Functions and Cardinality

Functions.

- Informally, we can think of a function as a machine, where the input objects are put into the top, and for each input, the machine spits out one output.

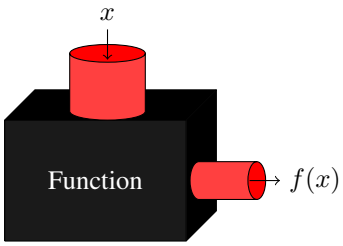

Figure 1: Function as a machine.

- For example, if a function is given by the formula $f(x) = x^2$, then the machine takes the real numbers as inputs. If we put $a = 3$ into the machine, then it will spit out $f(a) = 9$.

Definition. Let A and B be sets.

- A *function* (or *map*) f from A to B, denoted $f : A \to B$, is a subset $F \subseteq A \times B$ such that for each $a \in A$, there is one and only one pair of the form (a, b) in F.
- The set A is called the *domain* of the function, and
- the set B is called the *codomain* of the function. ▲

- Note that the definition consists of three things:
  - a domain,
  - a codomain, and
  - a subset of the product of the domain and codomain satisfying a certain condition.
- For two functions to be considered equal, they need to have all three of these things be the same. For example, the function
  - $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2 + 1$ for all $x \in \mathbb{R}$ is not the same function as
  - $g : \mathbb{R} \to \mathbb{R}^+$ given by $g(x) = x^2 + 1$ for all $x \in \mathbb{R}$, even though they both have the same formula and the same domain.
- Notice that in our definition of a function we did not mention the term "$f(a)$" – our definition was written entirely in terms of sets.

- We could instead use our intuitive "rule of assignment" approach that we are used to, and define $f(a) = b$ where b is the unique element in B such that the pair (a, b) is in F.
- Be careful not to write phrases such as "let $f(x)$ be a function."
  - If $f : A \to B$ is a function, then the name of the function is "f", not "$f(x)$."
  - The symbol "$f(x)$" denotes the *value* of the function f at the element x in the domain, and so $f(x)$ is an element of the codomain.
- Thus when we define a function by writing "let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = \cos x$ for all $x \in \mathbb{R}$," the phrase "for all $x \in \mathbb{R}$" crucial.

Definition. Let A and B be sets, and let $S \subseteq A$.

- A *constant map* $f : A \to B$ is any function of the form $f(x) = b$ for all $x \in A$, where $b \in B$ is some fixed element.
- The *identity map* on A is the function $1_A : A \to A$ defined by $1_A(x) = x$ for all $x \in A$.
- The *inclusion map* from S to A is the function $j : S \to A$ defined by $j(x) = x$ for all $x \in S$.
- If $f : A \to B$ is a map, the *restriction* of f to S denoted $f|_S$, is the map $f|_S : S \to B$ defined by $f|_S = f(x)$ for all $x \in S$.
- If $g : S \to B$ is a map, an *extension* of g to A is any map $G : A \to B$ such that $G|_S = g$.
- The *projection maps* from $A \times B$ are the functions $\pi_1 : A \times B \to A$ and $\pi_1 : A \times B \to B$ defined by $\pi_1((a, b)) = a$ and $\pi_2((a, b)) = b$ for all $(a, b) \in A \times B$. Projection maps

$$\pi_i : A_1 \times \cdots \times A_n \to A_i$$

  for any finite collection of sets $A_1, \ldots, A_n$ can be defined similarly. ▲

Example 1.

1. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = \sin x$ for all $x \in \mathbb{R}$. Then the restriction of f to $\mathbb{Q}$ is the map $f|_{\mathbb{Q}} : \mathbb{Q} \to \mathbb{R}$ defined by $f|_{\mathbb{Q}}(x) = \sin x$ for all $x \in \mathbb{Q}$.
2. Let $X = \{a, b, c, d\}$, let $Y = \{a, b\}$, and let $Z = \{1, 2, 3\}$. Let $f : Y \to Z$ be defined by $f(a) = 3$ and $f(b) = 2$. We now define two maps $g, h : X \to Z$ as follows.
   - Let $g(a) = 3$, let $g(b) = 2$, let $g(c) = 1$ and let $g(d) = 3$.
   - Let $h(a) = 3$, let $h(b) = 1$, let $h(c) = 2$ and let $h(d) = 3$.

   Then the map g is an extension of f, since $g|_Y = f$, but the map h is not an extension of f since $h(b) \neq f(b)$.

Image and Inverse Image.

Definition. Let $f : A \to B$ be a function. For each $P \subseteq A$, let $f_*(P)$ be defined by

$$f_*(P) = \{b \in B \mid b = f(p) \text{ for some } p \in P\} = \{f(p) \mid p \in P\}$$

For each $P \subseteq A$, the set $f_*(P)$ is called the *image* of P under f. The *range* of the function f is the set $f_*(A)$. The range is also known as the *image* of f. ▲

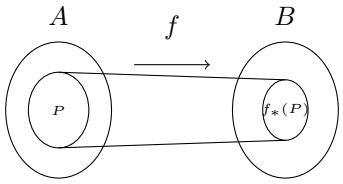

Figure 2: The image of P under f.

- Note that the notation $f_*(P)$ is not standard. It is common to abuse notation and write $f(P)$.

Definition. Let $f : A \to B$ be a function. For each $Q \subseteq B$, let $f^*(Q)$ be defined by

$$f^*(Q) = \{a \in A \mid f(a) = q \text{ for some } q \in Q\} = \{a \in A \mid f(a) \in Q\}$$

For each $Q \subseteq B$, the set $f^*(Q)$ is called the *inverse image* of Q under f. ▲

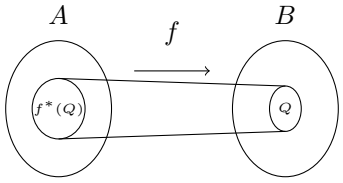

Figure 3: The inverse image of Q under f.

Composition and Inverse Functions.

- The notation $f^*(Q)$ for the inverse image is not standard. Commonly notation is abused and we write $f^{-1}(Q)$ but this can cause confusion between
  - the notion of an inverse image (which is a subset of the set A and always exists) and
  - an inverse function denoted f^{-1} (which may not exist).
- The notation $f^{-1}(Q)$ can also be misleading as f_* and f^* do not "cancel each other out" as we can see in the following example.

Example 2. Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = x^2$ for all $x \in \mathbb{R}$.

- It is straightforward to compute that $f_*([0,3]) = [0,9]$ and $f^*([-4,4]) = [-2,2]$.
- Hence $f^*(f_*([0,3])) = f^*([0,9]) = [-3,3]$ and we see that $f^*(f_*([0,3])) \neq [0,3]$.
- Similarly $f_*(f^*([-4,4])) = f_*[-2,2] = [0,4]$ and we see that $f_*(f^*([-4,4])) \neq [-4,4]$. ♦
- The following theorem establishes some basic properties of images and inverse images.

Theorem 1. *Let $f : A \to B$ be a function. Let $C, D \subseteq A$, and let $S, T \subseteq B$.*

1. *$f_*(\emptyset) = \emptyset$ and $f^*(\emptyset) = \emptyset$.*
2. *$f^*(B) = A$.*
3. *$f_*(C) \subseteq S$ iff $C \subseteq f^*(S)$.*
4. *If $C \subseteq D$ then $f_*(C) \subseteq f_*(D)$.*
5. *If $S \subseteq T$,then $f^*(S) \subseteq f^*(T)$.*

Definition. Let $f : A \to B$ and $g : B \to C$ be functions. The *composition* of f and g is the function $g \circ f : A \to C$ defined by

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

for all $x \in A$. ▲

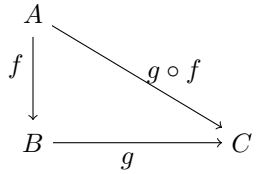

Figure 4: The composition of f and g.

- An example of the use of composition of maps is coordinate functions. In multivariable calculus maps into $\mathbb{R}^n$ are written in terms of coordinate functions. We can generalize this idea to arbitrary sets.

Definition. Let $A, A_1, \ldots, A_n$ be sets, for some positive integer n, and let $f : A \to A_1 \times \cdots \times A_n$ be a function. For each $i \in \{1, \ldots, n\}$, let $f_i : A \to A_i$ be defined by $f_i = \pi_i \circ f$, where $\pi_i : A_1 \times \cdots \times A_n \to A_i$ is the projection map. The functions $f_1, \ldots, f_n$ are the *coordinate functions* of f. ▲

- In the definition above, we see that $f(x) = (f_1(x), \ldots, f_n(x)) \in A_1 \times \cdots \times A_n$ for all $x \in A$.

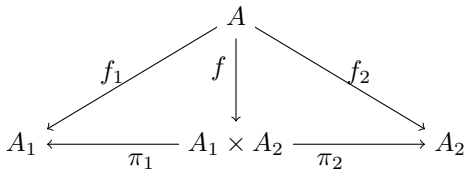

Figure 5: The coordinate functions of $f : A \to A_1 \times A_2$

Example 3.

- Let $f : \mathbb{R}^2 \to \mathbb{R}^3$ be given by

$$f((x, y)) = (xy, \sin x^2, x + y^3)$$

for all $(x, y) \in \mathbb{R}^2$.

- The three coordinate functions of f are $f_1, f_2, f_3 : \mathbb{R}^2 \to \mathbb{R}$ given by

$$f_1((x, y)) = xy, \quad f_2((x, y)) = \sin x^2, \quad \text{and} \quad f_3((x, y)) = x + y^3$$

for all $(x, y) \in \mathbb{R}^2$. ♦

Theorem 2. *Let $f : A \to B$, let $g : B \to C$ and let $h : C \to D$ be functions.*

1. $(h \circ g) \circ f = h \circ (g \circ f)$ *(Associative Law).*
2. $f \circ 1_A = f$ *and* $1_B \circ f = f$ *(Identity Law).*

- We next look at whether functions have inverses under composition.

Definition. Let $f : A \to B$ and $g : B \to A$ be functions.

- The function g is a *right inverse* for f if $f \circ g = 1_B$.
- The function g is a *left inverse* for f if $g \circ f = 1_A$.
- The function g is an *inverse* for f if it is both a right inverse and a left inverse. ▲

Theorem 3. *Let $f : A \to B$ be a function.*

1. *If f has an inverse, then the inverse is unique.*
2. *If f has a right inverse g and a left inverse h, then $g = h$; hence f has an inverse.*
3. *If f has an inverse g, then g has an inverse, which is f.*

Proof of (1). Suppose that $g, h : B \to A$ are both inverses of f. We will show that $g = h$. By hypothesis on g and h, we know that

$$f \circ g = 1_B \quad \text{and} \quad h \circ f = 1_A.$$

Now,

$$g = 1_A \circ g = (h \circ f) \circ g = h \circ (f \circ g) = h \circ 1_B = h. \qquad \blacksquare$$

Injectivity, Surjectivity and Bijectivity.

- Are there any simple criteria by which to check whether a function has a left inverse, right inverse or both?
- To answer this question we need to define some properties which any given map may or may not have.

Definition. Let $f : A \to B$ be a function.

- The map f is *injective* (or *one-to-one*) if $x \neq y$ implies $f(x) \neq f(y)$ for all $x, y \in A$ Equivalently, f is injective if $f(x) = f(y)$ implies $x = y$ for

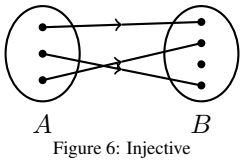

Figure 6: Injective

all $x, y \in A$.

- The map f is *surjective* (or *onto*) if for every $b \in B$, there exists some $a \in A$ such that $f(a) = b$. Equivalently, f is surjective if $f_*(A) = B$ (the

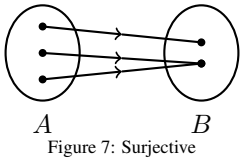

Figure 7: Surjective

range of f is equal to the codomain).

- The map f is *bijective* if it is both injective and surjective. ▲

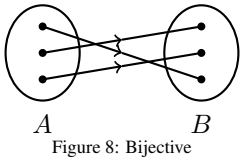

Figure 8: Bijective

Example 4.

1. Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = x^2$ for all $x \in \mathbb{R}$. Then f is neither injective nor surjective.
2. Let $g : \mathbb{R}^+ \to \mathbb{R}$ be given by $g(x) = x^2$ for all $x \in \mathbb{R}^+$. Then g is injective but not surjective.

3. Let $h : \mathbb{R} \to \mathbb{R}^+$ be given by $h(x) = x^2$ for all $x \in \mathbb{R}$. Then h is surjective but not injective.

4. Let $k : \mathbb{R}^+ \to \mathbb{R}^+$ be given by $k(x) = x^2$ for all $x \in \mathbb{R}^+$. Then k is injective and surjective.

- With these definitions, we can now describe simple tests of whether a function has a left inverse, a right inverse or both.

Theorem 4. *Let $f : A \to B$ and $g : B \to C$ be functions.*

1. *If f and g are both injective, then so is $g \circ f$.*
2. *If f and g are both surjective, then so is $g \circ f$.*
3. *If f and g are both bijective, then so is $g \circ f$.*

Theorem 5. *Let A and B be non-empty sets, and let $f : A \to B$ be a function.*

1. *The function f has a right inverse iff f is surjective.*
2. *The function f has a left inverse iff f is injective.*
3. *The function f has an inverse iff f is bijective.*

Example 5. We go back to our simple example.

1. Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = x^2$ for all $x \in \mathbb{R}$. Since f is neither injective nor surjective it has no type of inverse.

2. Let $g : \mathbb{R}^+ \to \mathbb{R}$ be given by $g(x) = x^2$ for all $x \in \mathbb{R}^+$. Since g is injective but not surjective, it has a left inverse, but no right inverse or inverse.

   - To find a left inverse, go back to the definition. A function $p : \mathbb{R} \to \mathbb{R}^+$ is a left inverse for g if $p \circ g = 1_{\mathbb{R}^+}$.
   - Using the definitions of composition and the identity map we can write this is as: a function p is a left inverse for g if

     $$p(g(x)) = x$$

     for all $x \in \mathbb{R}^+$.

2. - Now use the definition of g to write p is a left inverse for g if

     $$p(x^2) = x$$

     for all $x \in \mathbb{R}^+$. Thus $p(x) = \sqrt{x}$ for all $x \in \mathbb{R}^+$.

   Note that our definition of the inverse only resricted the value of the funtion p for positive real numbers. For $x \in (-\infty, 0]$, we can define the function to take any value we want. So two left inverses are $p, q : \mathbb{R} \to \mathbb{R}^+$, given by

$$p(x) = \begin{cases} \sqrt{x}, & \text{if } x > 0 \\ 1 & \text{if } x \leq 0, \end{cases} \qquad q(x) = \begin{cases} \sqrt{x}, & \text{if } x > 0 \\ \sin x & \text{if } x \leq 0, \end{cases}$$

They are both left inverses because

$$(p \circ g)(x) = p(g(x)) = \sqrt{x^2} = x$$

for all $x \in \mathbb{R}^+$ and similarly for q.

3. Let $h : \mathbb{R} \to \mathbb{R}^+$ be given by $h(x) = x^2$ for all $x \in \mathbb{R}$. Since h is surjective but not injective, it has a right inverse, but no left inverse or inverse. Two right inverses $r, s : \mathbb{R}^+ \to \mathbb{R}$ are given by $r(x) = \sqrt{x}$ and $s(x) = -\sqrt{x}$ for all $x \in \mathbb{R}^+$. They are both right inverses because

$$(h \circ r)(x) = h(r(x)) = (\sqrt{x})^2 = x.$$

for all $x \in \mathbb{R}^+$ and similarly for s.

4. Let $k : \mathbb{R}^+ \to \mathbb{R}^+$ be given by $k(x) = x^2$ for all $x \in \mathbb{R}^+$. Since k is injective and surjective, it has an inverse. The inverse is the function $t : \mathbb{R}^+ \to \mathbb{R}^+$ given by $t(x) = \sqrt{x}$ for all $x \in \mathbb{R}^+$.

Cardinality of Sets.

- We now look at the sizes of sets. It is obvious that finite sets come in different sizes, but what about infinite sets?
- How do we determine whether two sets have the "same size"?
  - The idea of the size of a finite set is intuitive - just count the number of elements.
  - The notion of "size" for an infinite set is less obvious.
  - Turns out to discuss whether two sets have the "same size" we do not need to first figure out the size of each set.
- A simple example illustrates the idea.
  - Suppose a group of people want to stay at a hotel, with each person in a separate room. The hotel manager only wants to take the group if it completely fills the hotel. We need to find out whether the right number of rooms are vacant.
  - There are two ways to proceed.
  - We could count the number of people and the number of free rooms, then see if the two numbers are the same.
  - Or, we could make a list of people, a list of free rooms and go down the lists and match up each person to a distinct vacant room. If all people and all rooms are taken care of then everyone is happy.The first method works only when everything is finite but the second works even if the number of people and number of rooms are infinite.

- The following definition formalizes this idea using bijective maps.

Definition. Let A and B be sets. We say that A and B have the *same cardinality*, written $A \sim B$, if there is a bijective map $f : A \to B$. ▲

Lemma 1. *Let A, B and C be sets.*

1. *$A \sim A$.*
2. *If $A \sim B$ then $B \sim A$.*
3. *If $A \sim B$ and $B \sim C$, then $A \sim C$.*

Definition. Let $n \in \mathbb{N}$. We let $[\![1, n]\!]$ denote the set $\{x \in \mathbb{N} \mid 1 \leq x \leq n\}$. ▲

Definition.

- A set is *finite* if it is either the empty set or it has the same cardinality as $[\![1, n]\!]$ for some $n \in \mathbb{N}$.
- A set is *infinite* if it is not finite.
- A set is *countably infinite* if it has the same cardinality as $\mathbb{N}$.
- A set is *countable* if it is finite or countably infinite.
- A set is *uncountable* if it is not countable. ▲

Definition. Let A be a finite set. We define the *cardinality* of A, denoted $|A|$, as follows. If $A = \emptyset$, we let $|A| = 0$. If $A \neq \emptyset$, we let $|A| = n$, where $A \sim [\![1, n]\!]$. ▲

- Could it happen that a set has the same cardinality as both $[\![1, n]\!]$ and $[\![1, m]\!]$ for different positive integers n and m?
- The next lemma ensures that this cannot happen and so the above definition is sound.

Lemma 2. *Let $n, m \in \mathbb{N}$. Then $[\![1, n]\!] \sim [\![1, m]\!]$ iff $n = m$.*

Corollary 1. *Let A and B be finite sets. Then $A \sim B$ iff $|A| = |B|$.*

- The corollary tells us that the approach of pairing up elements of two sets (using a bijective map) and that of counting the number of elements in two sets and comparing the numbers give the same result.
- We next list some properties of the cardinalities of finite sets.

Theorem 6. *Let A and B be finite sets.*

1. *If $X \subseteq A$, then X is finite.*
2. *If $X \subseteq A$, then $|A| = |X| + |(A \setminus X)|$.*
3. *If $X \subset A$, then $|X| < |A|$.*

4. *If $X \subset A$, then $X \not\sim A$.*

Corollary 2. *Let A be a set. Then A is infinite iff it contains an infinite subset.*

Proof.

- Suppose A contains an infinite subset and that A is finite. But then by (1), any subset X of A is finite, a contradiction.
- Suppose A is infinite and it does not contain an infinite subset. But A is an infinite subset of A, a contradiction. ■
- There is one thing we need to clear up.
  - Earlier we defined "countably infinite" which suggests that such a set is infinite.
  - But we defined "countably infinite" and "infinite" separately.
  - We have to prove that countably infinite sets are indeed infinite.

Theorem 7.

1. *The set $\mathbb{N}$ is infinite.*
2. *A countably infinite set is infinite.*

Proof of (2). Let B be a countably infinite set. Then $B \sim \mathbb{N}$.

- Suppose that B is finite. Then $B \sim [\![1, n]\!]$ for some $n \in \mathbb{N}$.
- Since $\sim$ is an equivalence relation (lemma 1) we deduce that $\mathbb{N} \sim [\![1, n]\!]$ a contradiction to part (1) of this theorem.

Thus B is infinite. ■

- We now look at uncountable sets. To show that there exists such sets we look at the cardinality of the power set.

Example 6. Let $A = \{1, 2\}$. Then $\mathcal{P}(A) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$. Thus $A \not\sim \mathcal{P}(A)$.

- This result generalizes.

Theorem 8. *Let A be a set. Then $A \not\sim \mathcal{P}(A)$.*

Corollary 3. *The set $\mathcal{P}(\mathbb{N})$ is uncountable.*

Proof. By the above theorem, we know that $\mathcal{P}(\mathbb{N}) \not\sim \mathbb{N}$, and so $\mathcal{P}(\mathbb{N})$ is not countably infinite. To show that $\mathcal{P}(\mathbb{N})$ is not countable, we need to show that it is not finite.

- Suppose $\mathcal{P}(\mathbb{N})$ is finite. Let $T = \{\{n\} \mid n \in \mathbb{N}\} \subseteq \mathcal{P}(\mathbb{N})$.
- Since a subset of a finite set is finite (theorem 6 1), it follows that T is finite.
- But clearly $T \sim \mathbb{N}$, which implies that $\mathbb{N}$ is finite, a contradiction. ■
- Putting our results together, we see that any set is precisely one of finite, countably infinite or uncountable, and that there exist sets of each type.

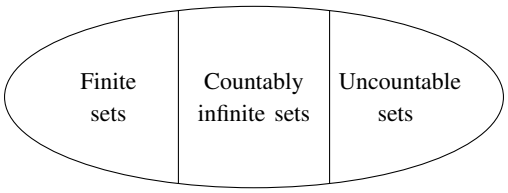

Figure 9: Types of sets.

Cardinality of the Number Systems.

- We now use the results we have seen to discuss the cardinality of the standard number systems, i.e. the natural numbers, integers, rational numbers, real numbers and complex numbers.

Theorem 9. *The set $\mathbb{Z}$ is countably infinite.*

Proof.

- We need to show that $\mathbb{N} \sim \mathbb{Z}$, i.e. $\mathbb{Z}$ has the same cardinality as $\mathbb{N}$.
- By definition of cardinality, we need to show there is a bijective map $f : \mathbb{N} \to \mathbb{Z}$.
- Let f be given by

$$f(n) = \begin{cases} \frac{n}{2}, & \text{if } n \text{ is even} \\ -\frac{n-1}{2}, & \text{if } n \text{ is odd.} \end{cases}$$

- We need to show that this map is bijective i.e. that it is injective and surjective.
- First we show that f is injective. Consider any $x, y \in \mathbb{N}$ such that $f(x) = f(y)$. We want to show that $x = y$. We must consider two cases.
  - If $f(x) = f(y) > 0$, then $x = 2f(x)$ and $y = 2f(y)$ so that $x = y$.
  - If $f(x) = f(y) \leq 0$, then $x = -2f(x) + 1$ and $y = -2f(y) + 1$ so that $x = y$.

  Thus f is injective.
- Now we show f is surjective. Let $z \in \mathbb{Z}$. We need to show there exists an $n \in \mathbb{N}$ such that $f(n) = z$.
  - If $z \geq 0$, we can let $n = 2z$ and then $f(n) = z$.
  - If $z < 0$, we can let $n = -2z + 1$ and then $f(n) = z$.
- We have shown that the map $f : \mathbb{N} \to \mathbb{Z}$ is bijective, and so $\mathbb{N} \sim \mathbb{Z}$. ■
- If we think of the real number line, the integers sit "discretely" in $\mathbb{R}$ – there are 'gaps" between the integers.

- In contrast, the rational numbers are "dense" in $\mathbb{R}$ – between any two real numbers, we can always find a rational number.
- It would appear that there are "more rational numbers than integers". The following theorem shows that this intuition is not correct.

Theorem 10. *The set $\mathbb{Q}$ is countably infinite.*

- This theorem tells us that the elements of $\mathbb{Q}$ can be "lined up" in order like the elements of $\mathbb{N}$, though not necessarily according to increasing size.
- Cantor came up with a way of lining up the elements of $\mathbb{Q}$.
- Follow the path of the arrows, and drop every fraction that is equal to one that has already been encountered.

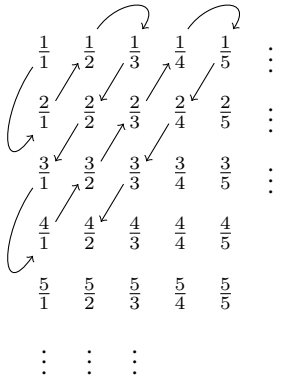

Figure 10: Cantor's method for lining up the elements of $\mathbb{Q}$.

- The following theorem tells us that $\mathbb{R} \not\sim \mathbb{N}$. (In fact $\mathbb{R} \sim \mathcal{P}(\mathbb{N})$.)

Theorem 11. *The set $\mathbb{R}$ is uncountable.*

- The set of irrational numbers is the set of all real numbers that are not rational, i.e. $\mathbb{R} \setminus \mathbb{Q}$. The following theorem should not be take as obvious, since not every uncountable set has the same cardinality as $\mathbb{R}$. (For example, $\mathcal{P}(\mathbb{R}) \not\sim \mathbb{R}$.)

Theorem 12. *The set of irrational numbers has the same cardinality as $\mathbb{R}$.*

- The next theorem, which says that the n-fold Cartesian product of R has the same cardinality as $\mathbb{R}$ for every $n \in \mathbb{N}$, seems slightly counterintuitive.

Theorem 13. *Let $n \in \mathbb{N}$. Then $\mathbb{R}^n \sim \mathbb{R}$.*

Mathematical Induction.

- Mathematical induction is a very useful method for proving certain kinds of statements.
- It is used to prove statements of the form $(\forall n \in \mathbb{N})(P(n))$, where $P(n)$ is some statement involving n.

Theorem 14 (Principle of Mathematical Induction)**.** *Let $G \subseteq \mathbb{N}$. Suppose that*

1. $1 \in G$*;*
2. *if $n \in G$, then $n + 1 \in G$.*

Then $G = \mathbb{N}$.

- When using the PMI, we do not need to show directly that $n \in G$, nor that $n + 1 \in G$. We only need to show that $n \in G$ implies $n + 1 \in G$.

Example 7. We will show that $8^n - 3^n$ is divisible by 5 for all $n \in \mathbb{N}$.

- We start by defining a set G by

$$G = \{n \in \mathbb{N} \mid 8^n - 3^n \text{ is divisible by } 5\}.$$

  If we can show that $G = \mathbb{N}$, then it follows that $8^n - 3^n$ is divisible by 5 for all $n \in \mathbb{N}$. We will use the PMI to show that $G = \mathbb{N}$.

- First, note that $G \subseteq N$ by definition.
- To use the PMI, we first need to show that $1 \in G$. Observe that $8^1 - 3^1 = 5$ is divisible by 5, so that indeed $1 \in G$ and part (1) of the statement of the PMI holds.
- To show part (2), we assume that $n \in G$. We then need to deduce that $n + 1 \in G$. Since $n \in G$, we know that $8^n - 3^n$ is divisible by 5. This means that there is some $k \in \mathbb{Z}$ such that $8^n - 3^n = 5k$. To show that $n + 1 \in G$, we are need to show that $8^{n+1} - 3^{n+1}$ is divisible by 5. We can use our inductive hypothesis that $8^n - 3^n$ is divisible by 5 in this proof. Now

$$\begin{aligned} 8^{n+1} - 3^{n+1} &= 8 \cdot 8^n - 3 \cdot 3^n \\ &= (5 \cdot 8^n + 3 \cdot 8^n) - 3 \cdot 3^n \\ &= 5 \cdot 8^n + 3(8^n - 3^n) \\ &= 5 \cdot 8^n + 3(5k) \\ &= 5(8^n + 3k). \end{aligned}$$

  Hence $8^{n+1} - 3^{n+1}$ is divisible by 5, and so $n + 1 \in G$. We have shown part (2) of the PMI holds. The PMI now implies that $G = \mathbb{N}$, and the result is proved. ♦

- We can make proof by induction less cumbersome by avoiding mentioning the set G explicitly.
- Suppose we are trying to show the statement $P(n)$ holds for all $n \in \mathbb{N}$.
  - First state that we are trying to prove the statement $P(n)$ for all $n \in \mathbb{N}$ by induction.
  - Then show that $P(1)$ holds.
  - Finally assume that $P(n)$ holds (the inductive hypothesis) and use this to show that $P(n+1)$ holds. (This is called the inductive step).

Example 8. We will prove that all cats are the same color, i.e. that the statement "in any collection of n cats, all the cats have the same colour," is true for all $n \in \mathbb{N}$. Since there are only finitely many cats in the world, it will then follow that all cats have the same colour.

- First suppose $n = 1$. It is obviously true that in any collection of one cat, all cats have the same colour.
- Now suppose the result is true for n. This is our inductive hypothesis: in any collection of n cats, all the cats have the same colour.
- We need to show that the result is true for $n + 1$.
  - Let $\{C_1, \ldots, C_{n+1}\}$ be a collection of $n + 1$ cats.
  - The set $\{C_1, \ldots, C_n\}$ has n cats and so by hypothesis all cats in this set have the same colour. Also, the set $\{C_2, \ldots, C_{n+1}\}$ has n cats, so all cats in this set have the same colour.
  - It follows, then, that cats C_n and C_{n+1} have the same colour.
  - Combining this fact with the observation that $C_1, \ldots, C_n$ have the same colour, it follows that $C_1, \ldots, C_{n+1}$ all have the same colour.

  We have thus proved the inductive step. Hence all cats have the same colour! ⧫

- There are also some variants of the PMI, which can be useful, which I will list for completeness.

Theorem 15 (Principle of Mathematical Induction – Variant 1)**.** *Let $G \subseteq \mathbb{N}$ and let $k_0 \in \mathbb{N}$. Suppose that*

1. *$k_0 \in G$;*
2. *if $n \in \mathbb{N}$ is such that $n \geq k_0$ and $n \in G$, then $n + 1 \in G$.*

Then $\{n \in \ \mid n \geq k_0\} \subseteq G$.

- This variant is useful when we wish to prove that a statement $P(n)$ is true for all natural numbers n such that $n \geq k_0$, for some given natural number k_0.
- For example, PMI-V1 can be used to prove "if $n \in \mathbb{N}$ and $n \geq 5$, then $4^n > n^4$."

Theorem 16 (Principle of Mathematical Induction – Variant 2)**.** *Let $G \subseteq \mathbb{N}$. Suppose that*

1. *$1 \in G$;*
2. *if $n \in \mathbb{N}$ and $\{i \in \mathbb{N} \mid 1 \leq i \leq n\} \subseteq G$, then $n + 1 \in G$.*

Then $G = \mathbb{N}$.

- When using PMI-V2, the inductive step involves showing that if the desired statement is assumed to hold for all values in $\{1, \ldots, n\}$, then it holds for $n + 1$.

Theorem 17 (Principle of Mathematical Induction – Variant 3)**.** *Let $G \subseteq \mathbb{N}$ and let $k_0 \in \mathbb{N}$. Suppose that*

1. *$k_0 \in G$;*
2. *if $n \in \mathbb{N}$ is such that $n \geq 0$ and $\{i \in \mathbb{N} \mid k_0 \leq i \leq n\} \subseteq G$, then $n + 1 \in G$.*

Then $\{n \in \mathbb{N} \mid n \geq k_0\} \subseteq G$.

- For example, PMI-V3 can be used to prove that "for $n \in \mathbb{N}$ such that $n \geq 2$, either n is prime or it is the product of finitely many prime numbers."