Basic Set Theory

Stéphane Dupraz

1 Sets

The notion of **set** or **collection** is a primitive notion that we take in its everyday meaning. A set is characterized by what objects it contains; we call them its **elements** or **members**. Two sets that have the same members are the same, and are said to be equal.

- To say that the object x **belongs to**, or **is an element of** the set X, we note $x \in X$.
- When describing a set through an exhaustive listing of its elements, we use curly brackets:

  $X = \{1, \pi, \text{Economics}, \text{Columbia University}\}$.
- The curly brackets can also be used to define a set through the properties that its elements satisfy:

  $X = \{\text{integers } n \text{ such that } n^2 = 4\}$
- We allow for the possibility that a set contains no element. We call this set the **empty set**, and note it $\emptyset$.
- Note that it is meaningful to talk about a set of sets (we usually call it a collection of sets).

1.1 Inclusion

- We say that A **is included in** B, or that A **is a subset of** B, and note $A \subseteq B$ if every member of A is a member of B: $x \in A$ implies $x \in B$. Any set is a subset of itself.
- It follows that two sets A and B are equal if and only if $A \subseteq B$ and $B \subseteq A$.
- We say that A is a **proper subset** of B, and note $A \subset B$ if A is a subset of B but is not equal to B.
- The inclusion is **transitive** : if $A \subseteq B$ and $B \subseteq C$, then $A \subseteq C$.
- The set of all subsets of a set X is called the **power set** of X, and noted $P(X)$.

1.2 Union and Intersection

- The **intersection** of two sets A and B, noted $A \cap B$ is the set of all elements that belong to both A and B: $x \in A \cap B$ iff $x \in A$ and $x \in B$.
- If no element lies in both A and B, $A \cap B$ is still defined: defined to be the empty set $A \cap B = \emptyset$. We say that A and B are **disjoint**.
- The intersection is commutative: $A \cap B = B \cap A$.
- The intersection is associative $(A \cap B) \cap C = A \cap (B \cap C)$.
- $A \cap B = A$ iff $A \subseteq B$.
- More generally, we can talk about the intersection of any collection of sets, possibly infinite. If the A_i are sets indexed by i, their intersection, noted $\cap_i A_i$, is the set of all elements that belong to all A_i.
- The **union** of two sets A and B, noted $A \cup B$ is the set of all elements that belong to either A or B (possibly both): $x \in A \cap B$ iff $x \in A$ or $x \in B$.
- The union is commutative $A \cup B = B \cup A$.
- The union is associative $(A \cup B) \cup C = A \cup (B \cup C)$.
- $A \cup B = A$ iff $B \subseteq A$.
- More generally, we can talk about the union of any collection of sets, possibly infinite. If the A_i are sets indexed by $i \in I$, their union, noted $\cup_i A_i$, is the set of all elements that belong to at least one A_i.
- A **partition** of a set A is a collection of sets $A_i, i \in I$ such that their union is equal to A, $A = \cup_i A_i$, and the A_i are pairwise disjoint: for all $i \neq j$, $A_i \cap A_j = \emptyset$.
- The intersection is distributive wrt. the union: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
- The union is distributive wrt. the intersection: $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

1.3 Difference and complement

- Given two sets A and B, the **set difference** $A - B$ or $A \backslash B$ is the set of all x that belong to A but not to B.

- Given a set U and a subset A of U, the **complement** of A in U is $U \backslash A$. Most often, the set U is implicit and we note the complement A^c.
- We note $x \notin A$ for $x \in A^c$ ($x \in U$ implicitly).
- A and A^c form a partition of U: $A \cap (A^c) = \emptyset$ and $A \cup A^c = U$.
- $(A^c)^c = A$.
- **De Morgan's law**: $(A \cup B)^c = A^c \cap B^c$ and $(A \cap B)^c = A^c \cup B^c$.
- More generally, $(\cup_i A_i)^c = \cap_i A_i^c$ and $(\cap_i A_i)^c = \cup_i A_i^c$.

Venn diagrams—drawing sets as potatolike figures in the plane–are useful to represent inclusions, intersections, unions, etc. of sets.

1.4 Cartesian Product

- The **cartesian product** of two sets A and B, noted $A \times B$, is the set of all ordered pairs (a, b) such that $a \in A$ and $b \in B$: $A \times B = \{(a, b) / a \in A, b \in B\}$.
- More generally, we can talk about the cartesian product of any collection of sets, possibly infinite. If the A_i are sets indexed by $i \in I$, their cartesian product, noted $\prod_i A_i$, is the set of arrays $(a_i)_i$ such that for all i, $a_i \in A_i$.

2 Functions

2.1 Definition

Definition 2.1. *Let X and Y be two sets. A **function** or **map** or **mapping** f from X to Y associates to every element $x \in X$ an element $y \in Y$, noted $f(x)$. We note:*

$$f : X \to Y$$
$$x \mapsto f(x)$$

*We call X the **domain** of f and Y the **codomain** of f.*
*We call x an **argument** of f, and $f(x)$ the **image** or **value** of f at x.*
*We call the subset of $X \times Y$ $\{(x, y)/x \in X, y = f(x)\}$ the **graph** of f.*

2.2 Image and Inverse Image

From a function from X to Y, we can define two new functions on power sets: one from $P(X)$ to $P(Y)$, and one from $P(X)$ to $P(Y)$.

Definition 2.2.

- *We note f: $P(X) \to P(Y)$ the function that, to a subset S of X, associate the **image of** X **under** f, $f(S)$, defined as the subset of Y, $f(S) = \{f(x), s \in S\}$.*

  *The image of the whole domain X under f, $f(X)$, is called the **image** or **range** of f.*

- *We note f^{-1}: $P(Y) \to P(X)$ the function that, to a subset T of Y, associate the **inverse image of** X **under** f, $f^{-1}(T)$, defined as the subset of X, $f^{-1}(T) = \{x/f(x) \in T\}$.*

Note that, a bit confusingly, we use the same symbol f to designate the function on X and the function on $P(X)$. The proposition below indicates how the inclusion, the intersection, the union and the complement operators behave with f and f^{-1}.

Proposition 2.1.

1. *Inclusion:*

   - $S_1 \subseteq S_2 \Rightarrow f(S_1) \subseteq f(S_2)$

- $T_1 \subseteq T_2 \Rightarrow f^{-1}(T_1) \subseteq f^{-1}(T_2)$

2. Union:

- $f(S_1 \cup S_2) = f(S_1) \cup f(S_2)$
- $f^{-1}(T_1 \cup T_2) = f^{-1}(T_1) \cup f^{-1}(T_2)$

3. Intersection:

- $f(S_1 \cap S_2) \subseteq f(S_1) \cap f(S_2)$
- $f^{-1}(T_1 \cap T_2) = f^{-1}(T_1) \cap f^{-1}(T_2)$

4. Complement:

- $f^{-1}(S^c) = (f^{-1}(S))^c$.
- *No general result for* f.

In short, all that we would like to be true is true for the inverse image, but for the image, the intersection and the complement are frustrating.

2.3 Composite mapping

Definition 2.3. *Let X, Y, Z be sets. Let $f : X \to Y$ and $g : Y \to Z$ be mappings.*

*The **composite mapping** of f and g, noted $g \circ f$, is the mapping from X to Z defined as:*

$$\begin{aligned} g \circ f : X &\to Z \\ x &\mapsto (g \circ f)(x) = g(f(x)) \end{aligned}$$

The composition of mappings is an associative operator: $h \circ (g \circ f) = (h \circ g) \circ f$.

2.4 Injectivity, surjectivity, bijectivity, inverse function

Consider a mapping $f : X \to Y$, an element of the codomain $y \in Y$, and the inverse image of y, $f^{-1}(\{y\})$—we will abuse notations slightly and note it $f^{-1}(y)$. The function f is:

- injective if: for all $y \in Y$, $f^{-1}(y)$ contains at most one element,
- surjective if: for all $y \in Y$, $f^{-1}(y)$ contains at least one element,
- bijective if: for all $y \in Y$, $f^{-1}(y)$ contains exactly one element—if f is both injective and surjective.

Definition 2.4.

- *A mapping $f : X \to Y$ is **injective** or **one-to-one** iff two distinct elements of X have distinct images under f: for all $x, x' \in X$, $f(x) = f(x') \Rightarrow x = x'$.*
- *A mapping $f : X \to Y$ is **surjective**, or **onto** iff the image of f is all of its codomain Y: $f(X) = Y$.*
- *A mapping is **bijective** or a **one-to-one correspondence** iff it is both injective and surjective.*

The **identity mapping** from X to X is the mapping $Id_X : X \to X$, $Id_X(x) = x$. It is bijective.

Definition 2.5. *Let $f : X \to Y$ be a mapping.*

*f has an **inverse** if there exists a mapping $g : Y \to X$ such that $f \circ g = Id_X$ and $g \circ f = Id_Y$.*

If it exists, the inverse of f is necessarily unique, and is noted f^{-1}.

Note that, just as with the notation f, we use the same notation f^{-1} for the function on Y and the function on the power set $P(Y)$. Be careful that this time, f^{-1} is always defined on $P(Y)$, but is defined on Y only if f is bijective.

Proposition 2.2. *f is bijective if and only if f has an inverse.*

3 Relations

Definition 3.1. *A **relation** R from X to Y is a subset of the cartesian product $X \times Y$.*

*If $(a, b) \in R$, we say that a and b are **in relation** and note aRb.*

That's a bit abstract; you can think of a relation as a function which is allowed to be multi-valued and not necessarily everywhere defined. (Alternatively, a function $f : X \to Y$ is a relation that is single-valued and defined everywhere: for all $a \in X$, there exists a unique element $b \in Y$ such that a and b are in relation).

3.1 Equivalence Relations

Definition 3.2. *A **equivalence relation** on a set X, usually noted $\sim$, is a relation on X that satisfies the 3 following axioms:*

1. *Reflexivity: for all $a \in X$, $a \sim a$.*
2. *Symmetry: for all $a, b \in X$, if $a \sim b$ then $b \sim a$.*
3. *Transitivity: for all $a, b, c \in X$, if $a \sim b$ and $b \sim c$, then $a \sim c$.*

3.2 Orders

Definition 3.3. *A **pre-order** on a set X, usually noted $\leq$, is a relation on X that satisfies the 2 following axioms:*

1. *Reflexivity: for all $a \in X$, $a \leq a$.*
2. *Transitivity: for all $a, b, c \in X$, if $a \leq b$ and $b \leq c$, then $a \leq c$.*

If $\leq$ is a pre-order, we often note $a < b$ when $a \leq b$ and $a \neq b$.

Note that an equivalence relation is a pre-order that also satisfies the symmetry axiom. But the name *pre-order* comes from the fact that if we add the property of antisymmetry, we have a (partial) order.

Definition 3.4. *A **partial order** on a set X, usually noted $\leq$, is a pre-order on X that satisfies a third axiom:*

3. Antisymmetry: for all $a, b \in X$, if $a \leq b$ and $b \leq a$, then $a = b$.

*A set endowed with a partial order is called a **partially ordered set**.*

In a partially ordered set, not all elements need to be comparable, i.e. there can be elements a, b such that neither $a \leq b$ nor $b \leq a$. Partial orders that have the property that all pairs of elements are comparable are called total orders.

Definition 3.5. *A **total order** on a set X, usually noted $\leq$, is a partial order on X that satisfies a fourth axiom:*

4. Totality: for all $a, b \in X$, $a \leq b$ or $b \leq a$ (or both).

*A set endowed with a total order is called a **totally ordered set** or a **chain**.*

There are three main examples and applications of orders:

- The standard $\leq$ on $\mathbb{R}$ (and its subsets). Here, $\leq$ is a total order and $\mathbb{R}$ a totally ordered set.
- The inclusion $\subseteq$ on the power set $P(X)$ of any set X. Here, $\subseteq$ is a partial order and $P(X)$ a partially ordered set.
- In consumer theory, we define preferences through relations on the set of bundles of goods. We impose transitivity and totality to the relation, and call preferences that satisfy both axioms *"rational"*. As totality implies reflexivity, a rational preferences relation can be defined as a **total pre-order** (but it is not required to be antisymmetric, hence it is not necessarily an order).

3.3 Upper-bound, maximum and supremum on partially ordered sets

Let X be a partially ordered set with partial order $\leq$, and S a subset of X.

Definition 3.6.

- *$u \in X$ is an **upper-bound** of S iff for all $s \in S$, $s \leq u$. We say that X is **bounded above** (by u).*
- *$l \in X$ is an **lower-bound** of S iff for all $s \in S$, $s \geq l$. We say that X is **bounded below** (by l).*

Upper-bounds are usually not unique. An upper (lower) bound of S does not need to belong to S. If it does, we call it a maximum (minimum):

Definition 3.7.

- *$u \in X$ is a **maximum** of S iff $u \in S$ and u is an upper-bound of S.*
- *If a maximum of S exists, it is unique.*
- *$l \in X$ is a **minimum** of S iff $l \in S$ and l is a lower-bound of S.*
- *If a minimum of S exists, it is unique.*

The uniqueness follows from the antisymmetry axiom: if there exist two maxima u and u' of S, then $u \leq u'$ and $u' \leq u$, so that $u = u'$. A set may well have no maximum, even if it has upper-bounds. In this case we might want to distinguish one particular upper-bound. Consider $[0, 1)$ in $\mathbb{R}$. It has no maximum but all reals (weakly) greater than 1 are upper-bounds. We want to distinguish 1 among these upper-bounds.

Definition 3.8.

- *$u \in X$ is the **least upper-bound** or **supremum** of S iff it is the minimum of the set of upper-bounds of S, i.e. iff:*
  1. *u is an upper-bound of S.*
  2. *For all upper-bound v of S, $u \leq v$.*
- *$l \in X$ is the **greatest lower-bound** or **infimum** of S iff it is the maximum of the set of lower-bounds of S, i.e. iff:*
  1. *l is a lower-bound of S.*
  2. *For all lower-bound h of S, $l \geq h$.*

A set may well not have a least upper-bound either. But still, it is more likely to have an upper-bound than to have a maximum since:

Proposition 3.1.

- *If a maximum exists, then it is the least upper-bound.*
- *If a minimum exists, then it is the greatest lower-bound.*

Indeed, a maximum m is an upper-bound of S and, since $m \in S$, it is smaller than any upper-bound of S.

If $f : X \to Y$ is a function whose codomain Y is a partially ordered set, we define upper-bounds, the maximum, and the least upper-bound of f as the upper-bounds, the maximum, and the least upper-bound of its image $f(X)$. The arguments x such that $f(x)$ is the maximum of f are called the **maximizers** of f; the set of maximizers of f is noted $argmax(f)$.

If $f : X \to Y$ is a function whose both domain and codomain are partially ordered sets:

- f is **increasing** if $x \leq x'$ implies $f(x) \leq f(x')$
- f is **decreasing** if $x \leq x'$ implies $f(x) \geq f(x')$
- f is **strictly increasing** if $x < x'$ implies $f(x) < f(x')$
- f is **strictly decreasing** if $x > x'$ implies $f(x) > f(x')$

.

4 Common Numbers Sets

There are 5 common numbers sets, included in each other like Russian dolls:

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$$

- $\mathbb{N}$ is the set of **positive integers** or **natural numbers** $\mathbb{N} = \{0, 1, 2, 3, ...\}$ or excluding zero $\mathbb{N} = \{1, 2, 3, ...\}$. (Yes that's confusing, but both definitions are used).
- $\mathbb{Z}$ is the set of **integers** $\mathbb{N} = \{..., -2, -1, 0, 1, 2, 3, ...\}$.
- $\mathbb{Q}$ is the set of **rational numbers** $\mathbb{Q} = \{x/x = \frac{p}{q}, p, q \in \mathbb{Z}, q \neq 0\}$.

We spend a bit more time on the **real numbers** $\mathbb{R}$ and **complex numbers** $\mathbb{C}$.

4.1 Real numbers

We will not define/construct the real line. Instead, we state the key property of the real line:

Theorem 4.1.

- *Any non empty and bounded above subset of* $\mathbb{R}$ *has a least upper-bound.*
- *Any non empty and bounded below subset of* $\mathbb{R}$ *has a greatest lower-bound.*

What is not difficult to show it that the first statement is equivalent to the second one. The proof of the first statement depends on how the real line has been constructed/defined. It actually can be used not as a theorem, but as a definition axiom of the real line.

4.2 Complex numbers

The set of complex numbers $\mathbb{C}$ is at its core just the set $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$, the set of all ordered pairs of $\mathbb{R}$. Instead of noting $z = (a, b) \in \mathbb{R}^2$, we note $z = a + ib$, where i is the **imaginary unit**. We call a the **real part**, noted $Re(z)$, and b the **imaginary part** of $a + ib$, noted $Im(z)$. We see the real line as the subset of $\mathbb{C}$ whose numbers have imaginary part equal to zero. The set really becomes $\mathbb{C}$ once we define operations on it: an addition and a multiplication, like on $\mathbb{R}$. The operations are defined so that they generalize the operations on $\mathbb{R}$.

- The addition: $(a + ib) + (c + id) = (a + b) + i(b + d)$.
- The multiplication : we define $i^2 = -1$ so that $(a + ib)(c + id) = (ac - bc) + i(bc + ad)$.

We define the **conjugate** of a complex number $z = a + ib$ to be $\bar{z} = a - ib$.

We define the **modulus** of a complex number $z = a + ib$ to be $|z| = \sqrt{a^2 + b^2}$. The modulus extends the notion of absolute value $|x| = \max(x, -x)$ on $\mathbb{R}$.

We have that $z\bar{z} = |z|^2$.

5 Cardinality and Countability

5.1 Equal and higher cardinality

How to define a set as "bigger" than another? We do have a partial order defined on collections of sets—the inclusion $\subseteq$—so may want to use it to say that if $S \supseteq T$, then S is "bigger" than T. But the inclusion $\subseteq$ is only a partial order: in general, only for a few pairs of sets S and T do we have $S \subseteq T$ or $T \subseteq S$. Most of the time, the inclusion is not helpful to define one set as bigger than another.

For finite sets, we have a straightforward way to define the "size" of a set: the number of elements it contains. This allows to compare any two finite sets: in other words, it defines a total pre-order on a collection of finite sets. Moreover, it chimes with the partial ordering given by the inclusion whenever a set is included in another: if S and T are finite sets and if $T \subseteq S$, then T has fewer elements than S. It even extends to strict inclusion: if $T \subset S$, then T has strictly fewer elements than S.

So the notion of number of elements looks promising. But how to extend the notion to infinite sets? Should we consider that all infinite sets are all "as large" or is there a meaningful way to distinguish between different "sizes of infinity"? To answer this question, let us reflect on our notion of "size" for finite sets. When we say that a set S has n elements, we mean that we can count them from 1 to n: that there is a bijection from $[\![1, n]\!]$ onto S. If two finite sets S and T have the same size—if they have the same number of elements, say n—they are both in bijection with $[\![1, n]\!]$, so S and T are in bijection. If the finite set S is larger than the finite set T—if S has a larger number of elements than T—we can build an injection from T to S. Let us take stock and extend the notions of "having the same size" and "having a larger size" to all sets, possibly infinite, using bijections and injections. From now on, we replace the word size by the word cardinality.

Definition 5.1. *Let S and T two sets.*

- *S and T have **equal cardinality** (are **equinumerous**) if there exists a bijection between them.*
- *S has **higher cardinality** than T if there exists an injection from T onto S.*
- *S has **strictly higher cardinality** than T if it has a higher but not equal cardinality than T.*

To start with what remains intuitive with the notion of cardinality, let us state without proof two results that are true, but are actually not obvious to prove:

- If S has a higher cardinality than T, and T has a higher cardinality than S, then S and T have equal cardinality (Schröder-Bernstein theorem).[1]

[1] For the history and/or logic geeks in the room: the theorem is sometimes called the Cantor-Schröder-Bernstein theorem,

- Given two sets S and T, either S has higher cardinality than T, or T has higher cardinality than S.[2] In other words, because the higher-cardinality relation is also reflexive and transitive, we have defined a total pre-order on any collection of sets.

So we have defined a total pre-order. How does it compare with our earlier partial pre-order, the inclusion? At first, it looks good: is $S \supseteq T$, then S has higher cardinality than T. To see it, not that $f : T \to S : x \mapsto x$ is an injection. But let us note right now a disturbing property of cardinality: when dealing with infinite sets, strict inclusion no longer necessarily implies strictly lower cardinality, as it does for finite sets. In other words, defining size as cardinality, the whole is not necessarily larger than its parts. To see it, consider $\mathbb{N}$ and $\mathbb{N}^* = \mathbb{N} - \{0\}$; the function $f : \mathbb{N} \to \mathbb{N}^* : n \mapsto n + 1$ is a bijection. David Hilbert illustrated this funny property of infinity through the *paradox of the Infinite Hotel*. Consider a hypothetical hotel with $\mathbb{N}$ rooms, and suppose the hotel is full: it accommodates one guest in each of its rooms. One new customer shows up at the front desk and asks for a room. Asking the guest in room 1 to move to room 2, the guest in room 2 to move in room 3, etc., the hotel manager can free a room for his new guest.

5.2 Distinguishing between infinites

With this definition in hands, let us now investigate whether the different infinite sets that we know have equal cardinalities, or if some infinities were born more infinite than others. Out of the 5 common number sets, $\mathbb{N}$ is the one with the lowest cardinality (at least weakly) since it is included in the others. It defines the countable infinity.

Definition 5.2.

*A set S is **countably infinite** if it has the same cardinality as $\mathbb{N}$.*

*A set S is **countable** if it is finite or countably infinite.*

*A set S is **uncountable** if it is not countable.*

To prove that a set S is countable, we build a bijection from $\mathbb{N}$ to S: that is, we find a way to count all its elements as 1, 2, 3... Above, this is how we proved that the subset $\mathbb{N}^*$ of $\mathbb{N}$ is infinitely countable. We can prove this way that:

Proposition 5.1. *The finite union of countable sets is countable.*

because Cantor provided an earlier proof. But Cantor's proof relied on the axiom of choice, whereas the result does not require the axiom of choice.

[2] For the logicians still in the room: this result requires the axiom of choice.

Proof. We just need to prove it for the union of two countable sets, and the result will follow by induction. Consider two countable sets S_1 and S_2. If both are finite, their union is finite, hence countable. If only one is finite, we can use the same argument as when showing that $\mathbb{N}^*$ is countable. So we focus on the case where both S_1 and S_2 are infinitely countable. We can index the elements of S_1 as $(x_1^1, x_2^1, x_3^1...)$ and the elements of S_2 as $(x_1^2, x_2^2, x_3^2, ...)$. Now, the argument is best understood visually. Place both sequence as the two columns of a matrix with an infinite number of rows. Now, snake along the rows to define the sequence $(x_1^1, x_1^2, x_2^1, x_2^2, x_3^1, x_3^1, ...)$. This way, we count all the arguments of $S_1 \cup S_2$. (Note that if the sets are not disjoint, we are counting some elements several times; but this does not invalidate the proof: just skip the elements that are repeating themselves). □

Corollary 5.1. *$\mathbb{Z}$ is (infinitely) countable.*

Proof. Because $\mathbb{Z} = \mathbb{N} \cup (-\mathbb{N})$. □

The previous proposition can be generalized to be much stronger: actually a *countable* union of countable sets is countable.

Theorem 5.1. *The countable union of countable sets is countable.*

Proof. We need to extend the previous proof to an infinitely countable union. Consider an infinitely countable collection of set $\{S_1, S_2, ...\}$. For each set S_j, since S_j is countable, we can index its argument as $(x_1^j, x_2^j, x_3^j, ...)$. Place each sequence $(x_1^j, x_2^j, x_3^j, ...)$ as the j^{th} column of a matrix with infinitely many rows. The problem is that this time, the matrix has also infinitely many columns, so we cannot follow our previous proof. However, Georg Cantor, the late XIXth-century mathematician who founded set theory, got a smart idea: snake along the diagonals to define the sequence $(x_1^1, x_2^1, x_1^2, x_3^1, x_2^2, x_1^3, ...)$. This way, we count all the arguments of $\bigcup_n S_n$. (Again, some elements may repeat themselves; just drop them). □

Using this proposition, it is easy to show that $\mathbb{Q}$ is countable.

Corollary 5.2. *$\mathbb{Q}$ is countable.*

Proof. Partition $\mathbb{Q}$ by the value of the denominators of fractions: $\mathbb{Q} = \bigcup_{q \in \mathbb{N}, q \neq 0} Q_p$ with $Q_p = \{\frac{p}{q}, p \in \mathbb{N}\}$. Each Q_p is countable, so $\mathbb{Q}$ is a countable union of countable sets. □

What about the cartesian product of countable sets? This time, only *finite* unions of countable sets are countable.

Proposition 5.2. *The cartesian product of finitely many countable sets is countable.*

Proof. It is enough to prove it for the cartesian product of two sets; the result will then follow by induction. The proof is a corollary of the result on the countable union of countable sets. Let S and T be two countable sets. Note $(x_1, x_2, ...)$ the elements of S. Write the cartesian product as $S \times T = \bigcup_n T_n$, where $T_n = \{(x_n, y), y \in T\}$. Each T_n is countable, so $S \times T$ is a countable union of countable sets. □

But the cartesian product of a countable number of sets is not in general countable. Indeed, we can prove that a countable cartesian product of countable sets can result in a set of higher cardinality—establishing on the way the existence of "bigger infinities" than countable infinity. A proof of it is Cantor's diagonal argument, another of Cantor's famous proofs. Consider the easiest example of a countable cartesian product we can think of: the cartesian product S of $\{0, 1\}$ with itself "$\mathbb{N}$ times", noted $S = \{0, 1\}^{\mathbb{N}}$ (an element of S is a sequence of 0 and 1). By contradiction, assume it is countable. Then there exists a bijection from $\mathbb{N}$ to S; call it $(s^1, s^2, ..., s^n, ...)$, where each s^j is an element of S, i.e. a sequence of 0 and 1. Place the sequence s^j as the j^{th} column of an infinite matrix. Now, consider the diagonal of this matrix: it is a sequence of 0 and 1, that is an element of S; call it d. Define the sequence $s^* \in S$ such that for all i, s^*_i is the complement of d_i ($s^*_i = 0$ if $d_i = 1$, and conversely). But s^* is different from all $s^j, j \in \mathbb{N}$ since $s^j_j \neq s^*_j$, which means $s^* \notin S$, a contradiction.

Because it turns out that $\mathbb{R}$ is in bijection with $\{0, 1\}^{\mathbb{N}}$, we have established that $\mathbb{R}$ is uncountable.

Proposition 5.3. $\mathbb{R}$ *is uncountable.*

Proof. The proof that $\mathbb{R}$ is in bijection with $\{0, 1\}^{\mathbb{N}}$ is in two steps. First we show that $\mathbb{R}$ is in bijection with the interval $(0, 1)$, then that $(0, 1)$ is in bijection with $\{0, 1\}^{\mathbb{N}}$.

- For the first step, note that $x \mapsto \tan(\pi x - \frac{\pi}{2})$ works.
- For the second step, the idea is to use the binary expression of a real number: any real $x \in (0, 1)$ can be written as $x = \sum_{n=1}^{\infty} a_n 2^{-n}$, so that the sequence $(a_n)_n \in \{0, 1\}^{\mathbb{N}}$ characterizes x. There is a small difficulty because some numbers have actually two such representations: for instance $2^{-1} = \sum_{n=2}^{\infty} 2^{-n}$ are the same number. But we just need to agree not to use the second representation for such numbers.

□

In other words, although $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$ are "infinities of the same size", $\mathbb{R}$ is a "bigger infinity".

A last question: what about $\mathbb{R}^n$? It is easy to see that $\mathbb{R}^n$ has higher or equal cardinality than $\mathbb{R}$: the function from $\mathbb{R}$ to $\mathbb{R}^n$, $x \mapsto (x, 0, ..., 0)$ is injective. So in particular $\mathbb{R}^n$ is uncountable. But is $\mathbb{R}^n$ strictly "bigger" than $\mathbb{R}$? No:

Proposition 5.4. *$\mathbb{R}^n$ (and so $\mathbb{C}$) has the same cardinality as $\mathbb{R}$.*

Proof. By induction, it is enough to prove that $\mathbb{R}^2$ has the same cardinality as $\mathbb{R}$. Since $\mathbb{R}$ is in bijection with $S = \{0,1\}^{\mathbb{N}}$, $\mathbb{R}^2$ is in bijection with S^2. Hence it is enough to find a bijection between S^2 and S. Consider the function $f : S^2 \to S$ that associates to the couple $((a_n)_n, (b_n)_n)$ the interlaced sequence $(a_1, b_1, a_2, b_2, ...)$. It is a bijection. □