Chapter 1

Some Logic and Set Theory

This chapter will develop:

- fluency with the logic of propositions (order zero) and predicates (order one),
- familiarity with the axioms of Zermelo-Fraenkel set theory with Choice,
- a working understanding of relations, functions, and numbers.

It sets the foundation for the sequel, which begins with an introduction to general topology.

For §1.1 on propositional logic:

The organizing theme will be to view logical objects as **functions** defined with respect to a two element set of truth values

$$\mathcal{V} = \{\bot, \top\}$$

the truth values themselves are "nullary truth functions," the identity and negation operators "unary truth functions," and the classical logical connectives "binary truth functions."

This departs from the usual presentation, wherein truth values are assumed intuitive enough to be introduced without any context, identity is ignored altogether, and negation is lumped in with the connectives even though it has a fundamentally different character.

Functions are so important to mathematics, though, that this departure is warranted.

We dissect the distinction between a necessary and sufficient condition, a topic often unclear even to native English speakers. We give examples of logical deduction from a historic trio of axioms (plus a law of inference, modus ponens), such as deducing transitivity of implication. We show that under each set of logical connectives, there exists a structure that echoes throughout the rest of the book: the **bounded lattice**.

For §1.2 on predicate logic:

Probably the most important feature of this chapter is the featuring of Cauchy's ε–δ definition of the **limit** of a function, which motivates introducing variables, predicates, and quantifiers. We also cover the principle of mathematical **induction** and the related idea of **well-ordering**.

For §1.3 on Zermelo-Fraenkel with Choice:

Rather than treating the list as a set of commandments to memorize, we focus on what each axiom affords us. The Axiom of Choice is discussed in terms of three major schools of mathematical philosophy, with the Banach-Tarski phenomenon used as an illustrative example.

For §1.4 on Relation and Number:

Having touched on general binary relations in the previous section, we begin here with a focus on the homogeneous case, out of which springs preorderings, partial orderings (antisymmetric preorderings), and total orderings (partial orderings where everything is comparable). This perspective is worthwhile, e.g. since the real numbers $\mathbf{R}$ form a totally ordered lattice, with $\overline{\mathbf{R}}$ forming a bounded totally ordered complete lattice.

We cover functions in terms of their characterization as left-total and right-unique relations. This sets up invertibility as further satisfying right-totality and left-uniqueness, thus expressing set isomorphisms as exactly the bidirectionally total-and-unique relations. The capstone is a proof of the Schröder-Bernstein theorem via Knaster-Tarski, thus establishing the theorem as a corollary of a fixed-point result.

We introduce the integers $\mathbf{Z}$ and the rational numbers $\mathbf{Q}$ first from an algebraic perspective, then from an order-theoretic perspective. As a fun bonus, we cover negabinary expansions as a natural way to sequence the integers, then compose with unique factorization to sequence the rationals. This leads nicely into a presentation of Cantor's diagonal argument.

For an arbitrary ordered field Γ, we cover the ordered ring of Cauchy sequences $\kappa(\Gamma)$. We then provisionally define $\mathbf{R}$ to be $\kappa(\mathbf{Q})/\mathfrak{m}_0$, where $\mathfrak{m}_0$ denotes the maximal ideal of Cauchy sequences that tend towards zero. In the next chapter, we prove the completeness of $\mathbf{R}$ via nested Cauchy sequences, which bootstraps the proof of the existence of arbitrary metric space completions.

1.1 Propositional Logic

We begin with a definition that will be used thoughout.

Definition 1.1. A **function** f consists of a *rule of assignment*

for mapping from a **domain** X to a **codomain** Y, such that:

every element of the domain is paired with a unique element of the codomain.

To denote functions we write $f : X \to Y$ with $x \mapsto f(x)$ denoting the rule of assignment.

Now, since we haven't even defined sets (i.e. domains and codomains) yet, starting with functions might feel like getting ahead of ourselves. The compromise is that the functions we work with in this chapter will be defined on a two element set; this is much smaller and thus easier to understand than, say, the uncountable continuum $\mathbf{R}$, which needs a fair amount of set theory to even be precisely described.

This two element set is

$$\mathcal{V} = \{\bot, \top\}$$

where $\bot$ denotes *false* and $\top$ denotes *true*.

A note on finiteness for the careful:

Observe that $\mathcal{V}$ has two elements, $\mathcal{V}^2$ (all 2-tuples with entries in $\mathcal{V}$) has four elements, and in general $\mathcal{V}^n$ (n-tuples with entries in $\mathcal{V}$) has 2^n elements. (We define n-tuples formally in §1.3, but they may also be thought of informally as n things taken at once in a certain order.) A function going from $\mathcal{V}^n$ to $\mathcal{V}$ must make 2^n binary decisions (i.e. whether to set each function value to either $\bot$ or $\top$), so there are 2^{2^n} possible functions that could go from $\mathcal{V}^n$ to $\mathcal{V}$. Thus, everything involved is finite.

We will soon work with arbitrary (i.e. potentially infinite) sets, though. Here's why we did not just start with infinite sets: intuitively, one thinks of a set as an unordered collection of objects with no repeats. However, this naive conception can lead to logical disaster.

For instance, we have the following argument, known as *Russell's Paradox*:

> Consider the set of all sets that are not elements of themselves – call this set Ω.
>
> On the one hand, if Ω is an element of itself,
>
> > then it is (by definition of Ω) not an element of itself.
>
> On the other hand, if Ω is not an element of itself,
>
> > then it belongs with all the other sets that aren't elements of themselves – namely, in Ω.
>
> So Ω contains itself if and only if Ω does not contain itself – a contradiction.

We would like to exclude sets like Ω from all of mathematics. The working solution is called ZFC, a collection of axioms and axiom schemata that specify how sets ought to behave. An exploration of ZFC right now would be a distraction, though we'll get to it in a moment.

Thus, we restrict our attention to finite sets, where our intuition is trustworthy. The natural numbers are essentially a part of the language layer, as in people know of them regardless of whether they care to learn mathematics. Even nonhuman animals have been observed to count.

Truth Functions

We now begin our study of propositional logic.

Definition 1.2. A **proposition** (or *statement*) is a grammatically correct declarative sentence that can be assigned exactly one value from $\mathcal{V}$. That is, propositions are true xor false.

We often use the letters P, Q, R, S to denote propositions.

Small detail: truth functions don't work with propositions directly, but rather with their truth values.

Nullary Truth Functions

The set $\mathcal{V}^0 = \{()\}$ consists of exactly one 0-tuple (the only possible 0-tuple).

Thus we have two nullary truth functions:

$$\bot : \mathcal{V}^0 \to \mathcal{V} \quad \text{via} \quad () \mapsto \bot, \qquad\qquad \text{and} \qquad\qquad \top : \mathcal{V}^0 \to \mathcal{V} \quad \text{via} \quad () \mapsto \top.$$

That is, *the nullary truth functions are simply the truth values.*

Unary Truth Functions

The set $\mathcal{V}^1$ consists of two 1-tuples, corresponding to the two elements of $\mathcal{V}$.

To each 1-tuple there are two choices of output, spawning a total of four unary truth functions.

(1) $\bot(\mathsf{P}) := \bot$ (2) $\top(\mathsf{P}) := \top$ (3) $+(\mathsf{P}) := \mathsf{P}$ (4) $-(\mathsf{P}) := -\mathsf{P}$

We can describe the unary truth functions in tabular form:

P	$\bot$	$\top$	$+$	$-$
$\bot$	$\bot$	$\top$	$\bot$	$\top$
$\top$	$\bot$	$\top$	$\top$	$\bot$

Here are some comments on these functions.

- The functions (1) and (2) are exactly the nullary truth functions from before.
- The function (3) is called the **identity** map. In general, an identity map outputs its input unchanged.
- The function (4) is called **negation.** It is a consequence of the table above that

$$\mathsf{P} = -(-\mathsf{P}),$$

a fact known as *double negation*.

Binary Truth Functions

The set $\mathcal{V}^2$ consists of four 2-tuples:

$$(\bot, \bot), \quad (\bot, \top), \quad (\top, \bot), \quad (\top, \top).$$

Denoting an element of $\mathcal{V}^2$ by (P, Q), sixteen binary truth functions follow.

(1) $\bot(P, Q) := \bot$

(2) $\top(P, Q) := \top$

(3) $= (P, Q) := (P = Q)$

(4) $\neq (P, Q) := (P \neq Q)$

(5) $+P(P, Q) := +P$

(6) $-P(P, Q) := -P$

(7) $+Q(P, Q) := +Q$

(8) $-Q(P, Q) := -Q$

(9) $\leq (P, Q) := (P \leq Q)$

(10) $\geq (P, Q) := (P \geq Q)$

(11) $< (P, Q) := (P < Q)$

(12) $> (P, Q) := (P > Q)$

(13) $\wedge(P, Q) := (P \wedge Q)$

(14) $\vee(P, Q) := (P \vee Q)$

(15) $\uparrow (P, Q) := (P \uparrow Q)$

(16) $\downarrow (P, Q) := (P \downarrow Q)$

Note that while we use prefix notation to define these truth functions, in practice one uses infix notation.

We give their descriptions all at once:

P	Q	$\bot$	$\top$	$P = Q$	$P \neq Q$	$+P$	$-P$	$+Q$	$-Q$
$\bot$	$\bot$	$\bot$	$\top$	$\top$	$\bot$	$\bot$	$\top$	$\bot$	$\top$
$\bot$	$\top$	$\bot$	$\top$	$\bot$	$\top$	$\bot$	$\top$	$\top$	$\bot$
$\top$	$\bot$	$\bot$	$\top$	$\bot$	$\top$	$\top$	$\bot$	$\bot$	$\top$
$\top$	$\top$	$\bot$	$\top$	$\top$	$\bot$	$\top$	$\bot$	$\top$	$\bot$

P	Q	$P \leq Q$	$P \geq Q$	$P < Q$	$P > Q$	$P \wedge Q$	$P \vee Q$	$P \uparrow Q$	$P \downarrow Q$
$\bot$	$\bot$	$\top$	$\top$	$\bot$	$\bot$	$\bot$	$\bot$	$\top$	$\top$
$\bot$	$\top$	$\top$	$\bot$	$\top$	$\bot$	$\bot$	$\top$	$\top$	$\bot$
$\top$	$\bot$	$\bot$	$\top$	$\bot$	$\top$	$\bot$	$\top$	$\top$	$\bot$
$\top$	$\top$	$\top$	$\top$	$\bot$	$\bot$	$\top$	$\top$	$\bot$	$\bot$

Here are some comments on these functions.

- The functions (1) and (2) are exactly the nullary truth functions from before.
- Functions (3) **biconditional** and (4) **exclusive disjunction** have to do with whether the inputs agree:

$$(P \neq Q) = -(P = Q)$$

- Functions (5 - 8) are constructed by restricting to a single input and then applying a unary truth function.
- Functions (9) **implication**, (10) **reverse implication**, (11) **negated reverse implication**, and (12) **negated implication** are all *asymmetric* and *transitive* binary truth functions.
- Functions (13) **conjunction** aka "and" and (14) **inclusive disjunction** aka "or (possibly both)" satisfy both *de Morgan's laws* and the *distributive laws*:

$$(-P) \vee (-Q) = -(P \wedge Q) \qquad (-P) \wedge (-Q) = -(P \vee Q)$$

$$P \wedge (Q \vee R) = (P \wedge Q) \vee (P \wedge R) \qquad P \vee (Q \wedge R) = (P \vee Q) \wedge (P \vee R)$$

- Implication can be written in terms of negation and inclusive disjunction:

$$(P \leq Q) = (-P \vee Q)$$

- Functions (15) **negated conjunction** aka "nand" and (16) **negated inclusive disjunction** aka "nor" are interesting because each individually can generate all the other binary truth functions:

$$-P = P \uparrow P = P \downarrow P \qquad P \wedge Q = -(P \uparrow Q) \qquad P \vee Q = -(P \downarrow Q)$$

  We will return to this shortly.

Conditional Statements

In the proposition $P \leq Q$, we call P the **antecedent** and Q the **consequent**.

The antecedent suffices for the consequent, whereas the consequent necessitates the antecedent.

Amplifying,

“P is a sufficient condition for Q” holds exactly when “Q is a necessary condition for P.”

Be prepared to recognize the following forms of $P \leq Q$:

- “P implies Q” or “Q is implied by P”
- “if P then Q” or “Q if P”
- “only if Q then P” or “P only if Q”

The conditional $P \leq Q$ is the **converse** to $Q \leq P$ and the **contrapositive** of $-Q \leq -P$.

- The converse of the converse of a conditional is the original conditional.
- The contrapositive of the contrapositive of a conditional is the original conditional.

Logical Deduction

This refers to the process of starting with a set of assumptions and arriving at a conclusion after a finite number of steps. We care about truth functions and tabular proof because it is a quick way to get to the truth value of any proposition; we care about logical deduction because it is a microcosm not dissimilar to how mathematics actually functions.

Three Axioms + One Law of Inference

Formally, a propositional calculus can be thought of as a set of propositions, a set of logical connectives, a set of axioms, and a set of laws of inference. Here is a common starting point:

(1) $\vdash (P \leq (Q \leq P))$

(2) $\vdash ((P \leq (Q \leq R)) \leq ((P \leq Q) \leq (P \leq R)))$

(3) $\vdash ((-P \leq -Q) \leq (Q \leq P))$

(MP) $P, (P \leq Q) \vdash Q$

The first three lines are *axioms* where P, Q, etc. can be any propositions.

The remaining line is **modus ponens**, a *law of inference*. Laws of inference are distinct from conditionals due to operating one level above where conditionals are defined: see Lewis Carroll’s “What the Tortoise Said to Achilles” for an illustration of this distinction.

The $\vdash$ symbol is called a *turnstyle*, and any expression involving it should be read as

> "given what is left of the turnstyle, we have what is right of the turnstyle."

Laws of inference thus employ the turnstyle as a kind of meta-conditional; this is essential for describing logical consequences which themselves include conditional statements.

Using the Axioms and Modus Ponens

The first move is to turn those axioms into inferences via (MP):

(1–i) $$P \vdash (Q \leq P)$$

(2–i) $$(P \leq (Q \leq R)) \vdash ((P \leq Q) \leq (P \leq R))$$

(3–i) $$(-P \leq -Q) \vdash (Q \leq P)$$

Another application of (MP) to (3–i) transforms it into **modus tollens**:

(MT) $$(-P \leq -Q), Q \vdash P.$$

Proposition 1.3.

$$(P \leq Q), (P \leq (Q \leq R)) \vdash (P \leq R).$$

Proof. We start by assuming both $P \leq Q$ and $P \leq (Q \leq R)$. We then may infer that $(P \leq Q) \leq (P \leq R)$ via (2–i). Finally we get $P \leq R$ via modus ponens. ■

Proposition 1.4.

$$(P \leq Q), (Q \leq R) \vdash (P \leq R).$$

Proof. Start by assuming both $P \leq Q$ and $Q \leq R$. Then by (1–i), we get $P \leq (Q \leq R)$. Finally we get $P \leq R$ via Proposition 1.3. ■

Meredith's Sole Axiom

One of the surprising things about propositional calculus is it only needs one axiom to get off the ground:

(!) $$(((((A \leq B) \leq (-C \leq -D)) \leq C) \leq E) \leq ((E \leq A) \leq (D \leq A)))$$

This is known as *Meredith's sole axiom,* and from it one can derive the axioms given above.

Methods of Proof

There are roughly three ways one can approach a proof. We will use as our working example the fact that every integer is either even or odd. We can structure this in conditional form as follows:

n is an integer $\leq$ n is either even or odd.

The simplest (but often not easiest) way to prove Q given P is to show $P \leq Q$ and then use modus ponens:

> *If n is an integer, we may apply division with remainder, which states that for integers a and b there exists a unique integer quotient q and remainder r such that $a = bq + r$ and $0 \leq r < b$. The only two possiblilities for r in this case are 0 and 1; hence, n is either even or odd.*

This is called **direct proof**.

Another way to prove Q given P is to show $-Q \leq -P$ and then apply modus tollens:

> *Suppose n is neither even nor odd. Then $n + 1$ is neither even nor odd, so $n(n + 1)$ is not necessarily divisible by 2. Since for every integer m we must have $2 \mid m(m + 1)$, n is not necessarily an integer, i.e. it is not the case that n must be an integer.*

This is known as **proof by contrapositive**.

A third way to prove Q given P is to show that if one had P but also $-Q$, then disaster ensues:

> *Suppose n is neither even nor odd. Then neither n nor $n+1$ is divisible by 2, so $n(n+1)$ is not divisible by 2. This contradicts the fact that every product of consecutive integers is divisible by 2.*
>
> *If nm is a product of consecutive integers, then nm is divisible by 2?*
>
> *Indeed: suppose nm is a product of consecutive integers not divisible by 2. Then nm is odd, and this entails n and m are odd. But then odd numbers are at least two apart, contradicting our assumption that n and m were consecutive.*

This is called **proof by contradiction**.

We will encounter all three methods of proof throughout our study of math.

Use of the $\leq$ Symbol

Readers with previous logic experience may be wondering whatever happened to the → or $\implies$ symbol.

We now take a moment to explain.

Definition 1.5. A **bounded lattice** is a 6-tuple $(X, \wedge, \vee, \top, \bot, \leq)$

where X is a set, $\vee$ (the *join* or **supremum**) and $\wedge$ (the *meet* or **infimum**)

are functions from X^2 to X, and $\top$ and $\bot$ are special elements of X.

These objects satisfy the following conditions:

$$x \wedge (y \wedge z) = (x \wedge y) \wedge z \qquad \text{and} \qquad x \vee (y \vee z) = (x \vee y) \vee z$$

$$x \wedge y = y \wedge x \qquad \text{and} \qquad x \vee y = y \vee x$$

$$x \wedge (x \vee y) = x \qquad \text{and} \qquad x \vee (x \wedge y) = x.$$

$$x \wedge \top = x \qquad \text{and} \qquad x \vee \bot = x.$$

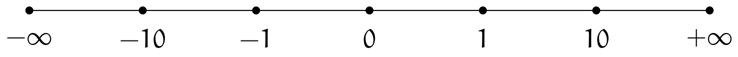
We may further define $x \leq y$ to mean $x = x \wedge y$ (or equivalently $y = x \vee y$).

A bounded lattice with a notion of complement $-x$ is called a **boolean lattice**.

The logic we are introducing can be understood in terms of boolean lattices.

This is why we use $X \leq Y$ instead of the more traditional $X \to Y$.

We note that $\mathcal{V}$ as a lattice admits a satisfying visualization. Consider the number line **R**, and add two endpoints, $-\infty$ placed to the left of all negative reals and $+\infty$ placed to the right of all positive reals:



the result is $\overline{\mathbf{R}}$, the extended real number line. We may thus think of

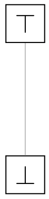
- $+\infty$ as a $\top$ and $-\infty$ as a $\bot$,
- $\wedge$ as taking the minimum of two extended reals, $\vee$ as taking the maximum,
- and negation as rotating the number line 180° about 0.

So $\overline{\mathbf{R}}$ is *also* a bounded lattice, with $\mathcal{V}$ a boolean sublattice. Note, however, that "−" does not logically complement elements in **R**. We will say more about lattices in §1.4.

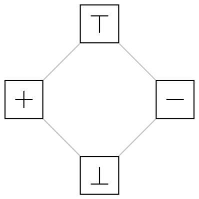
Lattices of Truth Functions

The n-ary truth functions also form a boolean lattice.

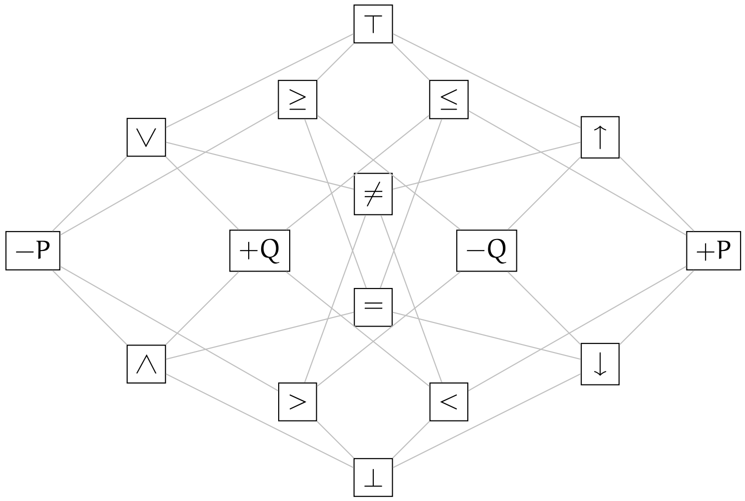
Here is $n = 0$:



Here is $n = 1$:



Here is $n = 2$:



Functional Completeness

We end this section on propositional logic with an application to electronics.

Definition 1.6. A set of binary truth functions is **functionally complete** if it generates all of the remaining binary truth functions.

For example, $\{-, \wedge, \vee\}$ is functionally complete, and by de Morgan's laws, so is $\{-, \wedge\}$ and $\{-, \vee\}$.

Now, note that

$$X \uparrow X = -X, \qquad (X \uparrow Y) \uparrow (X \uparrow Y) = X \wedge Y, \qquad (X \uparrow X) \uparrow (Y \uparrow Y) = X \vee Y$$

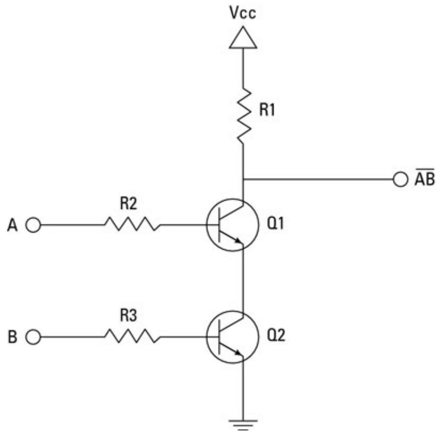
so $\{\uparrow\}$ is functionally complete!

Similarly, $\{\downarrow\}$ is functionally complete.

Classical Circuits: nand and nor Gates

Electrical engineers leverage this functional completeness: instead of having to stock and gates and or gates and not gates, they can work with nand gates exclusively.

Here is the circuit for a nand gate:



This gate is central to electronics design.

1.2 Predicate Logic

In this section we develop a formal language of quantified statements with variables which will serve as the basis for defining set theory, itself the foundation for building the remainder of our mathematical apparatus.

We will work within a *universal collection* of logical statements, denoted Ω. The question of how many elements Ω contains is tricky to answer, since our notions of quantity beyond the finite will come from defining sets rigorously. At the very least, we may distinguish between two variants of infinite, using the criterion of whether the elements of a given collection can be arranged in a sequence.

Definition 1.7. If X can be sequenced, we will say X is **countable** and denote X using the notation

$$X = (x_i)_{i \geq 0} = (x_0,\ x_1,\ x_2,\ \ldots).$$

All other infinite Y will be designated **uncountable** and denoted as follows:

$$Y = \{y_\iota\}_{\iota \in J}$$

Here, J is an *index set*, which can be thought of as some uncountable reference collection.

The usual countably infinite sets are:

$\mathbf{N}$ the *convex cone* of natural numbers (nonnegative integers),

$\mathbf{N}^{++}$ the *convex cone* of positive elements of $\mathbf{N}$,

$\mathcal{P}$ the prime elements of $\mathbf{N}^{++}$,

$\mathbf{Z}$ the *ordered ring* of rational integers,

and $\mathbf{Q}$ the *ordered field* of rational numbers (fractions).

The first three sets are clearly sequences, and we will show how to sequence the last two sets later on.

Some common *uncountably infinite* sets include:

$\mathbf{R}$ the *ordered field* of real numbers,

$(-\varepsilon, \varepsilon)$ the *open interval* of real numbers between $-\varepsilon$ and ε,

$[-\varepsilon, \varepsilon]$ the *closed interval* of real numbers between $-\varepsilon$ and ε,

$\mathbf{C}$ the *field* of complex numbers,

$\mathbf{D}$ the *open disk* of complex numbers of modulus less than 1,

and $\partial\mathbf{D}$ the *circle* of complex numbers of modulus 1.

For now, we will think of Ω as countable. As we shall see, countability is a fairly tame flavor of infinity.

Propositions and Infinity

The assumption that Ω is countable is admittedly a mathematical idealization, since once we define real numbers, we will be able to generate an uncountable number of propositions, such as

The real number x is rational.

Honestly though, mathematics in practice often requires a certain calm flexibility with notions of the infinite. For example, we will show that the set of all subsets $\mathcal{P}(X)$ of a set X is always strictly larger than X, which, if X is countable, implies a sequence of sets each infinitely larger than the ones before it,

$$X, \qquad \mathcal{P}(X), \qquad \mathcal{P}(\mathcal{P}(X)), \qquad \mathcal{P}(\mathcal{P}(\mathcal{P}(X))), \qquad \ldots$$

with nothing but philosophical discomfort to keep us from continuing the sequence. What's more, beyond-uncountable sets (such as $\mathbf{R}^{\mathbf{R}}$, the set of all functions from $\mathbf{R}$ to $\mathbf{R}$) routinely appear in the setting of functional analysis, which we visit in Chapter 4.

Context and the Symbol Grounding Problem

There's another philosophical abyss lurking in the background of propositional logic, this one far more insidious than the problem of producing really big numbers. Every proposition must be interpreted within a certain context: for example, the classic proposition

It is raining.

depends heavily on location, time, the number of raindrops coming down (see: sorites paradox), et cetera. But so far, **no one has yet figured out how to define context without falling back on propositions.** So the entire mathematical edifice is, to some degree, floating in midair. Depending on your personal philosophical stance, this is either irrelevant, devastating, or hilarious.

Syntax of Predicate Logic

If you're still reading, welcome to the club. We now specify what we mean by predicate logic.

Terms and Predicates

The words in our formal language are known as *terms*.

Definition 1.8. By a **term** we mean either:

- a **constant** c, i.e. some fixed object in Ω, or
- a **variable** x, i.e. a symbol representing any object in Ω, or
- a **function** $f : \Omega^n \to \Omega$.

We note that constants are nullary functions, and that the argument of a function could be another function. So we really should write $\Omega = \Omega_0$ for the universe of propositions, and then Ω_n for all the n-ary functions from Ω_0 to Ω_0. Keep in mind that variables can only refer to objects in Ω_0 in our formal language.

By an n-ary **predicate** we mean a function $P : \Omega^n \to \mathcal{V}$. That is, an n-ary predicate takes in n propositions and outputs a truth value. Predicates are the building blocks of formulas, which we discuss next.

Formulas and Quantifiers

The sentences in our formal language are known as *formulas*.

Definition 1.9. By a **formula** we mean either:

- an n-ary predicate applied to an n-tuple of terms, or
- a negated formula, or two formulas joined by a binary truth function, or
- $\forall x P(x)$ or $\exists x P(x)$ where x is a variable and $P(x)$ is a formula.

Here are some examples of formulas:

$$c_1, \quad \top, \quad P(c_1, c_2) \wedge Q, \quad \forall x R(x), \quad \exists x S(x) \vee P(c_3, c_4)$$

The symbols $\forall$ "for all" and $\exists$ "there exists" are called *quantifiers*.

Definition 1.10. By a **quantifer** we mean one of two symbols:

- the *universal quantifier* $\forall$, where the statement $\forall x P(x)$ may be read "$P(x)$ for all x,"
- the *existential quantifier* $\exists$, where the statement $\exists x P(x)$ may be read "there exists an x where $P(x)$."

Many logical malapropisms can be attributed to accidentally swapping quantifiers.

Negating Quantifiers

We record that

$$-(\forall x P(x)) = \exists x(-P)(x) \qquad \text{and} \qquad -(\exists y Q(y)) = \forall y(-Q)(y)$$

The above is basically de Morgan's laws: think of $\forall$ as behaving like $\wedge$, and $\exists$ as behaving like $\vee$.

This captures our intuition about what "for all" and "exists" mean in natural language: if something isn't true for everything, it is false for at least one thing; similarly, if there is nothing for which something is true, then for all things that something just be false.

It's merely terse notation, but these ideas appear so frequently that the brevity is warranted.

Quantifier Case Study: Cauchy's ε–δ Limit Definition

Here is an example where the quantifiers really matter. It looks forward a bit, but the example is so important that we preview it here. It is Cauchy's ε–δ definition of a limit: remember it and it will serve you well.

Definition 1.11. Let $f : \mathbf{R} \to \mathbf{R}$.

We say that the **limit** of f as x approaches c is L if

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \text{such that} \quad (|x-c| < \delta) \;\leq\; (|f(x) - L| < \varepsilon)$$

and write $\lim\limits_{x \to c} f(x) = L$.

This can be formalized in terms of games (yes, there is a mathematical theory of games): if Player A gives Player B an ε_0, Player B can then respond with a δ_0, and if Player A then sets $\varepsilon_1 = \delta_0$ and then proposes ε_1 to Player B, then Player B produces a corresponding δ_1 for Player A, and so on.

The Non-Logical ∈ Symbol

We need one more symbol, this time having to do with sets.

Definition 1.12. The symbol $\in$ denotes elementhood, i.e. $x \in X$ denotes the proposition

The set x is an element of X.

Elements of sets are always other sets.

The $\in$ symbol was orginally an ϵ (lower case Greek epsilon), but eventually became its own glyph.

Thus in our mathematical ontology we shall have two *types*: propositions and sets.

There is an entire theory of types, but getting into it now would be a distraction.

Sets: Key Structures

In order to accomodate sets, we won't be able to use a set as a universe. We need a *category*.

Definition 1.13. A **category** $\mathcal{C}$ consists of:

- a class of *objects* $\mathcal{C}_0$,
- for any two objects x, y a set of *arrows* $\mathcal{C}_1(x, y)$,
- for any three objects a, b, c an associative binary operation

  $$\mathcal{C}_1(a, b) \times \mathcal{C}_1(b, c) \to \mathcal{C}_1(a, c)$$

  called *composition of arrows*,
- for any object x an arrow $\mathrm{id}_x \in \mathcal{C}_1(x, x)$.

Category theory is the more modern way to think about mathematical foundations, but it largely serves the subject of algebra (in particular: algebraic topology, algebraic geometry, and algebraic number theory). Analysis can be done without category theory, so we do this in the interest of keeping things as simple as is reasonably achievable.

Still, when the temptation is irresistable, a category-flavored insight may appear down the line.

The Category Set of Sets

Just so that we have a definition on the record, we will say

$$\mathsf{Set} = \{x : x = x\}$$

defines the category of sets. This object, when taken alongside the universe of propostions Ω, may be called the *classical universe of mathematical discourse.* Classical here means no categories beyond Set.

This definition is a reference to the three classical laws of thought:

- Excluded middle: *every proposition is either true or false*.
- Noncontradiction: *no proposition is both true and false*.
- Identity: *every object is equal to itself*.

Note that excluded middle and noncontradiction are encoded in our definition of proposition.

The law of identity as used above is formal instead of informal, since we define a rigorous notion of what it means for two sets to be equal in the next section. It doesn't formally extend to objects that aren't sets, since we have no specification of the $=$ symbol for arbitrary objects.

On a related note, the $=$ symbol being used to denote the biconditional earlier doesn't formally count as a notion of "proposition equality" since it merely forms a new proposition from two given propositions. Instead we express that two propositions P and Q are logically equivalent by first establishing $P \vdash Q$ and then $Q \vdash P$.

The Void

If we can characterize anything with the assumption $x = x$,

we ought to be able to characterize nothing at all with the assumption $x \neq x$.

Definition 1.14. We call

$$\varnothing = \{x : x \neq x\}$$

the **empty set**.

The empty set is sometimes an element of another set, but no set is an element of the empty set:

$$\vdash (-(x \in \varnothing)).$$

There is exactly one function on $\varnothing$, the *empty function* which sends nothing nowhere.

The Natural Numbers and Successor Map

The *binary union*

$$x \cup y$$

of two sets x and y consists of all z such that either $z \in x$ or $z \in y$. It is a specific case of the arbitrary union, which we assert to exist in §1.3.

Definition 1.15 (von Neumann). We define the **successor** of any set X to be

$$X^{++} = X \cup \{X\}.$$

In particular, the set consisting of $\varnothing$ and all sets obtained via applying the successor to $\varnothing$ a finite number of times is called the **natural numbers.** Concretely, the first few natural numbers are denoted

$$\begin{aligned}
&0 = \varnothing \\
&1 = \varnothing^{++} = \{0\} \\
&2 = (\varnothing^{++})^{++} = \{0, 1\}, \\
&3 = ((\varnothing^{++})^{++})^{++} = \{0, 1, 2\}, \\
&4 = (((\varnothing^{++})^{++})^{++})^{++} = \{0, 1, 2, 3\} \\
&\quad\vdots
\end{aligned}$$

When restricted to the natural numbers, the successor map becomes a unary operation $(\cdot)^{++} : \mathbf{N} \to \mathbf{N}$.

A **sequence** taking values in a set X is a function from $\mathbf{N}$ to X.

Induction and Well-Ordering

As a consequence of asserting the existence of $\mathbf{N}$, we have the following schema (one axiom per proposition):

Axiom 1.16 (Induction schema). *If* $P(n)$ *satisfies*

$$P(0) \qquad \textit{and} \qquad \forall n \in \mathbf{N}(P(n) \leq P(n^{++})),$$

then $\forall n \in \mathbf{N}, \;\; P(n)$.

In plain English, this is saying that if one has a property that holds at zero, and one can show that the property holding at n implies the property holding at n^{++}, then the property holds for every natural number. We may restructure the axiom so that it is about sets instead of propositions, which is what we really want:

Axiom 1.17 (Set Induction). *If a subset* $X \subseteq \mathbf{N}$ *satisfies* $0 \in X$ *and*

$$(n \in X) \;\leq\; (n^{++} \in X)$$

for all $n \in \mathbf{N}$*, then* $X = \mathbf{N}$.

These are logically equivalent to the following:

Axiom 1.18 (Well-Ordering Principle). *Every nonempty subset* $X \subset \mathbf{N}$ *has a least element.*

Proof. Let X be a subset of $\mathbf{N}$ satisfying $0 \in X$ and

$$(n \in X) \;\leq\; (n^{++} \in X).$$

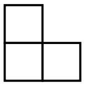
We proceed by contradiction: suppose $X \neq \mathbf{N}$. Then there is a nonnegative integer not in X, i.e. $\mathbf{N} \setminus X$ is nonempty. Then $\mathbf{N} \setminus X$ has a least element n. Note that $n \neq 0$, since $0 \in X$. Thus, $n > 0$, and since n is the least element not in X, $n - 1$ must be in X. But by assumption, $(n-1)^{++} = n \in X$, contradicting our assumption that $n \notin X$. Thus proves that the well-ordering principle implies the principle of induction.

Conversely, consider a nonempty subset $Y \subset \mathbf{N}$. If Y has just one element, then that element is the least element of Y. Now suppose the well ordering principle is true for all subsets of $\mathbf{N}$ with n elements, and suppose Y has n^{++} elements. Take $y \in Y$ and let z be the least element of $Y \setminus y$. Then $\min(\{y, z\})$ is the least element of Y. This proves that the principle of induction implies the well-ordering princple. ■

Confusingly, there is also a *Well-Ordering Theorem*, which is a statement equivalent to the Axiom of Choice (i.e. one may deduce the Well-Ordering Theorem from the Axiom of Choice and also deduce the Axiom of Choice from the Well-Ordering Theorem). The Well-Ordering Theorem states that *any* set admits a well-ordering akin to the well-ordering given by $\in$ on $\mathbf{N}$.

In particular, $\mathbf{R}$ can in theory be well-ordered, but the resulting well-order is so pathological that asserting its existence or non-existence very quickly leads to statements independent of the set axioms!

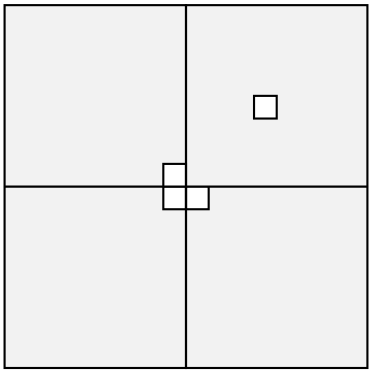
A *tromino* is a 2×2 square of unit (i.e. 1×1) squares, with a corner unit square missing:



Proposition 1.19. *Every* $2^n \times 2^n$ *grid of squares missing one unit square can be tiled with trominos.*

Proof. The claim holds for $n = 1$. So assume the claim holds for $n = m - 1$.

Divide a $2^m \times 2^m$ square with a unit square missing into four $2^{m-1} \times 2^{m-1}$ quadrants. Place a tromino in the center so that the missing square of the tromino aligns with the quadrant with the deleted square.



By our inductive hypothesis, each of these quadrants can be tiled with trominos, so the claim holds for $n = m$. By induction, the claim then holds for all natural numbers. ■

Induction Example: A Fibonacci Identity

Though the problem itself is fairly well known, the author initially encountered the solution in Loren C. Larson's book *Problem Solving Through Problems.* There, one can find all the induction practice one will ever need.

Let F_n denote the nth Fibonacci number, defined so that $F_0 = F_1 = 1$ and $F_{n+2} = F_{n+1} + F_n$.

Proposition 1.20. $F_{n+1}^2 + F_n^2 = F_{2n+1}$.

Remark. A hint that this is going to be trickier than expected is that $F_{2n+3} = F_{2n+2} + F_{2n+1}$, and there is no logical entry point for F_{2n+2} based on how the problem is currently stated. The technique is thus to simply encode F_{2n+2} into a separate identity, prove that, and then prove the original claim. Watch what happens, though.

Proof. Let's play around with the identities before proceeding formally.

Following the above, we need to show, given

$$F_m^2 + F_{m+1}^2 = F_{2m+1}, \tag{1.1}$$

we need to show

$$F_{m+1}^2 + F_{m+2}^2 = F_{2m+3} = F_{2m+1} + F_{2m+2}. \tag{1.2}$$

Subtract (1.2) from (1.1) to obtain:

$$F_{2m+2} = F_{m+2}^2 - F_m^2 = F_{m+1}^2 + 2F_mF_{m+1}. \tag{1.3}$$

If we can show (1.1) implies (1.3) via induction, then we would also get that (1.1) implies (1.2) via induction, which is our claim. But we derived (1.3) from (1.2), and a bit of experimentation confirms that the two indentities are relentlessly intertwined.

The solution is to stipulate that

$$\begin{array}{lll} P(n) & \text{represents the claim} & F_n^2 + F_{n+1}^2 = F_{2n+1} \\ Q(n) & \text{represents the claim} & 2F_nF_{n+1} + F_{n+1}^2 = F_{2n+2} \end{array}$$

We already know that $(P(m) \land Q(m)) \leq P(m+1)$. Now note that

$$\begin{aligned} 2F_{m+1}F_{m+2} + F_{m+2}^2 &= 2F_{m+1}(F_m + F_{m+1}) + F_{m+2}^2 \\ &= 2F_{m+1}^2 + 2F_mF_{m+1} + F_{m+2}^2 \\ &= 2(F_{2m+1} - F_m^2) + 2F_mF_{m+1} + F_{m+2}^2 \\ &= 2F_{2m+1} + F_{m+1}^2 - F_m^2 \\ &= F_{2m+1} + F_{m+1}^2 + F_{2m+2} - 2F_mF_{m+1} \\ &= F_{2m+3} + F_{m+1}^2 - 2F_mF_{m+1} \\ &= F_{2m+3} + F_{2m+2} = F_{2m+4}. \end{aligned}$$

This shows that $P(m) \land Q(m) \leq Q(m+1)$. The base cases aren't hard to verify.

Thus the induction we needed all along was neither $P(n)$ nor $Q(n)$ but $P(n) \land Q(n)$, which when proven via induction (as we essentially just did) gives both $P(n)$ and $Q(n)$ as corollaries. ■

Most instances of induction aren't this intricate. But some are.

1.3 The ZFC Axioms

It is now time to define sets properly. Having thoroughly prepared the reader, we deliver the axioms in pure symbols. The content of this chapter draws heavily from a similar treatment in *Set Theory* by Thomas Jech.

The theory ZF has six axioms and two axiom schemata:

- Extent
- Pairing
- Power set
- Union
- Induction
- Separation Schema
- Replacement Schema
- Foundation

In addition, we have the Axiom of Choice (C). Collectively, the axioms assert which sets can and cannot exist.

The axioms are the result of hardening the intuitive "collection of elements with no repeats" proto-definition into a robust specification which most mathematicians at least implicitly rely on.

Equipped with the appropriate definitions, one can prove from the axioms virtually any mathematical result. Of course, this doesn't usually happen; in practice one relies on an intuitive understanding of sets gained through careful study of their properties. Contrary to what some may think: to bake an apple pie, one does *not* first need to invent the universe. References exist. Other mathematicians exist. No one works in a vacuum.

On the other hand, there is something oddly satisfying about seeing the exact criteria that specify what a set is. It is kind of like reading through the ingredients on an energy drink, or knowing that one could sit down and scan through the entire terms and conditions of an end user license agreement. Knowing that something has been scrutinized to down to atoms can be a fascinating reward of its own. At the very least, a great many mathematicians seem to think so.

Sets Are Determined by Their Elements

Two sets are equal exactly when they have the same elements:

Axiom 1.21 (Extent).

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

As we shall see, a *set-isomorphism* (or **bijection**) is essentially a bidirectional function that tells us how to obtain any element of the first set from the second set and vice versa. Whenever we assert the equivalence of two sets via bijection, we are implicitly using the Axiom of Extent.

Existence of Unordered Pairs

For any pair of elements there exists a set containing exactly those elements:

Axiom 1.22 (Pairing).

$$\forall a \forall b \exists x \forall c((x \in c) \leq ((x = a) \vee (x = b))).$$

The **unordered pair** defined above is unique by extensionality. We write $\{a, b\}$ for any sets a and b. In the case $a = b$, we may write $\{a\}$, the singleton containing just a. Note that $a \neq \{a\}$. Note also that $\{a, b\} = \{b, a\}$.

Existence of Power Sets

We denote **subsets** by writing $y \subseteq x$ to mean

$$\forall z((z \in y) \leq (z \in x)),$$

which reads "anything in y must also be in x."

There exists a set of all subsets of any set x, and by asserting the existence of it, we get all the subsets at once.

Axiom 1.23 (Power set).

$$\forall x \exists p \forall y((y \in p) = (y \subseteq x)).$$

We call this set the **power set** of x and denote it $\mathcal{P}(x)$.

Existence of Unions

There exists for every set x a set $y = \bigcup x$ for which $z \in y$ exactly when z is an element of some element of x.

The set y is called the **union** of x.

Axiom 1.24 (Union).

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

In the case of $x = \{x_1, \ldots, x_n\}$ we write

$$\bigcup x = \bigcup_{i=1}^{n} x_i = x_1 \cup \cdots \cup x_n.$$

For the dual concept (intersection of a set), see the separation schema.

Existence of Inductive Sets

There exists a set containing the empty set that is stable under the successor function:

Axiom 1.25 (Induction).

$$\exists s(\varnothing \in s \wedge \forall x((x \in s) \leq (x^{++} \in s))).$$

We call such a set an **inductive set**. The smallest inductive set (i.e. the intersection of all the inductive sets) is **N**.

In particular, this axiom is the only thing that guarantees that any sets exist at all.

That is, at least one set exists because **N** exists, and every set in the universe of sets is thus essentially either a subset of **N** or a subset of one of its power sets. Look up "von Neumann hierarchy" for more precise details on this.

The set **N** is also sometimes denoted ω, for "first infinite ordinal." Ordinals and their arithmetic are just barely beyond the scope of this book; the interested reader is enouraged to explore further on their own.

Definiable Subclasses of Sets are Sets

The intersection of a set with a class forms another set:

Axiom 1.26 (Separation schema). *For* $\phi(z, p)$ *a formula, we have*

$$\forall x \forall p \exists y \forall z((z \in y) = (z \in x \wedge \phi(z, p))).$$

From separation, we get the existence of the empty set, and also the existence of intersections and complements.

$$\exists \varnothing, \qquad \bigcap x = \{u : \forall y(u \in y \in x)\}, \qquad x \setminus y = \{z : (z \in x) \wedge (z \notin y)\}$$

We also get the existence of subsets, but in a controlled way: one can only create a subset from an existing set.

This is different from the *comprehension schema* which falsely states that one can take subsets of any class.

Set-Images are Sets

The image of function whose domain is a set must be a set:

Axiom 1.27 (Replacement schema). *For* $\phi(x, y, p)$ *a formula, we have*

$$\forall x \forall y \forall z(\phi(x, y, p) \wedge \phi(x, z, p) \leq (y = z)) \leq \forall z \exists w \forall y(y \in w = (\exists x \in z)\phi(x, y, p)).$$

All Sets are Well Founded

There are no infinite membership chains and no cyclic membership loops:

every chain of elementhood has an $\in$-minimal element.

Axiom 1.28 (Foundation).

$$\forall s((s \neq \varnothing) \leq (\exists x \in s)(s \cap x \neq \varnothing))$$

In particular, $x \in x$ never happens.

Construction: Cartesian Products

Sometimes there is some confusion as to how functions, ordered pairs, and Cartesian products are defined, since they're usually defined in terms of each other. That is, arbitrary Cartesian products are sets of functions, but functions are defined to be collections of ordered pairs and hence subsets of a binary Cartesian product. Here we offer a possible way to clarify things: the key is to distinguish between "Kuratowski ordered pair" and "2–tuple."

Binary Proto-Products

First, we define a binary proto-product to be a collection of Kuratowski ordered pairs, whose first element is drawn from the first set of the proto-product and whose second element is drawn from the second set.

Definition 1.29. Let X and Y be sets. Define the **binary proto-product**

$$\langle\langle X,\ Y\rangle\rangle := \big\{\, \{\{x\},\{x,y\}\} : x \in X,\ y \in Y \,\big\}$$

This exists as a subset of $\mathcal{P}(\mathcal{P}(X \cup Y))$.

The prefix "proto-" suggests these objects are essentially temporary definitions: proto-products will be replaced by Cartesian products.

Spaces of Proto-Maps

We then use the proto-product to define proto-maps, which are kind of like functions, except built from proto-products instead of Cartesian products.

Definition 1.30. Let X and Y be sets. Define the **space of proto-maps**

$$\langle\langle X;\ Y\rangle\rangle := \big\{ f \subset \langle\langle X,\ Y\rangle\rangle : \forall x \in X\ \exists! y \in Y \quad \text{such that} \quad \{\{x\},\{x,y\}\} \in f \big\}.$$

This exists as a subset of $\mathcal{P}(\langle\langle X,\ Y\rangle\rangle)$.

We define proto-maps so that Cartesian products will have a function-like object that may be used to specify elements of the product.

Cartesian products are then collections of proto-maps from some index set.

Definition 1.31. Let $\{X_\iota\}_{\iota \in J}$ be a collection of indexed sets with index set J,

which exists as an element of $\langle\langle J; \mathcal{P}\left(\bigcup_\iota X_\iota\right)\rangle\rangle$,

i.e. the collection is a proto-map whose Kuratowski ordered pairs look like $\{\{\iota\}, \{\iota, X_\iota\}\}$.

Define the **arbitrary Cartesian product** to be a subset of the proto-map space

$$\prod_\iota X_\iota \subset \left\langle\left\langle J; \bigcup_\iota X_\iota \right\rangle\right\rangle$$

such that the elements are proto-maps $\phi : J \to \bigcup_\iota X_\iota$

whose Kuratowski ordered pairs look like $\{\{\iota\}, \{\iota, x_\iota\}\}$ where $x_\iota \in X_\iota$.

In particular, $n = \{0, \ldots, n-1\} \in \mathbf{N}$ may serve as an index set.

Definition 1.32. The **finite Cartesian product** is then

$$X_1 \times \cdots \times X_n = \prod_{i=1}^{n} X_i \subset \left\langle\left\langle n; \bigcup_{i=1}^{n} X_i \right\rangle\right\rangle.$$

Elements of a finite Cartesian product with n factors are called **n–tuples.**

For example,

$$X_1 \times X_2 \times X_3 = \prod_{i=1}^{3} X_i$$

has 3-tuples as elements, where each 3-tuple (x_1, x_2, x_3) is a function

$$\phi : 3 \to \bigcup_{i=1}^{3} X_i$$

such that $\phi(0) = x_1 \in X_1,\ \phi(1) = x_2 \in X_2,\ \phi(2) = x_3 \in X_3.$

Relations are Subsets of Cartesian Products

Definition 1.33. Define an **n-ary relation** to be any subset of $\prod_{i=1}^{n} X_i$.

The three major kinds of binary relations are *orderings, equivalence relations,* and **functions.**

The Axiom of Choice

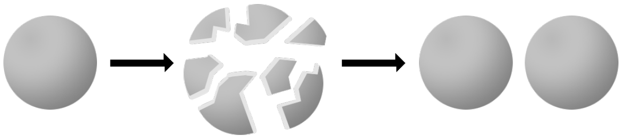
Given any collection of sets, each containing at least one element,
it is possible to construct a new set by choosing one element from each set.

Axiom 1.34 (Choice). *Every family of nonempty sets has a choice function.*

There are several equivalent formulations, one of which is:

Axiom 1.35 (Weak Tychonoff). *An arbitrary Cartesian product of nonempty sets is itself nonempty.*

The Banach-Tarski Phenomenon



Choice can lead to bizarre consequences, one of which is the Banach-Tarski phenomenon (pictured above):

> *Given a solid ball in 3-space, there is a decomposition of that ball into finitely many disjoint subsets such that those subsets can then be reassembled to form two identical copies of the original ball.*

The construction involves non-measurable sets, which we will cover in Chapter 10.

A nontrivial number of mathematicians do not accept the axiom of choice as valid. Then again, the same could be said of excluded middle. There is a philosphical cost to doing mathematics, in that one must include in one's assumptions hypotheses which lead to wild conclusions. What can we make, then, of a formal system that deduces results such as the Banach-Tarski phenomenon?

There are several approaches. Here are three popular ones:

Platonism Mathematics exists in a realm independent from physical reality, i.e. independent of the human minds which record it. Banach-Tarski is then a real fact which happens to disagree with our preconceived physical intuition.

Formalism Mathematics is but a formal game played with symbols, and Banach-Tarski an amusing yet valid result, provided one assumes choice.

Intuitionism Mathematics ought to be formed from constructed objects instead of abstract characterizations. Intuitionists reject choice, and so Banach-Tarski is moot to them.

Don't worry if you do not feel committed to any particular stance yet.
Sometimes the best answer one can give to a question is an honest "I don't know."

1.4 Relation and Number

In this section we cover the major kinds of binary relations and the standard number systems.

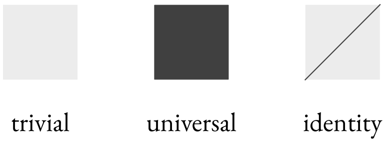
Homogeneous Binary Relations

Definition 1.36. A relation $R \subseteq X \times Y$ is *homogeneous* if $X = Y$.

Three examples of homogeneous relations $x_1 R x_2$ are:

- The *trivial relation* $R = \varnothing$, which holds for no $(x_1, x_2) \in X^2$;
- The *universal relation* $R = X^2$, which holds for any $(x_1, x_2) \in X^2$;
- The *identity relation* $R = \mathrm{id}$, which holds when $x_1 = x_2$.

The following illustration may be helpful.



Preorders are the most general type of ordered set.

A partial order are special kind of preorder; a total order is a special kind of partial order.

Preordered Sets

Definition 1.37. A homogeneous relation R is:

- *reflexive* when $\mathrm{id} \subseteq R$
- *transitive* when $(xRy \wedge yRz) \leq xRz$ for all $x, y, z \in X$.

A homogeneous relation that is both reflexive and transitive is called a **preorder**.

- A preorder is *symmetric* if $xRy = yRx$;

  a symmetric preorder is also known as an **equivalence relation**.
- A preorder is *antisymmetric* if $((xRy \wedge yRx) \leq (x = y))$;

  an antisymmetric preorder is also known as a **partial order**.

Definition 1.38. A set equipped with a partial order is called a **poset**.

Posets are the simplest environment in which one can reason about boundedness.

Definition 1.39. Let $(X, \leq)$ be a poset, $S \subseteq X$.

We say $m \in X$ is a *lower bound* of S

(and that S is *bounded below*)

if $m \leq s$ for every $s \in S$.

We say $M \in X$ is an *upper bound* of S

(and that S is *bounded above*)

if $s \leq M$ for every $s \in S$.

If S is bounded both below and above, then we say S is *bounded*.

We say $M \in X$ is the *meet* or *least upper bound* or **supremum** of S (denoted $\bigvee S$ or $\sup S$)

if M is an upper bound of S

and if for any upper bound M' of S,

we have $M \leq M'$.

We say $m \in X$ is the *join* or *greatest lower bound* or **infimum** of S (denoted $\bigwedge S$ or $\inf S$)

if m is a lower bound of S

and if for any lower bound m' of S,

we have $m' \leq m$.

A lattice is a special kind of poset.

Definition 1.40. A **lattice** is a poset stable under finite meets and joins;

a lattice is **complete** when stable under arbitary meets and joins.

We move from poset to poset via monotone (order preserving) and antitone (order reversing) maps.

Definition 1.41. Let $(X, \leq_X)$ and $(Y, \leq_Y)$ be partially ordered sets. A function

$$f : X \to Y$$

is **monotone** if for all $x, y \in X$ we have

$$(x \leq_X y) \;\leq\; (f(x) \leq_Y f(y)),$$

and **antitone** if for all $x, y \in X$ we have

$$(x \leq_X y) \;\leq\; (f(y) \leq_Y f(x)).$$

Note that the composition of an even number of antitone maps forms a monotone map.

The real numbers $\mathbf{R}$ form a **totally ordered lattice.**

Note, however, that even though $\mathbf{R}$ is complete in other senses (i.e. as a metric space), $\mathbf{R}$ is not lattice-complete becuase the sequence $(s_i)_i \subset \mathbf{R}$ given by $s_i = i$ has no supremum in $\mathbf{R}$.

Adding in a global upper bound $+\infty := \top$ and a global lower bound $-\infty := \bot$ fixes this problem.

As a result, the *extended real numbers*

$$\overline{\mathbf{R}} = \bigcup \{-\infty,\ \mathbf{R},\ +\infty\}$$

form a **complete totally ordered bounded lattice.**

Definition 1.42. An **interval** is a set of one of the following forms:

$$\begin{aligned}
(a, a) &:= \{x \in \mathbf{R} : a < x < a\} = \varnothing \\
(a, a] &:= \{x \in \mathbf{R} : a < x \leq a\} = \varnothing \\
[a, a) &:= \{x \in \mathbf{R} : a \leq x < a\} = \varnothing \\
[a, a] &:= \{x \in \mathbf{R} : a \leq x \leq a\} = \{a\}
\end{aligned}$$

$$\begin{aligned}
(a, b) &:= \{x \in \mathbf{R} : a < x < b\} \\
(a, b] &:= \{x \in \mathbf{R} : a < x \leq b\} \\
[a, b) &:= \{x \in \mathbf{R} : a \leq x < b\} \\
[a, b] &:= \{x \in \mathbf{R} : a \leq x \leq b\}
\end{aligned}$$

where $a < b \in \overline{\mathbf{R}}$. The latter four intervals contain at least two points and are thus called **nondegenerate**.

Intervals where neither a nor b is either $+\infty$ or $-\infty$ are called **bounded.**

In our new interval notation, we have:

$$\mathbf{R} = (-\infty, \infty) \qquad \text{and} \qquad \overline{\mathbf{R}} = [-\infty, +\infty].$$

Using the smooth maps

$$f(x) = \lambda_\varepsilon \arctan x, \qquad f^{-1}(y) = \tan\left(y/\lambda_\varepsilon\right), \qquad \lambda_\varepsilon = \varepsilon/(\pi/2) \in (0, \infty)$$

and interpreting the asympotic values in the obvious way, we get the **open** and **closed** bounded images

$$f_*(\mathbf{R}) = (-\varepsilon, \varepsilon) \qquad \text{and} \qquad f_*\left(\overline{\mathbf{R}}\right) = [-\varepsilon, \varepsilon].$$

This is particularly nice, as it implies the parametrized curves

$$\gamma : (-\varepsilon, \varepsilon) \to \Omega, \qquad \overline{\gamma} : [-\varepsilon, \varepsilon] \to \Omega$$

have images either diffeomorphic to $\mathbf{R}$ (in the open case) or homeomorphic to $\overline{\mathbf{R}}$ (in the closed case).

For example, one can think of γ as the world-line of a particle traveling through a model of space-time.

First let's revisit what it means for an subset $\sigma \subseteq X \times X$ to be an equivalence relation on X:

Definition 1.43. An **equivalence relation** is a relation on X that is:

- Reflexive: $\sigma(x, x)$ for all $x \in X$
- Symmetric: $\sigma(x, x') = \sigma(x', x)$ for all $x, x' \in X$
- Transitive: $\big(\sigma(x, x') \wedge \sigma(x', x'')\big) \leq \sigma(x, x'')$ for all $x, x', x'' \in X$

where $\sigma(x, x')$ is shorthand for the proposition $(x, x') \in \sigma$.

Proposition 1.44. *The set Σ_X of all equivalence relations on X forms a complete lattice.*

Proof. Let's identify the lattice components.

The partial order is:

$$(\ \sigma_1 \leq \sigma_2\ ) \qquad = \qquad \Big(\ \sigma_1(x, x') \leq \sigma_2(x, x') \quad \text{for all} \quad (x, x') \in X \times X\ \Big)$$

The meet (which always exists) is:

$$\bigwedge_\iota \sigma_\iota = \bigcap_\iota \sigma_\iota$$

The join (which always exists) is:

$$\bigvee_\iota \sigma_\iota = \left(\bigcup_\iota \sigma_\iota\right)^+$$

where R^+ is the intersection of all the transitive relations containing R as a subset (i.e. the *transitive closure*).

This lattice has a $\bot$ (identity) and a $\top$ (universal relation), making it a *bounded complete lattice*. ■

For a concrete example, take $X = \mathbf{R}^{\mathbf{R}}$, and define a series of equivalence relations σ_i where $\sigma_i(f, g)$ exactly when the first i terms of their Taylor series around $x = 0$ agree.

Here's a taste of what these relations are like:

$$\sigma_1(\exp(x) - 1, x), \qquad \sigma_2(\sin x, x), \qquad \sigma_3(\cos x, 1 + x^2/2)$$

These relations form a totally ordered sublattice (or *chain*):

$$\sigma_0 \leq \sigma_1 \leq \sigma_2 \leq \cdots \leq \tau$$

where two functions $f, g : \mathbf{R} \to \mathbf{R}$ are τ-equivalent exactly when they agree on some open neighborhood of 0.

Note that while τ is a strict upper bound of the σ_i relations, *it is not a supremum!* This is a fancy way of encoding the observation that two real functions can have the same Taylor series about a point without coinciding.

Functions and Invertibility

Now we shift focus to relations $R \subseteq X \times Y$ where X does not necessarily equal Y.

Definition 1.45. Let $R \subseteq X \times Y$ be a relation. We say R is

- *left-total* if for all $x \in X$ there is some $y \in Y$ such that xRy
- *right-total* if for all $y \in Y$ there is some $x \in X$ such that xRy
- *left-unique* if for all $y \in Y$, $(xRy \wedge x'Ry) \leq (x = x')$
- *right-unique* if for all $x \in X$, $(xRy \wedge xRy') \leq (y = y')$

We then say that R is

- a *partial function* if R is right-unique but not necessarily left-total,
- a *multivalued function* if R is left-total but not necessarily right-unique,
- a (*well-defined*) **function** if R is both left-total and right-unique.

Note that this definition of function is simply a more verbose restating of the definition previously given.

We can also use the above four criteria to define when a function can be undone.

Definition 1.46. If $f : X \to Y$ is a function, then we say that f is

- **injective** or **left-invertible** if f is also left-unique,
- **surjective** or **right-invertible** if f is also right-total,
- **bijective** or **invertible** if f is also both left-unique and right-total.

Related to the adjectives above are the following nouns:

- injective functions are also called *injections,*
- surjective functions are also called *surjections,*
- bijective functions are also called *bijections,* or in certain cases *set-isomorphisms.*

Definition 1.47 (More on left-invertibility and right-invertibility).

The **identity map** on a set S is the map $1_S : S \to S$ given by $s \mapsto s$.

Left-inverses (i.e. functions $\ell : Y \to X$ such that $\ell \circ f = 1_X$) are also called **retractions.**

Right-inverses (i.e. functions $r : Y \to X$ such that $f \circ r = 1_Y$) are also called **sections.**

One can send subsets forwards and backwards through functions.

Definition 1.48. Let $f : X \to Y$ be a function. There are functions

$$f_* : \mathcal{P}(X) \to \mathcal{P}(Y) \qquad \text{and} \qquad f^* : \mathcal{P}(Y) \to \mathcal{P}(X)$$

called **image** and **preimage**, respectively given by

$$f_*(A) = \{y \in Y : \exists x \in A \text{ such that } f(x) = y\} \quad \text{and} \quad f^*(B) = \{x \in X : f(x) \in B\}.$$

Note that image respects the order of composition whereas preimage reverses the order of composition:

$$(g \circ f)_* = g_* \circ f_* \qquad \text{and} \qquad (g \circ f)^* = f^* \circ g^*.$$

In slightly more ornate language, we say image is *covariant* and preimage is *contravariant*.

Proposition 1.49. *Let* $f : X \to Y$ *with* $S \subseteq Z \subseteq X$ *and* $T \subseteq W \subseteq Y$. *Then:*

1. $f_*(S) \subseteq f_*(Z)$.
2. $f^*(T) \subseteq f^*(W)$.
3. $Z \subseteq f^*(f_*(Z))$.
4. $f_*(f^*(W)) \subseteq W$.
5. *If* f *is injective then* $f^*(f_*(Z)) \subseteq Z$.
6. *If* f *is surjective then* $W \subseteq f_*(f^*(W))$.

Proof. We'll show all of these step by step so that the reader can get an idea of how these work. All of them are easy once one understands the general pattern.

1. Let $s \in f_*(S)$. Then by definition of image, there is some $s' \in S$ such that $f(s') = s$. But $s' \in Z$ since $S \subseteq Z$. By definition of image, there is some $s' \in Z$ such that $f(s') = s$, so $s \in f_*(Z)$.
2. Let $t \in f^*(T)$. Then by definition of preimage, $f(t) \in T$. But $f(t) \in W$ since $T \subseteq W$. By definition of preimage, $t \in f^*(W)$.
3. Let $z \in Z$. Then $f(z) \in f_*(Z)$ by definition of image, and $z \in f^*(f_*(Z))$ by definition of preimage.
4. Let $w \in f_*(f^*(W))$. Then by definition of image, there is some $w' \in f^*(W)$ such that $f(w') = w$, so $w \in W$ by definition of preimage.
5. Let f be injective and let $z \in f^*(f_*(Z))$. Then by definition of preimage, $f(z) \in f_*(Z)$. By definition of image, there is some $z' \in Z$ such that $f(z) = f(z')$. But by injectivity of f, we must have $z = z'$, so $z \in Z$.
6. Let f be surjective and let $w \in W$. Then $w \in Y$ since $W \subseteq Y$. By surjectivity of f, there is some $w' \in X$ such that $f(w') = w$. By definition of preimage, $w' \in f^*(W)$. By definition of image, $w \in f_*(f^*(W))$.

For practice, it might be helpful to do these on your own without looking at this page. ■

Functions and Set Operations

Here's how image and preimage behave with respect to union, intersection, and complement:

Proposition 1.50. *Let* f *be a function. Then*

$$f_*\left(\bigcup_{\iota\in J} X_\iota\right) = \bigcup_{\iota\in J} f_*(X_\iota) \qquad f^*\left(\bigcup_{\iota\in J} Y_\iota\right) = \bigcup_{\iota\in J} f^*(Y_\iota)$$

$$f_*\left(\bigcap_{\iota\in J} X_\iota\right) \subseteq \bigcap_{\iota\in J} f_*(X_\iota) \qquad f^*\left(\bigcap_{\iota\in J} Y_\iota\right) = \bigcap_{\iota\in J} f^*(Y_\iota)$$

$$f_*(X^c) \subseteq (f_*(X))^c \qquad f^*(Y^c) = (f^*(Y))^c$$

with equality in the intersection image case when f *is injective,*
and equality in the complement image case when f *is surjective.*

Check these if you feel the need to do so. The proofs are very similar to the proof of the previous proposition, with the only novelty being that one must invoke the definitions of intersection, union, and complement.

Note that by the above result, preimages are "nicer" than images. A more precise way of stating this is that the preimage map forms a homomorphism of lattices (i.e. preserves all of the lattice structure), whereas the image map does not form a homomorphism of lattices due to its failure to preserve intersections.

More on Functions and Equivalence Relations

Equivalance relations often organize sets in the following way.

Definition 1.51. Let X be a set, and $\sigma \subset X \times X$ an equivalence relation.

The **quotient set** X/σ is then the partition whose parts are equivalence classes of σ,

i.e. collections of elements of X which are all related to each other via σ.

The equivalence classes of a quotient are pairwise disjoint, and their union is all of X. We will not prove this next proposition, since it should be clear by now what's going on: the isomorphism is inevitably $\phi([x]_f) = f(x)$.

Proposition 1.52 (Canonical Decomposition in Set).

Every function $f : X \to Y$ *defines an equivalence relation* $\sim_f$ *where* $x \sim_f x'$ *exactly when* $f(x) = f(x')$.

Furthermore, we have a decomposition

$$X \xrightarrow{\pi} X/{\sim_f} \xrightarrow{\phi} f_*(X) \xrightarrow{\iota} Y$$

where π *is a surjection,* ϕ *is a bijection, and* ι *is an injection.*

Cantor's Theorem

Theorem 1.53. *Let* X *be a set. Then* $|\mathcal{P}(X)| > |X|$.

Proof. We will show that any map $f : X \to \mathcal{P}(X)$ cannot be surjective.

Suppose otherwise, and let $Y = \{x \in X : x \notin f(x)\}$. Then there exists $\xi \in X$ such that $f(\xi) = Y$. But by construction, $\xi \in Y$ if and only if $\xi \notin f(\xi) = Y$. This is a contradiction, so f cannot be surjective.

On the other hand, $g : X \to \mathcal{P}(X)$ given by $g(x) = \{x\}$ is injective, so

$$|\mathcal{P}(X)| > |X|$$

as desired. ■

The Schröder-Bernstein Theorem

Here we prove that two opposing injections between two sets is enough to establish a bijection between the sets. As we will see, this is, at its core, a result about fixed points of monotone functions on complete lattices.

Theorem 1.54 (Knaster-Tarski). *Let* L *be a complete lattice and* $f : L \to L$ *a monotone function. Then*

$$\alpha = \bigvee\{x \in L : x \leq f(x)\}$$

is a fixed point of f. *Further,* α *is the greatest fixed point of* f.

Proof. Let $H = \{x \in L : x \leq f(x)\}$. For all $x \in H$ we have $x \leq \alpha$, so $x \leq f(x) \leq f(\alpha)$. Thus $f(\alpha)$ is an upper bound of H, so that $\alpha \leq f(\alpha)$. By monotonicity of f we have $f(\alpha) \leq f(f(\alpha))$. So $f(\alpha) \in H$, i.e. $f(\alpha) \leq \alpha$. So α is a fixed point of f. If $f(\beta) = \beta$ then $\beta \in H$ and so $\beta \leq \alpha$. ■

Corollary 1.55 (Banach's Decomposition). *Let* X *and* Y *be sets with* $f : X \to Y$ *and* $g : Y \to X$. *Then there exist disjoint subsets* X_1 *and* X_2 *of* X *and* Y_1 *and* Y_2 *of* Y *such that* $f(X_1) = Y_1$, $g(Y_2) = X_2$, $X = X_1 \sqcup X_2$ *and* $Y = Y_1 \sqcup Y_2$.

Proof. The key observation is that $\mathcal{P}(X)$ and $\mathcal{P}(Y)$ form complete lattices. For any set S, let $\alpha_S : \mathcal{P}(S) \to \mathcal{P}(S)$ denote complement, i.e. $\alpha_S(T) = S \setminus T$. Then define

$$\varphi : \mathcal{P}(X) \to \mathcal{P}(X) \quad \text{via} \quad S \mapsto \alpha_X \circ g_* \circ \alpha_Y \circ f_*(S).$$

Since φ is the composition of two monotone functions and two antitone functions, it is itself monotone. Apply Knaster-Tarski to obtain a fixed point. ■

Theorem 1.56 (Schröder-Bernstein). *Let* X *and* Y *be sets and suppose there exist injective maps* $f : X \to Y$ *and* $g : Y \to X$. *Then there is a bijective map* $h : X \to Y$.

Proof. Let f and g as in the Banach decomposition theorem be injective. Then we may set

$$h : X \to Y$$

to be defined as f on A and g^{-1} on $X \setminus A$. This completes the proof. ■

Enumerating **Z** and **Q**

We now discuss bijections from the naturals to the integers and rationals (and why the reals are different).

One way to enumerate the integers is to note that every integer can be written as a sum of powers of -2 in a unique way. This is sometimes called *negative binary.*

For example, $-1 = 11_{-2}$ and $8 = 11000_{-2}$.

So we can get an enumeration of the integers by simply counting in base -2.

Here are the first few terms of that enumeration:

n	0	1	2	3	4	5	6	7	8	9	10	11	12
m_n	0	1	-2	-1	4	5	2	3	-8	-7	-10	-9	-4

Note how the enumeration alternates between 1 positive number, then 2 negative numbers, then 4 positive numbers, and so on. This gives us a bijection $f : \mathbf{N} \to \mathbf{Z}$.

One way to enumerate the rationals is to use our bijection from the previous enumeration combined with the fundamental theorem of arithmetic: send n to $f(n)$, then map all of the exponents e_i in the prime factorization of $f(n)$ to $f(e_i)$.

Here are the first few terms of that enumeration:

n	0	1	2	3	4	5	6	7	8	9	10	11	12
m_n	0	1	-2	-1	$\frac{1}{4}$	5	2	3	$-\frac{1}{2}$	-7	-10	$-\frac{1}{9}$	$-\frac{1}{4}$

This gives us a bijection $g : \mathbf{N} \to \mathbf{Q}$.

Cantor's Diagonal Argument

We already know the reals are impossible to enumerate due to Cantor's theorem, but there's a classic argument here that's worth getting into. Consider the real numbers between 0 and 1 represented as binary strings, and suppose we had an enumeration:

0) 0.**0**1010101101101011011010101010100101101011011...
1) 0.0**0**001101011111010001011010110110100101001010...
2) 0.01**1**01010011101011000100010101110110010010101...
3) 0.100**1**0101011110101110110101010011010100100100...
4) 0.1111**0**100101010100101010100111101010101001010...
5) 0.00000**0**00000000000111010101010101011010101010...

$\vdots$

Then, simply by flipping the nth bit (to the right of the decimal point) of the $n-1$th number, we obtain

ω) 0.**110011**011101001001010001010101010100100100110...

which is a number not on our list. This is **Cantor's diagonal argument.**

Number Systems: Algebra

In this subsection, we analyze the algebraic structure of **Z** and **Q**, temporarily ignoring the order structure.

Rings, Domains, and Fields

Definition 1.57. A **ring** is an ordered pair $((R, +, 0), \cdot)$ where:

1. The first element, an ordered triple $(R, +, 0)$, forms an abelian group:
   a) $a + b = b + a$ for all $a, b \in R$
   b) $(a + b) + c = a + (b + c)$ for all $a, b, c \in R$
   c) $a + 0 = a$ for all $a \in R$
   d) For every $a \in R$ there is a unique $-a \in R$ such that $a + (-a) = 0$
2. The second element is a binary operation $\cdot : R \times R \to R$ given by $\cdot(a, b) = ab$ satisfying:
   a) $(ab)c = a(bc)$
   b) $(a + b)(c + d) = ac + ad + bc + bd$

Rings are very general objects, so we qualify them further in order to keep sane:

Definition 1.58. Some miscellaneous ring terms:

- A **unital ring** is a triple $((R, +, 0), \cdot, 1)$ where $((R, +, 0), \cdot)$ forms a ring

  and where $1 \in R$ is such that $a1 = a = 1a$ for all $a \in R$.
- A ring $((R, +, 0), \cdot)$ is **commutative** if $ab = ba$ for all $a, b \in R$.
- The **zero ring** $((R, +, 0), \cdot, 1)$ is a commutative unital ring where $R = \{0\}$ (thus $1 = 0$)

  and where the operations $+$ and $\cdot$ are defined such that $0 + 0 = 0$ and $00 = 0$.
- A ring $((R, +, 0), \cdot)$ is said to have the **cancellation property** if

  $$\Big((a \neq 0) \wedge (ab = ac)\Big) \quad \leq \quad (b = c)$$

  for all $a, b, c \in R$.
- A nonzero commutative unital ring $((R, +, 0), \cdot, 1)$ with the cancellation property is called a **domain.**
- The **unit group** of a unital ring $((R, +, 0), \cdot, 1)$ is the triple $(R^\times, \cdot, 1)$

  where $R^\times \subset R$ is the subset of elements in R with a multiplicative inverse.
- A domain for which $R^\times = R \setminus \{0\}$ is called a **field.**

Notably, the integers **Z** form a domain and the rationals **Q** form a field.

As one might predict, rings form a subcategory of Set, which we'll call Ring; similarly, fields form a subcategory of Ring, which we'll call Field.

The arrows in Ring are maps that preserve both the additive group structure and the ring multiplication:

Definition 1.59. Let R and S be rings.

A **ring homomorphism** is a function $\varphi : R \to S$ such that:

$$\varphi(ab + cd) = \varphi(a)\varphi(b) + \varphi(c)\varphi(d).$$

If, furthermore, R and S are both unital, we also require $\varphi(1_R) = 1_S$.

A **ring isomorphism** is a bijective ring homomorphism.

Arrows in Field don't have a stricter definition, but nonetheless have stricter structure. For example, every field homomorphism $\varphi : K \to L$ is *injective*, so that we usually think of L as an **extension** of K and write $K \subset L$.

Definition 1.60. The **field of fractions** of a domain R is the quotient set

$$\mathrm{Frac}(R) := \big(R \times R \setminus \{0\}\big)/\sim$$

where $(a, b) \sim (c, d)$ exactly when $ac = bd$. As the name might suggest, this forms a field.

The archetypal example is of course $\mathbf{Q} = \mathrm{Frac}(\mathbf{Z})$.

Ideals of a Ring

Definition 1.61. An **ideal** I of a ring R is a subset of R closed under I–addition and R–multiplication.

- Note that $\{0\}$ forms an ideal, denoted (0).
- Similarly, R forms an ideal, denoted (1).
- Ideals are used to form **quotient rings** via the equivalence relation

$$(a \sim_I b) = a - b \in I.$$

- We say I is **prime** if $(ab \in I) \leq (a \in I \vee b \in I)$ for all $a, b \in R$.
  - The quotient of a ring by a prime ideal is an domain.
- We say I is **maximal** if $\big(I \subseteq X \subseteq R\big) \leq \big(X = I \ \wedge \ X = R\big)$ for any ideal X of R.
  - The quotient of a ring by a maximal ideal is a field.

Some examples and observations:

- For $\mathbf{Z}$, the maximal ideals are (p) for all primes $p \in \mathbf{Z}$, and the prime ideals are the maximal ideals plus (0).
- For a field (i.e. for $\mathbf{Q}$), the only maximal/prime ideal is (0).
- An example of a noncommutative ring is the set of $n \times n$ matrices $\mathrm{Mat}_n(R)$ with entries in a ring R. The unit group of $\mathrm{Mat}_n(R)$ is denoted $\mathrm{GL}(n, R)$.
- Note that noncommutative rings have *left-ideals* and *right-ideals,* something we probably won't explore further (as this gets very complicated very quickly). Also note that R can itself be a matrix ring; we may explore this later.
- Given an ordered field Γ, the set of all Cauchy sequences $\kappa(\Gamma)$ forms an ordered ring. The set of all Cauchy sequences tending to zero forms a maximal ideal $\mathfrak{m}_0$, so that $\kappa(\Gamma)/\mathfrak{m}_0$ forms a field. We give more details on this in the next subsection.
- In particular, the real numbers as an algebraic structure may be defined as $\mathbf{R} := \kappa(\mathbf{Q})/\mathfrak{m}_0$.
- An example of a nonzero commutative unital ring *without* the cancellation property is $C^\infty(\mathbf{R})$, the ring of smooth functions $f : \mathbf{R} \to \mathbf{R}$.

  Indeed, consider the smooth functions:

  $$f(x) = \begin{cases} \exp(-1/x) & x > 0 \\ 0 & \text{otherwise} \end{cases} \qquad \text{and} \qquad g(x) = \begin{cases} \exp(1/x) & x < 0 \\ 0 & \text{otherwise} \end{cases}$$

  Note that $f \neq 0$ and $g \neq 0$, yet $fg = 0$.

  We will revisit the ring $C^\infty(\mathbf{R})$ in the future when we study calculus and differential equations.

Polynomial Rings

Definition 1.62. The **center** of a ring R, denoted $Z(R)$, is the largest commutative subring of R.

An **R–algebra** is a pair (A, φ) where A is a ring and $\varphi : R \to Z(A)$ is a ring homomorphism.

The archetypal example of a R–algebra is the **ring of polynomials in one variable with coefficients in R,** denoted $R[\xi]$. We define $R[\xi, \eta] := (R[\xi])[\eta]$. If ξ is a domain, so is $R[\xi]$. However, when K is a field, the polynomial ring $K[\xi]$ is only a field in the most trivial circumstances.

Definition 1.63. The **complex numbers** can be formed via the quotient

$$\mathbf{C} := \frac{\mathbf{R}[\xi]}{(\xi^2 + 1)}.$$

Since $(\xi^2 + 1)$ is maximal in $\mathbf{R}[\xi]$, this forms a field.

Number Systems: Order

We now reintroduce the order structure onto $\mathbf{Z}$ (an ordered domain) and $\mathbf{Q}$ (an ordered field).

Definition 1.64. An **ordered ring** is a 6-tuple $(F, +, \cdot, 0, 1, <)$ such that:

- $(F, +, \cdot, 0, 1)$ forms a nonzero commutative unital ring
- for all $x \in F$, $x \not< x$
- for all $x, y, z \in F$, if $x < y$ and $y < z$ then $x < z$
- for all $x, y \in F$, one of the following hold:
$$x < y, \qquad y < x, \qquad x = y$$
- for all $x, y, z \in F$, if $x < y$ then $x + z < y + z$
- for all $x, y, z \in F$, if $x < y$ and $0 < z$ then $x \cdot z < y \cdot z$

That is, an ordered ring is a set in which one can add, subtract, multiply, and compare elements.

An **ordered field** is an ordered ring in which one can divide by nonzero elements.

The elements of an ordered ring can be divided into positive, negative, and 0.

This is enough to have an absolute value.

Definition 1.65. Let X be an ordered ring.

If $v : X \to X$ is such that:

- for all $x \in X$, $v(x) \geq 0$
- for all $x \in X$, $(v(x) = 0) = (x = 0)$
- for all $x, y \in X$, $v(xy) = v(x)v(y)$
- for all $x, y \in X$, $v(x + y) \leq v(x) + v(y)$

then v is called an **absolute value**.

Ordered Fields: Convex Cones

Definition 1.66. Let Γ be an ordered field.

The **positive convex cone** of Γ is simply its positive elements. We denote this cone by Γ^{+}.

The **nonnegative convex cone** of Γ is the union of its positive elements with 0. We denote this cone by $\Gamma_{\geq 0}$.

From this definition, we observe that $\mathbf{N} = \mathbf{Q}_{\geq 0} \cap \mathbf{Z}$ and $\mathbf{N}^{\mathbf{++}} = \mathbf{Q}^{+} \cap \mathbf{Z}$.

Ordered Fields: Cauchy Sequences

Throughout this subsubsection, let Γ be an ordered field. We denote the absolute value on Γ by $|\cdot|$.

Definition 1.67. Denote by Γ^+ the positive elements of Γ.

A sequence $(x_n)_n \subseteq \Gamma$ is **Cauchy** if for any $\varepsilon \in \Gamma^+$ there is an $N \in \mathbf{N}$ such that

$$n, n' \geq N \quad \text{implies} \quad |x_{n'} - x_n| < \varepsilon.$$

Denote by $\kappa(\Gamma)$ the set of all Cauchy sequences on Γ.

Because Γ is an ordered field, we may add and multiply sequences with entries in Γ pointwise,

i.e. by adding or multiplying their terms together.

Proposition 1.68. $\kappa(\Gamma)$ *is closed under addition.*

Proof. Let $(a_n)_n, (b_n)_n \in \kappa(\Gamma)$, and let $\varepsilon > 0$. Then $\varepsilon/2 > 0$ as well, so there exists N_a such that for all $n, n' \geq N_a$ we have

$$|a_n - a_{n'}| < \varepsilon/2,$$

and an N_b such that for all $n, n' \geq N_b$ we have

$$|b_n - b_{n'}| < \varepsilon/2.$$

Let $N = \max(\{N_a, N_b\})$. Then by the triangle inequality,

$$|(a_n + b_n) - (a_{n'} + b_{n'})| \leq |a_n - a_{n'}| + |b_n - b_{n'}| < \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

so $(a_n + b_n)_n \in \kappa(\Gamma)$. ■

Proposition 1.69. *If* $(x_n)_n$ *is Cauchy, then there is an* $M \in \Gamma^+$ *such that*

$$|x_n| < M$$

for all nonnegative n.

Proof. Let $\varepsilon = 1$. Then there exists an N such that for all $m, n \geq N$ we have $|x_n - x_m| < 1$. Let $M = \max(\{|x_0|, \ldots, |x_N|\}) + 1$. Clearly $|x_n| < M$ for all $n \leq N$. Now suppose $n > N$. Then

$$|x_n| = |x_n - x_N + x_N| \leq |x_n - x_N| + |x_N| < M$$

so in all cases $(x_n)_n$ is bounded by M. ■

Proposition 1.70. $\kappa(\Gamma)$ *is closed under multiplication.*

Proof. Let $(a_n)_n, (b_n)_n \in \kappa(\Gamma)$. By the previous proposition, there exist $M_a \in \Gamma^+$ bounding $(a_n)_n$ and $M_b \in \Gamma^+$ bounding $(b_n)_n$. Let $\varepsilon \in \Gamma^+$ so that $\lambda = \frac{\varepsilon}{M_a + M_b} \in \Gamma^+$ as well. Since $(a_n)_n$ is Cauchy, there exists $N_a \in \mathbf{N}$ such that for all $n, n' \geq N_a$, we have $|a_n - a_{n'}| < \lambda$ and also $N_b \in \mathbf{N}$ such that for all $n, n' \geq N_b$, we have $|b_n - b_{n'}| < \lambda$. Pick $N = \max(\{N_a, N_b\})$. Then

$$\begin{aligned} |a_n b_n - a_{n'} b_{n'}| &= |a_n b_n - a_n b_{n'} + a_n b_{n'} - a_{n'} b_{n'}| \\ &\leq |a_n||b_n - b_{n'}| + |a_n - a_{n'}||b_{n'}| \\ &< \lambda(M_a + M_b) = \varepsilon. \end{aligned}$$

Thus, $(a_n b_n)_n$ is Cauchy. ■

Definition 1.71. Say that $(a_n)_n \in \kappa(\Gamma)$ **tends to zero** if for every $\varepsilon \in \Gamma^+$ there is some $N_\varepsilon \in \mathbf{N}$ such that for all $n \geq N$, we have $|a_n| < \varepsilon$.

Say $(a_n)_n \sim_3 (b_n)_n$ if $(a_n - b_n)_n$ tends to zero.

Proposition 1.72. *The Cauchy sequences tending to zero form a maximal ideal of* $\kappa(\Gamma)$.

Proof. The idea is once one throws in a Cauchy sequence tending toward some nonzero number (say 1), one can then scale that Cauchy sequence by any number $r \in \Gamma$ to get a Cauchy sequence tending toward r. ■

Complex Numbers: Absence of Total Order

Suppose it were possible to impose a total order on $\mathbf{C}$. Then every nonzero element of $\mathbf{C}$ would be either positive or negative. Let's check i, the imaginary unit. Suppose i were positive. Then $i > 0$. But positive numbers form a convex cone, i.e. if one squares a positive number, it ought to stay positive. However, $i^2 = -1 < 0$. Now suppose i were negative. Then $-i$ would be positive. But squaring this supposedly positive number yields

$$(-i)^2 = (-1)^2 i^2 = i^2 = -1 < 0.$$

Since we can't even sort i as either positive or negative, there is no way we could sort all of $\mathbf{C}^\times$ into positive and negative subsets. So even though the complex numbers are algebraically closed and contain the real numbers $\mathbf{R}$ as a subfield, we do lose something when we go from $\mathbf{R}$ to $\mathbf{C}$, namely the ability to compare any two elements in a globally consistent way.

Though this is fairly obvious, it can lead to some bizzare consequences.

For example, in a real inner product space (see Chapter 3), a hyperplane basically bifurcates its complement into a part that's "above" the hyperplane and a part that's "below" the hyperplane. For complex inner product spaces, the complement of a hyperplane is actually *simply connected.* Thus, **complex hyperplanes don't have "sides:"** if they did, this would imply that $\mathbf{C}$ could be totally ordered!

Perhaps the reals aren't so pathological after all.