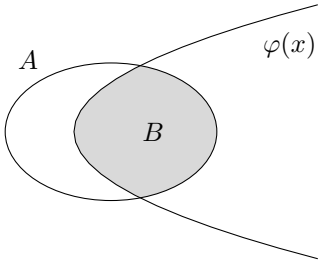
Proof. For an arbitrary w, $w \in z$ if and only if $w = x \vee w = y$. $w = x \vee w = y$ if and only if $w \in z'$. Therefore $w \in z \leftrightarrow w \in z'$. By the axiom of extensionality, $z = z'$. □

Since this set exists and is unique, we will denote it $\{x, y\}$. We denote $\{x\} \leftrightharpoons \{x, x\}$ and call it the singleton of x. $\forall y(y \in \{x\} \leftrightarrow y = x \vee y = x)$, which is to say $\forall y(y \in \{x\} \leftrightarrow y = x)$

Ax3 (Separation Schema): For any set A and any set-theoretical property $\varphi(x)$, which doesn't contain free instances of B, there exists a set B containing all elements of A that satisfy the property:

$$\forall A \exists B \forall x(x \in B \leftrightarrow x \in A \,\&\, \varphi(x))$$

The seperation schema is not an axiom itself, but rather a rule by which axioms may be identified. This is what matters when writing a proof: being able to recognize if something is an axiom.



There may not be a set that contains all x which satisfy $\varphi(x)$, but we can intersect this collection with any set and get a set.

Prop: There is only such set, i.e. if B and B' are sets and $\forall x(x \in B \leftrightarrow x \in A \,\&\, \varphi(x))$ and $\forall x(x \in B' \leftrightarrow x \in A \,\&\, \varphi(x))$, then $B = B'$.
Proof. Let x be arbitrary. $x \in B$ if and only if $x \in A \,\&\, \varphi(x)$, but $x \in A \,\&\, \varphi(x)$ if and only if $x \in B'$. Therefore $x \in B \leftrightarrow x \in B'$. By the axiom of extensionality, $B = B'$. □

Given this uniqueness, we will denote this set with the notation $\{x \mid x \in A \,\&\, \varphi(x)\}$. The notation $\{x \in A \mid \varphi(x)\}$ is commonly used, but we'll avoid it to emphasise the use of the axiom.

Examples

- Separating by $x = x$ yields the same set.
- What does separating by $x \neq x$ yield? $\emptyset$.

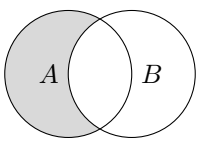
There is only one set x with the property $\forall y(y \notin x)$ and it is called the *empty set* and denoted $\emptyset$.

Prop: There is no set containing all sets.
Proof. Suppose for the sake of contradiction, that V is such that $\forall x(x \in V)$. Using the separation schema with the property $x \notin x$, we achieve $R \leftrightharpoons \{x \mid x \in V \,\&\, x \notin x\}$. However, since $x \in V$ is true for any x, $x \in R \leftrightarrow x \notin x$, which is Russel's paradox. □

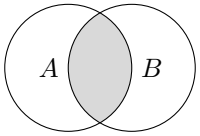
What if we have a property of more than one variable? For example, $y \in x$. Separate all x from A, which have some given y as an element.

$$\underline{\forall y} \forall A \exists B \forall x(x \in B \leftrightarrow x \in A \,\&\, y \in x)$$

Def (Set Difference): $A \setminus B \leftrightharpoons \{x \mid x \in A \,\&\, x \notin B\}$



Def (Intersection): $A \cap B \leftrightharpoons \{x \mid x \in A \,\&\, x \in B\}$



Def (Subset): $A \subseteq B \leftrightharpoons \forall x(x \in A \to x \in B)$

Prop: $\forall x(\emptyset \subseteq x)$

Prop: $\forall x(x \subseteq \emptyset \to x = \emptyset)$

Prop: For any sets A and B:

- $(A \subseteq B) \,\&\, (B \subseteq A) \leftrightarrow A = B$
- $A \setminus B \subseteq A$
- $A \cap B \subseteq B$

Notice how we've now introduced multiple new symbols to the language, but these are all defined in terms of the original two.

Can we define union in a similar way, do we know it exists from the axioms so far?

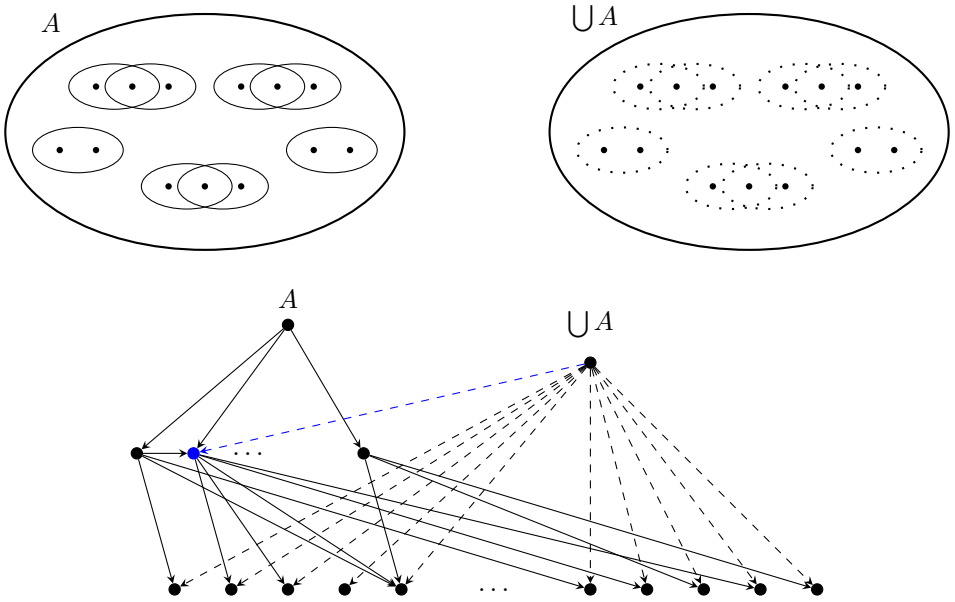
Using only the axioms so far, We can prove the existence of infinitely many sets (e.g., $\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \ldots$), but they all have at most two elements.

Ax4 (Union): For any set A, there exists a set B whose elements are the elements of the elements of A:

$$\forall A \exists B \forall x(x \in B \leftrightarrow \exists y(y \in A \,\&\, x \in y))$$

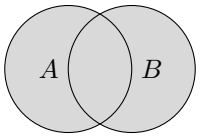
Prop: This set is unique.

We denote this set with $\bigcup A$.



Note that elements of A can also be elements of $\bigcup A$ (emphasized in blue on the figure), as long as they are also *elements of elements* of A. For example, consider $A \leftrightharpoons \{\emptyset, \{\emptyset\}\}$. In this case, $\emptyset \in A$ and $\emptyset \in \bigcup A$.

Def (Union of Two Sets): $A \cup B \leftrightharpoons \bigcup\{A, B\}$



Prop: $\forall x(x \in A \cup B \leftrightarrow x \in A \vee x \in B)$

Prop: $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$

Prop: $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

Prop: $A \cap B \subseteq A \subseteq A \cup B$

Can we define the intersection of a set of sets in a similar way to the union of a set, $\bigcup A$?

Prop: Let A be a nonempty set. Then, there exists a set B, containing exactly those objects which are elements of all elements of A.

$$\exists B \forall x(x \in B \leftrightarrow \forall y(y \in A \rightarrow x \in y))$$

Proof using separation.

Why is B unique?

Def (Intersection of A): We shall denote this set B with $\bigcap A$.

Now we can construct a set whose elements are exactly the sets a, b and c and more generally, $a_1, \ldots, a_n$:

$$\{a, b, c\} \leftrightharpoons \{a, b\} \cup \{c\}$$
$$\{a_1, \ldots, a_n\} \leftrightharpoons \{a_1\} \cup \cdots \cup \{a_n\}$$

We have achieved sets with more than two elements!

Lemma: $\{x, y\} = \{x, z\} \leftrightarrow y = z$

Note now that $\{x, y\} = \{y, x\} = \{x, y, y, x\}$ (proof).

Is it possible to prove the existence of an object where the order of x and y matters?

We'd like to be able to define $\langle x, y\rangle$ for all sets x and y, such that

$$\forall x \forall y \forall u \forall v(\langle x, y\rangle = \langle u, v\rangle \leftrightarrow x = u \,\&\, y = v)$$

Is that possible?

Def (Ordered Pair, Kuratowski): $\langle a, b\rangle \leftrightharpoons \{\{a\}, \{a, b\}\}$

Prop: This definition satisfies the above property (use the lemma).

The axioms presented so far can show the existence of sets based on *other sets* (pairs) or their *elements* (separation and union). In contrast, the next axiom can collect all *subsets* of a set into a set.

Ax5 (Power Set): For any set A, there exists the set of all subsets of A:

$$\forall A \exists B \forall x(x \in B \leftrightarrow x \subseteq A)$$

Is this set unique?

We denote this set by $\mathcal{P}(A)$.

Can $\mathcal{P}(A) = \emptyset$ for some set A?

Does there exist a set A such that $A \subseteq \mathcal{P}(A)$?

Does there exist a set A such that $\mathcal{P}(A) \subseteq A$? (to prove not, separate from A by the Russell property).

2.1 Cartesian Product, Relations, Functions

Let A and B be arbitrary sets. Is there a set, whose elements are exactly the ordered pairs with a first element from A and a second from B?

For any x and y, such that $x \in A$ and $y \in B$, we have that $\{x, y\} \subseteq A \cup B$ and so $\{x, y\} \in \mathcal{P}(A \cup B)$. Further, $\{x\} \subseteq A \cup B$ and so $\{x\} \in \mathcal{P}(A \cup B)$. Thus, $\{\{x\}, \{x, y\}\} \subseteq \mathcal{P}(A \cup B)$ and $\{\{x\}, \{x, y\}\} \in \mathcal{P}(\mathcal{P}(A \cup B))$.

Therefore:

$$\forall x \forall y (x \in A \,\&\, y \in B \to \langle x, y \rangle \in \mathcal{P}(\mathcal{P}(A \cup B))) \tag{1}$$

$\mathcal{P}(\mathcal{P}(A \cup B))$ also contains many sets, which are not the ordered pairs we need. Let's separate from $\mathcal{P}(\mathcal{P}(A \cup B))$ all $\langle x, y \rangle$, where $x \in A$ and $y \in B$:

Def (Cartesian Product): $A \times B \leftrightharpoons \{p \mid p \in \mathcal{P}(\mathcal{P}(A \cup B)) \,\&\, \exists x \exists y (x \in A \,\&\, y \in B \,\&\, p = \langle x, y \rangle)\}$

Prop: For any sets A and B

$$\forall x \forall y (\langle x, y \rangle \in A \times B \leftrightarrow x \in A \,\&\, y \in B)$$

(Prove using the definition of cartesian product for one direction and (1) for the other.)

Prop: For any sets A, B and C:

- $A \times B = \emptyset \leftrightarrow (A = \emptyset \vee B = \emptyset)$
- $A \neq \emptyset \,\&\, B \neq \emptyset \to -y(A \times B = B \times A \leftrightarrow A = B)$
- $A \times (B \cup C) = (A \times B) \cup (A \times C)$
- $A \times (B \cap C) = (A \times B) \cap (A \times C)$

Def (Binary Relation): A *binary relation* is a set of ordered pairs.

In other words, R is a binary relation if for some sets X and Y, $R \subseteq X \times Y$.

We will write

$$xRy \leftrightharpoons \langle x, y \rangle \in R$$

For example, if $\prec$ is a binary relation:

$$\prec \subseteq A \times B$$

then

$$x \prec y \leftrightharpoons \langle x, y \rangle \in \prec$$

Examples:

- $\{p \mid (\exists x \in \mathbb{N})(\exists y \in \mathbb{N})(p = \langle x, y \rangle \,\&\, x < y)\}$
- $\{p \mid (\exists x \in \mathbb{N})(\exists y \in \mathbb{N})(p = \langle x, y \rangle \,\&\, x \nmid y)\}$
- $\{p \mid (\exists x \in \mathbb{N})(\exists y \in \mathbb{N})(p = \langle x, y \rangle \,\&\, y = x + 1)\}$
- For any set A, the identity relation on that set is defined as:

$$id_A \leftrightharpoons \{p \mid p \in A \times A \,\&\, (\exists x \in A)(\langle x, x \rangle = p)\}$$

$$x \; id_A \; y \leftrightarrow x = y$$

- The empty relation: $\emptyset$, i.e. the identity relation on the empty set: $id_\emptyset$.

We will think of a binary relation in two ways:

- as a property which may hold for two given objects
- as non-deterministic transformation

Is it possible to collect all the first components of a binary relation into a set? What about the second components? Suppose R is a binary relation and $\langle x, y\rangle \in R$. This is the same as to say $\{\{x\}, \{x, y\}\} \in R$ and so $\{\{x\}, \{x, y\}\} \subseteq \bigcup R$. Thus, $\{x, y\} \in \bigcup R$ and so $\{x, y\} \subseteq \bigcup\bigcup R$, which means that

$$x \in \bigcup\bigcup R \tag{2}$$

and

$$y \in \bigcup\bigcup R \tag{3}$$

Def (Domain): If R is a binary relation, then the *domain* of R, $Dom(R)$, is the set of all first components of the ordered pairs in R:

$$x \in Dom(R) \leftrightarrow \exists p \exists y (p \in R \,\&\, p = \langle x, y\rangle)$$

Why does such a set exist to begin with? Because due to (2), we can separate it from $\bigcup\bigcup R$:

$$Dom(R) \leftrightharpoons \{x \mid x \in \bigcup\bigcup R \,\&\, \exists p \exists y (p \in R \,\&\, p = \langle x, y\rangle)\}$$

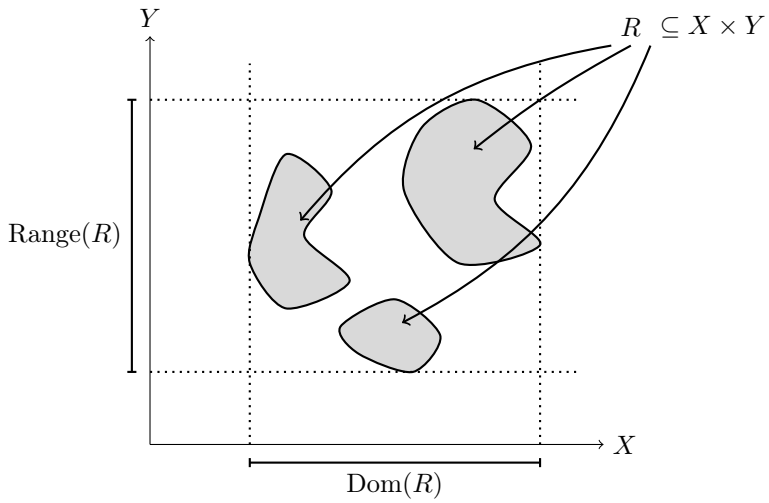
Similarly, due to (3), we can separate the second components as follows.

Def (Range): If R is a binary relation, then the *range* of R, $Range(R)$, is the set of all second components of the ordered pairs in R.

$$Range(R) \leftrightharpoons \{y \mid y \in \bigcup\bigcup R \,\&\, \exists p \exists x (p \in R \,\&\, p = \langle x, y\rangle)\}$$

$$x \in Dom(R) \leftrightarrow \exists p \exists y (p \in R \,\&\, p = \langle x, y\rangle)$$

If R is binary relation, then $R \subseteq X \times Y$ for some sets X and Y. $Dom(R) \subseteq X$ and $Range(R) \subseteq Y$.



$Dom(\emptyset) = \emptyset$ and $Range(\emptyset) = \emptyset$. $Dom(id_R) = Range(id_R) = A$.

Def (Composition): Let R_1 and R_2 be relations.

$$R_2 \circ R_1 \leftrightharpoons \{p \mid p \in Dom(R_1) \times Range(R_2) \,\&\, \exists x \exists y \exists z (p = \langle x, z\rangle \,\&\, \langle x, y\rangle \in R_1 \,\&\, \langle y, z\rangle \in R_2)\}$$

In other words,

$$\langle x, y\rangle \in R_2 \circ R_1 \leftrightarrow \exists z (\langle x, y\rangle \in R_1 \,\&\, \langle y, z\rangle \in R_2)$$

or with infix notation,

$$x(R_2 \circ R_1)y \leftrightarrow \exists z (xR_1y \,\&\, yR_2z)$$

Prop: Let R, S and T be arbitrary relations.

- $id_{Range(R)} \circ R = R \circ id_{Dom(R)} = R$
- $\emptyset \circ R = R \circ \emptyset = \emptyset$
- $R \circ (S \circ T) = (R \circ S) \circ T$
- $R \circ (S \cup T) = (R \circ S) \cup (R \circ T)$
- $R \circ (S \cap T) \subseteq (R \circ S) \cap (R \circ T)$

Is it true that $R \circ (S \cap T) = (R \circ S) \cap (R \circ T)$?
What about $R \circ (S \setminus T) = (R \circ S) \setminus (R \circ T)$?

Def (Inversion): Given a relation R,

$$R^{-1} \leftrightharpoons \{p \mid p \in Range(R) \times Dom(R) \,\&\, \exists x \exists y (p = \langle x, y \rangle \,\&\, \langle y, x \rangle \in R)\}$$

In other words,

$$xR^{-1}y \leftrightarrow yRx$$

Prop: Let R and S be relations and A be a set.

- ${id_A}^{-1} = id_A$
- $(R^{-1})^{-1} = R$
- $(R \cup S)^{-1} = R^{-1} \cup S^{-1}$
- $(R \cap S)^{-1} = R^{-1} \cap S^{-1}$
- $(R \circ S)^{-1} = S^{-1} \circ R^{-1}$

We will come back to relations when we talk about well-orders and perhaps also introduce some other kinds of relations like equivalence relations.

We have considered only binary relations thus far and those will be enough for the needs of this course, but general relations can of course be easily defined as follows.

Def (Tuple): We can extend ordered pairs to tuples:

$$\begin{aligned} \langle a, b, c \rangle &\leftrightharpoons \langle \langle a, b \rangle, c \rangle \\ \langle a, b, c, d \rangle &\leftrightharpoons \langle \langle a, b, c \rangle, d \rangle \\ &\vdots \\ \langle a_1, \ldots, a_{n+1} \rangle &\leftrightharpoons \langle \langle a_1, \ldots, a_n \rangle, a_{n+1} \rangle \end{aligned}$$

$\langle a_1, \ldots, a_n \rangle$ is called an *n-tuple*.

We define

$$A_1 \times \cdots \times A_{n+1} \leftrightharpoons (A_1 \times \cdots \times A_n) \times A_{n+1}$$

Thus, $A_1 \times \cdots \times A_n$ is the set of all tuples $\langle a_1, \ldots, a_n \rangle$, such that $a_1 \in A_1 \,\&\, \ldots \,\&\, a_n \in A_n$.

We also let $A^n \leftrightharpoons \underbrace{A \times \cdots \times A}_{n \text{ times}}$

Def (n-ary Relation): An *n-ary relation* is a set of n-tuples.

R is a relation *on* X if $R \subseteq X^n$.

Def (Function): Suppose A and B are sets. f is a function from A to B means that $f \subseteq A \times B$ and $Dom(f) = A$, $Range(f) \subseteq B$ and

$$\forall x \forall y \forall y' (\langle x, y \rangle \in f \,\&\, \langle x, y' \rangle \in f \rightarrow y = y')$$

We denote this by $f\colon A \to B$ or by $Func(f)$.

For any $x \in Dom(f)$, there exists a unique y, such that $\langle x, y\rangle \in f$. Therefore, we can denote this y as $f(x)$. We may sometimes, more rarely, also denote this y as f_x.

Prop: If $f\colon A \to B$ and $g\colon B \to C$, then $g \circ f\colon A \to C$.

Is it true that if $f\colon A \to B$, then $f^{-1}\colon Range(f) \to A$?

Def (Function Restriction): Suppose $f\colon A \to B$ and X is a set (typically $X \subseteq A$, but not necessarily).

$$f \restriction X \leftrightharpoons f \cap (X \times Range(f))$$

Prop: If $f\colon A \to B$ and $X \subseteq A$, then $f \restriction X\colon X \to B$

Prop: $f \restriction X = f \circ id_X$

Def (Function Extension): Suppose f and g are functions. g is an extension of f means that $f \subseteq g$. That is to say

$$Dom(f) \subseteq Dom(g)\ \&\ (\forall x \in Dom(f))(f(x) = g(x))$$

Def (Image): For a given function $f\colon A \to B$ and a set X, we define the image of X by f:

$$f[X] \leftrightharpoons \{y \mid y \in B\ \&\ (\exists x \in X)(f(x) = y)\}$$

Is it true that $f[X \cup Y] = f[X] \cup f[Y]$? What about $f[X \cap Y] = f[X] \cap f[Y]$?

Def (Inverse Image): For any function $f\colon A \to B$ and a set Y, we define the inverse image of Y by f:

$$f_{-1}[Y] \leftrightharpoons \{x \mid x \in A\ \&\ f(x) \in Y\}$$

2.2 The Remaining Axioms

We won't need to use the remaining axioms just yet, but we will list them out for the sake of completeness and return to them later.

Before we introduce the axiom of infinity, let's consider whether the axioms so far might be contradictory. We can in fact reduce the consistency of all axioms aside from the axiom of infinity (even the ones not yet introduced) to the conisistency of the the natural numbers. For any two natural numbers n and m,

$$n \in m \leftrightharpoons \text{“}m \text{ has a 1 in the } n\text{-th position of its binary representation”}$$

Def (Successor): If x is a set, then $S(x) \leftrightharpoons x \cup \{x\}$.

The key properties of $S(x)$ are that $x \subseteq S(x)$ and $x \in S(x)$. $S(x)$ is the set whose elements are the elements of x, but which also contains x itself as an element.

Note that as long as $x \notin x$, $S(x) \neq x$.

Ax6 (Infinity): There exists an infinite set.

$$\exists A(\emptyset \in A\ \&\ (a \in A \to S(a) \in A))$$

If we take A_0 to be a witness for the above. Since

$$\emptyset \in A_0$$

we have

$$S(\emptyset) = (\emptyset \cup \{\emptyset\}) = \{\emptyset\} \in A_0$$

and therefore

$$S(\{\emptyset\}) = (\{\emptyset\} \cup \{\{\emptyset\}\}) = \{\emptyset, \{\emptyset\}\} \in A_0$$

and therefore

$$S(\{\emptyset, \{\emptyset\}\}) = (\{\emptyset, \{\emptyset\}\} \cup \{\{\emptyset, \{\emptyset\}\}\}) = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} \in A_0$$

and so on. Note that all these sets are different.

Let us denote $0 \leftrightharpoons \emptyset$. Further, let $1 \leftrightharpoons S(0)$, $1 = \{0\}$. Let $2 = S(1)$, i.e. $2 = 0 \cup \{1\} = \{0, 1\}$. Continuing, $3 \leftrightharpoons S(2)$ and $3 = \{0, 1, 2\}$. In general,

$$n + 1 \leftrightharpoons S(n)$$

and

$$n + 1 = \{0, 1, \ldots, n\}$$

In these terms, the axiom of infinity states that there exists a set containing $0, 1, 2, \ldots$.

We'll come back to the infinity axiom when we introduce ordinal numbers.

Ax7 (Regularity): Every non-empty set A has an $\in$-minimal element, i.e. this is an element of A which itself contains none of the elements of A.

$$\forall A(A \neq \emptyset \to (\exists u \in A)(\forall x \in A)(x \notin u))$$

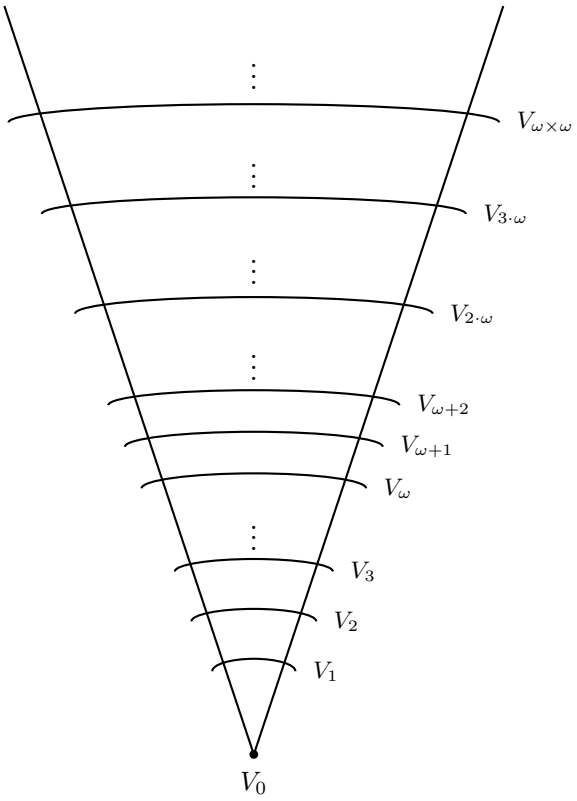
Equivalently,

$$\forall A(A \neq \emptyset \to (\exists u \in A)(u \cap A = \emptyset))$$

Prop: Some consequences of this are:

- $\forall x(x \notin x)$
- $\forall x \forall y \neg(x \in y \,\&\, y \in x)$
- $\forall x \forall y \forall z \neg(x \in y \,\&\, y \in z \,\&\, z \in x)$
- More generally:
$$\forall x_0 \forall x_1 \ldots \forall x_n \neg(x_0 \in x_1 \in \cdots \in x_n \in x_0)$$
- Although we can't yet formulate it with a formula, it's not possible to have an unlimited chain of sets like so:
$$\cdots \in x_2 \in x_1 \in x_0$$

The axiom of regularity shows that certains kinds of sets do not exist (like the ones listed above). This doesn't lead to contradictions and is not needed for the results in this course and in fact for the development of all of ordinary mathematics. However, it is very useful in the metamathematics of set theory, in the construciton of models. It allows us to assign ranks to all sets and arrange them in a cumulative hierarchy.



Ax8 (Choice): For any set A of nonempty sets, there exists a *choice function* f that is defined on A and maps each element x of A to an element of x.

$$\forall A((\forall x \in A)(x \neq \emptyset) \to \exists f(Func(f) \,\&\, Dom(f) = A \,\&\, (\forall x \in A)(f(x) \in x)))$$

The axiom of choice is consistent with and independent from the other axioms. The second of these results earned Paul Cohen the Fields Medal. The first was proven by Kurt Gödel some time before that.

We will come back to the axiom of choice later, once we have introduced well-orders, ordinals and cardinals. We'll first use it to show that any set can be well-ordered (Zermelo's Theorem). If we have time, we'll look at some of its equivalent formulations like Zorn's Lemma and maybe some surprising consequences.

Def (Set-Theoretical Operation): A set-theoretical operation is a formula $\varphi(x, y)$, such that

$$\forall x \forall y \forall y' (\varphi(x, y) \,\&\, \varphi(x, y') \to y = y')$$

Note that we do not require that $\forall x \exists y\, \varphi(x, y)$.

Examples:

- $\varphi(x, y) \leftrightharpoons y = x \cup \{x\}$
- $\varphi(x, y) \leftrightharpoons y = x \times x$

Ax9 (Replacement Schema): If $\varphi(x, y)$ is a set-theoretical operation and A is a set, then there exists the set

$$\{y \mid x \in A \,\&\, \varphi(x, y)\}$$

In other words,

$$\forall A \exists B \forall y (y \in B \leftrightarrow (\exists x \in A)\varphi(x, y))$$

This allows us to write

$$\{\varphi(x) \mid x \in A\}$$

The replacement schema is powerful and we'll do our best to use it sparingly, only when we really need it. We'll come back to it most notably when we prove the transfinite recursion theorem.

In fact, the separation schema follows from the replacement schema. To see this, for a given set-theoretical property $\varphi(x)$, simply replace it by the set-theoretical operation $\psi(x, y) \leftrightharpoons y = x \,\&\, \varphi(x)$. We chose to mention the separation schema separately to allow us to use the replacement schema only for a limited number of results.

3 Cardinality

Def (Injection): A function $f\colon A \to B$ is an *injection*, denoted $f\colon A \rightarrowtail B$, if

$$(\forall x \in A)(\forall x' \in A)(f(x) = f(x') \to x = x')$$

Def (Surjection): A function $f\colon A \to B$ is an *surjection*, denoted $f\colon A \twoheadrightarrow B$, if $Range(f) = B$.

Def (Bijection): If $f\colon A \rightarrowtail B$ and $f\colon A \twoheadrightarrow B$, then f is a *bijection*, denoted $f\colon A \rightarrowtail\!\!\!\!\twoheadrightarrow B$.

Prop:

- $id_A\colon A \rightarrowtail\!\!\!\!\twoheadrightarrow A$.
- If $f\colon A \rightarrowtail B$, then $f^{-1}\colon Range(A) \rightarrowtail\!\!\!\!\twoheadrightarrow A$.
- If $f\colon A \rightarrowtail\!\!\!\!\twoheadrightarrow B$, then $f^{-1}\colon B \rightarrowtail\!\!\!\!\twoheadrightarrow A$.
- If $f\colon A \rightarrowtail B$ and $g\colon B \rightarrowtail C$, then $g \circ f\colon A \rightarrowtail C$.
- If $f\colon A \rightarrowtail\!\!\!\!\twoheadrightarrow B$ and $g\colon B \rightarrowtail\!\!\!\!\twoheadrightarrow C$, then $g \circ f\colon A \rightarrowtail\!\!\!\!\twoheadrightarrow C$.

Def (Non-Strict Cardinal Inequality): $\overline{\overline{A}} \leq \overline{\overline{B}} \leftrightharpoons \exists f(f\colon A \rightarrowtail B)$

Def (Cardinal Equality): $\overline{\overline{A}} = \overline{\overline{B}} \leftrightharpoons \exists f(f\colon A \rightarrowtail\!\!\!\!\twoheadrightarrow B)$

Def (Strict Cardinal Inequality): $\overline{\overline{A}} < \overline{\overline{B}} \leftrightharpoons \overline{\overline{A}} < \overline{\overline{B}} \,\&\, \overline{\overline{A}} \neq \overline{\overline{B}}$

Prop:

- $\overline{\overline{A}} = \overline{\overline{A}}$
- $\overline{\overline{A}} = \overline{\overline{B}} \to \overline{\overline{B}} = \overline{\overline{A}}$
- $\overline{\overline{A}} = \overline{\overline{B}} \,\&\, \overline{\overline{B}} = \overline{\overline{C}} \to \overline{\overline{A}} = \overline{\overline{C}}$
- $\overline{\overline{A}} \leq \overline{\overline{A}}$
- $\overline{\overline{A}} \leq \overline{\overline{B}} \,\&\, \overline{\overline{B}} \leq \overline{\overline{C}} \to \overline{\overline{A}} \leq \overline{\overline{C}}$

Theorem (Cantor): For any set X, $\overline{\overline{X}} \neq \overline{\overline{\mathcal{P}(X)}}$.

Proof. Suppose, for the sake of contradiction, that $h\colon X \rightarrowtail\!\!\!\!\twoheadrightarrow \mathcal{P}(X)$.

$$C \leftrightharpoons \{x \in X \mid x \notin h(x)\}$$

$C \subseteq X$, so $C \in \mathcal{P}(X)$. Let $c \leftrightharpoons h^{-1}(C)$. Then, $h(c) = C$ and we have the contradiction:

$$c \in C \leftrightarrow c \notin h(c) \leftrightarrow c \notin C$$

□

Prop: $\overline{\overline{X}} < \overline{\overline{\mathcal{P}(X)}}$

Proof. Let $f(x) = \{x\}$. $f\colon X \rightarrowtail \mathcal{P}(X)$, so $\overline{\overline{X}} \leq \overline{\overline{\mathcal{P}(X)}}$. Combining this with Cantor's theorem, we get $\overline{\overline{X}} < \overline{\overline{\mathcal{P}(X)}}$. □

Lemma: For any set A and any $X \in A$, $X \subseteq \bigcup A$

Def (Fixed Point): For any function f, x is a fixed point of f exactly when $f(x) = x$.

The following theorem is widely used in the theoretical treatment of the semantics of programming languages and in the theory of computation. We state it in a slightly less general form than usual, but the structure of the proof is the same.

We can think of building a structure recursively, step by step, where F represents the recursive step. Reaching a fixed point X, such that $F(X) = X$, would mean that this process has stabilized.

Theorem (Knaster-Tarski): Suppose A is a set, $F\colon \mathcal{P}(A) \to \mathcal{P}(A)$ and F is monotonically increasing, i.e. for any $X \subseteq A$ and $Y \subseteq A$

$$X \subseteq Y \to F(X) \subseteq F(Y)$$

Then there exists a greatest (w.r.t. $\subseteq$) fixed point of F, i.e.

$$(\exists M \in \mathcal{P}(A))(F(M) = M \,\&\, (\forall N \in \mathcal{P}(A))(F(N) = N \to N \subseteq M))$$

Proof. Let

$$I \leftrightharpoons \{X \mid X \in \mathcal{P}(A) \,\&\, X \subseteq F(X)\}$$

Note that since $\emptyset \in I$, $I \neq \emptyset$ and also $\bigcup I \subseteq A$. Further, note that $(X \in I) \to (F(X) \in I)$, because if $X \subseteq F(X)$, i.e. $X \in I$, by monotonicity, $F(X) \subseteq F(F(X))$, so $F(X) \in I$ as well.

We will show that $\bigcup I$ is the desired fixed point.

To show that $\bigcup I \subseteq F(\bigcup I)$, suppose $x \in \bigcup I$ is arbitrary. There must exist $X \in I$, such that $x \in X$. Since $X \in I$, we have $X \subseteq F(X)$ and so $x \in F(X)$. By the lemma, since $X \in I$, we get $X \subseteq \bigcup I$, so by using the monotonicity of F, we have $F(X) \subseteq F(\bigcup I)$. Thus $x \in F(X) \subseteq F(\bigcup I)$ or $x \in F(\bigcup I)$. Since x was chosen to be an arbitrary element of $\bigcup I$, we have

$$\bigcup I \subseteq F(\bigcup I)$$

By applying monotonicity to the above, we achieve $F(\bigcup I) \subseteq F(F(\bigcup I))$, which shows that $F(\bigcup I) \in I$. By the lemma, $F(\bigcup I) \subseteq \bigcup I$ and therefore,

$$F(\bigcup I) = \bigcup I$$

To show that this is the greatest fixed point, consider any $X_0 \in \mathcal{P}(A)$, such that $F(X_0) = X_0$. $X_0 \subseteq F(X_0)$, so $X_0 \in I$ and by the lemma, $X_0 \subseteq \bigcup I$. □

Lemma: If $f\colon A \rightarrowtail B$ and $X \subseteq A$, then $f \restriction X\colon X \rightarrowtail\!\!\!\!\!\rightarrow f[X]$

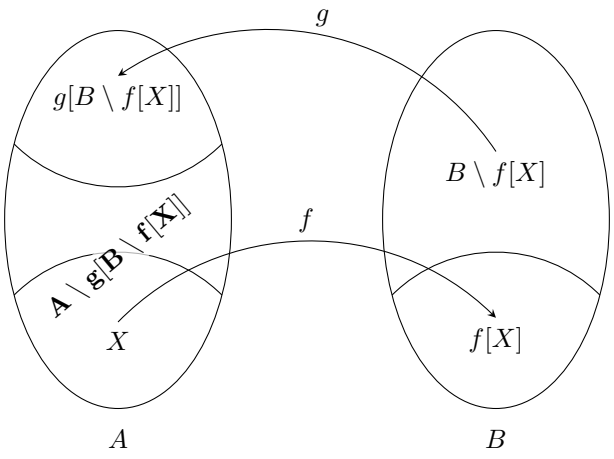
Lemma: Suppose $A \cap B = \emptyset$ and $A' \cap B' = \emptyset$. If $f\colon A \rightarrowtail\!\!\!\!\!\rightarrow A'$ and $g\colon B \rightarrowtail\!\!\!\!\!\rightarrow B'$, then

$$f \cup g\colon A \cup B \rightarrowtail\!\!\!\!\!\rightarrow A' \cup B'$$

Theorem (Cantor-Bernstein): For any sets A and B, if $\overline{\overline{A}} \leq \overline{\overline{B}}$ and $\overline{\overline{B}} \leq \overline{\overline{A}}$, then $\overline{\overline{A}} = \overline{\overline{B}}$.

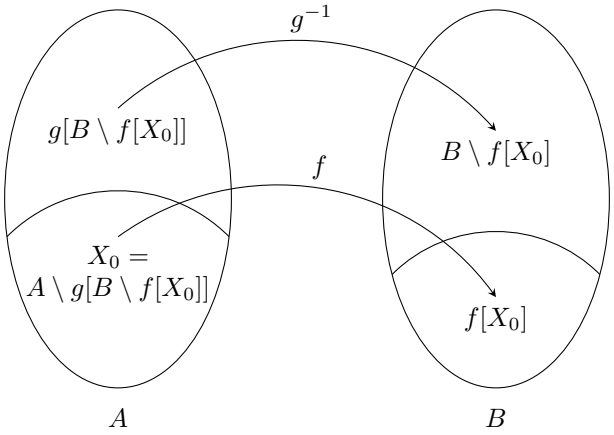
Proof. Let $f\colon A \rightarrowtail B$ and $g\colon B \rightarrowtail A$ be witnesses for the cardinal inequalities. Consider the function $F\colon \mathcal{P}(A) \to \mathcal{P}(A)$:

$$F(X) = A \setminus g[B \setminus f[X]]$$



Suppose $X \subseteq Y \subseteq A$. Then $f[X] \subseteq f[Y]$ and so $(B \setminus f[Y]) \subseteq (B \setminus f[X])$. Further, $g[B \setminus f[Y]] \subseteq g[B \setminus f[X]]$ and finally $A \setminus g[B \setminus f[X]] \subseteq A \setminus g[B \setminus f[Y]]$, i.e. $F(X) \subseteq F(Y)$.

By the Knaster-Tarski theorem, F has a fixed point. Let X_0 be a witness for that: $F(X_0) = X_0$, i.e. $g[B \setminus f[X_0]] = A \setminus X_0$



By the first lemma,

$$f \restriction X_0 \colon X_0 \rightarrowtail\!\!\!\!\!\rightarrow f[X_0]$$

and

$$g \restriction (B \setminus f[X_0]) \colon (B \setminus f[X_0]) \rightarrowtail\!\!\!\!\!\rightarrow g[B \setminus f[X_0]]$$

However, $g[B \setminus f[X_0]] = A \setminus X_0$, so

$$g \restriction (B \setminus f[X_0]) \colon (B \setminus f[X_0]) \rightarrowtail\!\!\!\!\!\rightarrow A \setminus X_0$$

Then

$$(g \restriction (B \setminus f[X_0]))^{-1} \colon A \setminus X_0 \rightarrowtail\!\!\!\!\!\rightarrow g[B \setminus f[X_0]]$$

and by the second lemma,

$$(f \restriction X_0) \cup (g \restriction (B \setminus f[X_0]))^{-1} \colon A \rightarrowtail\!\!\!\!\!\rightarrow B$$

□

Prop: $\overline{\overline{A}} < \overline{\overline{B}} \,\&\, \overline{\overline{B}} < \overline{\overline{C}} \to \overline{\overline{A}} < \overline{\overline{C}}$

Prop: $\overline{\overline{A}} = \overline{\overline{C}} \,\&\, A \subseteq B \subseteq C \to \overline{\overline{A}} = \overline{\overline{B}}$

The notation we use to denote comparisons by cardinality suggests that for each set A, there is an object $\overline{\overline{A}}$ which represents the cardinality of that set. This will indeed turn out to be true, as we will see later in the course. and we'll need the axiom of choice to find the cardinality of some sets.

For now, let us see that a naive approach will not work. Is it possible to collect all sets of a given cardinality into a set? If this is the case, this set can nicely represent a given cardinality. Unfortunately, this is the case only for the empty set, as the following proposition shows.

Prop: For any set A, there exists a set C, such that $\forall x(x \in C \leftrightarrow \overline{\overline{x}} = \overline{\overline{A}})$, if and only if $A = \emptyset$.

Proof. Suppose $A = \emptyset$. Any bijection from or to $\emptyset$ would itself be empty (having an empty domain), therefore $\emptyset$ is the only set with that cardinality and $\{\emptyset\}$ satisfies the requirement.

Conversely, let $A \neq \emptyset$ and suppose for the sake of contradiction that C is such that $\forall x(x \in C \leftrightarrow \overline{\overline{x}} = \overline{\overline{A}})$.

We'll show that $\bigcup C$ is the set of all sets, which we've shown is impossible. Let p be an arbitrary set.

If $p \in A$, then $p \in \bigcup C$ (since $A \in C$).

If $p \notin A$, then consider some $q \in A$ ($A \neq \emptyset$). $\overline{\overline{(A \setminus \{q\}) \cup \{p\}}} = \overline{\overline{A}}$, because

$$id_{A\setminus\{q\}} \cup \{\langle q, p\rangle\} \colon A \rightarrowtail\!\!\!\!\!\rightarrow (A \setminus \{q\}) \cup \{p\}$$

and thus, $(A \setminus \{q\}) \cup \{p\} \in C$. Further, $p \in ((A \setminus \{q\}) \cup \{p\})$ and so $p \in \bigcup C$. In both cases, $p \in \bigcup C$ for an arbitrary p. □

We'll show one more proof of the property to illustrate a technique that will prove useful later on.

Proof. Let $A \neq \emptyset$, suppose $\forall x(x \in C \leftrightarrow \overline{\overline{x}} = \overline{\overline{A}})$ and let p be arbitrary.

We shall call $A \times \{p\}$ the *coloring of A in the color p*. The elements of $A \times \{p\}$ are all ordered pairs of the kind $\langle a, p\rangle$, where $a \in A$.

$f \colon A \rightarrowtail\!\!\!\!\!\rightarrow A \times \{p\}$, where $f(a) = \langle a, p\rangle$, for $a \in A$. Thus $\overline{\overline{A \times \{p\}}} = \overline{\overline{A}}$ and $A \times \{p\} \in C$.

Since $A \neq \emptyset$, let $q \in A$. $\langle q, p\rangle \in A \times \{p\}$ and so $\langle q, p\rangle \in \bigcup C$. Substituting in Kuratowski's definition, $\{\{q\}, \{q, p\}\} \in \bigcup C$. Therefore $\{q, p\} \in \bigcup\bigcup C$ and $p \in \bigcup\bigcup\bigcup C$ and $\bigcup\bigcup\bigcup C$ is the set of all sets, a contradiction. □

3.1 Cardinal Operations

We have already mentioned in a lemma that if $A \cap B = \emptyset$, $A' \cap B' = \emptyset$, $\overline{\overline{A}} = \overline{\overline{A'}}$ and $\overline{\overline{B}} = \overline{\overline{B'}}$, then

$$\overline{\overline{A \cup B}} = \overline{\overline{A' \cup B'}}$$

If we look at set union of disjoint sets as an *addition operation* on cardinals, this shows its correctness: sets can be replaced by sets with equal cardinalities and the cardinality of the sum is maintained.

Suppose two cardinalities need to be added, but the particular sets we want to add are not disjoint ($A \cap B \neq \emptyset$). By using the coloring technique, we can find two other sets with the same cardinalities, which *are* disjoint: $A \times \{0\}$ and $B \times \{1\}$.

$$(A \times \{0\}) \cap (B \times \{1\}) = \emptyset$$

and

$$\overline{\overline{A}} = \overline{\overline{A \times \{0\}}} \,\&\, \overline{\overline{B}} = \overline{\overline{B \times \{1\}}}$$

Due to the correctness of cardinal addition, these new witnesses for this cardinality can be added together without issue.

Correctness also holds for the cartesian product, which can be looked upon as *cardinal multiplication*:

Prop: If $\overline{\overline{A}} = \overline{\overline{A'}}$ and $\overline{\overline{B}} = \overline{\overline{B'}}$, then

$$\overline{\overline{A \times B}} = \overline{\overline{A' \times B'}}$$

Proof. Let $f\colon A \rightarrowtail\!\!\!\!\twoheadrightarrow A'$ and $g\colon B \rightarrowtail\!\!\!\!\twoheadrightarrow B'$ be witnesses for $\overline{\overline{A}} = \overline{\overline{A'}}$ and $\overline{\overline{B}} = \overline{\overline{B'}}$ respectively. Define

$$h \leftrightharpoons \{p \mid p \in (A \times B) \times (A' \times B') \,\&\, \exists a \exists b (p = \langle\langle a, b\rangle, \langle f(a), g(b)\rangle\rangle)\}$$

If any $a \in A$ and $b \in B$, there clearly exists a unique $\langle f(a), g(b)\rangle \in A' \times B'$, so $h\colon A \times B \to A' \times B'$ and $h(\langle a, b\rangle) = \langle f(a), g(b)\rangle$.

Let $\langle a', b'\rangle \in A' \times B'$. $\langle f^{-1}(a'), g^{-1}(b')\rangle \in A \times B$ and $h(\langle f^{-1}(a'), g^{-1}(b')\rangle) = \langle f(f^{-1}(a')), g(g^{-1}(b'))\rangle = \langle a', b'\rangle$, so $Range(h) = A' \times B'$. Thus $h\colon A \times B \twoheadrightarrow A' \times B'$.

Suppose $a_1 \in A$, $a_2 \in A$, $b_1 \in B$ and $b_2 \in B$ and $h(\langle a_1, b_1\rangle) = h(\langle a_2, b_2\rangle)$. That means that $\langle f(a_1), g(b_1)\rangle = \langle f(a_2), g(b_2)\rangle$, so $f(a_1) = f(a_2)$ and $g(b_1) = g(b_2)$. By the bijectivity of f and g, $a_1 = a_2$ and $b_1 = b_2$ and so $\langle a_1, b_1\rangle = \langle a_2, b_2\rangle$. Thus $h\colon A \times B \rightarrowtail A' \times B'$. □

For the *cardinal exponentiation* operation, we will first define set exponentiation.

Def (Set Exponentiation): If A and B are sets, B^A denotes the set of all functions $f\colon A \to B$, which can be achieved from the cartesian product by separation:

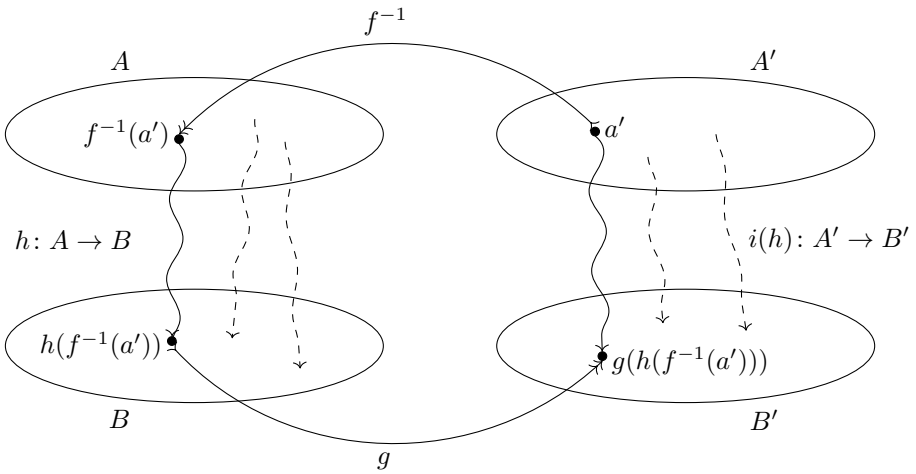
$$B^A \leftrightharpoons \{f \mid (f \in A \times B) \,\&\, (f\colon A \to B)\}$$

The correctness here is as follows:

Prop: If $\overline{\overline{A}} = \overline{\overline{A'}}$ and $\overline{\overline{B}} = \overline{\overline{B'}}$, then

$$\overline{\overline{B^A}} = \overline{\overline{B'^{A'}}}$$

Proof. To show this, we need to find the appropriate bijection $i\colon B^A \rightarrowtail\!\!\!\!\twoheadrightarrow B'^{A'}$. This means that for some $h\colon A \to B$, i will give us $h(i)\colon A' \to B'$. Suppose $f\colon A \rightarrowtail\!\!\!\!\twoheadrightarrow A'$ and $g\colon B \rightarrowtail\!\!\!\!\twoheadrightarrow B'$ are witnesses for $\overline{\overline{A}} = \overline{\overline{A'}}$ and $\overline{\overline{B}} = \overline{\overline{B'}}$ respectively.



Let

$$i \leftrightharpoons \{\langle h, h'\rangle \mid \langle h, h'\rangle \in B^A \times B'^{A'} \,\&\, (\forall a' \in A')(h'(a') = g(h(f^{-1}(a'))))\}$$

i is a function, since if for some h, h' and h'', $\langle h, h'\rangle \in i$ and $\langle h, h''\rangle \in i$, then for all $a' \in A'$,

$$h'(a') = g(h(f^{-1}(a'))) = h''(a')$$

so $h' = h''$, since $Dom(h') = Dom(h'') = A'$.

To show $Dom(i) = B^A$, consider that for any $h\colon A \to B$, there indeed exists $h'\colon A' \to B'$, such that $h'(a') = g(h(f^{-1}(a')))$ for all $a' \in A'$. This is because we can separate it from $A' \times B'$ and because $f^{-1} : A' \to A$, $h : A \to B$ and $g : B \to B'$.

Clearly, $Range(i) \subseteq B'^{A'}$, since $i \subseteq B^A \times B'^{A'}$. To prove that $Range(i) = B'^{A'}$, let $h'\colon A' \to B'$. Consider $h\colon A \to B$, such that for any $a \in A$, $h(a) = g^{-1}(h'(f(a)))$. h is well defined in this way because we can separate it from $A \times B$ and because $f : A \to A'$, $h' : A' \to B'$ and $g^{-1} : B' \to B$. For any $a' \in A'$, $f^{-1}(a') \in A$, so we can use the given definition of h to substitute $h(f^{-1}(a'))$ for $g^{-1}(h'(f(f^{-1}(a'))))$ in

$$i(h)(a') = g(h(f^{-1}(a'))) = g(g^{-1}(h'(f(f^{-1}(a'))))) = h'(a')$$

so $i(h) = h'$, since $i(h)(a') = h'(a)$ for all $a' \in A'$ and $Dom(i(h)) = Dom(h') = A'$. Thus $h' \in Range(i)$ and $i\colon B^A \twoheadrightarrow B'^{A'}$.

To show that i is injective, consider $h_1 \in B^A$ and $h_2 \in B^A$ and suppose $i(h_1) = i(h_2)$. Using this equality and by the definition of i, for all $a' \in A'$,

$$g(h_1(f^{-1}(a'))) = i(h_1)(a') = i(h_2)(a') = g(h_2(f^{-1}(a')))$$

so

$$g(h_1(f^{-1}(a'))) = g(h_2(f^{-1}(a')))$$

Let $a \in A$ be arbitrary. $f(a) \in A'$, so by the property above,

$$g(h_1(f^{-1}(f(a)))) = g(h_2(f^{-1}(f(a))))$$

Simplifying this by the fact that f is a bijection, we get $g(h_1(a)) = g(h_2(a))$. Since g is a bijection, $h_1(a) = h_2(a)$. This is for any $a \in A = Range(h_1) = Range(h_2)$, so $h_1 = h_2$. Therefore $i\colon B^A \rightarrowtail B'^{A'}$.

□

Prop: For any set A, $\overline{\overline{\mathcal{P}(A)}} = \overline{\overline{2^A}}$.

Proof. For any $X \subseteq A$, consider

$$f_X^A \leftrightharpoons ((A \setminus X) \times \{0\}) \cup (X \times \{1\})$$

If $p \in f_X^A$, $p \in (A \setminus X) \times \{0\}$ or $p \in X \times \{1\}$. Either way, $p \in A \times \{2\}$, so $f_X^A \subseteq A \times \{2\}$.

For any $a \in A$, if $a \in X$, $\langle a, 1\rangle \in X \times \{1\}$ and if $a \notin X$, $\langle a, 0\rangle \in (A \setminus X) \times \{0\}$. Either way, $a \in Dom(f_X^A)$, so $Dom(f_X^A) = A$.

Suppose $\langle a, n\rangle \in f_X^A$ and $\langle a, n'\rangle \in f_X^A$. If $n = 0$, then $a \notin X$ and so $n' = 0$. If $n = 1$, then $a \in X$ and $n' = 1$.

Thus, $f_X^A\colon A \to \{2\}$ and $f_X^A \in 2^A$.

This allows us to define

$$g \leftrightharpoons \{p \mid (p \in \mathcal{P}(A) \times 2^A)\ \&\ \exists X \exists f (p = \langle X, f\rangle\ \&\ f = ((A \setminus X) \times \{0\}) \cup (X \times \{1\}))$$

Clearly, $g \subseteq \mathcal{P}(A) \times 2^A$ and further $Dom(g) = \mathcal{P}(A)$, since $((A \setminus X) \times \{0\}) \cup (X \times \{1\})$ exists for any $X \in \mathcal{P}(A)$ and is an element of 2^A as shown above.

If $\langle X, f\rangle \in g$ and $\langle X, f'\rangle \in g$, $f = f_X^A = f'$, so $g\colon \mathcal{P}(A) \to 2^A$.

If $\langle X, f\rangle \in g$ and $\langle X', f\rangle \in g$, then $X = f^{-1}[\{1\}] = X'$. Thus $g\colon \mathcal{P}(A) \rightarrowtail 2^A$.

Finally, if $f \in 2^A$, consider that $g(f^{-1}[\{1\}]) = f$ and so $g\colon \mathcal{P}(A) \twoheadrightarrow 2^A$. □

Thus, Cantor's theorem can be formulated as $\overline{\overline{X}} < \overline{\overline{2^X}}$, for any X.

Prop: The following hold for any sets A, B, and C.

- $\overline{\overline{A \times B}} = \overline{\overline{B \times A}}$
- $\overline{\overline{(A \times B) \times C}} = \overline{\overline{A \times (B \times C)}}$
- $\overline{\overline{A^C \times B^C}} = \overline{\overline{(A \times B)^C}}$

  Proof. Let us define h by separating it from $(A^C \times B^C) \times (A \times B)^C$. Suppose $\langle f, g\rangle \in A^C \times B^C$, i.e. $f\colon C \to A$ and $g\colon C \to B$. For any $c \in C$, clearly $\langle f(c), g(c)\rangle \in A \times B$, so let $h(\langle f, g\rangle)(c) \leftrightharpoons \langle f(c), g(c)\rangle$. So defined, $h(\langle f, g\rangle)\colon C \to A \times B$ and so $h\colon (A^C \times B^C) \to (A \times B)^C$

  Suppose $h(\langle f, g\rangle) = h(\langle f', g'\rangle)$, for some $f, f' \in A^C$ and $g, g' \in B^C$. Let $c \in C$ be arbitrary. $\langle f(c), g(c)\rangle = h(\langle f, g\rangle)(c) = h(\langle f', g'\rangle)(c) = \langle f'(c), g'(c)\rangle$. Thus $f(c) = f'(c)$ and $g(c) = g'(c)$ for all $c \in C$, so $f = f'$ and $g = g'$ and $\langle f, g\rangle = \langle f', g'\rangle$, so $h\colon (A^C \times B^C) \rightarrowtail (A \times B)^C$

  For any ordered pair $\langle x, y\rangle$, define $Fst(\langle x, y\rangle) \leftrightharpoons x$ and $Snd(\langle x, y\rangle) \leftrightharpoons y$.

  Suppose $i \in (A \times B)^C$, i.e. $i\colon C \to A \times B$. We will define f and g by separating them from A^C and B^C respectively. For any $c \in C$, $i(c) \in A \times B$. We separate f and g so that $f(c) = Fst(i(c)) \in A$ and $g(c) = Snd(i(c)) \in B$. Since these were uniquely chosen, $f\colon C \to A$ and $g\colon C \to B$.

  To prove that $i = h(\langle f, g\rangle)$, take any $c \in C = Dom(i) = Dom(h(\langle f, g\rangle))$. $h(\langle f, g\rangle)(c) = \langle f(c), g(c)\rangle = \langle Fst(i(c)), Snd(i(c))\rangle = i(c)$. Thus, $h\colon (A^C \times B^C) \rightarrowtail\!\!\!\!\rightarrow (A \times B)^C$ □

- If $A \cap B = \emptyset$, then $\overline{\overline{C^{A \cup B}}} = \overline{\overline{C^A \times C^B}}$.
- $\overline{\overline{C^{A \times B}}} = \overline{\overline{(C^B)^A}}$ (currying)

  Proof. Let us separate i from $C^{A \times B} \times (C^B)^A$

  $$i \leftrightharpoons \{\langle f, g\rangle \mid f \in C^{A \times B} \,\&\, g \in (C^B)^A \,\&\, (\forall a \in A)(\forall b \in B)(f(\langle a, b\rangle) = g(a)(b))\}$$

  Let $f \in C^{A \times B}$ be arbitrary. $f\colon A \times B \to C$. We will separate g from $A \times C^B$. Let

  $$g \leftrightharpoons \{\langle a, h\rangle \mid a \in A \,\&\, h \in C^B \,\&\, (\forall b \in B)(h(b) = f(\langle a, b\rangle))\}$$

  Let $a \in A$ be arbitrary. We will find a function $h\colon B \to C$, such that $(\forall b \in B)(h(b) = f(\langle a, b\rangle)$ by separating it from $B \times C$. Let

  $$h \leftrightharpoons \{\langle b, c\rangle \mid b \in B \,\&\, c \in C \,\&\, c = f(\langle a, b\rangle)\}$$

  For any $b \in B$, $f(\langle a, b\rangle) \in C$ and if for some $b \in B$, $\langle b, c\rangle \in h$ and $\langle b, c'\rangle \in h$, then $c = f(\langle a, b\rangle) = c'$, so indeed $h\colon B \to C$ and $h(b) = f(\langle a, b\rangle)$.

  Given all this, $\langle a, h\rangle \in g$ and $a \in A$ was chosen arbitrarily, so $Dom(g) = A$. If for some $h' \in C^B$, $(\forall b \in B)(h'(b) = f(\langle a, b\rangle))$, then for any $b \in B$, $h'(b) = f(\langle a, b\rangle) = h(b)$, so since $Dom(h) = Dom(h') = B$, $h = h'$. Thus, $g\colon A \to C^B$ or $g \in (C^B)^A$. Further, $g(a) = h$, so $g(a)(b) = h(b) = f(\langle a, b\rangle)$.

  Given all this $\langle f, g\rangle \in i$. $f \in C^{A \times B}$ was chosen without restriction, so $Dom(i) = C^{A \times B}$. If for some $g' \in (C^B)^A$, $\langle f, g'\rangle \in i$, then for any $a \in A$, for any $b \in B$, $g'(a)(b) = f(\langle a, b\rangle) = g(a)(b)$, thus $g'(a) = g(a)$ and so $g' = g$. Thus $i\colon C^{A \times B} \to (C^B)^A$ and

  $$(\forall f \in C^{A \times B})(\forall a \in A)(\forall b \in B)(i(f)(a)(b) = f(\langle a, b\rangle))$$

  Suppose that $j \in (C^B)^A$, $j\colon A \to C^B$. Let

  $$f \leftrightharpoons \{\langle\langle a, b\rangle, c\rangle \mid a \in A \,\&\, b \in B \,\&\, c = j(a)(b)\}$$

For any $\langle a, b\rangle \in A \times B$, $\langle\langle a, b\rangle, j(a)(b)\rangle \in f$. If for some $c \in C$, $\langle\langle a, b\rangle, c\rangle \in f$, then $c = j(a)(b)$. Therefore $f \in C^{A\times B}$ and $(\forall a \in A)(\forall b \in B)(f(\langle a, b\rangle) = j(a)(b))$.

For any $a \in A$, for any $b \in B$, $i(f)(a)(b) = f(\langle a, b\rangle) = j(a)(b)$, so $i(f)(a) = j(a)$ and $i(f) = j$. Therefore $i\colon C^{A\times B} \twoheadrightarrow (C^B)^A$.

Suppose for some $f \in C^{A\times B}$ and some $f' \in C^{A\times B}$, $i(f) = i(f')$. For any $a \in A$, for any $b \in B$, $i(f)(a)(b) = f(\langle a, b\rangle) = i(f')(a)(b)$, so $i(f)(a) = i(f')(a)$, so $i(f) = i(f')$. Therefore

$$i\colon C^{A\times B} \rightarrowtail (C^B)^A$$

□

- If $\overline{\overline{A}} \leq \overline{\overline{B}}$, then $\overline{\overline{A^C}} \leq \overline{\overline{B^C}}$.
- If $A \neq \emptyset$ and $\overline{\overline{A}} \leq \overline{\overline{B}}$, then $\overline{\overline{C^A}} \leq \overline{\overline{C^B}}$.
  Why is $A \neq \emptyset$ a necessary condition?
- $\overline{\overline{A^0}} = \overline{\overline{1}}$.
- $\overline{\overline{1^A}} = \overline{\overline{1}}$.
- If $A \neq \emptyset$, then $\overline{\overline{0^A}} = \overline{\overline{0}}$.