PUMaC 2025 Power Round: The Continuum Hypothesis

Zongshu Wu

Fall 2025

Rules and Reminders

1. Your solutions should be turned in by **5pm EST on Thursday, November 20th**. You will submit the solutions through Gradescope. The instructions describing how to log into Gradescope will be sent to the coaches. The deadline for submission is clearly visible on the Gradescope site once you enroll in the course.

   Please make sure you submit your work in time, as no late submissions will be accepted. Please do not submit your work using email or in any other way. If you have questions about Gradescope, please post them on Piazza.

   You may either typeset the solutions in LaTeX or write them by hand. We strongly encourage you to typeset the solutions. This way, the proofs will often be more clear, and you will be less likely to lose points. You might want to use the Solutions Template we posted, or some of the LaTeX resources listed in point 2.

   In case your solutions are handwritten, then the cover sheet (the last page of this document) should be the first page of your submission.

   Each page should have on it the **team number** (not team name) and **problem number**. This number can be found by logging in to the coach portal and selecting the corresponding team. Solutions to problems may span multiple pages. Please put them in order when submitting your solutions.

2. You may resubmit several times before the due date, but **only your final submission will be graded** (and you may not submit any work after the deadline). The last version of the Power Round solutions that we receive from your team will be graded. Moreover, you must submit a PDF. No other file type will be graded. For those new and interested in LaTeX, check out Overleaf as well as its online guides. If you do not know the specific command for a math symbol, check out Detexify or TeX.StackExchange.

3. Do not submit identifying information aside from your team number.

4. When submitting to Gradescope, assign the solutions to the correct problems on the Gradescope submission outline. Failure to do this WILL result in a point deduction, as it creates a ton of extra work for us on the back-end.

5. On any problem, you may use without proof any result that is stated earlier in the test, as well as any problem from earlier in the test, even if it is a problem that

your team has not solved. These are the only results you may use. In particular, to solve a problem, you may not cite the subsequent ones.

The problems are graded separately, so you may not cite parts of your proof of other problems. If you wish to use a lemma in multiple problems, please reproduce the statement and proof in each problem.

6. When a problem asks you to "find", "show", or "prove" a result, a formal proof is expected, in which you justify each step you take, either by using a method from earlier or by proving that everything you do is correct. When a problem asks you to "explain", an informal explanation suffices.

7. All problems are numbered as "Problem $x.y.z$", where $x.y$ is the subsection number, and z is the the number of the problem within the subsection. Each problem's point value is stated on the problem, and can also be found on the cover sheet.

8. Teams whose members use English as a foreign language may use dictionaries for reference.

9. **You may NOT use any references, such as books or electronic resources, except those specified in points 2 and 8. You may NOT use computer programs, calculators, AI chatbots, or any other computational aids.**

10. You may ask questions about the test on our Piazza forum. On the forum, you may ask a public or private question. If you ask a public question, all other teams will be able to see it. Therefore, **if a public question reveals all or part of your solution to a Power Round question, your team's Power Round score will be penalized severely**. If your question might reveal aspects of your solution, please ask it as a private question. On the other hand, if you are sure that your question does not spoil anything, then we encourage you to make your question public, so that everybody can see it.

    We will post important clarifications on Piazza, and these clarifications will also be emailed to coaches.

11. **Communication with humans outside your team of 8 students about the content of these problems is prohibited.** Of course, asking questions on Piazza is the exception, and is allowed.

Introduction and Advice

In this Power Round, we will dive into the world of **axiomatic set theory**, which is the rigorous foundation for all of mathematics. We will ask and answer fundamental questions, such as "what is a set?" If you think this question is trivial, it is *not*: if you're not careful, you run into all kinds of logical paradoxes.

Building on the foundations, we will investigate the famous Continuum Hypothesis, which essentially asks: how large is the set $\mathbb{R}$ of real numbers? The answer turns out to be quite surprising: there is no way for us to know for sure what it is!

A large part of the difficulty in this Power Round will arise from the rigor required when working with the formal concepts. Since we need to put everything on completely solid footing, things that might seem obvious can often be quite nontrivial to prove. So, it is important to make sure that your logic is airtight.

Here is some further advice with regard to the Power Round:

- **Read the text of every problem!** Many important ideas are included in the problems and may be referenced later on. In addition, some of the theorems you are asked to prove are useful or even necessary for later problems. Even if you don't solve a problem, you can assume its results for future problems.
- **Make sure you understand the definitions!** A lot of the definitions are not easy to grasp; don't worry if it takes you a while to fully understand them. If you don't, then you will not be able to do the problems. Feel free to ask clarifying questions about the definitions on Piazza.
- **Don't make stuff up!** On problems that ask for proofs, you will receive more points if you demonstrate legitimate and correct intuition than if you fabricate something that *looks* rigorous just for the sake of having "rigor".
- **Check Piazza often!** Clarifications will be posted there. If you have a question, it is possible that it has already been asked and answered in a Piazza thread. If not, you can ask it, as long as you don't ask a public question that reveals any part of your solution to a problem.
- **Don't cheat!** As stated in Rules and Reminders, you may NOT use any references such as books or electronic resources (unless otherwise specified). If you cheat, you will be disqualified and banned from PUMaC, your school may be disqualified, and relevant external institutions may be notified of any misconduct.

Good luck, and have fun!

– Zongshu Wu, *Power Round Czar*

We would like to acknowledge and thank many individuals and organizations for their support; without their help, this Power Round (and the entire competition) could not exist. Please refer to the solutions for the Power Round for full acknowledgments and references.

Contents

Notation

- iff: if and only if.
- $\neg$: $\neg\varphi$ means "φ is false".
- $\wedge$: $\varphi \wedge \psi$ means "φ and ψ".
- $\vee$: $\varphi \vee \psi$ means "φ or ψ".
- $\implies$: $\varphi \implies \psi$ means "φ implies ψ".
- $\iff$: $\varphi \iff \psi$ means "φ iff ψ".
- $\forall$: $\forall x\, \varphi(x)$ means "$\varphi(x)$ holds for all x".
- $\exists$: $\exists x\, \varphi(x)$ means "$\varphi(x)$ holds for some x".
- $x \in X$ means "x is an element of X" or "X contains x".
- $(\forall x \in X)\, \varphi(x)$ means $\forall x\, (x \in X \implies \varphi(x))$, "$\varphi(x)$ holds for all $x \in X$".
- $(\exists x \in X)\, \varphi(x)$ means $\exists x\, (x \in X \wedge \varphi(x))$, "$\varphi(x)$ holds for some $x \in X$".
- $\exists!$: $\exists! x\, \varphi(x)$ is short for $\exists x\, (\varphi(x) \wedge \forall y\, (\varphi(y) \implies x = y))$, "$\varphi(x)$ holds for exactly one x". Similarly, $(\exists! x \in X)\, \varphi(x)$ is short for $\exists! x\, (x \in X \wedge \varphi(x))$, "$\varphi(x)$ holds for exactly one $x \in X$".
- $\subseteq$: $x \subseteq y$ is short for $(\forall z \in x)\, z \in y$.
- $\{F(x) \in X : \varphi(x)\}$ is short for $\{y \in X : \exists x\, (y = F(x) \wedge \varphi(x))\}$.

A Note on Rigor

In this Power Round, you may freely use any of the rules of logic, so there is no need to be pedantic about that. Furthermore, for problems that ask you to prove things about meta-mathematical objects (such as formulas), you may use informal arguments. After all, we do not give rigorous definitions for meta-mathematical objects, so being rigorous in your proofs is not even possible.

However, you must be very rigorous when proving things about sets. This is especially important in the first section, where we build everything up from the foundations. You are not allowed to write things like $\{x\}$ or $\{x, y, z\}$ or $X \times Y$ before they are introduced (unless you define them yourself and prove that they work)!

1 Zermelo-Fraenkel Set Theory (9 problems, 50 points)

At first glance, the mathematical concept of a *set* could not be simpler: it is simply any collection of things. However, after some careful thought, things start to break down. In 1901, Bertrand Russell considered the set

$$R = \{x : x \notin x\},$$

consisting of all sets that don't contain themselves. So, does R contain itself? Well, by definition, R contains R if and only if R does not contain R! This paradox, known as *Russell's paradox*, indicates that something is wrong with our naive notion of sets.

The only way to resolve this paradox is to declare that R does not exist – that there is no set consisting of precisely those sets that don't contain themselves. In other words, we need to part ways with the idea that any collection of things can be a set.

Since we can't make the assumption that any collection forms a set, we instead make a different (and much more complicated) list of assumptions in order to work with sets. These are known as the axioms of *Zermelo-Fraenkel set theory* (ZFC), named after Ernst Zermelo and Abraham Fraenkel, who developed it in the early 20th century.

1.1 Logical Formulas (10 points)

Before we get started, we need to say some words about how mathematical logic works in the world of axiomatic set theory. Read this part carefully!

In set theory, sets are the only kind of mathematical object. Everything is a set. To talk about sets, we use *formulas*.

Definition 1.1.1 — A *formula* (in set theory) is a statement about sets, involving some number of *variables*, which represent sets. If a formula explicitly depends on a variable, then the variable is called *free*. Otherwise, the variable is called *bound*. If a formula has no free variables, then it is called a *sentence*.

The expression $\varphi(x_1, x_2, \ldots, x_n)$ refers to some formula whose free variables are among $x_1, x_2, \ldots, x_n$. (It might be the case that some of the variables x_i are bound, or do not occur in the formula at all.)

By itself, this definition probably doesn't make a whole lot of sense, so we give many examples to illustrate what it means.

- $x \in y$ is a formula with two free variables, x and y. The meaning of this statement depends on what sets we substitute in place of x and y.
- $\exists x\,(x = y)$ is a formula with one free variable, y, and one bound variable, x. The meaning of this statement depends on y, but it does not depend on x, because x is just a dummy variable without an actual value assigned to it. Instead, we say that we are *quantifying* over x using the symbol $\exists$.
- $\forall x\,(\neg x \in x)$ has no free variables, so it is a sentence. Similarly, we are quantifying over x using the symbol $\forall$.
- $x \in y \wedge \forall x\,(y \in x)$ is... erm... is x free or bound?? It is free in its first occurrence, but bound in its later occurrences. Such cursed situations are technically allowed in logic, but for obvious reasons, it is a very bad idea to write formulas like this, so we will assume that this never happens.

Any formula can be written in terms of the symbols $=$ (equality), $\in$ (set membership), and the logical symbols $\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow, \forall, \exists$ defined in the Notation section. Usually, we will abbreviate formulas using other symbols, such as $\subseteq$. For instance, we can abbreviate $(\forall z \in x)\, z \in y$ as $x \subseteq y$.

Notice that when performing such abbreviations, bound variables can disappear, but the free variables don't change. In the example above, after we abbreviate the formula, the bound variable z disappears, but the free variables are x and y regardless.

Problem 1.1.1 (10 points)

Show that every formula is equivalent to one that only uses $=, \in, \neg, \wedge, \exists$.

Formulas are **not** considered mathematical objects in set theory, because they are not sets! In particular, in a formula, we can never quantify over a formula, so something like "there exists a formula such that ..." cannot be written as a formula. Instead, formulas are *meta*-mathematical objects – overlords that govern the mathematical world of sets.

1.2 The Axioms of ZFC (30 points)

We now describe the axioms of ZFC, a list of sentences that we assume when we prove anything about sets. They are written in natural language here, but it is possible (you can try) to write them using only $=, \in, \neg, \wedge, \vee, \Rightarrow, \Leftrightarrow, \forall, \exists$. But of course, that would be quite cumbersome, so we won't do that here.

Axiom (Extensionality)

Two sets x, y are equal iff $z \in x \iff z \in y$ for any set z.

Axiom (Pairing)

Given two sets x, y, there exists a set $\{x, y\}$ such that $z \in \{x, y\}$ iff $z = x$ or $z = y$.

Axiom (Union)

Given a set X, there exists a set $\bigcup X$ such that $y \in \bigcup X$ iff $y \in x$ for some $x \in X$.

Axiom (Power Set)

Given a set X, there exists a set $\mathcal{P}(X)$ such that $A \in \mathcal{P}(X)$ iff $A \subseteq X$.

Axiom (Separation)

Let $\varphi(x, p_1, \ldots, p_n)$ be a formula. Given a set X and some parameters $p_1, \ldots, p_n$, there exists a set Y such that $x \in Y$ iff $x \in X$ and $\varphi(x, p_1, \ldots, p_n)$. This set Y is denoted $\{x \in X : \varphi(x, p_1, \ldots, p_n)\}$.

The axiom of separation is not a single axiom; instead, it is an *axiom schema*, which means that it consists of infinitely many axioms, one for every formula $\varphi(x, p_1, \ldots, p_n)$. (In particular, ZFC has infinitely many axioms.)

Before we introduce the rest of the axioms (three axioms and one axiom schema), we take some time to see what we can do already with what we have.

Problem 1.2.1 (5 points)
Given a nonempty set X, prove that there exists a set $\bigcap X$ such that $y \in \bigcap X$ iff $y \in x$ for all $x \in X$.

Problem 1.2.2 (5 points)
Given two sets x, y, prove the existence of the following sets:

(a) The set $\{x\}$, such that $z \in \{x\}$ iff $z = x$;

(b) The set $x \cup y$, such that $z \in x \cup y$ iff $z \in x$ or $z \in y$;

(c) The set $x \cap y$, such that $z \in x \cap y$ iff $z \in x$ and $z \in y$.

In particular, given $x_1, \ldots, x_n$, we may form the set $\{x_1, \ldots, x_n\} = \{x_1\} \cup \cdots \cup \{x_n\}$, such that $y \in \{x_1, \ldots, x_n\}$ iff $y = x_i$ for some i.

Next, we formally define the notion of ordered pairs in terms of sets. This definition is due to Kazimierz Kuratowski.

Definition 1.2.1 — For two sets x, y, define the *ordered pair* $(x, y) = \{\{x\}, \{x, y\}\}$.

Problem 1.2.3 (5 points)
Prove that $(x, y) = (z, w)$ iff $x = z$ and $y = w$.

Problem 1.2.4 (5 points)
Given two sets X, Y, show that we can form the set $X \times Y$ of ordered pairs (x, y) where $x \in X$ and $y \in Y$, called the *Cartesian product* of X and Y.

Ordered pairs allow us to define what relations and functions are.

Definition 1.2.2 — A *relation* is a set R consisting of ordered pairs. We usually write $x\,R\,y$ as a shorthand for $(x, y) \in R$.

A relation f is a *function* if, for any set x, there is at most one set y such that $(x, y) \in f$. If such a y exists, then we write $f(x) = y$.

Problem 1.2.5 (5 points)
Let R be a relation. Show that we can form the sets

$$\mathrm{dom}(R) = \{x : \exists y\,(x, y) \in R\} \quad \text{and} \quad \mathrm{ran}(R) = \{y : \exists x\,(x, y) \in R\},$$

called the *domain* and *range* of R, respectively.

Definition 1.2.3 — A relation R is said to be *on* X if $R \subseteq X \times X$. A function f is said to be *on* X if $\text{dom}(f) = X$. A function f is said to be *from* X *to* Y, written $f : X \to Y$, if $\text{dom}(f) = X$ and $\text{ran}(f) \subseteq Y$.

Definition 1.2.4 — A function $f : X \to Y$ is called *injective/surjective/bijective*, or a(n) *injection/surjection/bijection*, if for any $y \in Y$, there is at most one/at least one/exactly one $x \in X$ such that $f(x) = y$.

Definition 1.2.5 — Let f be a function. The *restriction* of f to X is the function $f{\restriction}X = \{(x, y) \in f : x \in X\}$. The *image* of X under f is $f''X = \text{ran}(f{\restriction}X)$.

These definitions and results might be familiar from "normal math" – all we did was make everything rigorous using our set-theoretic framework. We now introduce the rest of the axioms of ZFC.

Axiom (Infinity)
There exists a set $\varnothing$ that contains nothing. Furthermore, there exists a set I such that $\varnothing \in I$, and if $x \in I$, then $x \cup \{x\} \in I$.

Axiom (Replacement)
Let $\varphi(x, y, p_1, \ldots, p_n)$ be a formula, and fix some parameters $p_1, \ldots, p_n$. If

$$\varphi(x, y, p_1, \ldots, p_n) \wedge \varphi(x, z, p_1, \ldots, p_n) \implies y = z$$

holds for all x, y, z, then given any set X, there exists a set Y such that $y \in Y$ iff $\varphi(x, y, p_1, \ldots, p_n)$ for some $x \in X$.

Axiom (Regularity)
Any nonempty set X contains an element x, called an $\in$-*minimal element*, such that $y \notin x$ for any $y \in X$.

Axiom (Choice)
Let X be a set. If all elements of X are nonempty, then there exists a function f on X, called a *choice function*, such that $f(x) \in x$ for all $x \in X$.

Just like the axiom of separation, the axiom of replacement is also an axiom schema consisting of infinitely many axioms.

Problem 1.2.6 (5 points)
Prove that no set contains itself, and no two sets contain each other.

1.3 Classes (10 points)

Not every collection of sets is a set. As you have already seen in the previous problems, when we want to define a set like $X \times Y$, we can't just say

$$X \times Y = \{(x, y) : x \in X \wedge y \in Y\};$$

instead, we need to prove that such a desired collection of sets actually exists using the axioms of ZFC. But what if we still want to talk about arbitrary collections of sets? For this, we introduce the notion of a *class*.

Definition 1.3.1 — Let $\varphi(x, p_1, \ldots, p_n)$ be a formula. Given parameters $p_1, \ldots, p_n$, we shall sometimes write $x \in \{x : \varphi(x, p_1, \ldots, p_n)\}$ in place of $\varphi(x, p_1, \ldots, p_n)$. The expression $\{x : \varphi(x, p_1, \ldots, p_n)\}$ is called a *class*.

Intuitively, a class $C = \{x : \varphi(x, p_1, \ldots, p_n)\}$ is the "collection" of sets satisfying the property $\varphi(x, p_1, \ldots, p_n)$. Classes are **not** sets; we introduce them simply because they are a convenient and intuitive notational shorthand.

Definition 1.3.2 — Let C and D be two classes. We say that C is a *subclass* of D, denoted $C \subseteq D$, if $x \in C \implies x \in D$ for all x. We say that C is *equal* to D, denoted $C = D$, if $x \in C \iff x \in D$ for all x.

Definition 1.3.3 — A class C *is a set* if, for some set X, we have $C = \{x : x \in X\}$ (that is, $x \in C \iff x \in X$ for all x). A class that is not a set is called a *proper class*.

By the axiom of extensionality, if such a set X exists, then it must be unique. Notice that saying "C is a set" is an abuse of terminology – classes are not actually sets! We do this because it just makes everything more convenient. If you pay attention to what you are doing, then there shouldn't be any issues.

In fact, if $C = \{x : x \in X\}$, then we will pretend as if C and X are the same thing. Under this convention, every set X "is" a class (namely, the class $\{x : x \in X\}$), and we can rephrase the axiom of separation as: *any subclass of a set is a set.*

Definition 1.3.4 — The *universe* is the class $V = \{x : x = x\}$ of all sets.

Problem 1.3.1 (5 points)
Show that the universe V is a proper class.

Many (but not all!) of the concepts we defined for sets work for classes as well. You should be able to guess how the definitions go before looking at them:

Definition 1.3.5 — Let C and D be classes. Define the following classes:

$$C \cup D = \{x : x \in C \vee x \in D\}, \quad \bigcup C = \{x : (\exists X \in C)\, x \in X\},$$
$$C \cap D = \{x : x \in C \wedge x \in D\}, \quad \bigcap C = \{x : (\forall X \in C)\, x \in X\},$$
$$C \times D = \{(x, y) : x \in C \wedge y \in D\}.$$

In fact, if C is nonempty, then $\bigcap C$ is always a set. This can be proved similarly to Problem 1.2.1. (For the empty class, we have $\bigcap \varnothing = V$.)

Definition 1.3.6 — A *class relation* is a class of ordered pairs. For a class relation R, we write $x\,R\,y$ to mean $(x, y) \in R$. A class relation *on* a class C is a subclass of $C \times C$. Given a class relation R, define the classes

$$\mathrm{dom}(R) = \{x : \exists y\,(x, y) \in R\} \quad \text{and} \quad \mathrm{ran}(R) = \{y : \exists x\,(x, y) \in R\}.$$

A class relation F is a *class function* if, for any set x, there is at most one set y such that $(x, y) \in F$. A class function F is *on* a class C if $\mathrm{dom}(F) = C$, and *from* C *to* D, written $F : C \to D$, if $\mathrm{dom}(F) = C$ and $\mathrm{ran}(F) \subseteq D$. The *restriction* of a class function F to a class C is the class function $F{\restriction}C = \{(x, y) \in F : x \in C\}$, and the *image* of C under F is $F''C = \mathrm{ran}(F{\restriction}C)$.

The following problem is a convenient rephrasing of the axiom of replacement.

Problem 1.3.2 (5 points)

Let F be a class function, and suppose that $\mathrm{dom}(F)$ is a set. Prove that $\mathrm{ran}(F)$ is a set, and conclude that F is also a set.

1.4 Philosophical Discussion: The Meta Theory

In the previous problems, you used the axioms of ZFC to prove many statements – that is, sentences. However, if you look carefully, you might notice that some of the problem statements *aren't* sentences. Notably, Problem 1.3.2 starts by picking an arbitrary class function F, which is not allowed in a sentence (as classes are not sets). So, what did you actually do by solving the problem?

Recall that the axioms of separation and replacement are actually axiom schemata:[1] they consist of infinitely many axioms. Similarly, we may think of solving Problem 1.3.2 as proving infinitely many sentences at once: for every class F, you prove the sentence that if F is a class function and $\mathrm{dom}(F)$ is a set, then $\mathrm{ran}(F)$ and F are sets.

That is, the statement of Problem 1.3.2 is a *meta-mathematical* statement, instead of a *mathematical* statement (i.e. a formula). To clarify this distinction, we introduce the terms *base theory* and *meta theory*. Sets live in the base theory, and when we prove sentences, we are working in the base theory. In contrast, meta-mathematical objects, like formulas or classes, live in the meta theory, and reasoning about them constitutes working in the meta theory.

For most problems in this Power Round, the distinction between the base theory and the meta theory can be mostly handwaved away. However, if you are not careful, you might still make mistakes! It is especially important to keep this in mind in the later sections, as there will be a blend of mathematical and meta-mathematical concepts.

Finally, if you are worried about the abuse of terminology where we say that certain classes "are" sets, rest assured that this will not cause any problems. Every time such an abuse of terminology occurs, it is always possible to rewrite things such that the abuse does not occur, often with the expense of making everything more cumbersome.

[1] The plural form of "schema".

2 Ordinals (17 problems, 180 points)

Ordinals are one of the most important concepts in set theory. Intuitively, they give us a way of counting past infinity. The natural numbers $0, 1, 2, \ldots$ are ordinals,[2] but beyond that, we have the ordinal ω, the smallest infinite ordinal. After that, we have $\omega + 1$, then $\omega + 2$, and so on, then $\omega + \omega = \omega \cdot 2$, and on and on and on...

As sets, each ordinal α is, intuitively, the set of ordinals smaller than α. For instance, since there are no ordinals less than 0, we have $0 = \varnothing$. Next, we have $1 = \{0\} = \{\varnothing\}$, and $2 = \{0, 1\} = \{\varnothing, \{\varnothing\}\}$, and then $\omega = \{0, 1, 2, \ldots\}$, and $\omega + 1 = \{0, 1, 2, \ldots, \omega\}$, and so on and so forth. This informal idea will be made rigorous below.

2.1 The Basics of Ordinals (75 points)

Definition 2.1.1 — A class x is *transitive* if any element of x is a subclass of x.

In other words, if x is transitive, then $z \in y$ and $y \in x$ imply $z \in x$. For example, the sets $\varnothing$ and $\{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}\}$ are transitive, but $\{\{\varnothing\}\}$ is not transitive.

Definition 2.1.2 — A set α is an *ordinal* if α is transitive, and every element of α is also transitive. The class of ordinals is denoted Ord.

For example, $\varnothing$ and $\{\varnothing, \{\varnothing\}\}$ are ordinals, but $\{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}\}$ (a transitive set) is not an ordinal, because it contains an element $\{\{\varnothing\}\}$ which is not transitive.

Problem 2.1.1 (5 points)

Show that any element of an ordinal is an ordinal.

Problem 2.1.2 (10 points)

Let C be a class of ordinals. Show that if C is a set, then $\bigcup C$ is an ordinal, and show that if C is nonempty, then $\bigcap C$ is an ordinal. Conclude that if α and β are ordinals, then $\alpha \cup \beta$ and $\alpha \cap \beta$ are also ordinals.

Problem 2.1.3 (20 points)

Let α be an ordinal, and let x, y be distinct elements of α. Prove that either $x \in y$ or $y \in x$. (Hint: use the axiom of regularity.)

Problem 2.1.4 (15 points)

Let α and β be ordinals. Prove that $\alpha \subseteq \beta$ iff $\alpha \in \beta$ or $\alpha = \beta$.

Problem 2.1.5 (10 points)

Let α and β be distinct ordinals. Prove that either $\alpha \in \beta$ or $\beta \in \alpha$.

[2] In set theory, 0 is considered a natural number.

Using the previous two problems, we can easily see that for any two ordinals α and β (not necessarily distinct), we have $\alpha \subseteq \beta$ or $\beta \subseteq \alpha$.

Definition 2.1.3 — Let α and β be ordinals. We write $\alpha < \beta$ or $\beta > \alpha$ for $\alpha \in \beta$, and write $\alpha \leq \beta$ or $\beta \geq \alpha$ for $\alpha \subseteq \beta$.

The results that we have shown so far imply that this method of comparing ordinals works exactly as you'd expect it to. For example, we have $\alpha \leq \beta$ iff $\alpha < \beta \vee \alpha = \beta$ iff $\alpha \not> \beta$. We can also combine inequalities: for instance, if $\alpha < \beta < \gamma$, then $\alpha < \gamma$, and if $\alpha \leq \beta \leq \gamma$, then $\alpha \leq \gamma$. From now on, you may freely use the basic properties of ordinal comparison without proof.

Problem 2.1.6 (10 points)

Let C be a class of ordinals. Show that

(a) If C is a set, then $\bigcup C$ is the smallest ordinal which is greater than or equal to all elements of C.

(b) If C is nonempty, then $\bigcap C$ is the smallest element of C.

Definition 2.1.4 — Let C be a class of ordinals. If C is a set, then its *supremum* is $\sup C = \bigcup C$. If C is nonempty, then its *minimum* is $\min C = \bigcap C$.

The axiom of regularity implies that any nonempty *set* of ordinals contains a smallest element, but the previous problem generalizes this statement to any nonempty *class* of ordinals, and also explicitly tells us what the minimum is!

Definition 2.1.5 — The *successor* of an ordinal α is the set $S(\alpha) = \alpha \cup \{\alpha\}$.

Problem 2.1.7 (5 points)

Prove that $S(\alpha)$ is the least ordinal greater than α.

Using the successor function, we can define

$$\begin{aligned}
0 &= \varnothing, \\
1 = S(0) &= \{\varnothing\}, \\
2 = S(1) &= \{\varnothing, \{\varnothing\}\}, \\
3 = S(2) &= \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}, \\
4 = S(3) &= \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}\},
\end{aligned}$$

and so on. It quickly becomes unwieldy to expand everything out completely: using the notation above, writing the number n in full requires $2^{n+1} - 1$ symbols. Figure 1 shows how tedious it can be even for relatively small numbers.

Warning: this is *not* a rigorous definition of a natural number! You might think that we can define a natural number as "the result of applying S finitely many times to 0", but this is circular logic, because we need natural numbers in order to formalize what "finite" means. In the next subsection, we will define natural numbers properly.

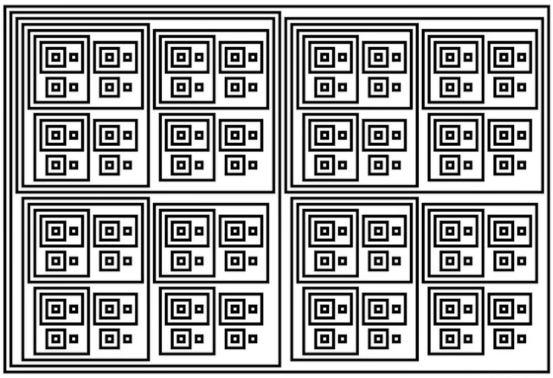


Figure 1: A visual representation of the number 8.

2.2 Induction and Recursion (60 points)

In this subsection, we will define the natural numbers, and rigorously justify induction and recursion. In "normal" math, we say "recursive definition" and "inductive definition" interchangeably, but in fact, recursion does **not** trivially follow from induction, and we need to do some work to justify recursion. (See Problem 2.2.6.)

Definition 2.2.1 — An ordinal α is called a *successor ordinal* if $\alpha = S(\beta)$ for some ordinal β. A *limit ordinal* is a nonzero ordinal which is not a successor ordinal.

The ordinal $0 = \varnothing$ is the only ordinal which is neither a successor nor a limit ordinal, just like how 1 is the only positive integer which is neither prime nor composite.

Problem 2.2.1 (5 points)

Show that a nonzero ordinal α is a limit ordinal iff $\alpha = \sup\{\beta : \beta < \alpha\}$. (Note that $\{\beta : \beta < \alpha\}$ is another way of writing the set α.)

Definition 2.2.2 — An ordinal n is a *natural number* if every ordinal k less than or equal to n is either 0 or a successor ordinal.

It is not hard to see that if n is a natural number, then $S(n)$ is a natural number, and every $k \leq n$ is a natural number. We now prove that mathematical induction works.

Problem 2.2.2 (10 points)

Let C be a class. Suppose that $0 \in C$, and if $n \in C$ for a natural number n, then $S(n) \in C$. Prove that C contains all natural numbers.

Problem 2.2.3 (10 points)

Prove that we can form a set ω such that $n \in \omega$ iff n is a natural number, and show that ω is the least limit ordinal.

In fact, we can prove a vast generalization of the principle of mathematical induction, known as *transfinite induction*, which works for all ordinals.

Problem 2.2.4 (5 points)

Let C be a class. Suppose that if α is an ordinal, and $\beta \in C$ for every $\beta < \alpha$, then $\alpha \in C$. Prove that C contains all ordinals.

Often, transfinite induction is split into three cases, depending on whether α is 0, a successor ordinal, or a limit ordinal.

Problem 2.2.5 (10 points)

Let C be a class. Suppose that

- $0 \in C$;
- If $\alpha \in C$, then $S(\alpha) \in C$;
- If α is a limit ordinal, and $\beta \in C$ for all $\beta < \alpha$, then $\alpha \in C$.

Prove that C contains all ordinals.

We can use transfinite induction to recursively define class functions on Ord. If we want to define a class function F on Ord, then it suffices to define $F(\alpha)$ in terms of the values of $F(\beta)$ for $\beta < \alpha$. This is known as *transfinite recursion*, and the next problem will ask you to justify it rigorously.

Problem 2.2.6 (15 points)

Let G be a class function on the universe V. Find a class function F on Ord such that $F(\alpha) = G(F{\restriction}\alpha)$ for every ordinal α, and show that any two such class functions are equal. (Note that $F{\restriction}\alpha$ is a set by Problem 1.3.2.)

Just like transfinite induction, we usually split transfinite recursion into three cases: zero, successor, and limit. To give a simple example, let α be an ordinal, and let G_α be the class function on V defined as

$$G_\alpha(f) = \begin{cases} \alpha & \text{if } f = \varnothing, \\ S(f(\beta)) & \text{if } f : S(\beta) \to \text{Ord}, \\ \sup(\text{ran}(f)) & \text{if } f : \beta \to \text{Ord for a limit ordinal } \beta, \\ \varnothing & \text{otherwise.} \end{cases}$$

Applying transfinite recursion, we get a class function F_α on Ord such that $F_\alpha(0) = \alpha$, $F_\alpha(S(\beta)) = S(F_\alpha(\beta))$, and $F_\alpha(\beta) = \sup\{F_\alpha(\gamma) : \gamma < \beta\}$ for limit ordinals β. Finally, we denote $\alpha + \beta = F_\alpha(\beta)$. We've just defined ordinal addition!

This definition may be cleanly summarized as follows:

Definition 2.2.3 — Define the *sum* $\alpha + \beta$ of two ordinals recursively as

- $\alpha + 0 = \alpha$,
- $\alpha + S(\beta) = S(\alpha + \beta)$,
- $\alpha + \beta = \sup\{\alpha + \gamma : \gamma < \beta\}$, if β is a limit ordinal.

Let's look at some examples. Firstly, we have

$$1 + 1 = 1 + S(0) = S(1 + 0) = S(1) = 2.$$

And in general, we have $\alpha + 1 = S(\alpha)$, and $\alpha + 2 = S(S(\alpha))$, etc., by definition. These facts are intuitive, but sometimes, weird things can happen. For instance,

$$1 + \omega = \sup\{1 + n : n < \omega\} = \omega \neq \omega + 1,$$

so ordinal addition is not commutative! (Ordinal addition *is* associative: $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ for all ordinals α, β, γ, but this is quite tricky to prove.)

Problem 2.2.7 (5 points)
Prove that if β is a limit ordinal, then $\alpha + \beta$ is also a limit ordinal.

Similarly, we can define ordinal multiplication.

Definition 2.2.4 — Define the *product* $\alpha \cdot \beta$ of two ordinals recursively as

- $\alpha \cdot 0 = 0$,
- $\alpha \cdot S(\beta) = \alpha \cdot \beta + \alpha$,
- $\alpha \cdot \beta = \sup\{\alpha \cdot \gamma : \gamma < \beta\}$, if β is a limit ordinal.

For example, we have $\alpha \cdot 1 = \alpha \cdot S(0) = \alpha \cdot 0 + \alpha = 0 + \alpha = \alpha$. Next, $\alpha \cdot 2 = \alpha + \alpha$, and then $\alpha \cdot 3 = \alpha + \alpha + \alpha$, and so on, by definition.

Just like ordinal addition, ordinal multiplication is associative: $(\alpha \cdot \beta) \cdot \gamma = \alpha \cdot (\beta \cdot \gamma)$, but not commutative: we have $\omega \cdot 2 = \omega + \omega$, but $2 \cdot \omega = \sup\{2n : n < \omega\} = \omega$. Ordinal multiplication also satisfies a distributive law: $\alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma$. However, it is *not* always true that $(\alpha + \beta) \cdot \gamma = \alpha \cdot \gamma + \beta \cdot \gamma$. (Take $\alpha = \beta = 1$ and $\gamma = \omega$.)

We can go further: ordinal exponentiation may be defined in a similar way as ordinal multiplication. But we won't go into that in this Power Round.

2.3 Well-Orders (45 points)

In addition to letting us "count past infinity", ordinals are useful in set theory because they allow us to quantify a special kind of ordering, called a *well-order*.

Definition 2.3.1 — Let X be a set. A relation $<$ on X is a *partial order* if

(1) $x < x$ is false for every $x \in X$, and

(2) $x < y$ and $y < z$ implies $x < z$ for all $x, y, z \in X$.

The relation $<$ is a *well-order* if, in addition, we have

(3) $x < y$ or $x = y$ or $y < x$ for all $x, y \in X$, and

(4) Any nonempty $A \subseteq X$ contains some m such that $x \not< m$ for all $x \in A$.

A *poset* (short for *partially ordered set*) is a pair $(X, <)$, where $<$ is a partial order on X. A poset $(X, <)$ is a *well-ordered set* if $<$ is a well-order on X.

For example, consider the relation $\{(A, B) \in \mathcal{P}(X) \times \mathcal{P}(X) : A \subsetneq B\}$ on $\mathcal{P}(X)$, which we will abbreviate as just "$\subsetneq$". Then, $(\mathcal{P}(X), \subsetneq)$ is a poset. But, if X has more than 1 element, then $\subsetneq$ is not a well-order on $\mathcal{P}(X)$, because it fails condition (3).

For another (very important) example, let α be an ordinal, and consider the relation $\{(x, y) \in \alpha \times \alpha : x \in y\}$ on α, which we will abbreviate as just "$\in$". Then, the problems in Section 2.1 tell us that $(\alpha, \in)$ is a well-ordered set.

Definition 2.3.2 — Let $(X, <_X)$ and $(Y, <_Y)$ be two well-ordered sets. A function $f : X \to Y$ is called an *order-isomorphism* if it is a bijection, and for all $x, y \in X$, if $x <_X y$, then $f(x) <_Y f(y)$. If there exists an order-isomorphism $f : X \to Y$, then we say that $(X, <_X)$ and $(Y, <_Y)$ are *isomorphic*, denoted $(X, <_X) \cong (Y, <_Y)$, or, if the context is clear, simply $X \cong Y$.

You can think of "isomorphic" as meaning "basically the same". For instance, consider the following two well-ordered sets:

$$(X, <_X) = (\{a, b, c\}, \{(a, b), (a, c), (b, c)\}),$$

where $a <_X b <_X c$, and

$$(Y, <_Y) = (\{p, q, r\}, \{(p, q), (r, p), (r, q)\}),$$

where $r <_Y p <_Y q$. Then, the structures of the two well-ordered sets are pretty much identical: in both cases, we have a chain of three elements in increasing order. Indeed, the function $f : X \to Y$ given by $f(a) = r$, $f(b) = p$, and $f(c) = q$ is an order-isomorphism, and thus, the two well-ordered sets are isomorphic.

Problem 2.3.1 (10 points)

Let $(X, <_X)$, $(Y, <_Y)$, and $(Z, <_Z)$ be well-ordered sets. Show that

(a) $X \cong X$.

(b) If $X \cong Y$, then $Y \cong X$.

(c) If $X \cong Y$ and $Y \cong Z$, then $X \cong Z$.

Problem 2.3.2 (15 points)

Let $(X, <)$ be a well-ordered set. Prove that there exists a unique ordinal α, called the *order type* of $(X, <)$, such that $(X, <) \cong (\alpha, \in)$.

Problem 2.3.3 (20 points)

Prove that for any set X, there exists a bijection f from X to some ordinal α, and conclude that there exists a well-order on X. (Hint: use the axiom of choice.)

This problem establishes the *well-ordering theorem*: every set can be well-ordered. It was first proven by Ernst Zermelo in 1904. The axiom of choice (often abbreviated as AC) is crucial in proving the well-ordering theorem. If your solution didn't use it, then it is wrong! In fact, if we work in ZF, which is ZFC without choice, then we can prove that

AC is *equivalent* to the well-ordering theorem. (Problem 2.3.3 establishes one direction: AC implies the well-ordering theorem. You are welcome to try the other direction, but we won't need this result.)

Many set theorists see the well-ordering theorem as somewhat unintuitive. It states that *all* sets can be well-ordered, including, for instance, $\mathbb{R}$.[3] How would you well-order $\mathbb{R}$? It's not something you can write down explicitly (without using AC), and the field of *descriptive set theory*, which studies $\mathbb{R}$ from a set-theoretic perspective, tells us that such a well-order would have bizarre properties. There is a famous joke by Jerry Bona:

> *The axiom of choice is obviously true, the well-ordering principle obviously false, and who can tell about Zorn's lemma?*

(*Zorn's lemma* is another result equivalent to the axiom of choice, and it is used in many areas of mathematics, but it has a rather complicated statement.)

2.4 Philosophical Discussion: Natural Numbers

In this section, we spent a lot of effort defining what natural numbers are, and making sure that the logic is completely airtight. However, if you look carefully, you may notice that we have actually been secretly using natural numbers since the very beginning of this Power Round! For instance, when we wrote "let $\varphi(x, p_1, \ldots, p_n)$ be a formula", we were invoking the concept of natural numbers, as n is a natural number. Did we commit the error of using a concept before defining it?

Don't worry – we didn't. When we write something like $\varphi(x, p_1, \ldots, p_n)$, the number n lives in the meta theory, instead of the base theory. It is a "meta natural number", if you will. So, we were using meta natural numbers in the meta theory, before defining natural numbers in the base theory. This is not circular reasoning!

But things still feels a bit suspicious. If we need meta natural numbers to formalize natural numbers, then we can ask: where do the meta natural numbers come from? We would need a "meta meta theory" to formalize the meta natural numbers, and a "meta meta meta theory" to formalize that, and so on. Turtles all the way down. And the issue is not just with natural numbers. If we want to formalize logic, then we would need a logical system in which to do so.

It seems that an infinite regress is unavoidable. In practice, logicians and set theorists deal with this problem by ignoring it. After all, we have to start *somewhere*. Instead of getting stuck in a fruitless cycle of formalization, we choose the meta theory as our starting point, and take it for granted. (In particular, there will be no such thing as a "meta meta theory".) From there, we can specify the basic rules of logic, and list out the axioms of the base theory ZFC.

In fact, we can go further, and formalize a copy of ZFC inside the base theory! To do logic within our set-theoretic framework, we encode each formula φ as a natural number $\ulcorner\varphi\urcorner$, called the *Gödel number* of φ (named after Kurt Gödel), and formalize all of the rules of logic within the base theory. Finally, we write down the Gödel numbers of the axioms of ZFC. The resulting set of Gödel numbers is called the *coded theory*.

The coded theory can be thought of as a copy of ZFC, inside the base theory ZFC, and one level "below" the base theory. In an abuse of notation, the coded theory is usually also denoted ZFC, but we shall write ZFC_c to avoid confusion.

[3] We won't give a precise definition of $\mathbb{R}$ in this Power Round.

3 Cardinals (19 problems, 180 points)

Infinite sets behave quite differently than finite sets, as famously illustrated in David Hilbert's Grand Hotel. Imagine a hotel with infinitely many rooms, numbered $0, 1, 2, \ldots$, each occupied by a guest, say, Room n is occupied by Guest n. The hotel is full, and yet, it can accommodate more guests: by moving Guest n to Room $n+1$ for all n, Room 0 is left vacant for a new guest, say Guest ω, to move into. In other words, the two sets $\omega = \{0, 1, 2, \ldots\}$ and $\omega + 1 = \{0, 1, 2, \ldots, \omega\}$ have the same "size", in some sense, even though ω is a proper subset of $\omega + 1$.

But are there any infinite sets that have a larger size than ω? In 1874, Georg Cantor answered this question in the affirmative: he proved that the set $\mathbb{R}$ of real numbers has strictly more elements than ω. If a guest for every *real* number came to Hilbert's Hotel, then the hotel would not be able to accommodate everyone.

More precisely, the size of a set X is measured by its *cardinality* $|X|$, a special kind of ordinal called a *cardinal*. For example, the cardinality of $\{a, b, c\}$ is 3, because $\{a, b, c\}$ has 3 elements, and the cardinalities of ω and $\omega + 1$ are $\aleph_0 = \omega$, which is the smallest infinite cardinal. After $\aleph_0$, the next cardinal is $\aleph_1$, and then $\aleph_2$, and after infinitely many of these, we reach $\aleph_\omega$. In fact, there is a cardinal $\aleph_\alpha$ for every ordinal α.

All of the informal ideas above will be made fully rigorous in what follows.

3.1 The Basics of Cardinals (60 points)

Definition 3.1.1 — Two sets X and Y are *equinumerous*, denoted $X \approx Y$, if there exists a bijection $f : X \to Y$.

For example, we have already seen that ω and $\omega + 1$ are equinumerous. It is also not hard to show that ω and $\omega \cdot 2$ are equinumerous: we can define a bijection $f : \omega \cdot 2 \to \omega$ via $f(n) = 2n$ and $f(\omega + n) = 2n + 1$, where $n < \omega$.

Problem 3.1.1 (5 points)

Let X, Y, Z be sets. Show that

(a) $X \approx X$.

(b) If $X \approx Y$, then $Y \approx X$.

(c) If $X \approx Y$ and $Y \approx Z$, then $X \approx Z$.

The well-ordering theorem (Problem 2.3.3) tells us that any set X is equinumerous to at least one ordinal, so we can use ordinals to measure the sizes of sets.

Definition 3.1.2 — The *cardinality* of a set X, denoted $|X|$, is the smallest ordinal α such that $X \approx \alpha$. An ordinal κ is a *cardinal* if $|\kappa| = \kappa$ (in other words, if κ is not equinumerous to a smaller ordinal).

A set X is *finite* if $|X| < \omega$ and *infinite* if $|X| \geq \omega$. An infinite set X is *countable* if $|X| = \omega$, and *uncountable* if $|X| > \omega$.

The cardinality of a set is always a cardinal. Indeed, we have $X \approx |X|$, so if $|X|$ were equinumerous to some smaller ordinal α, then X would also be equinumerous to α, but this would contradict the definition of $|X|$.

Problem 3.1.2 (5 points)
Prove that every infinite cardinal is a limit ordinal.

Problem 3.1.3 (20 points)
Let X and Y be sets. Prove that

(a) $|X| \le |Y|$ iff there exists an injection $f : X \to Y$.

(b) $|X| \ge |Y|$ iff there exists a surjection $f : X \to Y$.

(c) $|X| = |Y|$ iff there exists a bijection $f : X \to Y$ (that is, $X \approx Y$).

Problem 3.1.4 (10 points)
Prove that every natural number is a cardinal.

Problem 3.1.5 (5 points)
Prove that if X is a set of cardinals, then $\sup X$ is a cardinal. In particular, show that ω is a cardinal.

The previous problem ensures that countable sets exist: the set ω is countable, since $|\omega| = \omega$. Of course, $\omega + 1$ and $\omega \cdot 2$ are also countable. The existence of uncountable sets follows from *Cantor's theorem*, named after Georg Cantor.

Theorem 3.1.3 (Cantor)
If X is a set, then $|X| < |\mathcal{P}(X)|$.

Proof. Suppose for the sake of contradiction that $|X| \ge |\mathcal{P}(X)|$. Then by Problem 3.1.3, there exists a surjective function $f : X \to \mathcal{P}(X)$. Now consider the set

$$A = \{x \in X : x \notin f(x)\} \in \mathcal{P}(X).$$

We have $A = f(x)$ for some $x \in X$. But then $x \in A$ iff $x \notin A$, a contradiction. □

In particular, $\mathcal{P}(\omega)$ is uncountable. It is not too hard to show that $\mathbb{R}$ is equinumerous to $\mathcal{P}(\omega)$, so $\mathbb{R}$ is also uncountable. (Our proof of this is not Cantor's original 1874 proof; instead, it is essentially equivalent to another proof he gave in 1891.)

Cantor's theorem implies that there is no largest cardinal: indeed, if κ is a cardinal, then the cardinal $|\mathcal{P}(\kappa)|$ is always strictly larger than κ. Thus, we define:

Definition 3.1.4 — Define the *aleph numbers* $\aleph_\alpha$ (also denoted ω_α) recursively as

- $\aleph_0 = \omega$,
- $\aleph_{\alpha+1}$ is the smallest cardinal greater than $\aleph_\alpha$,
- $\aleph_\alpha = \sup\{\aleph_\beta : \beta < \alpha\}$, if α is a limit ordinal.

Problem 3.1.6 (15 points)
Prove that every infinite cardinal is equal to $\aleph_\alpha$ for some ordinal α.

3.2 Cardinal Arithmetic (60 points)

Before we begin, here is a basic definition.

Definition 3.2.1 — Let I be a set and C be a class. A *family* of elements of C, indexed by I, is a function $x : I \to C$. In this context, we write x_i in place of $x(i)$ and $(x_i)_{i\in I}$ in place of x.

In elementary school, the basic operations of arithmetic are introduced by studying what we would call the cardinality of finite sets. For example, the equality $1 + 1 = 2$ is usually interpreted to mean that if two disjoint sets $\{a\}$ and $\{b\}$ have cardinality 1, then the union $\{a\} \cup \{b\} = \{a, b\}$ has cardinality 2. Building on this idea, we define the sum of any number of cardinals.

Definition 3.2.2 — Let $(X_i)_{i\in I}$ be a family of sets. Their *union* $\bigcup_{i\in I} X_i$ is defined as $\bigcup \operatorname{ran}(X) = \{x : (\exists i \in I)\, x \in X_i\}$, and their *disjoint union* $\bigsqcup_{i\in I} X_i$ is defined as the union $\bigcup_{i\in I} X_i \times \{i\}$. If we only have two sets X, Y, then we write their disjoint union as $X \sqcup Y = (X \times \{0\}) \cup (Y \times \{1\})$.

For a family $(\kappa_i)_{i\in I}$ of cardinals, their *sum* $\sum_{i\in I} \kappa_i$ is defined as $\left|\bigsqcup_{i\in I} \kappa_i\right|$. If we only have two cardinals κ, λ, then we write their sum as $\kappa + \lambda = |\kappa \sqcup \lambda|$.

Intuitively, the disjoint union takes a family $(X_i)_{i\in I}$ of sets, and replaces each set X_i with $X_i \times \{i\}$ to ensure that the sets are disjoint, before taking the union.

Problem 3.2.1 (10 points)
Show that if $(X_i)_{i\in I}$ is a family of pairwise disjoint sets (that is, $X_i \cap X_j = \varnothing$ for any $i \neq j \in I$), then $\left|\bigcup_{i\in I} X_i\right| = \sum_{i\in I} |X_i|$. (Hint: if you think this is trivial, you are probably missing something.)

Cardinal addition satisfies many of the properties you would expect addition to satisfy. It is associative: $(\kappa + \lambda) + \mu = \kappa + (\lambda + \mu)$, commutative: $\kappa + \lambda = \lambda + \kappa$, increasing: $\lambda \leq \mu$ implies $\kappa + \lambda \leq \kappa + \mu$, and finally, $\kappa + 0 = \kappa$. These facts can be easily shown using the previous problem.

Warning: cardinal addition is **not** the same thing as ordinal addition, even though we use the symbol $+$ for both! For example, let's compute $\aleph_0 + 1$, where $+$ means cardinal addition. Since $|\omega| = \aleph_0$ and $|\{\omega\}| = 1$, and the sets ω and $\{\omega\}$ are disjoint, we have

$$\aleph_0 + 1 = |\omega \cup \{\omega\}| = |\omega + 1| = \aleph_0.$$

That is, $\aleph_0 + 1$ (where $+$ is cardinal addition) is not equal to $\omega + 1$ (where $+$ is ordinal addition). To prevent confusion, we will adopt the following convention:

- When thinking of $\aleph_\alpha = \omega_\alpha$ as the cardinality of some set (e.g. when doing cardinal arithmetic), we write $\aleph_\alpha$.
- When thinking of $\aleph_\alpha = \omega_\alpha$ as an ordinal that just so happens to be a cardinal (e.g. when doing ordinal arithmetic), we write ω_α (or simply ω if $\alpha = 0$).

Anyways, it should almost always be clear from context whether $+$ is supposed to stand for ordinal addition or cardinal addition.

Next, we turn our attention to multiplication. In elementary school arithmetic, the equality $2 \cdot 3 = 6$ means that if $\{a, b\}$ has cardinality 2 and $\{x, y, z\}$ has cardinality 3, then their Cartesian product $\{a, b\} \times \{x, y, z\} = \{(a, x), (a, y), (a, z), (b, x), (b, y), (b, z)\}$ has cardinality 6. We now define the product of any number of cardinals.

Definition 3.2.3 — Let $(X_i)_{i \in I}$ be a family of sets. The *Cartesian product* $\prod_{i \in I} X_i$ is defined as the set of families $(x_i)_{i \in I}$ such that $x_i \in X_i$ for all $i \in I$.

For a family $(\kappa_i)_{i \in I}$ of cardinals, their *product* $\prod_{i \in I} \kappa_i$ is defined as $\left|\prod_{i \in I} \kappa_i\right|$. If we only have two cardinals κ, λ, then we write their product as $\kappa \cdot \lambda = |\kappa \times \lambda|$.

Problem 3.2.2 (5 points)

Show that the Cartesian product $\prod_{i \in I} X_i$ is a set. Show that if X_i is nonempty for all $i \in I$, then $\prod_{i \in I} X_i$ is nonempty.

Unfortunately, we have to deal with several abuses of notation here. Firstly, $\prod_{i \in I} \kappa_i$ is used both for the Cartesian product of the cardinals and its cardinality.

Second, if we only have two sets X_0, X_1, then their Cartesian product as defined above, call it $X_0 \otimes X_1 = \prod_{i \in 2} X_i$, is not equal to the Cartesian product $X_0 \times X_1$ as defined in Section 1! Thankfully, there is a natural bijection $X_0 \otimes X_1 \to X_0 \times X_1$ sending $(x_i)_{i \in 2}$ to (x_0, x_1), so there is nothing to worry about.

Problem 3.2.3 (10 points)

Show that $\left|\prod_{i \in I} X_i\right| = \prod_{i \in I} |X_i|$. (The first $\prod$ is the Cartesian product of sets, and the second $\prod$ is the product of cardinals.)

In particular, it follows that $|X \times Y| = |X| \cdot |Y|$. Just like cardinal addition, cardinal multiplication satisfies many intuitive properties. It is associative: $(\kappa \cdot \lambda) \cdot \mu = \kappa \cdot (\lambda \cdot \mu)$, commutative: $\kappa \cdot \lambda = \lambda \cdot \kappa$, distributive: $\kappa \cdot (\lambda + \mu) = \kappa \cdot \lambda + \kappa \cdot \mu$, increasing: $\lambda \leq \mu$ implies $\kappa \cdot \lambda \leq \kappa \cdot \mu$, and finally, $\kappa \cdot 1 = \kappa$ and $\kappa \cdot 0 = 0$. Furthermore, the sets $\bigsqcup_{i \in Y} X$ and $X \times Y$ are *equal* (not just equinumerous), so $\sum_{i \in I} \kappa = \kappa \cdot |I|$.

Once again, it is important to stress that cardinal multiplication is **not** the same as ordinal multiplication. For example, observe that the Cartesian product $\omega \times 2$ and the ordinal $\omega \cdot 2$ are equinumerous: there is a bijection $\omega \times 2 \to \omega \cdot 2$ which sends (a, b) to $\omega \cdot b + a$. It immediately follows that

$$\aleph_0 \cdot 2 = |\omega \times 2| = |\omega \cdot 2| = \aleph_0.$$

Of course, $\omega \cdot 2 = \omega$ is not true under ordinal multiplication.

However, it turns out that for *finite* cardinals, cardinal arithmetic works exactly the same way as ordinal arithmetic.

Problem 3.2.4 (10 points)

Prove that $m +_o n = m +_c n$ and $m \cdot_o n = m \cdot_c n$ for all natural numbers $m, n \in \omega$, where $+_o, \cdot_o$ denote ordinal addition and multiplication, and $+_c, \cdot_c$ denote cardinal addition and multiplication.

On the other hand, cardinal addition and multiplication are kind of trivial for infinite cardinals, thanks to a theorem proven by Gerhard Hessenberg in 1906.

Theorem 3.2.4 (Hessenberg)
If κ is an infinite cardinal, then $\kappa \cdot \kappa = \kappa$.

Proof. We have $\kappa \cdot \kappa \geq \kappa \cdot 1 = \kappa$, so it suffices to show that $\kappa \cdot \kappa \leq \kappa$. Assume for the sake of contradiction that $\kappa \cdot \kappa > \kappa$ for some infinite cardinal κ. Then, we may take κ to be the *smallest* such cardinal.

We construct a well-order on $\kappa \times \kappa$ as follows:

$$\begin{aligned}(\alpha_1, \beta_1) < (\alpha_2, \beta_2) \iff &\max\{\alpha_1, \beta_1\} < \max\{\alpha_2, \beta_2\} \\ &\vee (\max\{\alpha_1, \beta_1\} = \max\{\alpha_2, \beta_2\} \wedge \alpha_1 < \alpha_2) \\ &\vee (\max\{\alpha_1, \beta_1\} = \max\{\alpha_2, \beta_2\} \wedge \alpha_1 = \alpha_2 \wedge \beta_1 < \beta_2).\end{aligned}$$

It is not hard to check that this is indeed a well-order. A visualization of this is shown below, where (α, β) is the cell on the αth row and the βth column.

	0	1	2	3	⋯	ω	⋯
0	0	1	4	9	⋯	ω	⋯
1	2	3	5	10	⋯	$\omega+1$	⋯
2	6	7	8	11	⋯	$\omega+2$	⋯
3	12	13	14	15	⋯	$\omega+3$	⋯
⋮	⋮	⋮	⋮	⋮		⋮	
ω	$\omega \cdot 2$	$\omega 2+1$	$\omega 2+2$	$\omega 2+3$	⋯	$\omega \cdot 3$	⋯
⋮	⋮	⋮	⋮	⋮		⋮	

Let ξ be the order type of $(\kappa \times \kappa, <)$. By assumption, we have $\xi \geq |\kappa \times \kappa| = \kappa \cdot \kappa > \kappa$, so suppose that the order-isomorphism $\xi \to \kappa \times \kappa$ maps κ to (α, β). In other words, the set $X = \{(\alpha', \beta') \in \kappa \times \kappa : (\alpha', \beta') < (\alpha, \beta)\}$ has order type κ under $<$. Note that α and β can't both be finite, as otherwise X would be finite.

Next, let $\delta = S(\max\{\alpha, \beta\})$, which is infinite but less than κ. Then $X \subseteq \delta \times \delta$, so the order type of $(\delta \times \delta, <)$ is at least κ. In particular, $|\delta| \cdot |\delta| = |\delta \times \delta| \geq \kappa > |\delta|$. However, we had defined κ to be the smallest infinite cardinal such that $\kappa \cdot \kappa > \kappa$. So since $|\delta|$ is an infinite cardinal smaller than κ, we have a contradiction. □

From this, we find that cardinal addition and multiplication with infinite cardinals is given by a very simple formula:

Problem 3.2.5 (5 points)
Show that if κ and λ are nonzero cardinals and at least one of them is infinite, then $\kappa + \lambda = \kappa \cdot \lambda = \max\{\kappa, \lambda\}$.

We also get a nice application to ordinal arithmetic.

Problem 3.2.6 (10 points)

Prove that if α and β are nonzero ordinals and at least one of them is infinite, then $|\alpha + \beta| = |\alpha \cdot \beta| = \max\{|\alpha|, |\beta|\}$.

Finally, we take a brief look at cardinal exponentiation.

Definition 3.2.5 — For two sets X, Y, the *hom-set* $\mathrm{Hom}(X, Y)$ (also denoted ${}^{X}Y$ or sometimes Y^X) is the set of functions from X to Y.

For two cardinals κ, λ, define $\kappa^\lambda = |\mathrm{Hom}(\lambda, \kappa)|$.

Note that $\mathrm{Hom}(X, Y) \subseteq \mathcal{P}(X \times Y)$, so $\mathrm{Hom}(X, Y)$ is a set. Furthermore, it is easy to show (similarly to Problems 3.2.1 and 3.2.3) that $|\mathrm{Hom}(X, Y)| = |Y|^{|X|}$.

Problem 3.2.7 (5 points)

Show that $|\mathcal{P}(X)| = 2^{|X|}$ for any set X.

Just like before, we list some easy properties of cardinal exponentiation. It satisfies the "distributive" identities $(\kappa \cdot \lambda)^\mu = \kappa^\mu \cdot \lambda^\mu$ and $\kappa^{\lambda+\mu} = \kappa^\lambda \cdot \kappa^\mu$ and $\kappa^{\lambda \cdot \mu} = (\kappa^\lambda)^\mu$, the "unit" identities $\kappa^0 = 1$ (in particular, $0^0 = 1$) and $1^\kappa = 1$ and $0^\kappa = 0$ for $\kappa > 0$, and is increasing: $\kappa \le \lambda$ implies $\kappa^\mu \le \lambda^\mu$, and $0 < \lambda \le \mu$ implies $\kappa^\lambda \le \kappa^\mu$. Finally, observe that the sets $\prod_{i \in X} Y$ and $\mathrm{Hom}(X, Y)$ are equal, so $\prod_{i \in I} \kappa = \kappa^{|I|}$.

Problem 3.2.8 (5 points)

Show that if $2 \le \kappa \le \lambda$ and λ is infinite, then $\kappa^\lambda = 2^\lambda$.

Cardinal exponentiation is much more mysterious than addition or multiplication. For finite cardinals, it is of course easy to compute, but for infinite cardinals, the very first nontrivial computation already leaves us stumped: what is $2^{\aleph_0}$?

Definition 3.2.6 — The *cardinality of the continuum* is the cardinal $\mathfrak{c} = 2^{\aleph_0}$.

We know, from previous problems, that $\mathfrak{c} = |\mathcal{P}(\omega)| = |\mathbb{R}|$. (The *continuum* refers to the set $\mathbb{R}$, which is "continuous", hence the name.) So which aleph number is $\mathfrak{c}$ equal to? The *continuum hypothesis* (CH) states that $\mathfrak{c} = \aleph_1$, a natural guess. But surprisingly, it is impossible to prove or disprove the continuum hypothesis using the axioms of ZFC! This strange phenomenon will be discussed shortly, and we will prove it by the end of this Power Round.

What about even larger cases, such as $2^{\aleph_1}$? We can make the following guess:

$$2^{\aleph_\alpha} = \aleph_{\alpha+1} \quad \text{for all ordinals } \alpha.$$

This is known as the *generalized continuum hypothesis* (GCH). Of course, GCH implies CH, but not the other way around. In particular, GCH is impossible to prove. In fact, it is also impossible to disprove.

It turns out that if we assume the generalized continuum hypothesis, then cardinal exponentiation becomes very nice, as we will see shortly.

3.3 Cofinality (60 points)

Suppose that $\mathfrak{c} = \aleph_\theta$. Since $\mathfrak{c}$ is uncountable, we have $\theta \geq 1$. Can we say anything else about θ? Not much, as it turns out, but we can use the concept of *cofinality* to rule out a few possibilities for θ. For instance, we will see that θ can't equal ω.

Definition 3.3.1 — Let α be a limit ordinal. For an ordinal β, a function $f : \beta \to \alpha$ is *cofinal* if $\sup\{f(\gamma) : \gamma < \beta\} = \alpha$. The *cofinality* of α is the least ordinal $\operatorname{cf}\alpha$ such that there exists a cofinal function $\operatorname{cf}\alpha \to \alpha$.

An infinite cardinal κ is *regular* if $\operatorname{cf}\kappa = \kappa$, and *singular* if $\operatorname{cf}\kappa < \kappa$.

If α is a limit ordinal, then by Problem 2.2.1, the identity function $\mathrm{id} : \alpha \to \alpha$ (where $\mathrm{id}(\beta) = \beta$ for $\beta < \alpha$) is cofinal, so the cofinality $\operatorname{cf}\alpha$ is well-defined and at most α. (If α is a successor ordinal, then no function $\beta \to \alpha$ is cofinal.)

Intuitively, the cofinality of α measures how easy it is to approach α. For example, the cardinal $\aleph_{\omega_1+\omega}$ is quite large, but it is very easy to approach. The function $\omega \to \aleph_{\omega_1+\omega}$ sending n to $\aleph_{\omega_1+n}$ is cofinal, so $\operatorname{cf}\aleph_{\omega_1+\omega} \leq \aleph_0$. That is, we can approach $\aleph_{\omega_1+\omega}$ using $\aleph_0$ ordinals. By Problem 3.3.1 below, the cofinality can't be finite, so $\operatorname{cf}\aleph_{\omega_1+\omega} = \aleph_0$.

On the other hand, the cardinal $\aleph_{\omega+1}$ is much smaller, but it is much more difficult to approach. In fact, by Problem 3.3.2 below, it is regular: $\operatorname{cf}\aleph_{\omega+1} = \aleph_{\omega+1}$, so we need $\aleph_{\omega+1}$ ordinals to approach $\aleph_{\omega+1}$.

Problem 3.3.1 (15 points)

Prove that $\operatorname{cf}\alpha$ is always a regular cardinal.

Problem 3.3.2 (10 points)

Let α be an ordinal. Prove that

(a) If $\alpha = 0$ or $\alpha = \beta + 1$, then $\aleph_\alpha$ is regular.

(b) If α is a limit ordinal, then $\operatorname{cf}\aleph_\alpha = \operatorname{cf}\alpha$.

In order to apply the concept of cofinality to study $\mathfrak{c}$, we need *König's theorem*, first shown by Gyula König in 1905.

Problem 3.3.3 (15 points)

Let $(\kappa_i)_{i\in I}$ and $(\lambda_i)_{i\in I}$ be families of cardinals such that $\kappa_i < \lambda_i$ for all $i \in I$. Then

$$\sum_{i\in I} \kappa_i < \prod_{i\in I} \lambda_i.$$

(Hint: if $\kappa_i = 1$ and $\lambda_i = 2$ for all $i \in I$, then we get $|I| < 2^{|I|}$, so König's theorem generalizes Cantor's theorem. Does the proof generalize as well?)

Problem 3.3.4 (10 points)

Prove that if κ is an infinite cardinal, then $\operatorname{cf} 2^\kappa > \kappa$.

In particular, this means that $\operatorname{cf} \mathfrak{c} > \aleph_0$. So if $\mathfrak{c} = \aleph_\theta$, then θ must either be a successor ordinal, or a limit ordinal with cofinality $> \aleph_0$. This rules out a few possibilities for what θ can be: for instance, θ can't equal ω, or $\omega \cdot 2$, or $\omega_1 + \omega$, or ω_ω, and so on.

We end with another application to cardinal exponentiation. For an infinite cardinal κ, let κ^+ denote the smallest cardinal greater than κ (that is, $\aleph_\alpha^+ = \aleph_{\alpha+1}$).

Problem 3.3.5 (10 points)

Assuming the generalized continuum hypothesis, prove that if κ and λ are infinite cardinals such that $\operatorname{cf} \kappa > \lambda$, then $\kappa^\lambda = \kappa$.

Theorem 3.3.2

Under the generalized continuum hypothesis, if κ and λ are infinite cardinals, then

(a) If $\lambda < \operatorname{cf} \kappa$, then $\kappa^\lambda = \kappa$.

(b) If $\operatorname{cf} \kappa \leq \lambda < \kappa$, then $\kappa^\lambda = \kappa^+$.

(c) If $\lambda \geq \kappa$, then $\kappa^\lambda = \lambda^+$.

Proof. (a) is Problem 3.3.5, and (c) follows from Problem 3.2.8: if $\lambda \geq \kappa$, then by GCH, we have $\kappa^\lambda = 2^\lambda = \lambda^+$. From now on, we assume that $\operatorname{cf} \kappa \leq \lambda < \kappa$.

It suffices to show that $\kappa^{\operatorname{cf} \kappa} > \kappa$. Indeed, if this is true, then we have $\kappa^\lambda \leq \kappa^\kappa = \kappa^+$ and $\kappa^\lambda \geq \kappa^{\operatorname{cf} \kappa} \geq \kappa^+$, so $\kappa^\lambda = \kappa^+$, and we would be done.

Let $s : \operatorname{cf} \kappa \to \kappa$ be cofinal, and for the sake of contradiction, suppose that $\kappa^{\operatorname{cf} \kappa} \leq \kappa$, so that there exists a surjective function $F : \kappa \to \operatorname{Hom}(\operatorname{cf} \kappa, \kappa)$. Next, define a function $f : \operatorname{cf} \kappa \to \kappa$ as follows: for $\xi < \operatorname{cf} \kappa$, define $f(\xi)$ as the smallest ordinal $\gamma < \kappa$ such that $\gamma \neq (F(\alpha))(\xi)$ for any $\alpha < s(\xi)$. (Such a γ always exists, as $|s(\xi)| < \kappa$.)

Since F is surjective, we have $f = F(\alpha)$ for some α. However, since s is cofinal, there exists some ξ such that $s(\xi) > \alpha$. Hence, $f(\xi) \neq (F(\alpha))(\xi)$, a contradiction. □

3.4 Interlude: Gödel's Incompleteness Theorems

In September 1930, David Hilbert gave a fiery speech for his retirement address at the Königsberg conference, ending with the words "Wir müssen wissen. Wir werden wissen!" (Translation: "We must know. We shall know!") These words were later engraved on his tombstone.

Hilbert believed that mathematics was *complete*: that it was possible to find a set of axioms for the entirety of mathematics, such that every statement can be either proved or disproved from the axioms. Of course, he also wanted to make sure that those axioms were *consistent*: that it was impossible to prove an obviously false statement.

But at the very same conference, a young Kurt Gödel announced his newest result: in any[4] consistent system of axioms, there will always statements that are neither provable nor disprovable – they are *independent* from the axioms. This is known as *Gödel's first incompleteness theorem*. Shortly thereafter, Gödel published this result, and with that, Hilbert's dream shattered.

We give a brief exposition of Gödel's work below.

[4] Actually, there are some additional technical conditions required: the set of axioms must be recursively enumerable, and the theory must be capable of describing the addition and multiplication of natural numbers. But all you need to know is that ZFC satisfies these conditions.

Definition 3.4.1 — Let φ be a sentence. We write $\mathsf{ZFC} \vdash \varphi$ if we can prove φ in ZFC. If $\mathsf{ZFC} \nvdash \varphi$ and $\mathsf{ZFC} \nvdash \neg\varphi$, then we say that φ is *independent* from ZFC.

Let $\bot$ denote the sentence $\exists x\,(x \neq x)$, which is clearly false. We say that ZFC is *consistent* if $\mathsf{ZFC} \nvdash \bot$.

Theorem 3.4.2 (Gödel, Rosser)
If ZFC is consistent, then there exists a sentence which is independent from ZFC.

The condition that ZFC is consistent is necessary. If we could prove a false sentence in ZFC, then we would be able to prove every possible sentence, and nothing would be independent (and mathematics would fall apart).

The construction that Gödel gave, the *Gödel sentence*, works by creating some clever interplay between the base theory and the coded theory. First, we construct a formula $\mathrm{Bew}(\ulcorner\varphi\urcorner)$, which states that $\ulcorner\varphi\urcorner$ can be proved in the coded theory ZFC_c. (The name of the formula is short for *Beweis*, which is German for "proof".) This formula satisfies the following *provability conditions*:

1. If $\mathsf{ZFC} \vdash \varphi$, then $\mathsf{ZFC} \vdash \mathrm{Bew}(\ulcorner\varphi\urcorner)$.
2. ZFC proves $\mathrm{Bew}(\ulcorner\varphi\urcorner) \implies \mathrm{Bew}(\ulcorner\mathrm{Bew}(\ulcorner\varphi\urcorner)\urcorner)$.
3. ZFC proves $(\mathrm{Bew}(\ulcorner\varphi\urcorner) \wedge \mathrm{Bew}(\ulcorner\varphi \Rightarrow \psi\urcorner)) \implies \mathrm{Bew}(\ulcorner\psi\urcorner)$.

The Gödel sentence G is then constructed, using a clever trick, so that ZFC proves the sentence $G \iff \neg\mathrm{Bew}(\ulcorner G\urcorner)$. It follows easily that $\mathsf{ZFC} \nvdash G$: if not, then we would have both $\mathsf{ZFC} \vdash \mathrm{Bew}(\ulcorner\varphi\urcorner)$ (by the first provability condition) and $\mathsf{ZFC} \vdash \neg\mathrm{Bew}(\ulcorner\varphi\urcorner)$ (by the definition of G), which contradicts the consistency of ZFC.

Unfortunately, Gödel was not able to show that $\mathsf{ZFC} \nvdash \neg G$ without introducing some additional assumptions. In 1936, J. Barkley Rosser got around this problem by slightly tweaking the definition of $\mathrm{Bew}(\ulcorner\varphi\urcorner)$, and thus he also gets credit for the theorem. We won't go into the details here.

In the paper that Gödel introduced his first incompleteness theorem, he also sketched a proof of his *second incompleteness theorem*, which we state below.

Definition 3.4.3 — The sentence $\mathrm{Con}(\mathsf{ZFC}_c)$ is defined as $\neg\mathrm{Bew}(\ulcorner\bot\urcorner)$.

Theorem 3.4.4 (Gödel)
If ZFC is consistent, then $\mathsf{ZFC} \nvdash \mathrm{Con}(\mathsf{ZFC}_c)$.

The proof is quite technical and provides little insight. After assuming for the sake of contradiction that $\mathsf{ZFC} \vdash \mathrm{Con}(\mathsf{ZFC}_c)$, it is essentially a matter of figuring out how to use the provability conditions given above to deduce that $\mathsf{ZFC} \vdash G$, which would produce a contradiction due to the first incompleteness theorem.

Contrary to popular belief, the sentence $\mathrm{Con}(\mathsf{ZFC}_c)$ does *not* state that ZFC (the base theory) is consistent. After all, "ZFC is consistent" is a *meta*-mathematical statement! Instead, the sentence $\mathrm{Con}(\mathsf{ZFC}_c)$ states that the *coded* theory ZFC_c is consistent, and it is important to keep in mind that ZFC_c is not the same as ZFC.

4 Models of Set Theory (16 problems, 180 points)

It's time to get even more *meta*!

In the last section, we discussed the phenomenon of a sentence being *independent* from ZFC. But barring contrived examples like the Gödel sentence, how would we show that something like the continuum hypothesis is independent?

Perhaps it is best to begin with a simpler example from arithmetic. Suppose that we start from some "axioms" about integers, say the following:

$$(x+y)+z = x+(y+z), \qquad x+y = y+x, \qquad x+0 = x, \qquad x+(-x) = 0,$$
$$(x\cdot y)\cdot z = x\cdot(y\cdot z), \qquad x\cdot 1 = x, \qquad x\cdot 0 = 0, \qquad x\cdot(y+z) = (x\cdot y)+(x\cdot z),$$

for all x, y, z. These are known as the *ring axioms*. Now, we ask: is it possible to prove or disprove the statement $0 \neq 1$ from the ring axioms? Well, of course we can't disprove it – it's true! But it turns out that we also can't prove it.

To see why, imagine an alien civilization that uses a bizarre number system, in which *all numbers are equal*. In particular, they believe that $0 = 1$. But they also believe that all of the ring axioms are true. If we somehow came up with a proof of $0 \neq 1$ using only the ring axioms, then the proof would work equally well for the aliens' number system, and they could use it to show that $0 \neq 1$. However, that is not true in the alien number system! Therefore, such a proof cannot exist.

What we've done here is construct a new setting (a new "number system") in which all the ring axioms hold, but in which $0 \neq 1$ fails. In general, a setting in which the ring axioms hold is called a *ring*. For example, we have a ring $\mathbb{Z}$ of integers, in which $0 \neq 1$ is true. But the aliens also have a ring $\{0\}$, consisting of just one number, in which $0 \neq 1$ is false. If a statement can be proved from the ring axioms, then it must be true in all rings, and since the statement $0 \neq 1$ is true in some rings but false in others, it must be independent from the ring axioms.

We can apply the same ideas to set theory. Let's define a *model of* ZFC as a "setting" in which the axioms of ZFC hold true. (This will be made rigorous below.) In order to prove that the continuum hypothesis is independent from ZFC, all we need to do is to exhibit a model of ZFC in which CH is true, and a model of ZFC in which CH is false. Of course, this is much easier said than done...

4.1 Relativization (30 points)

Before giving a rigorous definition of a model, we need to study some aspects of logical formulas in detail.

Definition 4.1.1 — The *universal quantifier* is the symbol $\forall$, and the *existential quantifier* is the symbol $\exists$. A *quantifier* is one of these two symbols.

Let φ be a formula, written formally (so no abbreviations). If a quantifier occurs as part of the expression $\forall x$ or $\exists x$ for some variable x, then that occurrence of the quantifier in φ is *unbounded*. If a quantifier occurs as part of $(\forall x \in X)$ or $(\exists x \in X)$ for some variables x and X, then that occurrence of the quantifier is *bounded*.

For example, let's write down the axiom of union formally:

$$\forall X\, \exists U\, \forall y\, (y \in U \iff (\exists x \in X)\, y \in x).$$

There are four quantifiers in this sentence: two universal quantifiers and two existential quantifiers. The first three quantifiers are unbounded, and the last one is bounded.

Every formula can be rewritten such that all quantifiers are unbounded: we can just replace all occurrences of $(\forall x \in X)\,\varphi(x)$ with $\forall x\,(x \in X \implies \varphi(x))$ and $(\exists x \in X)\,\varphi(x)$ with $\exists x\,(x \in X \wedge \varphi(x))$.

Definition 4.1.2 — Let φ be a formula. For a class M, the *relativization* of φ to M, denoted φ^M, is the formula obtained by writing φ such that all quantifiers are unbounded, and then replacing all instances of $\forall x$ with $(\forall x \in M)$ and all instances of $\exists x$ with $(\exists x \in M)$. If φ^M is true, then we say that φ is *true in* M.

For example, let φ be the formula $Y = \mathcal{P}(X)$, which is short for

$$\forall A\,(A \in Y \iff \forall x\,(x \in A \implies x \in X)).$$

After we relativize to M, the resulting formula φ^M is

$$(\forall A \in M)(A \in Y \iff (\forall x \in M)(x \in A \implies x \in X)).$$

Intuitively, φ^M states that "M thinks that φ is true", as M can only "see" sets that are in M. Anthropomorphism is quite common in set theory :)

Problem 4.1.1 (10 points)

Let M be a class. Prove that the axiom of regularity is true in M. Prove that if M is transitive, then the axiom of extensionality is true in M.

Relativization can cause a lot of wacky phenomena. For example, consider the formula $Y = \mathcal{P}(X)$ above. If $M = \{0, 1, \{2\}\}$, and $X = 0 = \varnothing$ and $Y = 1 = \{\varnothing\}$ are elements of M, then $Y = \mathcal{P}(X)$ is true: Y is the power set of X. However, the formula does not remain true when relativized to M. Take $A = \{2\} \in M$. Then, any $x \in M$ which is in A is also in X (vacuously!), and yet, $A \notin Y$.

Even worse, if we let $M = \{0, 1, \{0, 2\}\}$ and $X = 0$, then there are *two* distinct choices of $Y \in M$ such that $Y = \mathcal{P}(X)$ is true in M: namely $Y = 1 = \{0\}$ and $Y = \{0, 2\}$. The problem is that M can't "see" the set 2, so it can't distinguish between $\{0\}$ and $\{0, 2\}$. If M is transitive, then the axiom of extensionality is true in M, thus we don't have this problem: for given $X \in M$, there is at most one $Y \in M$ satisfying $(Y = \mathcal{P}(X))^M$. But there might be no such Y at all (take $M = \{0\}$ and $X = 0$).

Definition 4.1.3 — A formula $\varphi(x_1, \ldots, x_n)$ is *absolute* for a class M if, for every $x_1, \ldots, x_n \in M$, we have $\varphi^M(x_1, \ldots, x_n)$ iff $\varphi(x_1, \ldots, x_n)$.

Absoluteness is a nice property, so we would like to find situations in which it holds. Here is one of them:

Definition 4.1.4 — A formula is called Δ_0 if all of its quantifiers are bounded.

Problem 4.1.2 (10 points)

Prove that any Δ_0 formula is absolute for any transitive class. (Hint: when working with formulas, a common strategy is to induct on the length of the formula.)

If a formula $\varphi(x_1, \ldots, x_n)$ is logically equivalent to a Δ_0 formula $\psi(x_1, \ldots, x_n)$, then it is not hard to see that $\varphi(x_1, \ldots, x_n)$ is also absolute in any transitive class. (By logically equivalent, we mean that the sentence $\forall x_1 \cdots \forall x_n \, (\varphi(x_1, \ldots, x_n) \iff \psi(x_1, \ldots, x_n))$ can be proven logically, without using the axioms of ZFC.) For example,

- "$x \subseteq y$" is short for $(\forall z \in x)\, z \in y$, which is Δ_0.
- "$x = \varnothing$" is equivalent to $\neg(\exists y \in x)\, y = y$, which is Δ_0.
- "x is transitive" is short for $(\forall y \in x)\, y \subseteq x$, which is Δ_0.
- "α is an ordinal" is equivalent to $(\forall x \in \alpha)(x \subseteq \alpha \wedge x \text{ is transitive})$, which is Δ_0.
- "$y = x \cup \{x\}$" is equivalent to $x \subseteq y \wedge x \in y \wedge (\forall z \in y)(z \in x \vee z = x)$, which is Δ_0.
- "$z = \{x, y\}$" is equivalent to $x \in z \wedge y \in z \wedge (\forall w \in z)(w = x \vee w = y)$, which is Δ_0. In particular, "$y = \{x\}$" is equivalent to a Δ_0 formula.
- "$z = (x, y)$" is equivalent to $(\exists s \in z)(\exists p \in z)(z = \{s, p\} \wedge s = \{x\} \wedge p = \{x, y\})$, so it is equivalent to a Δ_0 formula.
- "f is a function $X \to Y$" is equivalent to $(\forall x \in X)(\exists y \in Y)((\exists p \in f)\, p = (x, y) \wedge (\forall z \in Y)((\exists q \in f)\, q = (x, z) \implies y = z)) \wedge (\forall p \in f)(\exists x \in X)(\exists y \in Y)\, p = (x, y)$. If that was too complicated to parse, feel free to skip it. Of course, this is also Δ_0.
- The following are all equivalent to Δ_0 formulas: $z = x \cup y$; $z = x \cap y$; $Y = \bigcup X$; $Y = \bigcap X$; $Z = X \times Y$; R is a relation; $X = \text{dom}(R)$; $X = \text{ran}(R)$; f is a function; f is an injection/surjection/bijection $X \to Y$.

Therefore, these formulas are all absolute in any transitive class.

Of course, not *every* formula is equivalent to a Δ_0 formula. First of all, sentences are never Δ_0: since all variables in the sentence are bound by quantifiers, the sentence must start with an unbounded quantifier (possibly after a $\neg$ symbol). Second of all:

Problem 4.1.3 (10 points)

Show that the statements "$Y = \mathcal{P}(X)$" and "$X \approx Y$" and "κ is a cardinal" are not logically equivalent to any Δ_0 formula.

4.2 Working in a Model (40 points)

And finally, we get to models of ZFC.

Definition 4.2.1 — A *model of* ZFC is a class M such that every axiom of ZFC is true in M.

The statement "M is a model of ZFC" cannot be written as a single formula. Instead, it consists of infinitely many formulas, one for each axiom of ZFC. So when we say that "ZFC proves that M is a model of ZFC", we really mean that for every axiom φ of ZFC, we have $\text{ZFC} \vdash \varphi^M$.

Problem 4.2.1 (5 points)

Show, in ZFC, that the universe V is a model of ZFC.

Lemma 4.2.2
Suppose that ZFC proves that M is a model of ZFC. If $\mathsf{ZFC} \vdash \varphi$, then $\mathsf{ZFC} \vdash \varphi^M$.

Proof. Suppose that we have a proof of φ from the axioms of ZFC, where each step is either a logical deduction or an axiom of ZFC. We relativize each step of the proof to M. The logical deductions remain valid, and by assumption, the relativized axioms are all provable in ZFC. Therefore, we get a proof of φ^M. □

In a model M of ZFC, relativization behaves much more nicely, and we avoid some of the pathological phenomena discussed in the previous subsection. For example, ZFC proves $\forall X\, \exists! Y\, (Y = \mathcal{P}(X))$, so it remains true in M as well: for all $X \in M$, there is a unique $Y \in M$ such that $(Y = \mathcal{P}(X))^M$. This Y is what M "thinks" the power set of X is, so we denote it as $\mathcal{P}^M(X)$.

In general, if we have a formula $\varphi(x_1, \ldots, x_n, y)$ such that

$$\mathsf{ZFC} \vdash \forall x_1 \cdots \forall x_n\, \exists! y\, \varphi(x_1, \ldots, x_n, y),$$

then we can define some operator F and write $y = F(x_1, \ldots, x_n)$ as an abbreviation of $\varphi(x_1, \ldots, x_n, y)$. Indeed, we have defined many operators like this. Now, given a model M of ZFC, we know by Lemma 4.2.2 that

$$\mathsf{ZFC} \vdash (\forall x_1 \in M) \cdots (\forall x_n \in M)(\exists! y \in M)\, \varphi^M(x_1, \ldots, x_n, y),$$

so for $x_1, \ldots, x_n \in M$, we shall write $y = F^M(x_1, \ldots, x_n)$ in place of $\varphi^M(x_1, \ldots, x_n, y)$, thus defining the relativized operator F^M. For example, we can relativize the power set operator (see above), and we also have things like $x \cup^M y$. Constants, such as $\varnothing$ and ω, can be treated as operators on $n = 0$ inputs, and so they can also be relativized to M in the same way, giving us $\varnothing^M$ and ω^M.

Similarly, if we have parameters $p_1, \ldots, p_n \in M$ and a class $C = \{x : \varphi(x, p_1, \ldots, p_n)\}$, then we can relativize C by defining $C^M = \{x \in M : \varphi^M(x, p_1, \ldots, p_n)\}$. For example, we have a class Ord^M of ordinals "according to M".

Definition 4.2.3 — Let M be a model of ZFC. For an operator F, defined so that $y = F(x_1, \ldots, x_n)$ is short for $\varphi(x_1, \ldots, x_n, y)$, we say that F is *absolute* for M if the formula $\varphi(x_1, \ldots, x_n, y)$ is absolute for M.

For parameters $p_1, \ldots, p_n \in M$ and a class $C = \{x : \varphi(x, p_1, \ldots, p_n)\}$, we say that C is *absolute* for M if the formula $\varphi(x, p_1, \ldots, p_n)$ is absolute for M.

In other words, an operator F is absolute for M iff $F^M(x_1, \ldots, x_n) = F(x_1, \ldots, x_n)$ for all $x_1, \ldots, x_n \in M$, and a class C is absolute for M iff $C^M = C \cap M$.

Of course, we can't expect absoluteness to always hold. For example, observe that by definition, $\varnothing^M$ doesn't contain any elements of M, but it might contain things not in M that M cannot "see", so $\varnothing^M$ is not necessarily empty.

However, if M is a *transitive* model of ZFC, then absoluteness does hold pretty often. In fact, we have a stronger version of Problem 4.1.2.

Definition 4.2.4 — A formula $\varphi(x_1, \ldots, x_n)$ is Δ_0^{ZFC} if there is some Δ_0 formula $\psi(x_1, \ldots, x_n)$ such that $\mathsf{ZFC} \vdash \forall x_1 \cdots \forall x_n\, (\varphi(x_1, \ldots, x_n) \iff \psi(x_1, \ldots, x_n))$. (That is: φ is provably equivalent to some Δ_0 formula.)

Problem 4.2.2 (5 points)
Let M be a transitive model of ZFC. Show that any Δ_0^{ZFC} formula is absolute for M.

Clearly, if a formula is logically equivalent to a Δ_0 formula, then it is Δ_0^{ZFC}. Looking back at our list in the previous subsection, there are a lot of absolute operators:

$$\varnothing^M = \varnothing, \quad \{x\}^M = \{x\}, \quad \{x,y\}^M = \{x,y\}, \quad (x,y)^M = (x,y),$$
$$\bigcup^M X = \bigcup X, \quad \bigcap^M X = \bigcap X, \quad x \cup^M y = x \cup y, \quad x \cap^M y = x \cap y,$$
$$X \times^M Y = X \times Y, \quad \mathrm{dom}^M(R) = \mathrm{dom}(R), \quad \mathrm{ran}^M(R) = \mathrm{ran}(R),$$

and absolute classes:

$$\mathrm{Ord}^M = \mathrm{Ord} \cap M, \quad \mathcal{P}^M(X) = \mathcal{P}(X) \cap M, \quad \mathrm{Hom}^M(X,Y) = \mathrm{Hom}(X,Y) \cap M.$$

And there are even more situations where absoluteness holds.

Problem 4.2.3 (10 points)
Show that the statements "α is a successor ordinal" and "α is a limit ordinal" and "α is a natural number" and "$\alpha = \omega$" are Δ_0^{ZFC}. Conclude that if M is a transitive model of ZFC, then $\omega^M = \omega$.

Problem 4.2.4 (5 points)
Let M be a transitive model of ZFC, and G be a class function on V such that G is absolute for M. By transfinite recursion, we obtain a (unique) class function F on Ord such that $F(\alpha) = G(F\restriction\alpha)$. Show that F is absolute for M.

That is: transfinite recursion preserves absoluteness. In particular, it is easily checked that ordinal addition and multiplication are absolute in any transitive model of ZFC.

The following instance of absoluteness is slightly trickier to show:

Problem 4.2.5 (15 points)
Prove that "x is a finite set" is absolute in any transitive model of ZFC.

4.3 The von Neumann Hierarchy (70 points)

As we have seen, the universe V is trivially a model of ZFC. But that's kind of useless. We now begin our quest of constructing nontrivial models of ZFC.

Definition 4.3.1 — The *von Neumann hierarchy* consists of the sets V_α (where α is an ordinal), which are defined recursively as

- $V_0 = \varnothing$,
- $V_{\alpha+1} = \mathcal{P}(V_\alpha)$,
- $V_\alpha = \bigcup_{\beta<\alpha} V_\beta$ if α is a limit ordinal.

The von Neumann hierarchy is named after John von Neumann, but it was actually first defined by Ernst Zermelo in 1930. Some people call it the *cumulative hierarchy*.

The first few levels of the hierarchy are $V_0 = \varnothing$, then $V_1 = \{\varnothing\}$, then $V_2 = \{\varnothing, \{\varnothing\}\}$, then $V_3 = \{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \{\varnothing, \{\varnothing\}\}\}$. After this, the levels quickly become impossible to write down: V_6 has 2^{65536} elements! But at least V_n is finite for $n < \omega$.

Problem 4.3.1 (10 points)

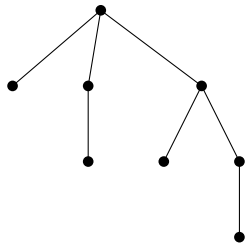
Let α, β be ordinals. Prove that

(a) V_α is transitive.

(b) If $\alpha \leq \beta$, then $V_\alpha \subseteq V_\beta$.

(c) $\alpha \in V_{\alpha+1}$ but $\alpha \notin V_\alpha$.

A nice way to visualize the elements of V_ω is to draw them as *trees*. To draw the tree for a set x, we first draw a root node. Then, we draw the trees for every element y of x, and connect the roots of these trees to the root node. (Note that the tree for $\varnothing$ is just a single root node.) For example, the tree for $3 = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$ is



If $x \in V_n$ for some natural number n, then every element of x must lie in V_k for some $k < n$, so by induction, it is not hard to show that the tree for x contains only finitely many nodes. Moreover, the "height" of the tree (the length of the longest branch) turns out to be the smallest m such that $x \in V_{m+1}$. For example, we have $3 \in V_4$ but $3 \notin V_3$, and indeed, the height of the tree above is 3.

In general, we can still draw the tree even if $x \notin V_\omega$, but then it would no longer have finitely many nodes, and our intuition breaks down.

Theorem 4.3.2

Every set is in V_α for some ordinal α.

In other words, the universe V is the union of V_α over all ordinals α. How do we prove this? The idea is that if there is a set which is not in any V_α, then we can try to find an $\in$-minimal set among all such sets, and then derive a contradiction.

Unfortunately, the axiom of regularity only applies to sets, not proper classes, so we don't get an $\in$-minimal set for free. Let's fix that right now:

Problem 4.3.2 (30 points)

Let C be a nonempty class. Prove that C has an $\in$-minimal element, that is, a set $x \in C$ such that $y \notin x$ for any $y \in C$. Using this, finish the proof of Theorem 4.3.2. (Hint: first try to prove this under the assumption that C contains a transitive set.)

Does the von Neumann hierarchy give us models of ZFC? Well, *almost*:

Problem 4.3.3 (20 points)
Let α be a limit ordinal greater than ω. Verify that all of the axioms of ZFC, with the possible exception of the axiom of replacement, are true in V_α.

However, it is hard to get V_α to satisfy the axiom of replacement. This axiom schema is very powerful: it's the whole reason we can do transfinite recursion, and we have seen that transfinite recursion allows us to do a lot of things, such as constructing really big ordinals. So we'd expect that α needs to be quite large.

Definition 4.3.3 — An infinite cardinal κ is a *strong limit cardinal* if $2^\lambda < \kappa$ for every cardinal $\lambda < \kappa$. A cardinal is *inaccessible* if it is uncountable, regular, and a strong limit cardinal.

Just how large are these beasts? Well, if $\aleph_\alpha$ is inaccessible, then first of all, α must be a limit ordinal (otherwise, if $\alpha = \beta + 1$, then we have $\aleph_\beta < \kappa$ but $2^{\aleph_\beta} \geq \kappa$). Next, since κ is regular, we have $\aleph_\alpha = \operatorname{cf} \aleph_\alpha = \operatorname{cf} \alpha \leq \alpha \leq \aleph_\alpha$. That is, $\aleph_\alpha = \alpha$ – it is a *fixed point* of the aleph function! The smallest such fixed point is denoted

$$\Phi(1, 0) = \sup\{\omega, \omega_\omega, \omega_{\omega_\omega}, \dots\},$$

but that has cofinality $\aleph_0$, and is nowhere near being regular, so inaccessible cardinals are much larger than that. In fact, in some sense, they are larger than anything we can ever write down explicitly.

Problem 4.3.4 (10 points)
Prove that if κ is an inaccessible cardinal, then V_κ is a model of ZFC.

4.4 The Constructible Universe (40 points)

In 1938, Kurt Gödel sent ripples through the mathematical community once again, by constructing a model of ZFC – the *constructible universe* L – in which the generalized continuum hypothesis holds. As you will later show, this means that if ZFC is consistent, then it cannot disprove CH.

Gödel realized the von Neumann hierarchy grows a bit too quickly to control – at level $\omega + 1$, we already have $|V_{\omega+1}| = 2^{|V_\omega|} = \mathfrak{c}$, and the sets V_α get large even more quickly after that. So, he tried to slow down this growth rate, so that we have more control over how cardinals behave.

Consider the set ω. It has uncountably many subsets, but not all of them are *definable*, because there are only countably many "definitions". Gödel defined a hierarchy similar to the von Neumann hierarchy, but instead of taking the power set at each step, we only take the set of all definable subsets. More precisely:

Definition 4.4.1 — Let X be a set. A subset $A \subseteq X$ is *definable* if there exists a natural number n, a family $(p_i)_{1 \leq i \leq n}$ of elements of X, and a Gödel number of a formula $\varphi(x, p_1, \dots, p_n)$, such that A is equal to $\{x \in X : \varphi^X(x, p_1, \dots, p_n)\}$. The set of definable subsets of X is denoted $\operatorname{Def}(X)$.

The awkward phrasing here is to emphasize that we are formalizing "A is a definable subset of X", which is naïvely a meta-mathematical concept, as a formula. The details of how this formalization is done are absolutely cursed, so we do not attempt to describe them. All you need to know is the following:

Lemma 4.4.2

Let $\varphi(x, p_1, \ldots, p_n)$ be a formula. Then for any set X and elements $p_1, \ldots, p_n \in X$, we have $\{x \in X : \varphi^X(x, p_1, \ldots, p_n)\} \in \mathrm{Def}(X)$.

This lemma (or more precisely, a *schema* of infinitely many lemmas, one for each φ) is not trivial! The number n mentioned in the lemma is a meta natural number, not an element of ω. The lemma also does not trivially imply that

Lemma 4.4.3

For any set X, the set $\mathrm{Def}(X)$ contains all finite subsets of X.

We are finally ready to define the constructible universe.

Definition 4.4.4 — The *constructible hierarchy* consists of the sets L_α (where α is an ordinal), which are defined recursively as

- $L_0 = \varnothing$,
- $L_{\alpha+1} = \mathrm{Def}(L_\alpha)$,
- $L_\alpha = \bigcup_{\beta<\alpha} L_\beta$ if α is a limit ordinal.

The *constructible universe* L is defined as the union of L_α over all ordinals α, and the elements of L are called *constructible sets*.

Problem 4.4.1 (5 points)

Prove that $L_\alpha \subseteq V_\alpha$ for any ordinal α, and $L_\alpha = V_\alpha$ if $\alpha \leq \omega$.

Problem 4.4.2 (15 points)

Let α, β be ordinals. Prove that

(a) L_α is transitive.

(b) If $\alpha \leq \beta$, then $L_\alpha \subseteq L_\beta$.

(c) $\alpha \in L_{\alpha+1}$ but $\alpha \notin L_\alpha$.

Therefore, L is transitive, and $\mathrm{Ord} \subseteq L$. Next, we show that L is a model of ZFC.

Theorem 4.4.5

The axioms of ZF (i.e. ZFC without the axiom of choice) are true in L.

Most of the proof is not too different to that of Problem 4.3.3, but there appear to be some annoying things with the axiom of separation. We omit the proof.

It is much harder to prove that the axiom of choice is true in L. Instead of proving it directly, we introduce another statement: the *axiom of constructibility*, which states that $V = L$, in other words, that every set is constructible.

Lemma 4.4.6

The axiom of constructibility, $V = L$, is true in L.

Sketch of proof. The first thing to show is that the operator Def is absolute for L. (This is not hard once we have a precise definition of Def, but we don't.) It then follows from Problem 4.2.4 that the constructible hierarchy is absolute for L (that is, $(L_\alpha)^L = L_\alpha$). Therefore, $L^L = L$, as desired. □

Thus, it suffices to show that the axiom of choice follows from $V = L$. Of course, we must work in ZF (without the axiom of choice) for this proof. Note that the axiom of choice is mainly used to deduce things about cardinals, so even without it, we can still define L and prove its basic properties.

Theorem 4.4.7

In ZF, the axiom of constructibility implies the axiom of choice.

Sketch of proof. It is possible, albeit extremely tedious, to construct a class relation $<_L$ on L which is a well-order. (We say that a class relation is a *well-order* if it satisfies the conditions in Definition 2.3.1 but with classes instead of sets.)

If $V = L$, then we have constructed a well-order of the entire universe. The axiom of choice then follows: if X is a set of nonempty sets, then we can define a choice function f by setting $f(x)$ to be the least element of x with respect to $<_L$. □

Therefore, the axiom of choice is true in L, and thus, L is a model of ZFC. It remains to show that GCH is true in L, and for that, we will make use of the following lemma. For an ordinal α, let α^+ denote the least cardinal greater than α.

Lemma 4.4.8

If $V = L$, then $\mathcal{P}(L_\alpha) \subseteq L_{\alpha^+}$ for all infinite ordinals α.

The proof is omitted, as it requires results that we have not introduced.

Problem 4.4.3 (10 points)

Show that if X is infinite, then $|\mathrm{Def}(X)| = |X|$. Conclude that $|V_\alpha| = |\alpha|$ for every infinite ordinal α. (Hint: recall that Gödel numbers are natural numbers.)

Problem 4.4.4 (10 points)

Prove that $V = L$ implies the generalized continuum hypothesis. Conclude that if ZFC is consistent, then $\mathsf{ZFC} \nvdash \neg\mathsf{CH}$.

5 Forcing (11 problems, 160 points)

So far, we have a wonderful proof that the continuum hypothesis cannot be disproved in ZFC. But what about the other part – that CH cannot be proved in ZFC? Can we also construct a model in which CH is false, just like how we constructed L?

Alas, this cannot work due to some complicated meta-mathematical reasons, and we need a new idea. In 1963, Paul Cohen introduced the technique of *forcing* (for which he won a Fields Medal): instead of constructing a model of ZFC from scratch, we assume that we already have a *countable* transitive model M, and we use it to construct a new model in which CH fails.

To explain some of the intuition behind forcing, we turn to arithmetic again. Say that we want to find a ring in which the statement $\exists x\,(x \neq 0 \wedge x \cdot x = 0)$ is true. So we start with the ring $\mathbb{Z}$, in which the statement is false, and we try to *force* the statement to be true by shoving an extra object ε into the ring and declaring that $\varepsilon \cdot \varepsilon = 0$. To be sure we still get a ring, we also throw in everything of the form $a + b\varepsilon$ for $a, b \in \mathbb{Z}$, and voilà – we indeed get a ring, denoted $\mathbb{Z}[\varepsilon]$. It consists of the "numbers" $a + b\varepsilon$ for $a, b \in \mathbb{Z}$, and it is the "smallest" ring containing all integers and also ε.

Doing something like this with a model of ZFC is of course much harder. (But that's why Cohen won a Fields Medal for figuring it out!) Say we start off with some countable transitive model M, and we want to find a new model N such that CH fails in N. Now notice that the failure of CH is equivalent to "there exist $\aleph_2$ distinct functions $\omega \to 2$", so we can try to *force* that to be true by shoving an extra object f into the model and declaring that f is an injective function $\omega_2^M \to \mathrm{Hom}^M(\omega, 2)$.

However, that doesn't quite work. In our example with rings, we could construct $\mathbb{Z}[\varepsilon]$ without any problems because we know *exactly* how arithmetic with ε should work. But for f, it's not clear at all how it would interact with the elements of M. If we aren't able to figure that out, then we won't be able to construct our model.

We are on the right track, though. To make this construction work, we first make a small alteration: instead of thinking about functions from ω_2^M to $\mathrm{Hom}^M(\omega, 2)$, we think about functions from $\omega_2^M \times \omega$ to 2, which are basically equivalent but more convenient to work with. Next, instead of trying to shove such a function in all at once, we instead insert many bits and pieces of such a function, and rig the construction in such a way that they assemble themselves correctly.

More precisely, we define a *finite partial function* from $\omega_2^M \times \omega$ to 2 to be a function p whose domain is a finite subset of $\omega_2^M \times \omega$, and whose range is a subset of 2. These will be our *forcing conditions*: oversimplifying a bit, a finite partial function $p : \omega_2^M \times \omega \to 2$ is supposed to force the existence of an actual function $f : \omega_2^M \times \omega \to 2$ such that $p \subseteq f$. So, if we have two finite partial functions p, q such that $p \subsetneq q$, then we can think of q as a "stronger" forcing condition than p.

Crucially, these finite partial functions turn out to be nice and simple enough that we can describe how they interact with M. By choosing just the right set of finite partial functions to throw into the model, the result will be very well-behaved, and we'll be able to force the failure of CH. We will describe all of the details below.

5.1 Names and Interpretation (40 points)

We start by describing how forcing works in full generality. We will return to the specific example of finite partial functions later.

Recall from Definition 2.3.1 that a *poset* is an ordered pair $\mathbb{P} = (P, <)$, where $<$ is a partial order on P. Our setup for forcing will involve a certain kind of poset, so we first say some general things about posets.

Definition 5.1.1 — Let $(P, <)$ be a poset. For $p, q \in P$, we write $p \leq q$ to mean $p < q \vee p = q$. An element $p \in P$ is a *largest element* if $q \leq p$ for all $q \in P$.

From this definition, we can easily prove the following facts: if $p, q, r \in P$, then $p \leq p$; if $p \leq q$ and $q \leq p$, then $p = q$; if $p \leq q$ and $q \leq r$, then $p \leq r$.

If a largest element exists in a poset, then it must be unique. Indeed, if p_1, p_2 are both largest elements, then $p_1 \leq p_2$ and $p_2 \leq p_1$, so $p_1 = p_2$.

Definition 5.1.2 — A *forcing notion* is a poset $\mathbb{P} = (P, <)$ with a largest element, denoted $1_{\mathbb{P}} \in P$. The elements of P are called *forcing conditions*. For $p, q \in P$, we say that p is *stronger* than q if $p < q$.

For the remainder of this subsection, let M be a countable transitive model of ZFC, and $\mathbb{P} = (P, <)$ be a forcing notion such that $\mathbb{P} \in M$. In particular, since M is transitive, this means that both the set P and the relation $<$ are in M.

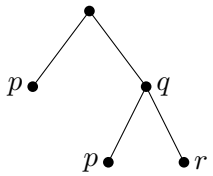
Our goal is to add a subset $G \subseteq P$ to our model M, to produce a bigger model $M[G]$. It turns out that there are certain restrictions on what subsets G we can use, but we'll worry about that later.

Definition 5.1.3 — Define the sets $M_\alpha^{\mathbb{P}}$, where α is an ordinal in M, recursively as

- $M_0^{\mathbb{P}} = \varnothing$,
- $M_{\alpha+1}^{\mathbb{P}} = \mathcal{P}^M(M_\alpha^{\mathbb{P}} \times P)$,
- $M_\alpha^{\mathbb{P}} = \bigcup_{\beta<\alpha} M_\beta^{\mathbb{P}}$ if α is a limit ordinal.

Lastly, define the set $M^{\mathbb{P}}$ of $\mathbb{P}$-*names* as the union of $M_\alpha^{\mathbb{P}}$ over all ordinals $\alpha \in M$.

For example, if $p, q, r \in P$, then $\{(\varnothing, p), (\{(\varnothing, p), (\varnothing, r)\}, q)\}$ is a $\mathbb{P}$-name, which lies in $M_3^{\mathbb{P}}$. Intuitively, a $\mathbb{P}$-name is a set, where everything in it is "labeled" with a forcing condition. Once again, we can use trees to visualize $\mathbb{P}$-names. This time, all the nodes of the tree, except the root, will be labeled with a forcing condition. For example, the tree for the $\mathbb{P}$-name above looks like



After we build these trees, we shall prune them.

Definition 5.1.4 — Let G be a subset of P containing $1_{\mathbb{P}}$. The *interpretation* of a $\mathbb{P}$-name τ, denoted τ^G, is defined as $\tau^G = \{\sigma^G : (\exists p \in G)\,(\sigma, p) \in \tau\}$ via transfinite recursion. Finally, let $M[G] = \{\tau^G : \tau \in M^{\mathbb{P}}\}$.

More precisely, the "transfinite recursion" involves recursively defining interpretation functions on $M_\alpha^{\mathbb{P}}$ for each ordinal $\alpha \in M$, and making sure that they all agree, which is routine but tedious. We omit the details.

For example, if we use the subset $G = \{q, r\}$, and $\tau = \{(\varnothing, p), (\{(\varnothing, p), (\varnothing, r)\}, q)\}$ is the $\mathbb{P}$-name used above, then the interpretation is $\tau^G = \{\{\varnothing\}\}$. Visually, if τ is drawn as a labeled tree, then to obtain τ^G, we destroy any node whose label is not in G (as well as the entire subtree below that node), and erase the labels on the remaining nodes.

Problem 5.1.1 (10 points)
Prove that $M \subseteq M[G]$ and $G \in M[G]$.

Problem 5.1.2 (15 points)
Prove that $M[G]$ is a countable transitive set containing the same ordinals as M.

Problem 5.1.3 (15 points)
Verify that the axioms of extensionality, pairing, union, infinity, and regularity are all true in $M[G]$.

If N is a transitive model of ZFC such that $M \subseteq N$ and $G \in N$, then it can be shown that interpretation is absolute for N. Hence, if $\tau \in M^{\mathbb{P}}$, then $\tau^G = (\tau^G)^N \in N$, which implies that $M[G] \subseteq N$.

That is, if $M[G]$ is a model of ZFC, then it really is the *smallest* model containing all elements of M and also G (although we won't need this fact). However, $M[G]$ might not be a model of ZFC! We will need to impose more complicated conditions on G for the remaining axioms to be true in $M[G]$.

5.2 The Forcing Relation (40 points)

As always, M is a countable transitive model of ZFC, and $\mathbb{P} = (P, <)$ is a forcing notion.

As we explained in the introduction to this section, the main obstacle we now face is to figure out a way to describe how the elements of $M[G]$ interact with each other. For our example with rings, arithmetic in $\mathbb{Z}[\varepsilon]$ can be reduced to arithmetic in $\mathbb{Z}$:

$$(a + b\varepsilon) + (c + d\varepsilon) = (a + c) + (b + d)\varepsilon,$$
$$(a + b\varepsilon) \cdot (c + d\varepsilon) = ac + (ad + bc)\varepsilon,$$

and because of this, it is possible to work in $\mathbb{Z}[\varepsilon]$, even if "ε" is not something that we have ever encountered before. Similarly, to work in $M[G]$, we want a way to describe the behavior of the elements of $M[G]$, using only things that M can "see".

So far, notice that we have not used the relation $<$ on P at all! Indeed, in general, it is not possible to describe the behavior of $M[G]$. It is only possible if G satisfies some technical conditions involving the relation $<$ on P.

Definition 5.2.1 — Two forcing conditions $p, q \in P$ are *compatible* if there exists some forcing condition $r \in P$ such that $r \leq p$ and $r \leq q$. A nonempty subset $G \subseteq P$ is a *filter* if the following conditions hold:

(1) If $p \in G$, then $q \in G$ for every $q \geq p$, and

(2) Any two elements of G are compatible.

In particular, any filter must contain $1_{\mathbb{P}}$ (by condition (1)).

Definition 5.2.2 — A subset $D \subseteq P$ is *dense* if, for any forcing condition $p \in P$, there exists a forcing condition $q \in D$ such that $q \leq p$. A filter $G \subseteq P$ is *generic* if it intersects all dense sets $D \subseteq P$ which lie in M (that is, $G \cap D \neq \varnothing$).

Problem 5.2.1 (15 points)
Prove that a generic filter exists.

Now, if G is a generic filter, then we *can* describe the behavior of $M[G]$, using a piece of black magic known as the *forcing relation*. If we have a forcing condition $p \in P$, some $\mathbb{P}$-names $\tau_1, \ldots, \tau_n \in M^{\mathbb{P}}$, and a formula $\varphi(\tau_1, \ldots, \tau_n)$, then we write

$$p \Vdash \varphi(\tau_1, \ldots, \tau_n),$$

read "p *forces* $\varphi(\tau_1, \ldots, \tau_n)$", to mean that if G is any generic filter containing p, then $\varphi(\tau_1^G, \ldots, \tau_n^G)$ is true in $M[G]$.

In general, a generic filter G will not be an element of M, so we might expect that M is unable to "understand" the forcing relation. However, there is a way to formalize the forcing relation to only mention elements of M! The formalization is given below, and is horribly complicated. Feel free to stare at it until it starts to make more sense, or skip it altogether – you won't need to use it to prove anything.

Definition 5.2.3 — Let $p \in P$ be a forcing condition, $\tau_1, \ldots, \tau_n \in M^{\mathbb{P}}$ be $\mathbb{P}$-names, and $\varphi(\tau_1, \ldots, \tau_n)$ be a formula, written such that it only uses $=, \in, \neg, \wedge, \exists$ (note that this is always possible by Problem 1.1.1) and has no bounded quantifiers.

We define the *forcing relation* $p \Vdash \varphi(\tau_1, \ldots, \tau_n)$ recursively as follows. A subset $D \subseteq P$ is *dense below* p if, for any $q \leq p$, there exists $r \in D$ such that $r \leq q$. Then

- $p \Vdash \tau_1 = \tau_2$ is defined as "for all $(\pi_1, s_1) \in \tau_1$, the set
$$\{q \in P : q \leq s_1 \implies (\exists (\pi_2, s_2) \in \tau_2)(q \leq s_2 \wedge q \Vdash \pi_1 = \pi_2)\}$$
is dense below p, and for all $(\pi_2, s_2) \in \tau_2$, the set
$$\{q \in P : q \leq s_2 \implies (\exists (\pi_1, s_1) \in \tau_1)(q \leq s_1 \wedge q \Vdash \pi_1 = \pi_2)\}$$
is dense below p".
- $p \Vdash \tau_1 \in \tau_2$ is defined as "the set $\{q \in P : (\exists (\pi, s) \in \tau_2)(q \leq s \wedge q \Vdash \pi = \tau_1)\}$ is dense below p".
- $p \Vdash \neg\varphi$ is defined as "$q \Vdash \varphi$ is false for every $q \leq p$".
- $p \Vdash \varphi \wedge \psi$ is defined as "$p \Vdash \varphi$ and $p \Vdash \psi$".
- $p \Vdash \exists \tau\, \varphi(\tau, \sigma_1, \ldots, \sigma_n)$ is defined as "$\{q \in P : (\exists \tau \in M^{\mathbb{P}})\, q \Vdash \varphi(\tau, \sigma_1, \ldots, \sigma_n)\}$ is dense below p".

The point is that every object mentioned in this definition is an element of M, so the forcing relation is really something that M can "see".

For the rest of this subsection, let G be some fixed generic filter. We promised that if $p \in G$ and $p \Vdash \varphi(\tau_1, \ldots, \tau_n)$, then $\varphi(\tau_1^G, \ldots, \tau_n^G)$ is true in $M[G]$. In fact, the situation is even nicer than that, because the converse also holds:

Theorem 5.2.4 (Fundamental theorem of forcing)
Let $\tau_1, \ldots, \tau_n \in M^{\mathbb{P}}$ and $\varphi(\tau_1, \ldots, \tau_n)$ be a formula. Then $\varphi(\tau_1^G, \ldots, \tau_n^G)$ is true in $M[G]$ iff there exists some forcing condition $p \in G$ such that $p \Vdash \varphi(\tau_1, \ldots, \tau_n)$.

Now, we can verify that the remaining axioms of ZFC (namely, separation, power set, replacement, and choice) are true in $M[G]$.

Problem 5.2.2 (25 points)
Prove that $M[G]$ is a countable transitive model of ZFC.

5.3 Adding Cohen Reals (80 points)

At last, we get to break the continuum hypothesis!

Definition 5.3.1 — Let I, J be sets. A *finite partial function* $I \to J$ is a function f such that $\mathrm{dom}(f)$ is finite, $\mathrm{dom}(f) \subseteq I$, and $\mathrm{ran}(f) \subseteq J$. (In other words, f is a function, and is a finite subset of $I \times J$.)

The set of finite partial functions $I \to J$ is denoted $\mathrm{Fn}(I, J)$. We put a partial order $<$ on $\mathrm{Fn}(I, J)$ via reverse inclusion: for $p, q \in \mathrm{Fn}(I, J)$, write $p < q$ iff q is a proper subset of p. In an abuse of notation, we shall also write $\mathrm{Fn}(I, J)$ to denote the resulting poset $(\mathrm{Fn}(I, J), <)$.

Notice that $\mathrm{Fn}(I, J)$ is a forcing notion: its largest element is the empty set (which is, of course, a finite partial function $I \to J$).

As before, let M be a countable transitive model of ZFC. Recall from Problem 4.2.5 that finiteness is absolute for M, so if $I, J \in M$, then $\mathrm{Fn}^M(I, J) = \mathrm{Fn}(I, J)$. To break the continuum hypothesis, we will use the forcing notion

$$\mathrm{Fn}(\omega_2^M \times \omega, 2) = \mathrm{Fn}(\omega_2 \times \omega, 2)^M \in M$$

that we alluded to in the introduction of this section. Finally, let $G \subseteq \mathrm{Fn}(\omega_2^M \times \omega, 2)$ be a generic filter.

Problem 5.3.1 (10 points)
Prove that $\bigcup G$ is a function $\omega_2^M \times \omega \to 2$ which lies in $M[G]$.

Problem 5.3.2 (15 points)
Construct an injective function $\omega_2^M \to \mathrm{Hom}^{M[G]}(\omega, 2)$ which lies in $M[G]$.

The $\aleph_2^M$-many functions $\omega \to 2$ that we have created are often called *Cohen reals*. Of course, they are not actually elements of $\mathbb{R}$ (in any sense), but set theorists like to call them "reals" because $\mathrm{Hom}(\omega, 2)$ is equinumerous to $\mathbb{R}$.

It looks like we're *very* close to proving that CH is false in $M[G]$! But we still need to show that $(\aleph_2)^{M[G]}$ and $\aleph_2^M$ are equal, and it turns out that this is quite tricky.

Definition 5.3.2 — Let $\mathbb{P}$ be any forcing notion. A subset $A \subseteq P$ is called a *strong antichain* if any two distinct forcing conditions $p, q \in A$ are incompatible. We say that $\mathbb{P}$ has the *countable chain condition* (or just the *c.c.c.*) if any strong antichain A is finite or countable (that is, $|A| \leq \aleph_0$).

Since $\mathrm{Fn}(\omega_2^M \times \omega, 2)$ is countable, it obviously has the countable chain condition. A more interesting question is whether or not "$\mathrm{Fn}(\omega_2^M \times \omega, 2)$ has c.c.c." is true in M. It turns out that this is true, and we will prove it below.

We start with the Δ-*system lemma*, introduced by Nikolai Shanin in 1946.

Theorem 5.3.3 (Shanin)

Let X be an uncountable set of finite sets. Then there exists an uncountable subset $Y \subseteq X$ and a finite set R, such that $A \cap B = R$ for any distinct $A, B \in Y$.

The set Y is called a Δ-*system*, and R is its *root*.

Proof. We first prove the result under the assumption that all elements of X have the same cardinality n. To do this, we induct on n. If $n = 1$, then we just take $Y = X$ and $R = \varnothing$. Now, assuming that the result holds for n, we prove it for $n + 1$.

For every set a, let $X_a = \{A \in X : a \in A\}$. If some X_a is uncountable, then we apply the inductive hypothesis to the uncountable set $X' = \{A \setminus \{a\} : A \in X_a\}$, to obtain an uncountable Δ-system $Y' \subseteq X'$ with root R'. We see that $Y = \{A' \cup \{a\} : A' \in Y'\}$ is an uncountable Δ-system with root $R = R' \cup \{a\}$.

If all X_a are finite or countable, then we define a family $(A_\alpha)_{\alpha<\omega_1}$ of elements of X recursively: given the values of A_β for $\beta < \alpha$, the union $\bigcup_{\beta<\alpha} A_\beta$ is countable, and the set $\{A \in X : (\exists \beta < \alpha)\, A \cap A_\beta \neq \varnothing\} = \bigcup\{X_a : a \in \bigcup_{\beta<\alpha} A_\beta\}$ is thus also countable, so we can pick A_α to be disjoint from the A_β for all $\beta < \alpha$. (This is formalized using the axiom of choice.) We get an uncountable Δ-system with root $R = \varnothing$.

This completes the induction, and now we prove the full theorem. For $n < \omega$, we let $X_{(n)} = \{A \in X : |A| = n\}$, so that $X = \bigcup_{n<\omega} X_{(n)}$. If all $X_{(n)}$ were countable, then X would be countable, so some $X_{(n)}$ has to be uncountable. We can then apply our result to $X_{(n)}$ to obtain an uncountable Δ-system. □

Problem 5.3.3 (15 points)

Prove that if $|J| \leq \aleph_0$, then $\mathrm{Fn}(I, J)$ has the countable chain condition.

In particular, $\mathrm{Fn}(\omega_2 \times \omega, 2)$ has the countable chain condition, so relativizing to M, it follows that "$\mathrm{Fn}(\omega_2^M \times \omega, 2)$ has c.c.c." is true in M. Next, we show that M is able to somewhat control the behavior of functions in $M[G]$.

Problem 5.3.4 (20 points)

Let $X, Y \in M$, and let $f : X \to Y$ be in $M[G]$ but not M. Prove that there exists some $F : X \to \mathcal{P}(Y)$ in M, such that $f(x) \in F(x)$ and $|F(x)|^M \leq \aleph_0$ for all $x \in X$. (Hint: use the forcing relation and the c.c.c. to control the behavior of f.)

With these results, we are finally able to break the continuum hypothesis!

Problem 5.3.5 (15 points)

Prove that "κ is a cardinal" is true in $M[G]$ iff it is true in M. Conclude that the continuum hypothesis is false in $M[G]$.

Are we done yet? Not 100% – we used a countable transitive model M to start with, and the existence of such a model is a very strong statement (stronger, in fact, than the consistency of ZFC). To finish, we need another result.

Theorem 5.3.4 (Reflection principle)

For any sentence φ, there is a countable transitive set X such that $\varphi^X \iff \varphi$.

Problem 5.3.6 (5 points)

Show that given a *finite* list of axioms of ZFC, there exists a countable transitive set in which these axioms are true.

This seems innocent enough – we can only make finitely many axioms true at a time, so what? But remember that *a proof only ever uses finitely many axioms*, so this result is actually very powerful, as we will see below.

Theorem 5.3.5 (Cohen)

If ZFC is consistent, then $\mathsf{ZFC} \nvdash \mathsf{CH}$.

Proof. Suppose that we have a proof of CH from the axioms of ZFC. We know that if we assume the existence of a countable transitive model M of ZFC, then we would be able to construct proofs of $\mathsf{CH}^{M[G]}$ (by Lemma 4.2.2) and $\neg\mathsf{CH}^{M[G]}$ (by Problem 5.3.5), and produce a contradiction.

Observe that producing this contradiction required only finitely many axioms of ZFC relativized to M. By Problem 5.3.6, there exists a countable transitive set X in which these axioms are true. We may repeat the proof above, but using X in place of M, to deduce a contradiction. This contradicts our assumption that ZFC is consistent. □

Therefore, if ZFC is consistent, then the continuum hypothesis is independent from ZFC. And this concludes the Power Round. Congratulations!

5.4 Epilogue: Towards Easton's Theorem

There are no problems in this subsection. Feel free to read it in your spare time.

Ever since Paul Cohen invented the technique of forcing in 1963, it has been applied to solve a myriad of problems in set theory. Most importantly, for our purposes, we now know a lot more about just how *badly* the generalized continuum hypothesis can fail to be true. The forcing techniques we have introduced are sufficient to force the statement $\mathfrak{c} \geq \aleph_2$, but we might ask: can we do better?

Indeed, it is not too hard to show that $\mathfrak{c}$ can be arbitrarily large. To do this, instead of using $\mathrm{Fn}(\omega_2 \times \omega, 2)$ as our forcing notion (relativized to our countable transitive model M, of course), we can use $\mathrm{Fn}(\kappa \times \omega, 2)$, where κ is *any* cardinal, say $\aleph_{\omega_1+3}$. Then, in the exact same way, we can add κ Cohen reals, and force $\mathfrak{c} \geq \kappa$.

Naturally, we then ask: can we force the continuum to be *equal* to $\aleph_2$, or some other cardinal? This is slightly trickier: when we are adding Cohen reals, we don't *just* add κ of them; due to cardinal arithmetic, we actually get

$$\mathfrak{c}^{M[G]} = (\mathfrak{c}^{\aleph_0})^{M[G]} \geq (\kappa^{\aleph_0})^{M[G]} \geq (\kappa^{\aleph_0})^{M},$$

and it can be shown that these are in fact equalities: that is, $\mathfrak{c} = \kappa^{\aleph_0}$ in $M[G]$. But the value of $\kappa^{\aleph_0}$ is pretty much impossible to compute, unless GCH is true in M. To fix this issue, we can simply *assume* that GCH is true in M. After all, we know (by Gödel) that if ZFC is consistent, then it can't disprove GCH, so this assumption is fair game by the reflection principle. By Problem 3.3.5, if $\operatorname{cf}\kappa > \aleph_0$, then we have $\kappa^{\aleph_0} = \kappa$, so we force $\mathfrak{c} = \kappa$. Of course, Problem 3.3.4 forbids us from going any further: if $\operatorname{cf}\kappa = \aleph_0$, then $\mathfrak{c}$ cannot be equal to κ. But still, this tells us that the cardinality of the continuum can be just about anything we want it to be!

Can we force something out of $2^{\aleph_1}$, or other instances of cardinal exponentiation? Of course, but we need to modify our argument. Let $\operatorname{Fn}(I, J, \lambda)$ be the set of functions f which are subsets of $I \times J$, but instead of requiring that f be finite, this time we require that $|f| < \lambda$. Instead of a *countable* chain condition, we now have a κ-chain condition, which states that every strong antichain has cardinality less than κ. The arguments also become more complicated. But it is worth it: we can do a lot of fun things like using the forcing notion $\operatorname{Fn}(\omega_4 \times \omega_1, 2, \omega_1)$ to force $2^{\aleph_1} = \aleph_4$.

And why not try to force multiple times, and force multiple things? By being clever about which forcing notions to use in which order, we can construct generic extensions in which $2^{\aleph_0} = \aleph_1$ and $2^{\aleph_1} = \aleph_3$ and $2^{\aleph_2} = \aleph_{\omega+17}$, or pretty much any combination of equations you want, as long as Problem 3.3.4 is not violated.

In 1970, William Easton substantially generalized all of this by proving his celebrated theorem: the powers of regular cardinals can be anything that isn't outright impossible (by Problem 3.3.4). More precisely:

Theorem 5.4.1 (Easton)

Let E be a class function, where $\operatorname{dom}(E)$ is a class of regular cardinals and $\operatorname{ran}(E)$ is a class of infinite cardinals. If E is non-decreasing, and $\operatorname{cf}(E(\kappa)) > \kappa$ for every κ in the domain of E, then the statement

$$2^\kappa = E(\kappa) \quad \text{for every } \kappa \in \operatorname{dom}(E)$$

is independent from ZFC (as long as ZFC is consistent).

So what about singular cardinals? These are somehow much more challenging to deal with, and Easton's proof completely failed to say anything about them. It wasn't until 1975 that Jack Silver proved the following:

Theorem 5.4.2 (Silver)

The smallest cardinal κ for which $2^\kappa > \kappa^+$, if it exists, cannot be a singular cardinal of uncountable cofinality.

These days, singular cardinals are still as mysterious as ever. Several strange results have been shown about them: for example, Saharon Shelah proved that if $\aleph_\omega$ is a strong limit cardinal, then $2^{\aleph_\omega} < \aleph_{\omega_4}$. But not much is known in general.