3. Relations and functions

Once the axioms of set theory are in place, we can review some elementary albeit very useful constructions that these axioms enable. To make out notations easier, we will henceforth abandon our convention that all sets be written using the capital letters.

3.1 Cartesian product.

Given two sets A and B, a natural object to consider is the set of pairs (x,y) with the first taken from A and the second from B. This is formalized using the notion of their *Cartesian product*,

$$A \times B := \big\{(x,y) \in \mathcal{P}(\mathcal{P}(A \cup B))\colon x \in A,\, y \in B\big\}, \tag{3.1}$$

in which (x,y) denotes an *ordered pair* that, in set theory, is defined as

$$(x,y) := \big\{\{x\},\{x,y\}\big\}, \tag{3.2}$$

(This is sometimes called the *Kuratowski pair*.) Here the set on the right exists by Pairset Axiom. An exercise in the use of Axiom of Extensionality shows:

Lemma 3.1 *Let A and B be sets. Then*

$$\forall x,\tilde x \in A\,\forall y,\tilde y \in B\colon\quad (x,y) = (\tilde x,\tilde y) \Leftrightarrow (x = \tilde x \wedge y = \tilde y). \tag{3.3}$$

In particular, the pair identifies its entries uniquely. The Cartesian product $A \times B$ is a set by Axiom of Separation.

The construction of the Cartesian product can naturally be iterated to construct the Cartesian product of three, four, etc sets. The problem is that order of operation matters, at least as far as the above definition is concerned. To demonstrate this, given three sets A, B and C and abbreviating $D := \mathcal{P}(\mathcal{P}(\mathcal{P}(\bigcup\{A,B,C\})))$, the above gives

$$A \times (B \times C) := \Big\{\big\{\{x\},\{\{y\},\{y,z\}\}\big\} \in D\colon x \in A \wedge y \in B \wedge z \in C\Big\} \tag{3.4}$$

while

$$\begin{aligned}&(A \times B) \times C\\ &\quad:= \Big\{\big\{\{\{x\},\{\{x\},\{x,y\}\},\{\{\{x\},\{\{x\},\{x,y\}\}\},z\}\big\} \in D\colon x \in A \wedge y \in B \wedge z \in C\Big\}\end{aligned} \tag{3.5}$$

Yet, both sets should intuitively give the set of all triplets (x,y,z) and thus describe the same object modulo identification. This identification requires proving that

$$\big\{\{x\},\{\{y\},\{y,z\}\}\big\} \mapsto \big\{\{\{x\},\{\{x\},\{x,y\}\},\{\{\{x\},\{\{x\},\{x,y\}\}\},z\}\big\}. \tag{3.6}$$

defines a bijection of (3.4) onto (3.5); thus showing that the Cartesian product is *associative*. Once this is done, we drop the parentheses and write the "result" as $A \times B \times C$.

Unfortunately, any specific use of the triple product still requires specifying which of the two sets we refer to, and the issues with the identification become only worse when more sets are involved in the product. So we will ultimately abandon this approach and come up with a unified definition of the Cartesian product of any number of sets that is void of these complications. For that we need the concept of a function which is in turn a special case of a relation, which is, however, based on the definition in (3.1).

3.2 Relations.

With the Cartesian product in hand, we can put foward:

Definition 3.2 *Given sets A and B, a* relation *on A and B is a subset $R \subseteq A \times B$. We say that $x \in A$ is in relation to $y \in B$, with the notation xRy, if the pair (x, y) lies in R, i.e.,*

$$xRy \;:=\; (x,y) \in R. \tag{3.7}$$

If $B = A$, we say that $R \subseteq A \times A$ is a relation on A.

An example of a relation is the set-inclusion $\subseteq$ on the power set $\mathcal{P}(A)$ of a set A. To describe this (and similar) relations formally, we give:

Definition 3.3 *Let A be a set and $R \subseteq A \times A$ a relation on A. We say that R is*

- reflexive*, if $\forall x \in A\colon\; xRx$,*
- antisymmetric*, if $\forall x, y \in A\colon\; xRy \wedge yRx \Rightarrow x = y$, and*
- transitive*, if $\forall x, y, z \in A\colon\; xRy \wedge yRz \Rightarrow xRz$.*

A relation which is reflexive, antisymmetric and transitive is called a partial order.

We leave it to homework for the reader to prove:

Lemma 3.4 *For any sets, the subset relation $\subseteq$ on $\mathcal{P}(A)$ is a partial order.*

Perhaps more familiar example of a relation that is reflexive, antisymmetric and transitive is the inequality $\leqslant$ on the number sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ (but not $\mathbb{C}$). As we will see when we construct these number sets from the foundations of set theory, the inequality relation $\leqslant$ will in fact be induced by $\subseteq$ applied to suitable sets.

In other to define our next frequent example of a relation, we put forward:

Definition 3.5 *A relation R on a set A is said to be* symmetric *if*

$$\forall x, y \in A\colon\;\; xRy \;\Leftrightarrow\; yRx. \tag{3.8}$$

Note that a symmetric relation need not be reflexive because we do not require xRx to actually hold. Nor are all pairs required to be in relation; all what symmetry says that if xRy then also yRx and *vice versa*. Also note that not being symmetric does not mean being antisymmetric, and *vice versa*. The role of symmetry is seen from:

Definition 3.6 *A relation $\sim$ on a set A that is reflexive, symmetric and transitive is called* equivalence. *For each $x \in A$, the set*

$$[x] := \{y \in A\colon y \sim x\} \tag{3.9}$$

is said to be the equivalence class of *x.*

The proof of the following lemma has been relegated to homework:

Lemma 3.7 *Let $\sim$ be an equivalence relation on a set A. Then*

$$\forall x, y \in A\colon [x] \cap [y] \neq \varnothing \;\Rightarrow\; [x] = [y] \tag{3.10}$$

This means that two equivalence classes are either disjoint or equal. Any element $z \in [x]$ is called a *representative* of $[x]$. In particular, x is a representative of $[x]$.

An example of an equivalence relation is the *equality* $=$, which is reflexive, symmetric and transitive. This is a very fine version of equivalence because (by Axiom of Extensionality) each equivalence class contains exactly one element; i.e., $\forall x \in A\colon [x] = \{x\}$. To give a more representative example, consider the following:

Lemma 3.8 *Writing $\mathbb{Z}$ for the set of integers and denoting by "$\cdot$" the standard operation of multiplication on $\mathbb{Z}$, let*

$$A := \{(m,n) \in \mathbb{Z} \times \mathbb{Z}\colon n \neq 0\} \tag{3.11}$$

and let $\sim$ be the relation on A defined as

$$(m,n) \sim (\widetilde{m},\widetilde{n}) \ := \ m \cdot \widetilde{n} = \widetilde{m} \cdot n \tag{3.12}$$

Then $\sim$ is an equivalence.

Proof. Symmetry and reflexivity are checked readily; all that needs a bit of work is transitivity. Assume $(m,n) \sim (\widetilde{m},\widetilde{n})$ and $(\widetilde{m},\widetilde{n}) \sim (\widehat{m},\widehat{n})$. This means

$$m \cdot \widetilde{n} = \widetilde{m} \cdot n \ \wedge \ \widetilde{m} \cdot \widehat{n} = \widehat{m} \cdot \widetilde{n}. \tag{3.13}$$

Multiplying the first equality by $\widehat{n}$ results in

$$m \cdot \widetilde{n} \cdot \widehat{n} = \widetilde{m} \cdot n \cdot \widehat{n} = \widetilde{m} \cdot \widehat{n} \cdot n = \widehat{m} \cdot \widetilde{n} \cdot n \tag{3.14}$$

where where we used commutativity and associativity of multiplication throughout and invoked the second equality in (3.13) in the last step. Since $\widetilde{n} \neq 0$, the fact that

$$\forall k,l,m \in \mathbb{Z}\colon \quad (k \neq 0 \ \wedge \ k \cdot l = k \cdot m) \ \Rightarrow \ l = m \tag{3.15}$$

implies $m \cdot \widehat{n} = \widehat{m} \cdot n$ meaning that $(m,n) \sim (\widehat{m},\widehat{n})$. □

The punchline of this example is that the equivalence class $[(m,n)]$ then contains all pairs $(\widetilde{m},\widetilde{n})$ such that, informally (because division is not defined on $\mathbb{Z}$), obey $\frac{\widetilde{m}}{\widetilde{n}} = \frac{m}{n}$. The set of equivalence classes thus provides a construction of the set of rationals $\mathbb{Q}$.

3.3 Functions.

We now move to a concept central to analysis:

Definition 3.9 *Let A and B be sets. A relation $F \subseteq A \times B$ satisfying*

$$\forall x \in A\, \forall y,z \in B\colon xFy \wedge xFz \Rightarrow y = z \tag{3.16}$$

is called a function. *We will use notation $F(x)$ for the unique $y \in B$ such that xFy.*

Moving to the convention that functions can be denoted by lowercase letters, we will naturally think of a function f as an assignment of a value in B to a value in A, with the notation $f\colon A \to B$. The relation $f \subseteq A \times B$ corresponding to a function f is then the *graph of f*. Not every $x \in A$ may appear as the first member of a pair in relation R; those that do are collected in the *domain* of R, i.e.,

$$\text{Dom}(R) := \{x \in A\colon (\exists y \in B\colon xRy)\} \tag{3.17}$$

Similarly, not every $y \in B$ is in relation with some $x \in A$; those that are form the *range* of R denoted as

$$\text{Ran}(R) := \{y \in B\colon (\exists x \in A\colon xRy)\}. \tag{3.18}$$

We will write Dom(f), resp., Ran(f) for the domain, resp., range of the relation that is a function f. We remark that one often writes $f\colon A \to B$ without necessarily requiring that $\mathrm{Dom}(f) = A$.

If $R \subseteq A \times B$ and $S \subseteq B \times C$ are two relations, then their *composition* RS is the relation defined by

$$\forall x \in A\ \forall y \in C\colon \quad x(RS)y\ :=\ \exists z \in B\colon xRz\ \wedge zSy. \tag{3.19}$$

If R and S are (graphs of) functions r and s, then their composition is also a function but (beware!) for functions the composition is written in reverse order; namely,

$$RS \text{ is the graph of } s \circ r \tag{3.20}$$

This is because functions "act" on the variable to the right while relations "act" on the variable to the left. Note also that for any transitive relation R we have $RR = R$.

The *inverse* R^{-1} of a relation R is the subset of $B \times A$ defined by

$$\forall x \in A\, \forall y \in B\colon \quad yR^{-1}x\ :=\ xRy \tag{3.21}$$

Note that RR^{-1} is an identity relation on Dom(R) (which is a subset of A) and $R^{-1}R$ is an identity relation on Ran(R) (which is a subset of B). If F is (the graph of) a function f, then F^{-1} is (the graph of) a function if and only if

$$\forall x, \tilde{x} \in A\, \forall y \in B\colon \big(y = f(x)\ \wedge\ y = f(y)\big)\ \Rightarrow\ x = \tilde{x} \tag{3.22}$$

The inverse function is then denoted by f^{-1}.

Given a function $f\colon A \to B$, for each $C \subseteq A$ we define the *image* $f(C)$ of C by

$$f(C) := \big\{y \in B\colon (\exists x \in C \cap \mathrm{Dom}(f)\colon y = f(x))\big\} \tag{3.23}$$

sometimes written simply as $\{f(x)\colon x \in C\}$ ignoring that $f(x)$ may not be defined for all $x \in C$. The image of the domain is then the range, $\mathrm{Ran}(f) = f(\mathrm{Dom}(f))$. Similarly, for all $D \subseteq B$ we define the *preimage* $f^{-1}(D)$ of D by

$$f^{-1}(D) := \big\{x \in A\colon x \in \mathrm{Dom}(f)\ \wedge\ f(x) \in D\big\} \tag{3.24}$$

which we at times write as $\{x\colon f(x) \in D\}$ ignoring domain restrictions. The preimage of the range is the domain, $f^{-1}(\mathrm{Ran}(f)) = \mathrm{Dom}(f)$. We caution the reader that the use of f^{-1} in the preimage map does not require the inverse f^{-1} of f to exist. Notwithstanding, if f^{-1} does exist, then $f^{-1}(D)$ is the image of D under f^{-1}.

While functions in analysis are generally of interest for their analytic properties, in the rest of mathematics functions are primarily used to identify sets with one another. For this we need the following concepts:

Definition 3.10 *A function $f\colon A \to B$ is*

- injective *if* $\mathrm{Dom}(f) = A$ *and* $\forall x, y \in A\colon f(x) = f(y) \Rightarrow x = y$
- surjective *if* $\mathrm{Ran}(f) = B$ *meaning* $\forall y \in B \exists x \in A\colon f(x) = y$,
- bijective *if it is both injective and surjective.*

Exhibiting a bijection between two sets puts these in *one-to-one* or *bijective correspondence*, which are simply different ways to talk about a bijection. For instance, the map (3.6) defines a bijection $f\colon (A \times B) \times C \to A \times (B \times C)$ that permits us to write this as $A \times B \times C$. A very fruitful use of above concepts is in the following insightful definition due to G. Cantor:

Definition 3.11 (Equinumerosity) *Sets A and B are said to be* equinumerous*, or are of the same* cardinality*, if there exists a bijection $f\colon A \to B$.*

Note that given a set of sets, equinumerosity is another example of equivalence relation. Indeed, reflexivity is provided by the identity map, symmetry by the inverse map (which uses that the inverse of a bijection is a bijection) and transitivity by composed maps (which uses that a composition of two bijections is a bijection). Related to this is:

Definition 3.12 *A set A is said to be* Dedekind infinite *if there is an injection $f\colon A \to A$ such that $\operatorname{Ran}(f) \neq A$.*

This is one way to define the notion of an infinite set, albeit one that is not generally used in mathematics. We will return to these concepts when we discuss the question of cardinality in more detail.

3.4 General Cartesian products.

Relying on the concept of a function, the notion of the Cartesian product can be generalized to products of arbitrary collections of sets. Such collections is typically written as $\{A_\alpha\colon \alpha \in I\}$, where α represents the *index* of A_α and I denotes the *index set*. While this concept is quite intuitive, the reader may wonder what this is in the formal language of the set theory. This comes in:

Definition 3.13 (Collections of sets) *Given sets I and A, a* collection $\{A_\alpha\colon \alpha \in I\}$ of subsets of A indexed by I *is the set* $\operatorname{Ran}(\phi)$ *for a map* $\phi\colon I \to \mathcal{P}(A)$ *such that* $\operatorname{Dom}(h) = I$. *Under this map, the sets in the collection are identified by $A_\alpha := \phi(\alpha)$.*

We note that, formally, $\phi \in \mathcal{P}(I \times \mathcal{P}(A))$ and so $\operatorname{Ran}(\phi)$ is indeed a set. (This is why we need all A_α's be subsets of one set.) We then put forward:

Definition 3.14 (General Cartesian product) *Given a set I and a collection $\{A_\alpha\colon \alpha \in I\}$ of sets (all of which have to be subsets of a given set) indexed by I,*

$$\bigtimes_{\alpha\in I} A_\alpha := \left\{ f \in \mathcal{P}\Big(I \times \bigcup_{\alpha\in I} A_\alpha\Big) \colon \textit{function} \wedge \operatorname{Dom}(f) = I \wedge (\forall \alpha \in I\colon f(\alpha) \in A_\alpha)\right\} \tag{3.25}$$

is the Cartesian product of the sets in $\{A_\alpha\colon \alpha \in I\}$.

The notation is easier to parse once we note that a function $f\colon I \to \bigcup_{\alpha\in I} A_\alpha$ is formally a subset of $I \times \bigcup_{\alpha\in I} A_\alpha$, which (by Unionset and Powerset axioms and our earlier construction of the Cartesian product) means that (3.25) is a set.

To see that (3.25) subsumes our earlier definition of the Cartesian product, note that a function f defined on two-point set $\{0, 1\}$ is determined by the pair of values $(f(0), f(1))$. The set of all functions with $f(0) \in A$ and $f(1) \in B$ is thus in a bijective correspondence with the set of all pairs (a, b) with $a \in A$ and $b \in B$. This identifies $A \times B$ with the set all functions $f\colon \{0, 1\} \to A \cup B$ satisfying $f(0) \in A$ and $f(1) \in B$, and thus the above general Cartesian product for $I := \{0, 1\}$, $A_0 := A$ and $A_1 := B$. We will henceforth not make a distinction between the two ways to define the Cartesian product of two sets.

A special case (mainly notation) of above definition is:

Definition 3.15 *Let I and A be sets. Then*

$$\bigtimes_{\alpha\in I} A := \{f \in \mathcal{P}(I\times A)\colon \text{ function}\} \tag{3.26}$$

is the Cartesian power *denoted, in short, by A^I.*

For instance, the set of all real-valued sequences form the set $\mathbb{R}^{\mathbb{N}}$ while $\mathbb{R}^{\mathbb{R}}$ denotes the set of all real-valued functions of one real-valued variable.

The Cartesian product of two non-empty sets is non-empty (which is witnessed by a pair (x,y) such that $x\in A$ and $y\in B$), and the same applies to Cartesian products three, four, etc sets. However, a perplexing issue arises once I is infinite where this argument can no longer be made because we have no way to string an infinite number of such statements together. In the ZFC theory, this is resolved by imposing yet another axiom:

- **Axiom of Choice:** For each nonempty set I and all collections $\{A_\alpha\colon \alpha\in I\}$ of sets satisfying $\forall\alpha\in I\colon A_\alpha\neq\emptyset$, there exists a function $f\colon I\to\bigcup_{\alpha\in I}A_\alpha$ such that

$$\text{Dom}(f) = I \;\wedge\; \forall\alpha\in I\colon f(\alpha)\in A_\alpha \tag{3.27}$$

  In short,

$$\forall I\,\forall\{A_\alpha\colon \alpha\in I\}\colon \Big(I\neq\emptyset \wedge \big(\forall\alpha\in I\colon A_\alpha\neq\emptyset)\Big) \Rightarrow \bigtimes_{\alpha\in I} A_\alpha\neq\emptyset \tag{3.28}$$

The name arises from the observation that a function $f\in\bigtimes_{\alpha\in I}A_\alpha$ gives us a *choice*, simultaneously for all $\alpha\in I$, of a representative $f(\alpha)\in A_\alpha$ provided, of course, $A_\alpha\neq\emptyset$. In sets with some underlying structure, this can sometimes be guaranteed constructively but that will not work in general which is why such an axiom is needed.

Following up on a question from class, here is one instance where Axiom of Choice is definitely NOT required:

Lemma 3.16 (Picking representatives from singletons) *Let $I\neq\emptyset$ and let $\{A_\alpha\colon \alpha\in I\}$ be a collection of subsets of a set A that are all singletons. Then there exists a function $f\colon I\to A$ such that* $\text{Dom}(f) = I$ *and $\forall\alpha\in I\colon f(\alpha)\in A_\alpha$. In short*

$$I\neq\emptyset \wedge \big(\forall\alpha\in I\,\exists x\in A\colon A_\alpha=\{x\}\big) \Rightarrow \bigtimes_{\alpha\in I} A_\alpha\neq\emptyset \tag{3.29}$$

Axiom of Choice is not required.

Proof. Let $\{A_\alpha \alpha\in I\}$ be as in the statement and let $\mathcal{F}$ be the set of all functions $f\colon I\to A$ such that $f(\alpha)\in A_\alpha$ for all $\alpha\in\text{Dom}(f)$. We write this as

$$\mathcal{F} := \Big\{F\subseteq I\times A\colon \big(\forall\alpha\in I\forall x\in A\colon \alpha F x \Rightarrow x\in A_\alpha\big)\Big\} \tag{3.30}$$

First we check that each $F\in\mathcal{F}$ is the graph of a function. Indeed, if $F\in\mathcal{F}$ and $\alpha\in\text{Dom}(F)$ then for any $x,y\in A$,

$$\alpha F x \wedge \alpha F y \Rightarrow x\in A_\alpha \wedge y\in A_\alpha \tag{3.31}$$

The fact that A_α is a singleton then forces $x=y$ and so F is the graph of a function. The same argument also implies that if $F,G\in\mathcal{F}$, then (switching to the notation for functions)

$$\forall\alpha\in\text{Dom}(F)\cap\text{Dom}(G)\colon F(\alpha)=G(\alpha) \tag{3.32}$$

Now define

$$\widehat{F} := \bigcup \mathcal{F} \tag{3.33}$$

Then (3.32) ensures that also $\widehat{F}$ is the graph of a function. It remains to show that $\text{Dom}(\widehat{F}) = I$. For this pick any $\alpha \in I$ and check that $F := \{(\alpha, x) \in I \times A : x \in A_\alpha\}$ is the graph of a function in $\mathcal{F}$. Hence, $F \subseteq \widehat{F}$, proving $\alpha \in \text{Dom}(\widehat{F})$. $\square$

A similar conclusion is obtained even without the requirement that the sets A_α be singletons provided there is a way to pick a "distinguished element" or a representative in each non-empty subset of A. Technically, this means that A admits a function $\phi \colon \mathcal{P}(A) \smallsetminus \{\emptyset\} \to A$ such that

$$\forall B \in \mathcal{P}(A) \smallsetminus \{\emptyset\} \colon\ \phi(B) \in B \tag{3.34}$$

That $\bigtimes_{\alpha \in I} A_\alpha$ is non-empty is demonstrated by applying the previous lemma for sets $A'_\alpha := \{\phi(A_\alpha)\}$ and then noting that $f \in \bigtimes_{\alpha \in I} A'_\alpha$ implies $f \in \bigtimes_{\alpha \in I} A_\alpha$.

As stated before, in sets A with some underlying structure a "distinguished element" can often be "picked" constructively, but this is not possible in general. The requirement that there be a function satisfying (3.34) is generally as strong as the Axiom of Choice itself. Indeed, one way to state the Axiom of Choice is by demanding that a function ϕ satisfying (3.34) exists for all non-empty sets A.

Notwithstanding the above discussion, mathematicians find the Axiom of Choice generally less acceptable than the rest of Zermelo's axioms, and so it is a good practice (to which we will adhere) to caution the reader whenever it is invoked.