XXI

Set Theory I: ZFC, Ordinals, and Cardinals

Part XXI: Contents

88 Interlude: Cauchy's functional equation and Zorn's lemma

This is an informal chapter on Zorn's lemma, which will give an overview of what's going to come in the last parts of the Napkin. It can be omitted without loss of continuity.

In the world of olympiad math, there's a famous functional equation that goes as follows:

$$f \colon \mathbb{R} \to \mathbb{R} \qquad f(x+y) = f(x) + f(y).$$

Everyone knows what its solutions are! There's an obvious family of solutions $f(x) = cx$. Then there's also this family of... uh... noncontinuous solutions (mumble grumble) pathological (mumble mumble) Axiom of Choice (grumble).

There's also this thing called Zorn's lemma. It sounds terrifying, because it's equivalent to the Axiom of Choice, which is also terrifying because why not.

In this post I will try to de-terrify these things, because they're really not as terrifying as they sound.

§88.1 Let's construct a monster

Let us just see if we can try and construct a "bad" f and see what happens.

By scaling, let's assume WLOG that $f(1) = 1$. Thus $f(n) = n$ for every integer n, and you can easily show from here that

$$f\left(\frac{m}{n}\right) = \frac{m}{n}.$$

So f is determined for all rationals. And then you get stuck.

None of this is useful for determining, say, $f(\sqrt{2})$. You could add and subtract rational numbers all day and, say, $\sqrt{2}$ isn't going to show up at all.

Well, we're trying to set things on fire anyways, so let's set

$$f(\sqrt{2}) = 2015$$

because why not? By the same induction, we get $f(n\sqrt{2}) = 2015n$, and then that

$$f\left(a + b\sqrt{2}\right) = a + 2015b.$$

Here a and b are rationals. Well, so far so good – as written, this is a perfectly good solution, other than the fact that we've only defined f on a tiny portion of the real numbers.

Well, we can do this all day:

$$f\left(a + b\sqrt{2} + c\sqrt{3} + d\pi\right) = a + 2015b + 1337c - 999d.$$

Perfectly consistent.

You can kind of see how we should keep going now. Just keep throwing in new real numbers which are "independent" to the previous few, assigning them to whatever junk we want. It feels like it *should* be workable. . .

In a moment I'll explain what "independent" means (though you might be able to guess already), but at the moment there's a bigger issue: no matter how many numbers we throw, it seems like we'll never finish. Let's address the second issue first.

§88.2 Review of finite induction

When you do induction, you get to count off 1, 2, 3, ... and so on. So for example, suppose we had a "problem" such as:

> Prove that the intersection of n open intervals is either $\varnothing$ or an open interval.

You can do this by induction easily: it's true for $n = 2$, and for the larger cases it's similarly easy.

But you can't conclude from this that *infinitely* many open intervals intersect at some open interval. Indeed, this is false: consider the intervals

$$(-1, 1), \quad \left(-\frac{1}{2}, \frac{1}{2}\right), \quad \left(-\frac{1}{3}, \frac{1}{3}\right), \quad \left(-\frac{1}{4}, \frac{1}{4}\right), \quad \dots$$

This *infinite* set of intervals intersects at a single point $\{0\}$!

The moral of the story is that induction doesn't let us reach infinity. Too bad, because we'd have loved to use induction to help us construct a monster. That's what we're doing, after all – adding things in one by one.

§88.3 Transfinite induction

Well, it turns out we can, but we need a new notion of number, the so-called *ordinal number*. I define these in their full glory in the first two sections of Chapter 90 (and curious readers are even invited to jump ahead to those two sections), but for this chapter I won't need that full definition yet.

Here's what I want to say: after all the natural numbers

$$0,\ 1,\ \dots,$$

I'll put a *new number* called ω, the first ordinal greater than all the natural numbers. After that there's more numbers called

$$\omega + 1,\ \omega + 2,\ \dots$$

and eventually a number called $\omega \cdot 2$.

The list goes on:

$$\begin{aligned}
&0, 1, 2, 3, \dots, \omega \\
&\quad \omega + 1, \omega + 2, \dots, \omega + \omega \\
&\quad \omega \cdot 2 + 1, \omega \cdot 2 + 2, \dots, \omega \cdot 3 \\
&\quad \vdots \\
&\quad \omega^2 + 1, \omega^2 + 2, \dots \\
&\quad \vdots \\
&\quad \omega^3, \dots, \omega^4, \dots, \omega^\omega \dots, \omega^{\omega^{\omega^\cdots}}
\end{aligned}$$

Pictorially, it kind of looks like this:

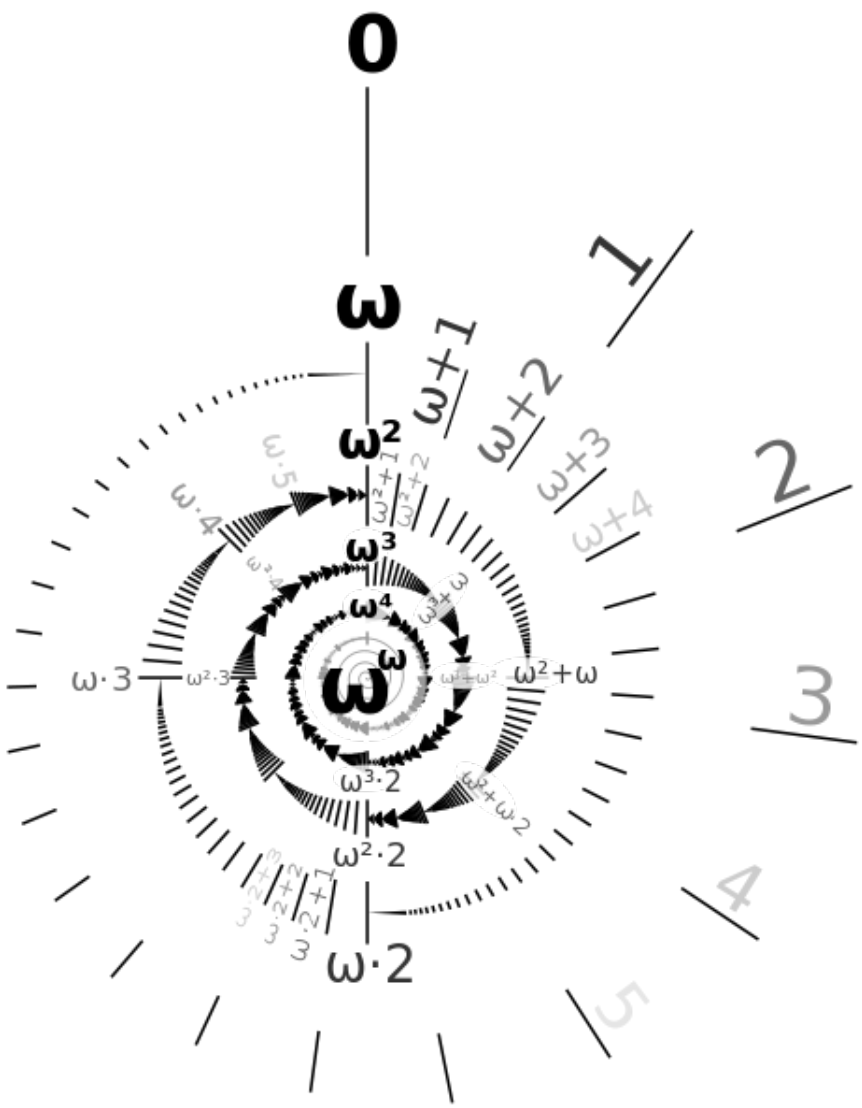


Image from [ca]

(Note that the diagram only shows an initial segment; there are still larger ordinals like $\omega^{\omega^\omega} + 1000$ and so on).

Anyways, in the same way that natural numbers "dominate" all finite sets, the ordinals dominate *all the sets*, in the following sense. Essentially, assuming the Axiom of Choice, it follows that for every set S there's some ordinal α which is larger than S (in a sense I won't make precise until later chapters).

But it turns out (and you can intuitively see) that as large as the ordinals grow, there is no *infinite descending chain*. Meaning: if I start at an ordinal (like $2\omega + 4$) and jump down, I can only take finitely many jumps before I hit 0. (To see this, try writing down a chain starting at $2\omega + 4$ yourself.) Hence, induction and recursion still work verbatim:

Theorem 88.3.1 (Transfinite induction)

Given a statement $P(-)$, suppose that

- $P(0)$ is true, and
- If $P(\alpha)$ is true for all $\alpha < \beta$, then $P(\beta)$ is true.

Then $P(\beta)$ is true.

Similarly, you're allowed to do recursion to define x_β if you know the value of x_α for all $\alpha < \beta$.

The difference from normal induction or recursion is that we'll often only do things like "define $x_{n+1} = \dots$". But this is not enough to define x_α for all α. To see this, try using our normal induction and see how far we can climb up the ladder.

Answer: you can't get ω! It's not of the form $n+1$ for any of our natural numbers n – our finite induction only lets us get up to the ordinals less than ω. Similarly, the simple $+1$ doesn't let us hit the ordinal $\omega \cdot 2$, even if we already have $\omega + n$ for all n. Such ordinals are called **limit ordinals**. The ordinal that *are* of the form $\alpha + 1$ are called **successor ordinals**.

So a transfinite induction or recursion is very often broken up into three cases. In the induction phrasing, it looks like

- (Zero Case) First, resolve $P(0)$.
- (Successor Case) Show that from $P(\alpha)$ we can get $P(\alpha+1)$.
- (Limit Case) Show that $P(\lambda)$ holds given $P(\alpha)$ for all $\alpha < \lambda$, where λ is a limit ordinal.

Similarly, transfinite recursion often is split into cases too.

- (Zero Case) First, define x_0.
- (Successor Case) Define $x_{\alpha+1}$ from x_α.
- (Limit Case) Define x_λ from x_α for all $\alpha < \lambda$, where λ is a limit ordinal.

In both situations, finite induction only does the first two cases, but if we're able to do the third case we can climb far above the barrier ω.

§88.4 Wrapping up functional equations

Let's return to solving our problem.

Let S_n denote the set of "base" numbers we have at the nth step. In our example, we might have

$$S_1 = \{1\}, \quad S_2 = \left\{1, \sqrt{2}\right\}, \quad S_3 = \left\{1, \sqrt{2}, \sqrt{3}\right\}, \quad S_4 = \left\{1, \sqrt{2}, \sqrt{3}, \pi\right\}, \quad \dots$$

and we'd like to keep building up S_i until we can express all real numbers. For completeness, let me declare $S_0 = \varnothing$.

First, I need to be more precise about "independent". Intuitively, this construction is working because

$$a + b\sqrt{2} + c\sqrt{3} + d\pi$$

is never going to equal zero for rational numbers a, b, c, d (other than all zeros). In general, a set X of numbers is "independent" if the combination

$$c_1x_1 + c_2x_2 + \cdots + c_mx_m = 0$$

never occurs for rational numbers $\mathbb{Q}$ unless $c_1 = c_2 = \cdots = c_m = 0$. Here $x_i \in X$ are distinct. Note that even if X is infinite, I can only take finite sums! (This notion has a name: we want X to be **linearly independent** over $\mathbb{Q}$; see the chapter on vector spaces for more on this!)

When do we stop? We'd like to stop when we have a set $S_{\text{something}}$ that's so big, every real number can be written in terms of the independent numbers. (This notion also has a name: it's called a $\mathbb{Q}$-basis.) Let's call such a set **spanning**; we stop once we hit a spanning set.

The idea that we can induct still seems okay: suppose S_α isn't spanning. Then there's some number that is independent of S_α, say $\sqrt{2015}\pi$ or something. Then we just add it to get $S_{\alpha+1}$. And we keep going.

Unfortunately, as I said before it's not enough to be able to go from S_α to $S_{\alpha+1}$ (successor case); we need to handle the limit case as well. But it turns out there's a trick we can do. Suppose we've constructed *all* the sets S_0, S_1, S_2, ..., one for each positive integer n, and none of them are spanning. The next thing I want to construct is S_ω; somehow I have to "jump". To do this, I now take the infinite union

$$S_\omega \stackrel{\text{def}}{=} S_0 \cup S_1 \cup S_2 \cup \ldots.$$

The elements of this set are also independent (why?).

Ta-da! With the simple trick of "union all the existing sets", we've just jumped the hurdle to the first limit ordinal ω. Then we can construct $S_{\omega+1}$, $S_{\omega+2}$, ..., once again – just keep throwing in elements. Then when we need to jump the next hurdle to $S_{2\omega}$, we just do the same trick of "union-ing" all the previous sets.

So we can formalize the process as follows:

1. Let $S_0 = \varnothing$.
2. For a successor stage $S_{\alpha+1}$, add any element to S_α to obtain $S_{\alpha+1}$.
3. For a limit stage S_λ, take the union $\bigcup_{\gamma<\lambda} S_\gamma$.

How do we know that we'll stop eventually? Well, the thing is that this process consumes a lot of real numbers. In particular, the ordinals get larger than the size of $\mathbb{R}$ (assuming Choice). Hence if we don't stop we will quite literally reach a point where we have used up every single real number. Clearly that's impossible, because by then the elements can't possibly be independent!

So by transfinite recursion, we eventually hit some S_γ which is spanning: the elements are all independent, but every real number can be expressed using it. Done!

§88.5 Zorn's lemma

Now I can tell you what Zorn's lemma is: it lets us do the same thing in any poset.

We can think of the above example as follows: consider all sets of independent elements. These form a partially ordered set by inclusion, and what we did was quite literally climb up a chain

$$S_0 \subsetneq S_1 \subsetneq S_2 \subsetneq \ldots.$$

It's not quite climbing since we weren't just going one step at a time: we had to do "jumps" to get up to S_ω and resume climbing. But the main idea is to climb up a poset until we're at the very top; in the previous case, when we reached the spanning set.

The same thing works verbatim with any partially ordered set $\mathbb{P}$. Let's define some terminology. A **local maximum** of the entire poset $\mathbb{P}$ is an element which has no other elements strictly greater than it. (Most authors refer to this as "maximal element", but I think "local maximum" is a more accurate term.)

Now a **chain of length γ** is a set of elements p_α for every $\alpha < \gamma$ such that $p_0 < p_1 < p_2 < \dots$. (Observe that a chain has a last element if and only if γ is a successor ordinal, like $\omega + 3$.) An **upper bound** to a chain is an element $\tilde{p}$ which is greater than or equal to all elements of the chain; In particular, if γ is a successor ordinal, then just taking the last element of the chain works.

In this language, Zorn's lemma states that

Theorem 88.5.1 (Zorn's lemma)

Let $\mathbb{P}$ be a nonempty partially ordered set. If every chain has an upper bound, then $\mathbb{P}$ has a local maximum.

Chains with length equal to a successor ordinal always have upper bounds, but this is not true in the limit case. So the hypothesis of Zorn's lemma is exactly what lets us "jump" up to define p_ω and other limit ordinals. And the proof of Zorn's lemma is straightforward: keep climbing up the poset at successor stages, using Zorn's condition to jump up at limit stages, and thus building a really long chain. But we have to eventually stop, or we literally run out of elements of $\mathbb{P}$. And the only possible stopping point is a local maximum.

If we want to phrase our previous solution in terms of Zorn's lemma, we'd say:

Proof. Look at the poset whose elements are sets of independent real numbers. Every chain $S_0 \subsetneq S_1 \subsetneq \dots$ has an upper bound $\bigcup S_\alpha$ (which you have to check is actually an element of the poset). Thus by Zorn, there is a local maximum S. Then S must be spanning, because otherwise we could add an element to it. □

So really, Zorn's lemma is encoding all of the work of climbing that I argued earlier. It's a neat little package that captures all the boilerplate, and tells you exactly what you need to check.

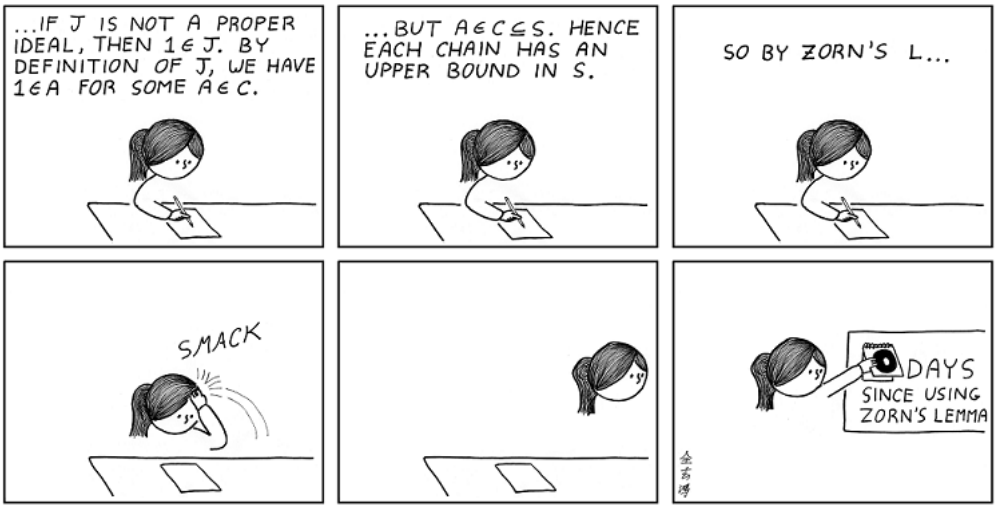


Image from [Go09]

One last thing you might ask: where is the Axiom of Choice used? Well, the idea is that for any chain there could be lots of $\tilde{p}$'s, and you need to pick one of them. Since you are making arbitrary choices infinitely many times, you need the Axiom of Choice. (Actually, you also need Choice to talk about cardinalities as in Theorem 88.3.1.) But really, it's nothing special.

§88.6 A few harder problems to think about

Problem 88A. Suppose $f\colon (0,\infty) \to \mathbb{R}$ satisfies $f(1) = 1$, $f(\pi) = \frac{22}{7}$, and

$$f(x+y) = f(x) + f(y)$$

for all $x, y > 0$. Given that $\pi \approx 3.14159265358979$, find, with proof and without a calculator, an example of a real number r such that $0 < r < 1$ and $f(r) > 9000$.

Problem 88B. Suppose $f\colon (0,\infty) \to \mathbb{R}$ satisfies

$$f(x+y) = f(x) + f(y)$$

for all $x, y > 0$. Label each of the following statements as true or false.

1. The function f can be extended to $f\colon \mathbb{R} \to \mathbb{R}$ while still satisfying Cauchy's functional equation.
2. If f is extended as in the previous statement, then f must be odd.
3. If $f(x) \geq 0$ for all $x > 0$, then f is linear.
4. If f is strictly increasing, then f is linear.
5. The function f is not a bijection.
6. It's possible for f to be injective but not linear.
7. It's possible for f to be surjective.
8. It's possible for f to be nonconstant and only take rational values.
9. It's possible for f to only take irrational values.

Problem 88C (Tukey's lemma)**.** Let $\mathcal{F}$ be a nonempty family of sets. Assume that for any set A, the set A is in $\mathcal{F}$ if and only if all its finite subsets are in $\mathcal{F}$.

Prove that there exists a maximal set $Y \in \mathcal{F}$ (i.e. Y not contained in any other set of $\mathcal{F}$).

89 Zermelo-Fraenkel with choice

Chapter 3.1 of [Le14] has a nice description of this.

§89.1 The ultimate functional equation

In abstract mathematics, we often define structures by what *properties* they should have; for example, a group is a set and a binary operation satisfying so-and-so axioms, while a metric space is a set and a distance function satisfying so-and-so axioms.

Nevertheless, these definitions rely on previous definitions. The colorful illustration of [Le14] on this:

- A *vector space* is an abelian group with...
- An *abelian group* has a binary operation such that...
- A *binary operation* on a set is...
- A *set* is ...

and so on.

We have to stop at some point, because infinite lists of definitions are bad. The stopping turns out to be a set, "defined" by properties. The trick is that we never actually define what a set is, but nonetheless postulate that these sets satisfy certain properties: these are the ZFC axioms. Loosely, ZFC can be thought of as the *ultimate functional equation*.

Before talking about what these axioms are, I should talk about the caveats.

§89.2 Cantor's paradox

Intuitively, a set is an unordered collection of elements. Two sets are equal if they share the same elements:

$$\{x \mid x \text{ is a featherless biped}\} = \{x \mid x \text{ is human}\}$$

(let's put aside the issue of dinosaurs).

As another example, we have our empty set $\varnothing$ that contains no objects. We can have a set $\{1, 2, 3\}$, or maybe the set of natural numbers $\mathbb{N} = \{0, 1, 2, \dots\}$. (For the purposes of set theory, 0 is usually considered a natural number.) Sets can even contain other sets, like $\{\mathbb{Z}, \mathbb{Q}, \mathbb{N}\}$. Fine and dandy, right?

The trouble is that this definition actually isn't good enough, and here's why. If we just say "a set is any collection of objects", then we can consider a really big set V, the set of all sets. So far no problem, right? We would have the oddity that $V \in V$, but oh well, no big deal.

Unfortunately, this existence of this V leads immediately to a paradox. The classical one is Russell's Paradox. I will instead present a somewhat simpler one: not only does V contain itself, *every subset* $S \subseteq V$ is itself an element of V (i.e. $S \in V$). If we let $\mathcal{P}(V)$ denote the **power set** of V (i.e. all the subsets of V), then we have an inclusion

$$\mathcal{P}(V) \hookrightarrow V.$$

This is bad, since:

Lemma 89.2.1 (Cantor's diagonal argument)

For *any* set X, it's impossible to construct an injective map $\iota\colon \mathcal{P}(X) \hookrightarrow X$.

Proof. Assume for contradiction ι exists.

Exercise 89.2.2. Show that if, ι exists, then there exists a surjective map $j\colon X \twoheadrightarrow \mathcal{P}(X)$. (This is easier than it appears, just "invert ι").

We now claim that j can't exist.

Let me draw a picture for j to give the idea first:

		x_1	x_2	x_3	x_4	x_5	$\dots$
x_1	$\overset{j}{\mapsto}$	**0**	1	1	0	1	$\dots$
x_2	$\overset{j}{\mapsto}$	1	**1**	0	1	1	$\dots$
x_3	$\overset{j}{\mapsto}$	0	1	**0**	0	1	$\dots$
x_4	$\overset{j}{\mapsto}$	1	0	0	**1**	0	$\dots$
x_5	$\overset{j}{\mapsto}$	0	1	1	1	**1**	$\dots$
$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Here, for each $j(x) \subseteq X$, I'm writing "1" to mean that the element is inside $j(x)$, and "0" otherwise. So $j(x_1) = \{x_2, x_3, x_5 \dots\}$. (Here the indices are ordinals rather than integers as X may be uncountable. Experts may notice I've tacitly assumed a well-ordering of X; but this picture is for motivation only so I won't dwell on the point.) Then we can read off the diagonal to get a new set. In our example, the diagonal specifies a set $A = \{x_2, x_4, x_5 \dots\}$. Then we "invert" it to get a set $B = \{x_1, x_3, \dots\}$.

Back to the formal proof. As motivated above, we define

$$B = \{x \mid x \notin j(x)\}.$$

By construction, $B \subseteq X$ is not in the image of j, which is a contradiction since j was supposed to be surjective. □

Now if you're not a set theorist, you could probably just brush this off, saying "oh well, I guess you can't look at certain sets". But if you're a set theorist, this worries you, because you realize it means that you can't just define a set as "a collection of objects", because then everything would blow up. Something more is necessary.

§89.3 The language of set theory

We need a way to refer to sets other than the informal description of "collection of objects".

So here's what we're going to do. We'll start by defining a formal *language of set theory*, a way of writing logical statements. First of all we can throw in our usual logical operators:

- $\forall$ means "for all"
- $\exists$ means "exists"
- $=$ means "equal"

- $X \implies Y$ means "if X then Y"
- $A \wedge B$ means "A and B"
- $A \vee B$ means "A or B"
- $\neg A$ means "not A".

Since we're doing set theory, there's only one more operator we add in: the inclusion $\in$. And that's all we're going to use (for now).

So how do we express something like "the set $\{1, 2\}$"? The trick is that we're not going to actually "construct" any sets, but rather refer to them indirectly, like so:

$$\exists S : x \in S \iff ((x = 1) \vee (x = 2)) .$$

This reads: "there exists an S such that x is in S if and only if either $x = 1$ or $x = 2$". We don't have to refer to sets as objects in and of themselves anymore — we now have a way to "create" our sets, by writing formulas for exactly what they contain. This is something a machine can parse.

Well, what are we going to do with things like 1 and 2, which are not sets? Answer:

Elements of sets are themselves sets.

We're going to make **everything** into a set. Natural numbers will be sets. Ordered pairs will be sets. Functions will be sets. Later, I'll tell you exactly how we manage to do something like encode 1 as a set. For now, all you need to know is that sets don't just hold objects; they hold other sets.

So now it makes sense to talk about whether something is a set or not: $\exists x$ means "x is a set", while $\nexists x$ means "x is not a set". In other words, we've rephrased the problem of deciding whether something is a set to whether it exists, which makes it easier to deal with in our formal language. That means that our axiom system had better find some way to let us show a lot of things exist, without letting us prove

$$\exists S \forall x : x \in S.$$

For if we prove this formula, then we have our "bad" set that caused us to go down the rabbit hole in the first place.

§89.4 The axioms of ZFC

I don't especially want to get into details about these axioms; if you're interested, read:

- https://blog.evanchen.cc/2014/11/13/set-theory-an-intro-to-zfc-part-1/
- https://blog.evanchen.cc/2014/11/18/set-theory-part-2-constructing-the-ordinals/

Here is a much terser description of the axioms, which also includes the corresponding sentence in the language of set theory. It is worth the time to get some practice parsing $\forall$, $\exists$, etc. and you can do so by comparing the formal sentences with the natural statement of the axiom.

First, the two easiest axioms:

- Extensionality is the sentence $\forall x \forall y \left((\forall a \left(a \in x \iff a \in y \right)) \implies x = y \right)$, which says that if two sets x and y have the same elements, then $x = y$.

- EmptySet is the sentence $\exists a : \forall x\ \neg(x \in a)$; it says there exists a set with no elements. By Extensionality this set is unique, so we denote it $\varnothing$.

The next two axioms give us basic ways of building new sets.

- Given two elements x and y, there exists a set a containing only those two elements. In machine code, this is the sentence Pairing, written
$$\forall x \forall y \exists a \quad \forall z,\ z \in a \iff ((z = x) \vee (z = y)).$$
By Extensionality this set a is unique, so we write $a = \{x, y\}$.
- Given a set a, we can create the union of the elements of a. For example, if $a = \{\{1,2\},\{3,4\}\}$, then $U = \{1,2,3,4\}$ is a set. Formally, this is the sentence Union:
$$\forall a \exists U \quad \forall x\ [(x \in U) \iff (\exists y : x \in y \in a)].$$
Since U is unique by Extensionality, we denote it $\cup a$.
- We can construct the **power set** $\mathcal{P}(x)$. Formally, the sentence PowerSet says that
$$\forall x \exists P \forall y (y \in P \iff y \subseteq x)$$
where $y \subseteq x$ is short for $\forall z (z \in y \implies z \in x)$. As Extensionality gives us uniqueness of P, we denote it $\mathcal{P}(x)$.
- Foundation says there are no infinite descending chains
$$x_0 \ni x_1 \ni x_2 \ni \dots.$$
This is important, because it lets us induct. In particular, **no set contains itself**.
- Infinity implies that $\omega = \{0, 1, \dots\}$ is a set.

These are all things you are already used to, so keep your intuition there. The next one is less intuitive:

- The **schema of restricted comprehension** says: if we are *given a set* X, and some formula $\phi(x)$ then we can *filter* through the elements of X to get a subset
$$Y = \{x \in X \mid \phi(x)\}.$$
Formally, given a formula ϕ:
$$\forall X \quad \exists Y \quad \forall y (y \in Y \iff y \in X \wedge \phi(y)).$$

Notice that we may *only* do this filtering over an already given set. So it is not valid to create $\{x \mid x \text{ is a set}\}$. We are thankful for this, because this lets us evade Cantor's paradox.

Abuse of Notation 89.4.1. Note that technically, there are infinitely many sentences, a Comprehension$_\phi$ for every possible formula ϕ. By abuse of notation, we let Comprehension abbreviate the infinitely many axioms Comprehension$_\phi$ for every ϕ.

There is one last schema called Replacement$_\phi$. Suppose X is a set and $\phi(x,y)$ is some formula such that for every $x \in X$, there is a *unique* y in the universe such that $\phi(x,y)$ is true: for example "$y = x \cup \{x\}$" works. (In effect, ϕ is defining a function f on X.) Then there exists a set Y consisting exactly of these images: (i.e. $f^{\text{img}}(X)$ is a set).

Abuse of Notation 89.4.2. By abuse of notation, we let Replacement abbreviate the infinitely many axioms Replacement$_\phi$ for every ϕ.

Remark 89.4.3 — What do we mean here that "for every $x \in X$, there is a *unique* y in the universe such that $\phi(x,y)$ is true"? How can we decide, given a formula ϕ, whether that statement is true, for Replacement$_\phi$ to be an axiom?

Turns out we cannot in general. But we don't need it! To circumvent the problem, for every $\phi(x,y)$, the axiom Replacement$_\phi$ states that

$$\text{“}\phi \text{ defines a function”} \implies \forall X \quad \exists Y \quad \text{“}Y = f^{\text{img}}(X)\text{”}.$$

In other words, the hypothesis that ϕ is a function is "folded in" the axiom Replacement$_\phi$ itself.

This will not really matter to us for now, but later on, it will matter in model theory, where we will state in Lemma 92.5.1 what it means for a model M to satisfy Replacement.

We postpone discussion of the Axiom of Choice momentarily.

§89.5 Encoding

Now that we have this rickety universe of sets, we can start re-building math. You'll get to see this more in the next chapter on ordinal numbers.

Definition 89.5.1. An **ordered pair** (x,y) is a set of the form

$$(x,y) \coloneqq \{\{x\}, \{x,y\}\}.$$

Note that $(x,y) = (a,b)$ if and only if $x = a$ and $y = b$. Ordered k-tuples can be defined recursively: a three-tuple (a,b,c) means $(a,(b,c))$.

Definition 89.5.2. A **function** $f\colon X \to Y$ is defined as a collection of ordered pairs such that

- If $(x,y) \in f$, then $x \in X$ and $y \in Y$.
- For every $x \in X$, there is a unique $y \in Y$ such that $(x,y) \in f$. We denote this y by $f(x)$.

Definition 89.5.3. The **natural numbers** are defined inductively as

$$\begin{aligned} 0 &= \varnothing \\ 1 &= \{0\} \\ 2 &= \{0,1\} \\ 3 &= \{0,1,2\} \\ &\vdots \end{aligned}$$

The set of all natural numbers is denoted ω.

Abuse of Notation 89.5.4. Yes, I'm sorry, in set theory 0 is considered a natural number. For this reason I'm using ω and not $\mathbb{N}$ since I explicitly have $0 \notin \mathbb{N}$ in all other parts of this book.

Et cetera, et cetera.

§89.6 Choice and well-ordering

The Axiom of Choice states that given a collection Y of nonempty sets, there is a function $g\colon Y \to \cup Y$ which "picks" an element of each member of Y. That means $g(y) \in y$ for every $y \in Y$. (The typical illustration is that Y contains infinitely many drawers, and each drawer (a y) has some sock in it.)

Formally, it is the sentence

$$\forall Y \left(\varnothing \notin Y \implies \exists g\colon Y \to \cup Y \text{ such that } \forall y \in Y \left(g(y) \in y\right).\right)$$

The tricky part is not that we can conceive of such a function g, but that in fact this function g is *actually a set*.

There is an equivalent formulation which is often useful.

Definition 89.6.1. A **well-ordering** $<$ of X is a strict, total order on X which has no infinite descending chains.

Well-orderings on a set are very nice, because we can pick minimal elements: this lets us do induction, for example.

Example 89.6.2 (Examples and non-examples of well-orderings)

(a) The natural numbers $\omega = \{0, 1, 2, \dots\}$ are well-ordered by $<$.

(b) The integers $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ are not well-ordered by $<$, because there are infinite descending chains (take $-1 > -2 > -3 > \dots$).

(c) The positive real numbers are not well-ordered by $<$, again because of the descending chain $\frac{1}{1} > \frac{1}{2} > \frac{1}{3} > \dots$.

(d) The positive integers are not well-ordered by the divisibility operation $|$. While there are no descending chains, there are elements which cannot be compared (for example $3 \nmid 5$, $5 \nmid 3$ and $3 \neq 5$).

Theorem 89.6.3 (Well-ordering theorem)

Assuming Choice, for every set we can place some well-ordering on it.

In fact, the well-ordering theorem is actually equivalent to the axiom of choice.

§89.7 Sets vs classes

Prototypical example for this section: The set of all sets is the standard example of a proper class.

We close the discussion of ZFC by mentioning "classes".

Roughly, the "bad thing" that happened was that we considered a set S, the "set of all sets", and it was *too big*. That is,

$$\{x \mid x \text{ is a set}\}$$

is not good. Similarly, we cannot construct a set

$$\{x \mid x \text{ is an ordered pair}\}.$$

The lesson of Cantor's Paradox is that we cannot create any sets we want; we have to be more careful than that.

Nonetheless, if we are given a set we can still tell whether or not it is an ordered pair. So for convenience, we will define a **class** to be a "concept" like the "class of all ordered pairs". Formally, a class is defined by some formula ϕ: it consists of the sets which satisfy the formula.

In particular:

Definition 89.7.1. The class of all sets is denoted V, defined by $V = \{x \mid x = x\}$. It is called the **von Neumann universe**.

A class is a **proper class** if it is not a set, so for example we have:

Theorem 89.7.2 (There is no set of all sets)

V is a proper class.

Proof. Assume not, and V is a set. Then $V \in V$, which violates Foundation. (In fact, V cannot be a set even without Foundation, as we saw earlier). □

Abuse of Notation 89.7.3. Given a class C, we will write $x \in C$ to mean that x has the defining property of C. For example, $x \in V$ means "x is a set".

It does not mean x is an element of V – this doesn't make sense as V is not a set.

§89.8 A few harder problems to think about

Problem 89A. Let A and B be sets. Show that $A \cap B$ and $A \times B$ are sets.

Problem 89B. Show that the class of all groups is a proper class. (You can take the definition of a group as a pair $(G, \cdot)$ where $\cdot$ is a function $G \times G \to G$.)

Problem 89C. Show that the axiom of choice follows from the well-ordering theorem.

Problem 89D†. Prove that actually, Replacement $\implies$ Comprehension.

Problem 89E (From Taiwan IMO training camp)**.** Consider infinitely many people each wearing a hat, which is either red, green, or blue. Each person can see the hat color of everyone except themselves. Simultaneously each person guesses the color of their hat. Show that they can form a strategy such that at most finitely many people guess their color incorrectly.

90 Ordinals

§90.1 Counting for preschoolers

In preschool, we were told to count as follows. We defined a set of symbols 1, 2, 3, 4, ... Then the teacher would hold up three apples and say:

> "One . . . two . . . three! There are three apples."

Image from [Ho]

The implicit definition is that the *last* number said is the final answer. This raises some obvious problems if we try to count infinite sets, but even in the finite world, this method of counting fails for the simplest set of all: how many apples are in the following picture?

Image from [Kr]

Answer: 0. There is nothing to say, and our method of counting has failed for the simplest set of all: the empty set.

§90.2 Counting for set theorists

Prototypical example for this section: $\omega + 1 = \{0, 1, 2, \dots, \omega\}$ *might work.*

Rather than using the *last* number listed, I propose instead starting with a list of symbols 0, 1, 2, ... and making the final answer the *first* number which was *not* said. Thus to count three apples, we would say

> "Zero . . . one . . . two! There are three apples."

We will call these numbers *ordinal numbers* (rigorous definition later). In particular, we'll *define* each ordinal to be the set of things we say:

$$\begin{aligned} 0 &= \varnothing \\ 1 &= \{0\} \\ 2 &= \{0, 1\} \\ 3 &= \{0, 1, 2\} \\ &\vdots \end{aligned}$$

In this way we can write out the natural numbers. You can have some fun with this, by saying things like

$$4 \coloneqq \{\{\}, \{\{\}\}, \{\{\}, \{\{\}\}\}, \{\{\}, \{\{\}\}, \{\{\}, \{\{\}\}\}\}\}.$$

In this way, we soon write down all the natural numbers. The next ordinal, ω,[1] is defined as

$$\omega = \{0, 1, 2, \dots\}$$

Then comes

$$\begin{aligned} \omega + 1 &= \{0, 1, 2, \dots, \omega\} \\ \omega + 2 &= \{0, 1, 2, \dots, \omega, \omega + 1\} \\ \omega + 3 &= \{0, 1, 2, \dots, \omega, \omega + 1, \omega + 2\} \\ &\vdots \end{aligned}$$

And in this way we define $\omega + n$, and eventually reach

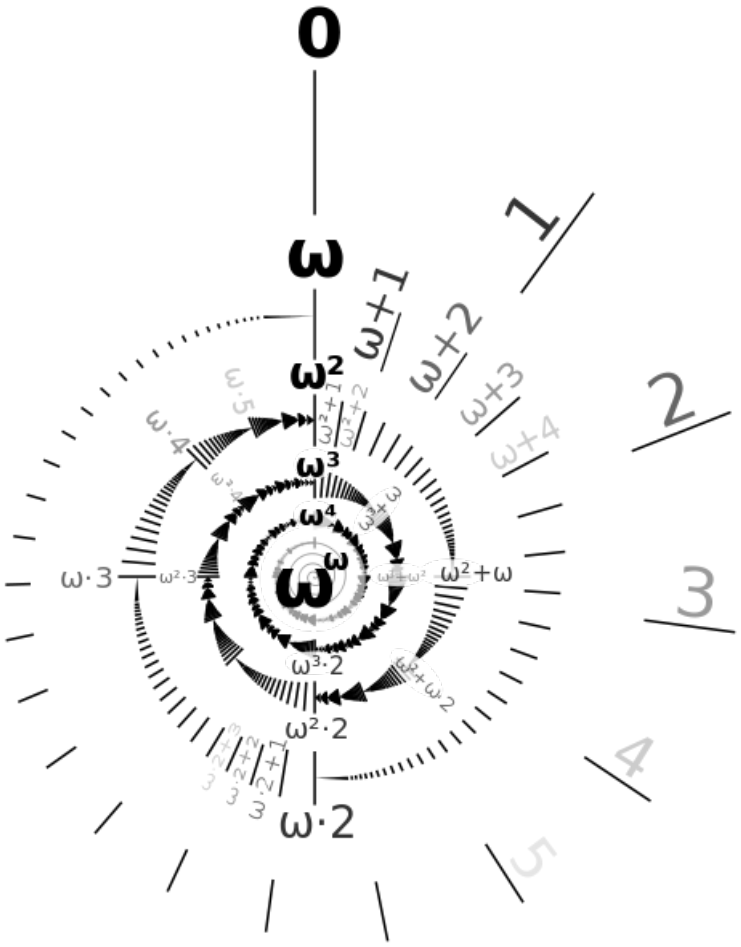
$$\begin{aligned} \omega \cdot 2 = \omega + \omega &= \{0, 1, 2 \dots, \omega, \omega + 1, \omega + 2, \dots\} \\ \omega \cdot 2 + 1 &= \{0, 1, 2 \dots, \omega, \omega + 1, \omega + 2, \dots, \omega \cdot 2\}. \end{aligned}$$

[1] As mentioned in the last chapter, it's not immediate that ω is a set; its existence is generally postulated by the Infinity axiom.

In this way we obtain

$$
\begin{aligned}
&0,\ 1,\ 2,\ 3,\ \ldots,\ \omega \\
&\quad \omega+1,\ \omega+2,\ \ldots,\ \omega+\omega \\
&\quad \omega\cdot 2+1,\ \omega\cdot 2+2,\ \ldots,\ \omega\cdot 3, \\
&\quad \vdots \\
&\quad \omega^2+1,\ \omega^2+2,\ \ldots \\
&\quad \vdots \\
&\quad \omega^3,\ \ldots,\ \omega^4,\ \ldots,\ \omega^\omega \\
&\quad \vdots \\
&\quad \omega^{\omega^{\omega^{\cdots}}}
\end{aligned}
$$

The first several ordinals can be illustrated in a nice spiral.



Remark 90.2.1 — You may think, well, why don't we define the ordinals like this

instead? It certainly looks shorter and simpler!

$$\begin{aligned} 0 &= \varnothing \\ 1 &= \{0\} \\ 2 &= \{1\} \\ 3 &= \{2\} \\ &\vdots \\ \omega &= \{0, 1, 2, \dots\} \\ \omega+1 &= \{\omega\} \\ &\vdots \end{aligned}$$

There are a few reasons why the usual definition is better.

- The "alternative definition" above is not uniform — an ordinal of the form $\alpha+1$ is defined by a singleton set, while the other ordinals are defined by a set that leads up to it.
- Comparison is simpler: for two ordinals α and β, $\alpha < \beta$ if and only if $\alpha \in \beta$.
- Cardinals will be simpler: the size of the set 5 is exactly 5.

Remark 90.2.2 (Digression) — The number $\omega^{\omega^{\omega^{\cdots}}}$ has a name, ε_0; it has the property that $\omega^{\varepsilon_0} = \varepsilon_0$. The reason for using "$\varepsilon$" (which is usually used to denote small quantities) is that, despite how huge it may appear, it is actually a countable set. More on that later.

§90.3 Definition of an ordinal

Our informal description of ordinals gives us a chain

$$0 \in 1 \in 2 \in \cdots \in \omega \in \omega+1 \in \dots.$$

To give the actual definition of an ordinal, I need to define two auxiliary terms first.

Definition 90.3.1. A set x is **transitive** if whenever $z \in y \in x$, we have $z \in x$ also.

Example 90.3.2 (7 is transitive)
The set 7 is transitive: for example, $2 \in 5 \in 7 \implies 2 \in 7$.

Question 90.3.3. Show that this is equivalent to: whenever $y \in x$, $y \subseteq x$.

Moreover, recall the definition of "well-ordering": a strict linear order with no infinite descending chains.

Example 90.3.4 ($\in$ is a well-ordering on $\omega \cdot 3$)

In $\omega \cdot 3$, we have an ordering

$$0 \in 1 \in 2 \in \cdots \in \omega \in \omega + 1 \in \cdots \in \omega \cdot 2 \in \omega \cdot 2 + 1 \in \ldots.$$

which has no infinite descending chains. Indeed, a typical descending chain might look like

$$\omega \cdot 2 + 6 \ni \omega \cdot 2 \ni \omega + 2015 \ni \omega + 3 \ni \omega \ni 1000 \ni 256 \ni 42 \ni 7 \ni 0.$$

Even though there are infinitely many elements, there is no way to make an infinite descending chain.

Exercise 90.3.5. (Important) Convince yourself there are no infinite descending chains of ordinals at all, without using the Foundation axiom.

Definition 90.3.6. An **ordinal** is a transitive set which is well-ordered by $\in$. The class of all ordinals is denoted On.

Question 90.3.7. Satisfy yourself that this definition works.

Example 90.3.8

- All of 0, 1, 2, $\ldots$, ω, $\omega + 1$, $\ldots$ defined above are ordinals.
- $\{3\}$ is not an ordinal — it's not transitive because $2 \in 3$, but $2 \notin \{3\}$.
- $\{0, 1, 2, \{0, 2\}\}$ is not an ordinal — the two elements 1 and $\{0, 2\}$ are not comparable.

We typically use Greek letters α, β, etc. for ordinal numbers.

Definition 90.3.9. We write

- $\alpha < \beta$ to mean $\alpha \in \beta$, and $\alpha > \beta$ to mean $\alpha \ni \beta$.
- $\alpha \le \beta$ to mean $\alpha \in \beta$ or $\alpha = \beta$, and $\alpha \ge \beta$ to mean $\alpha \ni \beta$ or $\alpha = \beta$,

Theorem 90.3.10 (Ordinals are strictly ordered)

Given any two ordinal numbers α and β, either $\alpha < \beta$, $\alpha = \beta$ or $\alpha > \beta$.

Proof. Surprisingly annoying, thus omitted. The key idea is that we can define $\min(\alpha, \beta) = \alpha \cap \beta$, then prove that this must be equal to either α or β. □

Theorem 90.3.11 (Ordinals represent all order types)

Suppose $<$ is a well-ordering on a set X. Then there exists a unique ordinal α such that there is a bijection $\alpha \to X$ which is order preserving.

Thus ordinals represent the possible *equivalence classes* of order types. Any time you have a well-ordered set, it is isomorphic to a unique ordinal.

We now formalize the "+1" operation we were doing:

Definition 90.3.12. Given an ordinal α, we let $\alpha+1 = \alpha \cup \{\alpha\}$. An ordinal of the form $\alpha+1$ is called a **successor ordinal**.

Definition 90.3.13. If λ is an ordinal which is neither zero nor a successor ordinal, then we say λ is a **limit ordinal**.

Example 90.3.14 (Successor and limit ordinals)
7, $\omega+3$, $\omega\cdot 2+2015$ are successor ordinals, but ω and $\omega\cdot 2$ are limit ordinals.

§90.4 Ordinals are "tall"

First, we note that:

Theorem 90.4.1 (There is no set of all ordinals)
On is a proper class.

Proof. Assume for contradiction not. Then On is well-ordered by $\in$ and transitive, so On is an ordinal, i.e. $\text{On} \in \text{On}$, which violates Foundation. □

Exercise 90.4.2 (Unimportant)**.** Give a proof without Foundation by considering $\text{On}+1$.

From this we deduce:

Theorem 90.4.3 (Sets of ordinals are bounded)
Let $A \subseteq \text{On}$. Then there is some ordinal α such that $A \subseteq \alpha$ (i.e. A must be bounded).

Proof. Otherwise, look at $\bigcup A$. It is a set. But if A is unbounded it must equal On, which is a contradiction. □

In light of this, every set of ordinals has a **supremum**, which is the least upper bound. We denote this by $\sup A$.

Question 90.4.4. Show that

(a) $\sup(\alpha+1) = \alpha$ for any ordinal α.

(b) $\sup \lambda = \lambda$ for any limit ordinal λ.

The pictorial "tall" will be explained in a few sections.

§90.5 Transfinite induction and recursion

The fact that $\in$ has no infinite descending chains means that induction and recursion still work verbatim.

Theorem 90.5.1 (Transfinite induction)
Given a statement $P(-)$, suppose that

- $P(0)$ is true, and
- If $P(\alpha)$ is true for all $\alpha < \beta$, then $P(\beta)$ is true.

Then $P(\alpha)$ is true for every ordinal α.

Theorem 90.5.2 (Transfinite recursion)
To define a sequence x_α for every ordinal α, it suffices to

- define x_0, then
- for any β, define x_β using only x_α for any $\alpha < \beta$.

The difference between this and normal induction lies in the *limit ordinals*. In real life, we might only do things like "define $x_{n+1} = \dots$". But this is not enough to define x_α for all α, because we can't hit ω this way. Similarly, the simple $+1$ doesn't let us hit the ordinal $\omega \cdot 2$, even if we already have $\omega + n$ for all n. In other words, simply incrementing by 1 cannot get us past limit stages, but using transfinite induction to jump upwards lets us sidestep this issue.

So a transfinite induction is often broken up into three cases. In the induction phrasing, it looks like

- (Zero Case) First, resolve $P(0)$.
- (Successor Case) Show that from $P(\alpha)$ we can get $P(\alpha + 1)$.
- (Limit Case) For λ a limit ordinal, show that $P(\lambda)$ holds given $P(\alpha)$ for all $\alpha < \lambda$.

Similarly, transfinite recursion is often split into cases too.

- (Zero Case) First, define x_0.
- (Successor Case) Define $x_{\alpha+1}$ from x_α.
- (Limit Case) Define x_λ from x_α for all $\alpha < \lambda$, where λ is a limit ordinal.

In both situations, finite induction only does the first two cases, but if we're able to do the third case we can climb above the barrier ω.

§90.6 Ordinal arithmetic

Prototypical example for this section: $1 + \omega = \omega \neq \omega + 1$.

To give an example of transfinite recursion, let's define addition of ordinals. Recall that we defined $\alpha + 1 = \alpha \cup \{\alpha\}$. By transfinite recursion, let

$$\begin{aligned} \alpha + 0 &= \alpha \\ \alpha + (\beta + 1) &= (\alpha + \beta) + 1 \\ \alpha + \lambda &= \bigcup_{\beta < \lambda} (\alpha + \beta). \end{aligned}$$

Here $\lambda \neq 0$.

We can also do this explicitly: The picture is to just line up α before β. That is, we can consider the set

$$X = (\{0\} \times \alpha) \cup (\{1\} \times \beta)$$

(i.e. we tag each element of α with a 0, and each element of β with a 1). We then impose a well-ordering on X by a lexicographic ordering $<_{\text{lex}}$ (sort by first component, then by second). This well-ordering is isomorphic to a unique ordinal.

Example 90.6.1 $(2 + 3 = 5)$

Under the explicit construction for $\alpha = 2$ and $\beta = 3$, we get the set

$$X = \{(0,0) < (0,1) < (1,0) < (1,1) < (1,2)\}$$

which is isomorphic to 5.

Example 90.6.2 (Ordinal arithmetic is not commutative)

Note that $1 + \omega = \omega$! Indeed, under the transfinite definition, we have

$$1 + \omega = \cup_n (1 + n) = 2 \cup 3 \cup 4 \cup \dots = \omega.$$

With the explicit construction, we have

$$X = \{(0,0) < (1,0) < (1,1) < (1,2) < \dots\}$$

which is isomorphic to ω.

Exercise 90.6.3. Show that $n + \omega = \omega$ for any $n \in \omega$.

Remark 90.6.4 — Ordinal addition is not commutative. However, from the explicit construction we can see that it is at least associative.

Furthermore, you can see that for small enough $\alpha \neq 0$, then $\alpha + \beta = \beta$ may happen; however, this does not happen on the other side — if $\beta < \gamma$, then $\alpha + \beta < \alpha + \gamma$.

Similarly, we can define multiplication in two ways. By transfinite induction:

$$\begin{aligned} \alpha \cdot 0 &= 0 \\ \alpha \cdot (\beta + 1) &= (\alpha \cdot \beta) + \alpha \\ \alpha \cdot \lambda &= \bigcup_{\beta < \lambda} \alpha \cdot \beta. \end{aligned}$$

We can also do an explicit construction: This time, the picture is to line up β copies, each copy contains α items. That is, $\alpha \cdot \beta$ is the order type of

$$<_{\text{lex}} \text{ applied to } \beta \times \alpha.$$

Example 90.6.5 (Ordinal multiplication is not commutative)
We have $\omega \cdot 2 = \omega + \omega$, but $2 \cdot \omega = \omega$.

Exercise 90.6.6. Prove this.

Exercise 90.6.7. Verify that ordinal multiplication (like addition) is associative but not commutative. (Look at $\gamma \times \beta \times \alpha$.)

Similar to ordinal addition, defining $\alpha \cdot (\beta + 1) = (\alpha \cdot \beta) + \alpha$ makes sure that if $\beta < \gamma$ then $\alpha \cdot \beta < \alpha \cdot \gamma$ — as long as $\alpha > 0$.

Exponentiation can also be so defined, though the explicit construction is less natural — since we will not use this definition in the rest of the book, you may ignore it.

For $\alpha = 0$, define $0^0 = 1$ and $0^\beta = 0$ for all $\beta > 0$. Otherwise:

$$\begin{aligned} \alpha^0 &= 1 \\ \alpha^{\beta+1} &= \alpha^\beta \cdot \alpha \\ \alpha^\lambda &= \bigcup_{\beta<\lambda} \alpha^\beta. \end{aligned}$$

Exercise 90.6.8. Verify that $2^\omega = \omega$.

§90.7 The hierarchy of sets

We now define the **von Neumann Hierarchy** by transfinite recursion.

Definition 90.7.1. By transfinite recursion, we set

$$\begin{aligned} V_0 &= \varnothing \\ V_{\alpha+1} &= \mathcal{P}(V_\alpha) \\ V_\lambda &= \bigcup_{\alpha<\lambda} V_\alpha \end{aligned}$$

By transfinite induction, we see V_α is transitive and that $V_\alpha \subseteq V_\beta$ for all $\alpha < \beta$.

Example 90.7.2 (V_α for $\alpha \leq 3$)
The first few levels of the hierarchy are:

$$\begin{aligned} V_0 &= \varnothing \\ V_1 &= \{0\} \\ V_2 &= \{0, 1\} \\ V_3 &= \{0, 1, 2, \{1\}\}. \end{aligned}$$

Notice that for each n, V_n consists of only finite sets, and each n appears in V_{n+1} for the first time. Observe that

$$V_\omega = \bigcup_{n\in\omega} V_n$$

consists only of finite sets; thus ω appears for the first time in $V_{\omega+1}$.

Question 90.7.3. How many sets are in V_5?

Definition 90.7.4. The **rank** of a set y, denoted $\text{rank}(y)$, is the smallest ordinal α such that $y \in V_{\alpha+1}$.

Example 90.7.5

$\text{rank}(2) = 2$, and actually $\text{rank}(\alpha) = \alpha$ for any ordinal α (problem later). This is the reason for the extra "+1".

Question 90.7.6. Show that $\text{rank}(y)$ is the smallest ordinal α such that $y \subseteq V_\alpha$.

It's not yet clear that the rank of a set actually exists, so we prove:

Theorem 90.7.7 (The von Neumann hierarchy is complete)

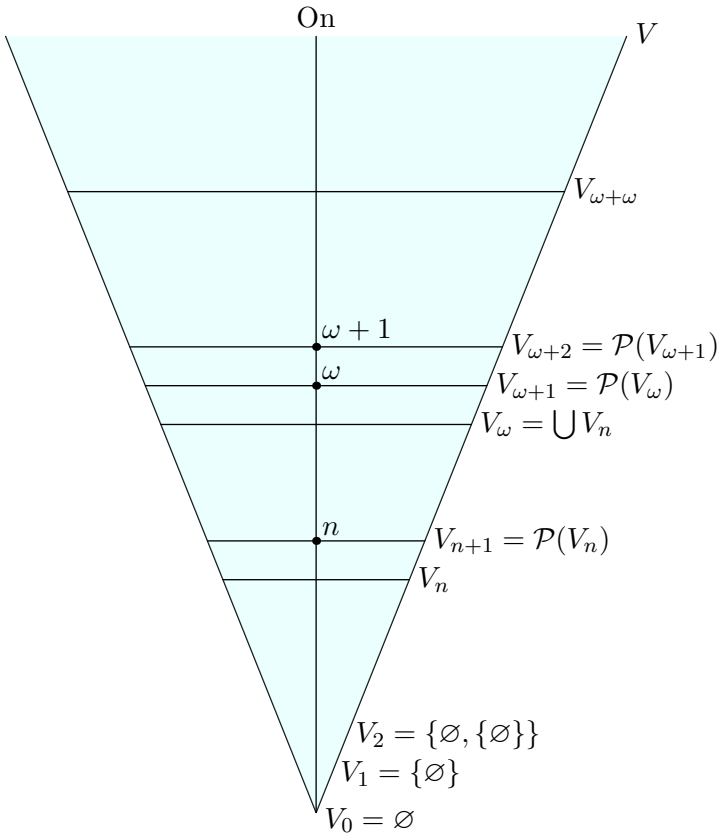
The class V is equal to $\bigcup_{\alpha \in \text{On}} V_\alpha$. In other words, every set appears in some V_α.

Proof. Assume for contradiction this is false. The key is that because $\in$ satisfies Foundation, we can take a $\in$-minimal counterexample x. Thus $\text{rank}(y)$ is defined for every $y \in x$, and we can consider (by Replacement) the set

$$\{\text{rank}(y) \mid y \in x\}.$$

Since it is a set of ordinals, it is bounded. So there is some large ordinal α such that $y \in V_\alpha$ for all $y \in x$, i.e. $x \subseteq V_\alpha$, so $x \in V_{\alpha+1}$. □

This leads us to a picture of the universe V:



We can imagine the universe V as a triangle, built in several stages or layers, $V_0 \subsetneq V_1 \subsetneq V_2 \subsetneq \dots$. This universe doesn't have a top: but each of the V_i do. However, the universe has a very clear bottom. Each stage is substantially wider than the previous one.

In the center of this universe are the ordinals: for every successor V_α, exactly one new ordinal appears, namely α. Thus we can picture the class of ordinals as a thin line that stretches the entire height of the universe. A set has rank α if it appears at the same stage that α does.

All of number theory, the study of the integers, lives inside V_ω. Real analysis, the study of real numbers, lives inside $V_{\omega+1}$, since a real number can be encoded as a subset of $\mathbb{N}$ (by binary expansion). Functional analysis lives one step past that, $V_{\omega+2}$. For all intents and purposes, most mathematics does not go beyond $V_{\omega+\omega}$. This pales in comparison to the true magnitude of the whole universe.

§90.8 A few harder problems to think about

Problem 90A. Prove that $\operatorname{rank}(\alpha) = \alpha$ for any α by transfinite induction.

Problem 90B (Online Math Open)**.** Count the number of transitive sets in V_5.

Problem 90C (Goodstein)**.** Let a_2 be any positive integer. We define the infinite sequence a_2, a_3, ... recursively as follows. If $a_n = 0$, then $a_{n+1} = 0$. Otherwise, we write a_n in base n, then write all exponents in base n, and so on until all numbers in the expression are at most n. Then we replace all instances of n by $n+1$ (including the

exponents!), subtract 1, and set the result to a_{n+1}. For example, if $a_2 = 11$ we have

$$\begin{aligned}
a_2 &= 2^3 + 2 + 1 = 2^{2+1} + 2 + 1 \\
a_3 &= 3^{3+1} + 3 + 1 - 1 = 3^{3+1} + 3 \\
a_4 &= 4^{4+1} + 4 - 1 = 4^{4+1} + 3 \\
a_5 &= 5^{5+1} + 3 - 1 = 5^{5+1} + 2
\end{aligned}$$

and so on. Prove that $a_N = 0$ for some integer $N > 2$.

91 Cardinals

An ordinal measures a total ordering. However, it does not do a fantastic job at measuring size. For example, there is a bijection between the elements of ω and $\omega+1$:

$$\begin{aligned} \omega+1 &= \{\ \omega\ 0\ 1\ 2\ \dots\ \} \\ \omega &= \{\ 0\ 1\ 2\ 3\ \dots\ \}. \end{aligned}$$

In fact, as you likely already know, there is even a bijection between ω and ω^2:

$+$	0	1	2	3	4	$\dots$
0	0	1	3	6	10	$\dots$
ω	2	4	7	11	$\dots$	
$\omega\cdot 2$	5	8	12	$\dots$		
$\omega\cdot 3$	9	13	$\dots$			
$\omega\cdot 4$	14	$\dots$				

So ordinals do not do a good job of keeping track of size. For this, we turn to the notion of a cardinal number.

§91.1 Equinumerous sets and cardinals

Definition 91.1.1. Two sets A and B are **equinumerous**, written $A \approx B$, if there is a bijection between them.

Definition 91.1.2. A **cardinal** is an ordinal κ such that for no $\alpha < \kappa$ do we have $\alpha \approx \kappa$.

Example 91.1.3 (Examples of cardinals)
Every finite number is a cardinal. Moreover, ω is a cardinal. However, $\omega+1$, ω^2, ω^{2015} are not, because they are countable.

Example 91.1.4 (ω^ω is countable)
Even ω^ω is not a cardinal, since it is a countable union

$$\omega^\omega = \bigcup_n \omega^n$$

and each ω^n is countable.

Question 91.1.5. Why must an infinite cardinal be a limit ordinal?

Remark 91.1.6 — There is something fishy about the definition of a cardinal: it relies on an *external* function f. That is, to verify κ is a cardinal I can't just look at κ itself; I need to examine the entire universe V to make sure there does not exist a bijection $f\colon \kappa \to \alpha$ for $\alpha < \kappa$. For now this is no issue, but later in model theory

this will lead to some highly counterintuitive behavior.

§91.2 Cardinalities

Now that we have defined a cardinal, we can discuss the size of a set by linking it to a cardinal.

Definition 91.2.1. The **cardinality** of a set X is the *least* ordinal κ such that $X \approx \kappa$. We denote it by $|X|$.

Question 91.2.2. Why must $|X|$ be a cardinal?

Remark 91.2.3 — One needs the well-ordering theorem (equivalently, choice) in order to establish that such an ordinal κ actually exists.

Since cardinals are ordinals, it makes sense to ask whether $\kappa_1 \le \kappa_2$, and so on. Our usual intuition works well here.

Proposition 91.2.4 (Restatement of cardinality properties)
Let X and Y be sets.

(i) $X \approx Y$ if and only if $|X| = |Y|$, if and only if there's a bijection from X to Y.

(ii) $|X| \le |Y|$ if and only if there is an injective map $X \hookrightarrow Y$.

Diligent readers are invited to try and prove this.

§91.3 Aleph numbers

Prototypical example for this section: $\aleph_0 = \omega$, *and* $\aleph_1$ *is the first uncountable ordinal.*

First, let us check that cardinals can get arbitrarily large:

Proposition 91.3.1
We have $|X| < |\mathcal{P}(X)|$ for every set X.

Proof. There is an injective map $X \hookrightarrow \mathcal{P}(X)$ but there is no injective map $\mathcal{P}(X) \hookrightarrow X$ by Lemma 89.2.1. □

Thus we can define:

Definition 91.3.2. For a cardinal κ, we define κ^+ to be the least cardinal above κ, called the **successor cardinal**.

This κ^+ exists and has $\kappa^+ \le |\mathcal{P}(\kappa)|$.
Next, we claim that:

Exercise 91.3.3. Show that if A is a set of cardinals, then $\cup A$ is a cardinal.

Thus by transfinite induction we obtain that:

Definition 91.3.4. For any $\alpha \in \mathrm{On}$, we define the **aleph numbers** as

$$\begin{aligned} \aleph_0 &= \omega \\ \aleph_{\alpha+1} &= (\aleph_\alpha)^+ \\ \aleph_\lambda &= \bigcup_{\alpha<\lambda} \aleph_\alpha. \end{aligned}$$

Thus we have the sequence of cardinals

$$0 < 1 < 2 < \cdots < \aleph_0 < \aleph_1 < \cdots < \aleph_\omega < \aleph_{\omega+1} < \ldots.$$

By definition, $\aleph_0$ is the cardinality of the natural numbers, $\aleph_1$ is the first uncountable ordinal, ...

We claim the aleph numbers constitute all the cardinals:

Lemma 91.3.5 (Aleph numbers constitute all infinite cardinals)
If κ is a cardinal then either κ is finite (i.e. $\kappa \in \omega$) or $\kappa = \aleph_\alpha$ for some $\alpha \in \mathrm{On}$.

Proof. Assume κ is infinite, and take α minimal with $\aleph_\alpha \geq \kappa$. Suppose for contradiction that we have $\aleph_\alpha > \kappa$. We may assume $\alpha > 0$, since the case $\alpha = 0$ is trivial.

If $\alpha = \overline{\alpha} + 1$ is a successor, then

$$\aleph_{\overline{\alpha}} < \kappa < \aleph_\alpha = (\aleph_{\overline{\alpha}})^+$$

which contradicts the definition of the successor cardinal.

If $\alpha = \lambda$ is a limit ordinal, then $\aleph_\lambda$ is the supremum $\bigcup_{\gamma<\lambda} \aleph_\gamma$. So there must be some $\gamma < \lambda$ with $\aleph_\gamma > \kappa$, which contradicts the minimality of α. □

Definition 91.3.6. An infinite cardinal which is not a successor cardinal is called a **limit cardinal**. It is exactly those cardinals of the form $\aleph_\lambda$, for λ a limit ordinal, plus $\aleph_0$.

§91.4 Cardinal arithmetic

Prototypical example for this section: $\aleph_0 \cdot \aleph_0 = \aleph_0 + \aleph_0 = \aleph_0$

Recall the way we set up ordinal arithmetic. Note that in particular, $\omega + \omega > \omega$ and $\omega^2 > \omega$. Since cardinals count size, this property is undesirable, and we want to have

$$\begin{aligned} \aleph_0 + \aleph_0 &= \aleph_0 \\ \aleph_0 \cdot \aleph_0 &= \aleph_0 \end{aligned}$$

because $\omega + \omega$ and $\omega \cdot \omega$ are countable. In the case of cardinals, we simply "ignore order".

The definition of cardinal arithmetic is as expected:

Definition 91.4.1 (Cardinal arithmetic)**.** Given cardinals κ and μ, define

$$\kappa + \mu := |(\{0\} \times \kappa) \cup (\{1\} \times \mu)|$$

and

$$\kappa \cdot \mu := |\mu \times \kappa|.$$

Question 91.4.2. Check this agrees with what you learned in pre-school for finite cardinals.

Abuse of Notation 91.4.3. This is a slight abuse of notation since we are using the same symbols as for ordinal arithmetic, even though the results are different ($\omega \cdot \omega = \omega^2$ but $\aleph_0 \cdot \aleph_0 = \aleph_0$). In general, I'll make it abundantly clear whether I am talking about cardinal arithmetic or ordinal arithmetic.

To help combat this confusion, we use separate symbols for ordinals and cardinals. Specifically, ω will always refer to $\{0, 1, \dots\}$ viewed as an ordinal; $\aleph_0$ will always refer to the same set viewed as a cardinal. More generally,

Definition 91.4.4. Let $\omega_\alpha = \aleph_\alpha$ viewed as an ordinal.

However, as we've seen already we have that $\aleph_0 \cdot \aleph_0 = \aleph_0$. In fact, this holds even more generally:

Theorem 91.4.5 (Infinite cardinals squared)
Let κ be an infinite cardinal. Then $\kappa \cdot \kappa = \kappa$.

Proof. Obviously $\kappa \cdot \kappa \geq \kappa$, so we want to show $\kappa \cdot \kappa \leq \kappa$.

The idea is to try to repeat the same proof that we had for $\aleph_0 \cdot \aleph_0 = \aleph_0$, so we re-iterate it here. We took the "square" of elements of $\aleph_0$, and then *re-ordered* it according to the diagonal:

	0	1	2	3	4	...
0	0	1	3	6	10	...
1	2	4	7	11	...	
2	5	8	12	...		
3	9	13	...			
4	14	...				

We'd like to copy this idea for a general κ; however, since addition is less well-behaved for infinite ordinals it will be more convenient to use $\max\{\alpha, \beta\}$ rather than $\alpha + \beta$. Specifically, we put the ordering $<_{\max}$ on $\kappa \times \kappa$ as follows: for (α_1, β_1) and (α_2, β_2) in $\kappa \times \kappa$ we declare $(\alpha_1, \beta_1) <_{\max} (\alpha_2, \beta_2)$ if

- $\max\{\alpha_1, \beta_1\} < \max\{\alpha_2, \beta_2\}$ or
- $\max\{\alpha_1, \beta_1\} = \max\{\alpha_2, \beta_2\}$ and (α_1, β_1) is lexicographically earlier than (α_2, β_2).

This alternate ordering (which deliberately avoids referring to the addition) looks like:

	0	1	2	3	4	...
0	0	1	4	9	16	...
1	2	3	5	10	17	...
2	6	7	8	11	18	...
3	12	13	14	15	19	...
4	20	21	22	23	24	...
⋮	⋮	⋮	⋮	⋮	⋮	⋱

Now we proceed by transfinite induction on κ. The base case is $\kappa = \aleph_0$, done above. Now, $<_{\max}$ is a well-ordering of $\kappa \times \kappa$, so we know it is in order-preserving bijection with some ordinal γ. Our goal is to show that $|\gamma| \leq \kappa$. To do so, it suffices to prove that for any $\overline{\gamma} \in \gamma$, we have $|\overline{\gamma}| < \kappa$.

Suppose $\overline{\gamma}$ corresponds to the point $(\alpha, \beta) \in \kappa \times \kappa$ under this bijection. If α and β are both finite then certainly $\overline{\gamma}$ is finite too. Otherwise, let $\overline{\kappa} = \max\{\alpha, \beta\} < \kappa$; then the number of points below $\overline{\gamma}$ is at most

$$|\alpha| \cdot |\beta| \le \overline{\kappa} \cdot \overline{\kappa} = \overline{\kappa}$$

by the inductive hypothesis. So $|\overline{\gamma}| \le \overline{\kappa} < \kappa$ as desired. □

From this it follows that cardinal addition and multiplication is really boring:

Theorem 91.4.6 (Infinite cardinal arithmetic is trivial)
Given cardinals κ and μ, one of which is infinite, we have

$$\kappa \cdot \mu = \kappa + \mu = \max\{\kappa, \mu\}.$$

Proof. The point is that both of these are less than the square of the maximum. Writing out the details:

$$\begin{aligned} \max\{\kappa, \mu\} &\le \kappa + \mu \\ &\le \kappa \cdot \mu \\ &\le \max\{\kappa, \mu\} \cdot \max\{\kappa, \mu\} \\ &= \max\{\kappa, \mu\}. \end{aligned}$$

□

§91.5 Cardinal exponentiation

Prototypical example for this section: $2^\kappa = |\mathcal{P}(\kappa)|$.

Definition 91.5.1. Suppose κ and λ are cardinals. Then

$$\kappa^\lambda := |\mathscr{F}(\lambda, \kappa)|.$$

Here $\mathscr{F}(A, B)$ is the set of functions from A to B.

Abuse of Notation 91.5.2. As before, we are using the same notation for both cardinal and ordinal arithmetic. Sorry!

In particular, $2^\kappa = |\mathcal{P}(\kappa)| > \kappa$, and so from now on we can use the notation 2^κ freely. (Note that this is totally different from ordinal arithmetic; there we had $2^\omega = \bigcup_{n\in\omega} 2^n = \omega$. In cardinal arithmetic $2^{\aleph_0} > \aleph_0$.)

I have unfortunately not told you what $2^{\aleph_0}$ equals. A natural conjecture is that $2^{\aleph_0} = \aleph_1$; this is called the **Continuum Hypothesis**. It turns out that this is *undecidable* – it is not possible to prove or disprove this from the ZFC axioms.

§91.6 Cofinality

Prototypical example for this section: $\aleph_0$, $\aleph_1$, ... *are all regular, but* $\aleph_\omega$ *has cofinality* ω.

Definition 91.6.1. Let λ be an ordinal (usually a limit ordinal), and α another ordinal. A map $f\colon \alpha \to \lambda$ of ordinals is called **cofinal** if for every $\overline{\lambda} < \lambda$, there is some $\overline{\alpha} \in \alpha$ such that $f(\overline{\alpha}) \ge \overline{\lambda}$. In other words, the map reaches arbitrarily high into λ.

Example 91.6.2 (Example of a cofinal map)

(a) The map $\omega \to \omega^\omega$ by $n \mapsto \omega^n$ is cofinal.

(b) For any ordinal α, the identity map $\alpha \to \alpha$ is cofinal.

Definition 91.6.3. Let λ be a limit ordinal. The **cofinality** of λ, denoted $\operatorname{cof}(\lambda)$, is the smallest ordinal α such that there is a cofinal map $\alpha \to \lambda$.

Question 91.6.4. Why must α be an infinite cardinal?

Usually, we are interested in taking the cofinality of a cardinal κ.

Pictorially, you can imagine standing at the bottom of the universe and looking up the chain of ordinals to κ. You have a machine gun and are firing bullets upwards, and you want to get arbitrarily high but less than κ. The cofinality is then the number of bullets you need to do this.

We now observe that "most" of the time, the cofinality of a cardinal is itself.[1] Such a cardinal is called **regular**.

Example 91.6.5 ($\aleph_0$ is regular)

$\operatorname{cof}(\aleph_0) = \aleph_0$, because no finite subset of $\aleph_0 = \omega$ can reach arbitrarily high.

Example 91.6.6 ($\aleph_1$ is regular)

$\operatorname{cof}(\aleph_1) = \aleph_1$. Indeed, assume for contradiction that some countable set of ordinals $A = \{\alpha_0, \alpha_1, \dots\} \subseteq \aleph_1$ reaches arbitrarily high inside $\aleph_1$. Then $\Lambda = \cup A$ is a *countable* ordinal, because it is a countable union of countable ordinals. In other words $\Lambda \in \aleph_1$. But Λ is an upper bound for A, contradiction.

On the other hand, there *are* cardinals which are not regular; since these are the "rare" cases we call them **singular**.

Example 91.6.7 ($\aleph_\omega$ is not regular)

Notice that $\aleph_0 < \aleph_1 < \aleph_2 < \dots$ reaches arbitrarily high in $\aleph_\omega$, despite only having $\aleph_0$ terms. It follows that $\operatorname{cof}(\aleph_\omega) = \aleph_0$.

We now confirm a suspicion you may have:

Theorem 91.6.8 (Successor cardinals are regular)

If $\kappa = \overline{\kappa}^+$ is a successor cardinal, then it is regular.

Proof. We copy the proof that $\aleph_1$ was regular.

Assume for contradiction that for some $\mu \le \overline{\kappa}$, there are μ sets reaching arbitrarily high in κ as a cardinal. Observe that each of these sets must have cardinality at most $\overline{\kappa}$. We take the union of all μ sets, which gives an ordinal Λ serving as an upper bound.

[1] Be careful — the cofinality of an *ordinal* is usually strictly less than itself. In fact, if the cofinality of an ordinal is itself, then that ordinal must be a cardinal.

The number of elements in the union is at most

$$\#\text{sets} \cdot \#\text{elms} \le \mu \cdot \overline{\kappa} = \overline{\kappa}$$

and hence $|\Lambda| \le \overline{\kappa} < \kappa$. □

§91.7 Inaccessible cardinals

So, what about limit cardinals? It seems that most of them are singular: if $\aleph_\lambda \neq \aleph_0$ is a limit cardinal (that is, λ is a limit ordinal), then the sequence $\{\aleph_\alpha\}_{\alpha\in\lambda}$ (of length λ) is certainly cofinal.

Example 91.7.1 (Beth fixed point)
Consider the monstrous cardinal

$$\kappa = \aleph_{\aleph_{\aleph_{\iddots}}}.$$

This might look frighteningly huge, as $\kappa = \aleph_\kappa$, but its cofinality is ω as it is the limit of the sequence

$$\aleph_0, \aleph_{\aleph_0}, \aleph_{\aleph_{\aleph_0}}, \dots$$

More generally, one can in fact prove that

$$\operatorname{cof}(\aleph_\lambda) = \operatorname{cof}(\lambda).$$

But it is actually conceivable that λ is so large that $\lambda = \aleph_\lambda$.

A regular limit cardinal other than $\aleph_0$ has a special name: it is **weakly inaccessible**. Such cardinals are so large that it is impossible to prove or disprove their existence in ZFC. It is the first of many so-called "large cardinals".

An infinite cardinal κ is a **strong limit cardinal** if

$$\forall \overline{\kappa} < \kappa \quad 2^{\overline{\kappa}} < \kappa$$

for any cardinal $\overline{\kappa}$. For example, $\aleph_0$ is a strong limit cardinal.

Question 91.7.2. Why must strong limit cardinals actually be limit cardinals? (This is offensively easy.)

Remark 91.7.3 — A limit cardinal can equivalently be defined as a nonzero cardinal κ such that

$$\forall \overline{\kappa} < \kappa \quad (\overline{\kappa})^+ < \kappa.$$

If you compare it with the definition of strong limit cardinals, you can see the parallel. (This remark also gives an answer to the previous question.)

A regular strong limit cardinal other than $\aleph_0$ is called **strongly inaccessible**.

§91.8 A few harder problems to think about

Problem 91A. Compute $|V_\omega|$.

Problem 91B. Prove that for any limit ordinal α, $\text{cof}(\alpha)$ is a *regular* cardinal.

Problem 91C* (Strongly inaccessible cardinals)**.** Show that for any strongly inaccessible κ, we have $|V_\kappa| = \kappa$.

Problem 91D (König's theorem)**.** Show that

$$\kappa^{\text{cof}(\kappa)} > \kappa$$

for every infinite cardinal κ.