Introductory topics in mathematics

Jia Xiaodong

May 9, 2022

This is an introduction to a few topics in modern mathematics.

1 Set theory

1.1 Preliminaries

This section will define some notions more concretely before we begin. We do not focus on the axioms, but will mention some of the more important ones. We will also touch on some of the later on, like the axiom of replacement and the axiom of choice.

Set existence axiom $\exists x\,(x = x)$.

Axiom of extensionality $\forall x \forall y\,[\forall z\,(z \in x \leftrightarrow z \in y) \to x = y]$.

Comprehension schema For each formula $\phi(x)$, $\forall z \exists y \forall x\,[x \in y \leftrightarrow (x \in z \wedge \phi(x))]$.

Pairing axiom We can form pairs: $\forall x \forall y \exists z\,(x \in z \wedge y \in z)$.

Union axiom We can form unions of sets.

Power axiom We can form power sets.

We know the union of two sets $X \cup Y$. What about arbitrary unions? For any set A, define

$$\bigcup A = \{b \mid \exists a\,(a \in A \wedge b \in a)\}.$$

In other words, the (big) union of A collects the members of members of A. The axiom of union guarantees that $\bigcup A$ is a set. For instance if $A = \{X, Y\}$ then $\bigcup A = X \cup Y$. We define intersections the same way for non-empty A:

$$\bigcap A = \{b \mid \forall a\,(a \in A \implies b \in a)\}.$$

1.2 Relations and functions

1.2.1 Ordered pairs

If you have two sets a and b, you can form a new set containing both a and b. This forms an unordered pair:

$$\{a, b\}$$

Definition 1.1. Let the *ordered pair* of a and b be denoted (a, b). It must satisfy the property $(a, b) = (c, d)$ iff $a = c$ and $b = d$. □

Theorem 1.1 (Hausdorff 1914)**.** *Choose two objects x, y that are not a or b. The following is a valid definition of an ordered pair:*

$$(a, b) = \{\{a, x\}, \{b, y\}\}$$

Proof.

($\Longrightarrow$): If

$$\{\{a, x\}, \{b, y\}\} = \{\{c, x\}, \{d, y\}\}$$

then it is quite obvious that $a = c$ and $b = d$.

($\Longleftarrow$): By substitution. ■

Theorem 1.2 (Wiener 1914)**.** *The following is a valid definition of an ordered pair:*

$$(a, b) = \{\{\{a\}, \varnothing\}, \{\{b\}\}\}$$

Proof. Similar to above. ■

The following definition is the one that is most often used today.

Theorem 1.3 (Kuratowski 1921)**.** *The following definition satisfies the defining property of ordered pairs:*

$$(a, b) = \{\{a\}, \{a, b\}\}.$$

Proof.
($\Longrightarrow$): Suppose that

$$\{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\}.$$

Any element on the left hand side must appear on the right hand side. We patiently consider all possible cases:

- If $\{a\} = \{c\}$
  - and if $\{a, b\} = \{c\} = \{a\}$, then $b = a = c$. Then the left hand side becomes $\{\{a\}\}$ and $d = a$.
  - otherwise if $\{a, b\} = \{c, d\} = \{a, d\}$,
    - and if $b \neq a$ then $d = b$.
    - otherwise $b = a$ and $\{a, b\} = \{a\} = \{a, d\}$ which means $a = b = d = c$.
- $\{a\} = \{c, d\}$

– then $a = c = d$. The right hand side reduces to $\{\{a\}\}$ and we gather $b = a$.

($\Longleftarrow$): By substitution. ■

Definition 1.2. The *Cartesian product* of A and B, denoted $A \times B$ is defined by

$$A \times B = \{(a, b) \mid a \in A \wedge b \in B\}$$

□

We can generalize the definitions. For example we can define $A \times B \times C = (A \times B) \times C$. We also use A^n to denote $A \times \cdots \times A$ n-times.

1.2.2 Relations and functions

Definition 1.3 (Relations)**.** A *binary relation* is a set of ordered pairs. Let R be a binary relation, then

- the *domain* of R is defined as $\text{dom}(R) = \{x \mid \exists y\,[(x, y) \in R]\}$
- and the *range* of R is defined as $\text{ran}(R) = \{y \mid \exists x\,[(x, y) \in R]\}$.

We write xRy or $R(x, y)$ instead of $(x, y) \in R$. □

Definition 1.4 (Functions)**.** A *function* f is a binary relation such that for every x there is at most one y for which $(x, y) \in f$:

$$(x, y) \in f \wedge (x, z) \in f \implies y = z.$$

We write $f(x) = y$ instead of $(x, y) \in f$. □

The word function was first used by Leibniz. Later on Dirichlet defined functions in terms of an arbitrary correspondence between variables representing numerical sets. This was to separate the concept of function from its analytical representation to prove Fourier's claim regarding expansion of functions as sums of trigonometric functions.

Contrast this definition of functions with the naive definition of functions as rules. This view stems from the view of functions as analytic expressions. However the rules are irrelevant and we only care about what set it forms. For example $f(n) = n + 1 - 1$ is the same as $f(n) = n$, because it describes the same set.

Definition 1.5. A *binary operation* on S is a function mapping a subset of $S \times S$ to S.

A is said to be *closed* under the operation f if

$$\forall x, y \in \text{dom}(f)\ (f(x, y) \in A).$$ □

This condition only has to hold when f is defined. For example, the rational numbers are closed under division.

Definition 1.6. Let R be a binary relation in A.

i. R is *antisymmetric* if $\forall a, b \in A, aRb \wedge bRa \implies a = b$.

ii. R is *asymmetric* if $aRb \implies \neg bRa$.

□

1.2.3 Orderings

Definition 1.7 (Partial orderings)**.** A binary relation R in A which is reflexive, antisymmetric, and transitive (RAT) is called a *partial ordering* of A.

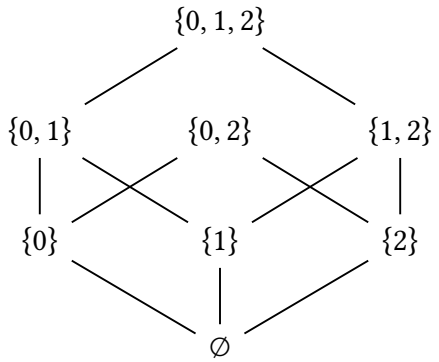
Then the pair (A, R) is called an partially ordered set (poset). We often use $\leq$ to denote a partial order. □

Definition 1.8 (Strict orders)**.** A binary relation R in A is a *strict* partial ordering if it is asymmetric and transitive. □

We often use $<$ to denote a strict partial ordering and $\leq$ to denote a non-strict partial ordering.

Example 1.1. $(\mathscr{P}(\{0,1,2\}), \subseteq)$ [1] is a partially ordered set.

Here is a hierarchy with some edges omitted for clarity.



It is reflexive since $x \subseteq x$. It is antisymmetric since if $x \subseteq y$ and $y \subseteq x$ then $x = y$. It is also transitive because $x \subseteq y$ and $y \subseteq z$ then $x \subseteq z$.

This also highlights why it is called a *partial* order. Certain elements cannot be compared, for example $\{0\}$ and $\{1\}$. ◊

Definition 1.9 (Linear orders)**.** An ordering $\leq$ (or $<$) of A is called *linear* or *total* if any two elements of A are comparable, i.e.

$$\forall x, y\,(x \leq y \vee y \leq x).$$

□

1.3 Natural numbers

Definition 1.10 (Successor)**.** Let 0 denote $\emptyset$, $1 = \{\emptyset\}$, $2 = \{\emptyset, \{\emptyset\}\} = \{0, 1\}$, and so on. For any set x, the *successor* of x, denoted by $S(x)$, is the set $x \cup \{x\}$. Some other common notations for the successor of x are x^+ and $x + 1$. □

Definition 1.11 (Inductive sets)**.** A set I is called *inductive* if

- $0 = \emptyset \in I$, and
- it is closed under successor, i.e. $\forall a \in I\,(S(a) \in I)$.

[1] Here $\mathscr{P}(A)$ denotes the power set of A.

□

The axiom of infinity basically says that an inductive set exists.

Definition 1.12 (Natural numbers)**.** The set of all natural numbers is defined as

$$\mathbb{N} = \{x \mid x \in I \text{ for every inductive set } I\}$$

□

In other words, a set x is a natural number iff it belongs to every inductive set.

Theorem 1.4. *$\mathbb{N}$ is an inductive set.*

Proof. First of all $\emptyset$ is in every inductive set so $\emptyset$ is also in $\mathbb{N}$. Next, if x is in any inductive set I then so is $S(x)$, which means $S(x) \in \mathbb{N}$ as well. ■

In fact, it can be seen that $\mathbb{N}$ is the smallest inductive set since it is effectively the intersection of all inductive sets.

Theorem 1.5 (Induction)**.** *Let $P(x)$ be a property. Assume that*

- *$P(0)$ holds.*
- *$\forall n \in \mathbb{N}\ [P(n) \implies P(n+1)]$.*

Then P holds for all natural numbers n.

Proof. Consider the set $A = \{n \in \mathbb{N} \mid P(n)\}$. Since $P(0)$, $0 \in A$. The inductive step also means that A is closed under successor. So A is inductive. As $\mathbb{N}$ is the smallest inductive set, $\mathbb{N} \subseteq A$. But by definition $A \subseteq \mathbb{N}$. Therefore $A = \mathbb{N}$. ■

1.3.1 Ordering on $\mathbb{N}$

Definition 1.13 (von Neumann)**.** The relation $<$ on $\mathbb{N}$ is defined by $m < n \iff m \in n$. □

Lemma 1.6. *This definition preserves the familiar properties of $<$:*

i. $\forall n \in \mathbb{N}\ (0 \leq n)$ [2]

ii. $\forall k, n \in \mathbb{N}\ (k < n+1 \iff k \leq n)$.

Proof.

i. Induction on n. Let $P(x)$ be the statement $0 \leq x$. Since $0 \leq 0$. $P(0)$ holds trivially. Suppose that $P(n)$ holds. By the definition of $\leq$, either $0 \in n$ or $0 = n$. In either case $0 \in S(n)$. Thus $0 < S(n)$ and $P(S(n))$ holds. By the induction principle, $P(n)$ holds for all $n \in \mathbb{N}$.

ii. If $k < S(n)$ then $k \in n \cup \{n\}$. Therefore either $k \in n$ or $k = n$. Similarly so for the other direction.

[2] Just to be clear here $0 \leq n$ means either $0 < n$ or $0 = n$.

■

Theorem 1.7. $(\mathbb{N}, <)$ *is a linearly ordered set.*

Proof. We check that it is a strict partial order, followed by linearity.

- Transitive: Induction on n. Let $P(x)$ be the statement

$$\forall k, m \in \mathbb{N}\ [(k < m \wedge m < x) \implies k < x].$$

  $P(0)$ is vacuously true. Suppose $P(n)$ holds. Consider $P(n+1)$. By lemma 1.6, $m < n+1$ means $m < n$ or $m = n$. In the first case, $k < n$ by the inductive hypothesis and so $k < n+1$, and in the second case, $k < n$ directly and therefore $k < n+1$ as well.

- Antisymmetric: Let $P(a, b)$ be the statement

$$a < b \implies b \not< a.$$

  We first perform induction on n and show $P(n, n)$ is true for all $n \in \mathbb{N}$. This part is simple and can be skipped. Next, by transitivity, if $m < n$ and $n < m$ then $n < n$ which is impossible, hence $m < n \implies n \not< m$.

- Linear: Let $P(x)$ be the statement

$$\forall m\ (m < x) \vee (m = x) \vee (x < m).$$

  $P(0)$ is true from lemma 1.6. Assume $P(n)$ is true. If $m < n$ or $m = n$ then by transitivity $m < n+1$. If $n < m$, we perform induction on m to show that $n+1 \leq m$. This is skipped.

■

Theorem 1.8. $\forall m, n\ [m < n \iff m+1 < n+1]$.

Proof. We prove the forward direction with induction. Let $P(x)$ be the statement that $\forall m\ [m < x \implies m+1 < x+1]$. $P(0)$ is vacuously true. Suppose $P(n)$ is true for some $n \in \mathbb{N}$. Consider $P(n+1)$. From lemma 1.6, $m < n+1 \implies m \leq n$. If $m < n$ by the inductive hypothesis $m+1 < n+1 < n+1+1$. If $m = n$ then $m+1 = n+1 < n+1+1$. In both cases $P(n+1)$ is true.

The converse is true by lemma 1.6. If $m+1 < n+1$, then either $m < m+1 < n$, or $m < m+1 = n$. ■

Theorem 1.9 (Strong induction)**.** *Let $P(x)$ be a property. Assume that for all $n \in \mathbb{N}$,*

$$[\forall k < n, P(k)] \implies P(n)$$

Then P holds for all $n \in \mathbb{N}$.

Proof. Let $Q(x)$ be the statement $\forall k < x\ [P(k)]$. $Q(0)$ is vacuously true. Suppose $Q(n)$ is true for some n. So $P(k)$ holds for all $k < n$. By our supposition, this means that $P(n)$ is true. Therefore $P(k)$ holds for all $k < n+1$, and $Q(n+1)$ is true. ■

1.3.2 Recursion on $\mathbb{N}$

Theorem 1.10 (Recursion on $\mathbb{N}$)**.** *Let A be a set, $a \in A$ and $g : A \times \mathbb{N} \to A$. Then there exists an unique function $f : \mathbb{N} \to A$ such that*

- $f(0) = a$ *and*
- $\forall n \in \mathbb{N}\ [f(n+1) = g(f(n), n)]$

Proof. A function $t : (m+1) \to A$ is called an m-step computation if $t(0) = a$ and for all k such that $0 \leq k < m$, $t(k+1) = g(t(k), k)$.

Let F be the set of all m-step computations for some $m \in \mathbb{N}$. Let $f = \bigcup F$.

We show that f satisfies all conditions by proving the following series of claims.

1. f is a function. This amounts to showing that for any $t_1, t_2 \in F$ are *compatible*, that is, if $x \in \mathrm{dom}(t_1) \cap \mathrm{dom}(t_2)$, then $t_1(x) = t_2(x)$. Assume $\mathrm{dom}(t_1) = m \in \mathbb{N}$, $\mathrm{dom}(t_2) = n \in \mathbb{N}$, and say $n \leq m$, then $n \subseteq m$ [3]. Therefore it suffices to show that $t_1(k) = t_2(k)$ for all $k < n$.

   Perform induction on k. When $k = 0$, $t_1(k) = t_2(k) = a$. Suppose $k + 1 < n$ and $t_1(k) = t_2(k)$. Then
   $$t_1(k+1) = g(t_1(k), k) = g(t_2(k), k) = t_2(k+1).$$

2. The second claim is that $\mathrm{dom}(f) = \mathbb{N}$ and $\mathrm{ran}(f) \subseteq A$. Firstly, $\mathrm{dom}(f) \subseteq \mathbb{N}$ and $\mathrm{ran}(f) \subseteq A$, since the same is true for all $t \in F$. To show that $\mathrm{dom}(f) \supseteq \mathbb{N}$ it suffices to prove that for every $n \in \mathbb{N}$ there is an n-step computation t.

   Perform induction on n. Clearly the 0-step computation exists. Assume that t is an n-step computation, then define $t^+ : (n+1) + 1 \to A$ by
   $$t^+(k) = \begin{cases} t(k) & \text{if } k \leq n \\ g(t(n), n) & \text{if } k = n+1 \end{cases}$$
   Then t^+ is an $(n+1)$-step computation.

3. f satisfies recursive equations (i) and (ii). First of all $f(0) = a$ since $t(0) = a$ for all $t \in F$. For every $n \in \mathbb{N}$, let t be a $(n+1)$-step computation. Then $f(n) = t(n)$ and $f(n+1) = t(n+1)$. Therefore
   $$f(n+1) = t(n+1) = g(t(n), n) = g(f(n))$$

4. Finally, we show the uniqueness of f. Let f_1 and f_2 both satisfy the theorem. We will perform induction on n to show that $f_1(n) = f_2(n)$ for all $n \in \mathbb{N}$. $f_1(0) = f_2(0) = a$ by definition. Suppose $f_1(n) = f_2(n)$, then
   $$f_1(n+1) = g(f_1(n), n) = g(f_2(n), n) = f_2(n+1).$$

■

Theorem 1.11 (Recursion with parameters)**.** *Let $a : P \to A$ and $g : A \times A \times \mathbb{N} \to A$ be functions. There exists an unique function $f : P \times \mathbb{N} \to A$ such that*

[3] It is easy to see that this is true from the definition of the successor function

- $\forall p \in P\,[f(p,0) = a(p)]$
- $\forall n \in \mathbb{N}, \forall p \in P\,[f(p,n+1) = g(p, f(p,n), n)]$

Proof. The proof is similar as for the case without parameters. ■

Corollary 1.11.1 (Addition on $\mathbb{N}$)**.** *There is an unique binary operation $A : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ such that*

- $\forall m \in \mathbb{N}\,[A(m,0) = m]$.
- $\forall m, n \in \mathbb{N}\,[A(m, S(n)) = S(A(m,n))]$.

We usually write $m+n$ instead of $A(m,n)$. We will also try to convince ourselves that this definition of addition is just like the addition familiar to us.

Lemma 1.12. $\forall n, m \in \mathbb{N}\,[n + S(m) = S(n) + m]$

Proof. Perform induction on m. When $m = 0$,

$$n + S(0) = S(n+0) = S(n) = S(n) + 0$$

Assume the claim is true for some m. Then

$$\begin{aligned} n + S(S(m)) &= S(n + S(m)) \\ &= S(S(n) + m) \\ &= S(n) + S(m) \end{aligned}$$

using the definition of addition. This establishes the claim. ■

Theorem 1.13. *The operation $+$ is commutative.*

Proof. Let $P(x)$ be the proposition

$$\forall m \in \mathbb{N}\ (x + m = m + x).$$

$P(0)$ holds since by definition $0 + m = m$. We just have to show that $m + 0 = m$ as well.

Let $Q(y)$ be the proposition,

$$0 + y = y.$$

$Q(0)$ holds by definition. Assume $Q(m)$. Then

$$\begin{aligned} 0 + S(m) &= S(0 + m) \\ &= S(m) \\ &= S(m) + 0 \end{aligned}$$

Therefore $Q(m)$ holds for all $m \in \mathbb{N}$ by the induction principle.

Now, let us continue. Assume $P(n)$. Then

$$\begin{aligned} m + S(n) &= S(m + n) \\ &= S(n + m) \\ &= n + S(m) \\ &= S(n) + m \end{aligned}$$

using lemma 1.12 for the last step. Thus $P(S(n))$ holds, and so by the inductive principle P holds for all $n \in \mathbb{N}$. ■

Theorem 1.14. *The operation $+$ is associative.*

Proof. Let $P(x)$ be the proposition

$$\forall k, m\, [(k + m) + x = k + (m + x)].$$

$P(0)$ is clearly true. Assume the $P(n)$ holds for some $n \in \mathbb{N}$. Consider $P(S(n))$.

$$\begin{aligned} (k + m) + S(n) &= S((k + m) + n) && \text{(Definition of addition)} \\ &= S(k + (m + n)) && \text{(Inductive hypothesis)} \\ &= k + S(m + n). && \text{(Definition of addition)} \end{aligned}$$

Thus $P(S(n))$ also holds. ■

Theorem 1.15 (Multiplication in $\mathbb{N}$)**.** *There is an unique binary operation $M : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ such that*

- $\forall n \in \mathbb{N}\, [M(n, 0) = 0]$.
- $\forall m, n \in \mathbb{N}\, [M(m, n + 1)] = M(m, n) + m]$.

Again, we will write $m \cdot n$ instead of $M(m, n)$.

Theorem 1.16. *The operation $\cdot$ is commutative.*

Theorem 1.17. *The operation $\cdot$ is associative.*

Theorem 1.18. *The operation $\cdot$ is distributive over addition.*

Proof. Let $P(x)$ be the proposition

$$\forall m, n\, [x \cdot (m + n) = x \cdot m + x \cdot n].$$

$P(0)$ is obviously true. Suppose $P(k)$ holds for some $k \in \mathbb{N}$. Then consider $P(S(k))$.

$$\begin{aligned} S(k) \cdot (m + n) &= k \cdot (m + n) + (m + n) && \text{(Defn. of mult., associativity)} \\ &= k \cdot m + k \cdot n + m + n && \text{(Inductive hypothesis)} \\ &= S(k) \cdot m + S(k) \cdot n && \text{(Defn. of mult., associativity).} \end{aligned}$$

Thus $P(k + 1)$ holds as well. ■

There is no division operation because we run into trouble with the rational numbers. However we can use multiplication to perform something similar to division.

Theorem 1.19 (Euclidean division)**.** *If m and d are non-zero natural numbers then there exist $q, r \in \mathbb{N}$ such that*

$$m = (d \cdot q) + r \qquad r < d.$$

Proof. We perform induction on $m \in \mathbb{N}$. Let P be the statement as above. Then $P(0)$ is true because $0 = 0 \cdot 0 + 0$. Suppose $P(n)$ is true for some $n \in \mathbb{N}$. Consider $P(n + 1)$. We can express $n + 1$ in the required form:

$$n + 1 = \begin{cases} d \cdot q + r + 1, & \text{if } r + 1 < d \\ d \cdot (q + 1), & \text{if } r + 1 = d \end{cases}$$

These are the only possible cases since $r < d$ by the inductive hypothesis, and using theorem 1.8 we get $r + 1 < d + 1$ and lemma 1.6 gives us the two cases.

In the first case both sides are clearly equivalent. In the second case, expanding and substituting d shows that it is equal as well. So $P(n + 1)$ is also true. ■