Every Reference in “Finite Simple Group (of Order Two)”

(And Possibly, Math (But No Promises))

CJ Quines

July–August 2022

These are notes for a Summer HSSP 2022° class aimed at high schoolers. Its ambitious goal is to cover every reference used in the Klein Four’s song, “Finite Simple Group (of Order Two)”. Because of that ambitious goal, we will throw rigorous math out the window, and everything will be at the mercy of intuition and incorrect explanations.

These notes will be more detailed than explanations we’ll have in class, but in class I can draw pictures and answer questions much more easily.

1 Functions (July 16)

If I know my high school curriculum right, then it’s likely you know some of these.

relation, well-defined. *Also: ill-defined.* Given two sets X and Y, a (binary) **relation** is a set of some ordered pairs (x, y), where $x \in X$ and $y \in Y$. In other words, a relation over X and Y is a subset of $X \times Y$.

This is a concept you’ve probably heard about before, we’re now just giving it a name. Many things in math are relations. Chances are, if you use the phrase “x is (something) of y” or “x is (something) to y”, you’re describing a relation. For example:

- “is a factor of” is a relation over the sets $\mathbb{Z}$ and $\mathbb{Z}$. It’s the set of all (x, y) such that x is a factor of y. The pairs $(2, 4)$ and $(-4, 8)$ are in the relation, while $(3, 8)$ and $(0, 3)$ aren’t.
- “is the square root of” is another relation over the sets $\mathbb{Z}$ and $\mathbb{Z}$. It’s the set of all (x, y) such that $x = y^2$, which includes $(2, 4)$ and $(-2, 4)$. As we can guess from here, it’s a common case that X and Y are the same—in which case, we can drop Y, and just say it’s a relation over X.
- “is double” is a third relation over $\mathbb{Z}$. It’s the set of all (x, y) such that $x = 2y$. Note that not every x has a y that relates to it—that doesn’t stop it from being a relation.
- “is equal to” is another relation over $\mathbb{Z}$. It can also be a relation over $\mathbb{R}$, or over $\mathbb{Q}$, or a relation between $\mathbb{Z}$ and $\mathbb{R}$, and so on. These are all different relations, because they have different sets. It’s not enough to say something is a relation, you have to say which set it’s a relation over.
- “has the Social Security number” is a relation over the set of people and nine-digit numbers. It’s the set of all (x, y) such that x has the Social Security number y. Like the previous example, not every person has a Social Security number, but it’s still a relation.

All relations we’ve talked about are **well-defined**: there’s a unique way to interpret it. The opposite of well-defined is **ill-defined**. An example of an ill-defined relation is “is friends with” over the set of people. There’s no clear definition of what friends means. There are more subtle examples of ill-defined relations. Consider the relation “has the last digit”, over the real numbers. That’s not well-defined: what’s the last digit of $\frac{1}{11}$?

Remark 1.1. Even if you only considered terminating decimals, "has the last digit" still isn't well-defined. For example, 1 can also be written as $0.999\ldots$.

domain, image. *Also: codomain, preimage.* Given a relation over X and Y, we call X the **domain** of the relation and Y the **codomain** of the relation.

Let's consider the relation "has the Social Security number". We defined its domain as the set of all people, and its codomain as the set of all nine-digit numbers. But not all nine-digit numbers° are Social Security numbers, like 000-00-0000. The codomain of this relation is thus larger than its "active" codomain, or the nine-digit numbers that are actually *used*.

The image is the set of the codomain's elements that are "used". More precisely, the **image** of a relation is the set of all y such that (x, y) is in the relation, for some x.

Remark 1.2. You might have heard the word "range" before. We'll never use that word, because it's ambiguous: does it mean codomain or image?

You might be wondering what the "opposite" of an image is. What's the set of all x such that (x, y) is in the relation, for some y? There's no widely agreed name for this, but we'll call it the **preimage**. The preimage of the "has the Social Security number" relation is the set of all people who have Social Security numbers.

Exercise 1.3. What are the domains and images of the relations we gave as examples?

function. You're probably familiar with the concept of a function as a machine: it takes an input, and produces an output. More precisely, a **function** from X to Y is a relation over X and Y, such that for any x, there is *exactly* one y such that (x, y) is in the relation. Exactly one means that:

- it can't be less than one. The relation "is double" over $\mathbb{Z}$ is not a function, because, for example, there's no y such that $(1, y)$ is in the relation.
- it can't be more than one. The relation "is a factor of" over $\mathbb{Z}$ is not a function, because, for example, both $(1, 2)$ and $(1, 3)$ are in the relation.

Of the examples we mentioned earlier, "is the square root of" over $\mathbb{Z}$ and "is equal to" over $\mathbb{Z}$ were functions.

Functions appear so often that we have special notation for them. We often represent functions with letters like f, and say $f : X \to Y$. Instead of saying (x, y) is in the function, we say $f(x) = y$. That way, we can think of f as the "machine": it takes an input, x, and returns an output, y. We'll talk about functions with this language moving forward.

Note that this notation is only *well-defined* for functions:

- for the relation "is double" over $\mathbb{Z}$, the notation isn't defined for some xs. What's $f(3)$?
- for the relation "is a factor of" over $\mathbb{Z}$, the notation is ambiguous for some xs. What's $f(1)$?

This shows us another meaning of the term *well-defined*. A notation is well-defined if it means something, and exactly one thing, for any way you can write it.

Exercise 1.4. Of the relations we talked about that weren't functions, some of them could be "made into" functions. For example, "has the Social Security number" can be turned into a function, if we considered it as a relation over a different domain and image. Which of the relations can be turned to functions this way?

Exercise 1.5. Two of the lines of the song are "But lately our relation's not so well-defined / And I just can't function without you". How are these two lines related to each other? Do you find this funny?

one-to-one. *Also: transpose, injective, surjective, bijective, inverse function.* Here's one important difference between "is the square root of" and "is equal to". The function "is the square root of" over $\mathbb{Z}$ isn't a function the "other way around".

Given any relation over X and Y, we can construct a new relation over Y and X, by flipping the (x, y)s to (y, x)s. We call this the **transpose** of the relation. When is the transpose of a function still a function? Remember that for a relation to be a function, it needs to follow two rules:

- Every x has to be related to *at most one* y. Because $(2, 4)$ and $(-2, 4)$ are in the relation, then both $(4, 2)$ and $(4, -2)$ are in the transpose relation.

- Every x has to be related to *at least one* y. There's no x such that $(x, -1)$ is in the relation. Thus, there's no y such that $(-1, y)$ is in the transpose relation.

A function that *does* follow these two rules for its transpose relation has a special name. If a function's transpose follows the first rule, we call the original function **injective**. If a function's transpose follows the second rule, we call the original function **surjective**. Another word for injective is **one-to-one**.

A function that is both injective and surjective is called **bijective**. That means that its transpose is also a function, which we call its **inverse**. In a bijective function, every x is related to exactly one y, and every y is related to exactly one x.

Exercise 1.6. Stop and think about these definitions! If you've seen these words before, then they're probably different definitions than what you're used to. In that case, convince yourself they're the same definition.

2 Group theory (July 16–23)

group, associative, identity, order. *Also: operation, inverse, subgroup.* A (binary) **operation** over a set G is a function that takes two elements of G and returns another element of G. Some examples are $+$ and $\times$ over $\mathbb{Q}$. While we could write them with the function notation of $+(2, 3) = 5$, we write them with the symbol in between instead, like $2 + 3 = 5$.

A **group** consists of a set and an operation with some properties. Some examples:

- The operation $+$ over $\mathbb{Z}$ forms a group.
    - It's **associative**, meaning $a + (b + c) = (a + b) + c$.
    - There's an **identity**, 0, which means $a + 0 = 0 + a = a$.

  - There's also an **inverse** for each a, called $-a$, which means $a + (-a) = (-a) + a = 0$.
- The operation × over the non-zero rational numbers $\mathbb{Q}^*$ forms a group.
  - It's *associative*, meaning $a \times (b \times c) = (a \times b) \times c$.
  - There's an *identity*, 1, which means $a \times 1 = 1 \times a = a$.
  - There's also an *inverse* for each a, called a^{-1}, which means $a \times a^{-1} = a^{-1} \times a = 1$.

We'll name the first group $\mathbb{Z}^+$ and the second group $\mathbb{Q}^\times$. Read out, these are "the additive group of integers" and "the multiplicative group of rationals".

Here are some things that are not groups. Why?

- The operation − over $\mathbb{Z}$.
- The operation × over (all) the rational numbers $\mathbb{Q}$.
- The operation × over $\mathbb{Z}$.

Here are some things that are groups. Convince yourself that they are groups.

- The operation "addition, then divide by n and take the remainder" over the set $\{0, 1, \ldots, n-1\}$ is a group. We call this $\mathbb{Z}/n\mathbb{Z}$, or "the additive group of integers modulo n".
  - We'll write its operation as +, because it's kinda like addition. In the group $\mathbb{Z}/5\mathbb{Z}$, $2 + 4 = 1$. But note that this is a different operation than normal addition. When we want to emphasize the difference, we'll use different symbols.
  - Another name for this group is "the cyclic group of order n". The **order** of a group is the number of elements in its set, and the set of $\mathbb{Z}/n\mathbb{Z}$ has n elements.

Remark 2.1. Note that we write $\mathbb{Z}/n\mathbb{Z}$ for both the set and the group. Unfortunately, it's common to write the group and the set using the same symbols. This can be confusing, but we'll try to make it clear what we mean.

- The operation "multiplication, then divide by n and take the remainder" over the set $\{1, 2, \ldots, p-1\}$ is a group, if p is a prime. We call this $(\mathbb{Z}/p\mathbb{Z})^\times$, or the "multiplicative group of integers modulo p".
  - It's not obvious that every element has an inverse, but it's true! For example, if $p = 7$, then you can check that $2 \times 4 = 3 \times 5 = 6 \times 6 = 1$.

Exercise 2.2. Why does p have to be prime?

- Consider this sheet of paper. Rotate it however you want, as long as it stays in portrait. Each rotation can be an element of a set. There are four rotations: don't rotate, turn upside-down, flip, and flip and turn upside-down. This is a group with the operation "apply the second rotation, and then the first one".
  - It might be weird thinking of a set whose elements are rotations. But you can think of the previous groups as having elements that also "apply" to something. For example, an element n of $\mathbb{Z}^+$ can be thought of as adding $2n$ to 42. The identity doesn't change the number you're working with, and inverses "return" to the same number.

> **Remark 2.3.** This is called a *group action*, because its elements "act" on an object. I think it's the right way to think of many groups. All groups have a group action.

- This group has many names. This is called the "dihedral group of order 4", or D_4. This is also called the "group of symmetries of a rectangle". This is also known as the Klein four-group. The Klein Four, the people who wrote the song we're studying, is named after this group.

> **Exercise 2.4.** What is the group of symmetries of a square? How many elements does it have?

A **subgroup** of a group is a subset of its elements that forms a group under the same operation. For example, a subgroup of $\mathbb{Z}$ is $3\mathbb{Z}$, the group with operation $+$ and elements $\{\dots, -3, 0, 3, \dots\}$. Another subgroup is the trivial group, the group with only the element 0. (There's not much choice for what the operation should be!)

kernel, quotient. *Also: homomorphism, isomorphism.* We care about how groups talk to each other. In fact, we care about this far more than groups themselves.

Consider the function $f : \mathbb{Z} \to \mathbb{Z}/3\mathbb{Z}$, that's "divide by 3 then take the remainder". For example, $f(5) = 2$, and $f(7) = 1$. Let's consider the groups $\mathbb{Z}^+$ and $\mathbb{Z}/3\mathbb{Z}$. For clarity, we'll write the operation of the first group as $+$, and the operation of the second group as $\oplus$.

How does f interact with our two groups? Well, a group is about its operations, so let's see how it affects the two operations $+$ and $\oplus$. Consider an operation like $5 + 7$. One way to apply f is to do $f(5 + 7)$. Another way is $f(5) \oplus f(7)$. In the first case, we get $f(12) = 0$. In the second case, we get $2 \oplus 1 = 0$. And we get the same result!

In a sense, f is a function that *maintains the operation*. You could also say that it *commutes* with the operation: it doesn't matter whether you apply $+$ and then f, or f and then $\oplus$. If we have an group with operation $+$ over set G, and a group with operation $\oplus$ over set H, then a **homomorphism** $f : G \to H$ is a function such that $f(a + b) = f(a) \oplus f(b)$.

An important kind of homomorphism is an **isomorphism**, which is a homomorphism that is also a bijection. We say that two groups that have an isomorphism are actually the same group. An example is how $\mathbb{Z}$ is isomorphic to $3\mathbb{Z}$, with the isomorphism being "multiply by three".

> **Exercise 2.5.** Convince yourself that $\mathbb{Z}/6\mathbb{Z}$ has an isomorphism to $(\mathbb{Z}/7\mathbb{Z})^\times$.

We now define **kernel** and **quotient**, but I already have a nice writeup about this called Canonical decomposition and the first isomorphism theorem°, so I won't repeat it here.

> **Exercise 2.6.** Two of the lines of the song are "I'm living in the kernel of a rank-one map / From my domain its image looks so blue". If you're in the kernel, why would your image "look blue"? Do you find this funny?

simple group, finite. *Also: normal, abelian.* Let's say $f : G \to H$ is a homomorphism. We've talked about the set of elements $\ker f$. Not only is it a subset of the elements of G, it's actually a *subgroup* of G!

Exercise 2.7. Convince yourself that $\ker f$ is a subgroup of G. Does the operation stay in $\ker f$? Does it have an identity? Does it have inverses? The only information we have about $\ker f$ is that it's the kernel of a homomorphism, but homomorphisms are a lot of information.

Not only is $\ker f$ a subgroup of G, but it's a special kind of subgroup called a normal subgroup. A **normal subgroup** is a subgroup that is the kernel of some homomorphism. Non-normal subgroups exist, but not in **abelian groups**, a group where $a \times b = b \times a$ for all a and b.

Remark 2.8. The smallest example of a non-normal subgroup is in the group of symmetries of an equilateral triangle, D_6. Any subgroup of order two is a non-normal subgroup. One way to check this is to try to look for a homomorphism $D_6 \to \mathbb{Z}/3\mathbb{Z}$.

A **simple group** is a non-trivial group whose only normal subgroups are the trivial group and itself. If this sounds a lot like the definition of "prime", then you're right. The cyclic groups of prime order are all simple groups.

A group is **finite** if its order is finite. In a sense, a finite simple group is like a prime number, in that they're the "building blocks" of finite groups. If a finite group isn't simple, then it has a simple subgroup. The counterpart of the statement that all positive integers have a unique prime factorization would be the Jordan–Hölder theorem.

Finite simple groups are pretty deep. One of the big projects of mathematics, spanning roughly from the 1950s to the 2010s, was to find all the finite simple groups. We know that the cyclic groups of prime order are simple, and in fact, they're the only finite simple abelian groups.

Exercise 2.9. The refrain of the song mentions a "finite simple group of order two". How many groups of order two are there? Try to construct these groups, by starting with two elements, and going over the possibilities for what the operation could be.

3 Metric topology (TBD)

open. *Also: metric space, metric, subspace.* If a group consists of a set and an operation, a metric space consists of a set and a **metric**. A metric over a set M is a function $d : M \times M \to \mathbb{R}_{\geq 0}$. The metric function should be interpreted as the "distance" between two elements in M, which we call *points*.

A **metric space** consists of a set and a metric with some properties:

- It is symmetric, so $d(x, y) = d(y, x)$.
- It is positive definite, so $d(x, y) = 0$ if and only if $x = y$.
- It has the triangle inequality, so $d(x, y) + d(y, z) \geq d(x, z)$.

Here are some things that are metric spaces. Convince yourself that they are metric spaces.

- The normal distance function over $\mathbb{R}^2$ gives a metric space. The metric is

$$d((x_1, x_2), (y_1, y_2)) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.$$

  - In fact, any $\mathbb{R}^n$ is a metric space with this distance. Further, any subset of $\mathbb{R}^n$ can be made into a metric space, with the same distance.
  - In the case of $\mathbb{R}$, the metric becomes "absolute difference".
  - A subspace of $\mathbb{R}$ is $\mathbb{Q}$, with the same metric of absolute difference.

> **Remark 3.1.** We'll overload notation, just like we did with groups, by writing $\mathbb{R}^2$ for both the set of points and the metric space, but we'll be claer which is which.

- The taxicab distance function over $\mathbb{R}^2$ also gives a metric space, where the distance function is

$$d((x_1, x_2), (y_1, y_2)) = |x_1 - x_2| + |y_1 - y_2| .$$

- There's the discrete metric, where given any set, define $d(x, y) = 1$ if $x \neq y$ and $d(x, y) = 0$ if $x = y$.

> **Remark 3.2.** This is a specific case of a metric space on a connected simple graph, where the metric is the length of the shortest path between two vertices.

- Let p be a prime. There's the p-adic metric over $\mathbb{Q}$, where $d(x, y) = p^{-i}$, and i is defined such that $p^i(x - y)$, in simplest form, has neither numerator nor denominator divisible by p. For example, when $p = 2$, the 2-adic distance between $\frac{3}{8}$ and $\frac{1}{4}$ is 8.

> **Exercise 3.3.** Check that the triangle inequality is satisfied for the p-adic metric.

> **Remark 3.4.** Although the last two examples of metrics don't feel really "distance-y", you can still think of a metric as a distance and follow along pretty well. You can get pretty far thinking about metric spaces just by considering $\mathbb{R}^2$. Compare this to groups, where we didn't even discuss non-abelian groups that much, which are completely different from abelian groups.

A **subspace** of a metric space is a subset of its points, with the same metric. This is always a metric space, unlike subgroups.

Finally, an **open subset** of a metric space is a subset of its points such that, for any x in the subset, all points that are "close enough" to x are in the subset too. This is a kind of ϵ thing: if you want to show me that something is an open set, I'll name a point x in the open set, and you give me a distance ϵ, such that all points of distance at most ϵ to x are in the subset too. One typical example to think of is a circle in $\mathbb{R}^2$ without its boundary. Another are the open intervals in $\mathbb{R}$.

> **Exercise 3.5.** What are the open sets in the discrete metric space?

dense. *Also: converge, closed, closure.* There are also these things called closed sets. A closed subset is, unlike what the name suggests, *not* the opposite of an open subset. There are, in fact, sets that are both closed and open. And most sets are *neither* closed nor open! I want to say this now, before we get into definitions, because it's that important.

Consider an infinite sequence of points in a metric space, say, $x_1, x_2, \ldots$. We say that the sequence **converges** to x, for some x also in the metric space, if the sequence gets permanently as close to x as we want to go. You can think of this as another ϵ thing. If you want to show me a sequence converges to x, I'll name a distance ϵ, and you'll have to give me an N, such that all of $x_N, x_{N+1}, \ldots$ are distance at most ϵ to x.

As an example, think of the sequence $3, 3.1, 3.14, 3.141, \ldots$. Considered as a sequence over $\mathbb{R}$, this converges to π. If I say an ϵ like 0.01, you could say, "well, all the things from $3.14, 3.141, \ldots$ have distance at most 0.01 to π." But it doesn't converge over $\mathbb{Q}$!

Now, a **closed subset** of a metric space is a subset of its points such that, for any sequence of points in the subset that does converge, the point it converges to is in the subset. Typical examples are circles in $\mathbb{R}^2$, with their boundary, and closed intervals in $\mathbb{R}$.

Exercise 3.6. Think about the sets $[0, 1)$ and $\varnothing$ in $\mathbb{R}$. Are they open, closed, neither, both? Now that I've given you reasons why these names are bad, here's the only reason they're good: a subset is open if its complement is closed, and vice versa. Convince yourself this is true by drawing some pictures.

Another way to think about closed subsets is through its closure. Say you did take a subset, find all the convergent sequences, and take the set of points they converge to. That set is the **closure** of the original subset, the smallest closed set containing a subset. A set is closed if it's its own closure.

Finally, a set of a metric space is **dense** if its closure is the whole metric space. The typical example is that $\mathbb{Q}$ is dense in the metric space $\mathbb{R}$. To see why every point in $\mathbb{R}$ is the result of a convergent sequence, think of the π example we had earlier. The much harder part to prove is that these are the *only* points that sequences in $\mathbb{Q}$ converge to, although depending on how you define $\mathbb{R}$, you could say that it's the closure of $\mathbb{Q}$.

Exercise 3.7. One line of the song mentions "My heart was open but too dense." We just mentioned that $\mathbb{Q}$ is dense in $\mathbb{R}$, but is it open in $\mathbb{R}$?

continuous, mirror pair. *Also: homeomorphic.* Groups talk to each other through homomorphisms. Metric spaces talk to each other through **continuous functions**. Homomorphisms maintain the group operation: a function f is a homomorphism if $f(a + b) = f(a) \oplus f(b)$. Continuous functions maintain convergence: a function f is continuous if $x_1, x_2, \ldots$ converges to x means that $f(x_1), f(x_2), \ldots$ converges to $f(x)$.

Two groups that are the same are isomorphic. Two metric spaces that are the same are **homeomorphic**. An isomorphism is a bijective homomorphism. A homeomorphism is a bijective continuous function. . . whose inverse function is also continuous. As an example, the square's boundary is homeomorphic to a circle's boundary. The old joke is that donuts and coffee cups are homeomorphic. An object is always homeomorphic to its **mirror pair**—which just means what you think it means.

Remark 3.8. The "inverse function is also continuous" is an important condition. For example, consider $[0, 1)$ and a circle's boundary: there's a bijective continuous function, but the inverse isn't continuous.

path, smooth. A continuous function p from $[0, 1]$ to a metric space is called a **path** in that metric space. It should line up with your intuition about what a path is: a curve connecting one point, $p(0)$, to another, $p(1)$. The fact that it's a continuous means that the path doesn't jump.

The song mentions a "path to love" that's "never smooth". I don't think there's actually a definition for what a smooth path is. The most common definition of smooth comes from calculus. A function is smooth if you can take its derivative infinitely many times. There's a notion of taking a derivative of a path, called the *metric derivative*, which is the instantaneous distance traveled at a given point, but I've never heard of it outside this one Wikipedia article I read.

Exercise 3.9. The second line of the song says that, even though the narrator's path to love isn't smooth, it's continuous. If you know what a derivative is, convince yourself that being continuous is a far cry from being smooth. Is it funny yet?

simply connected. *Also: clopen, connected, path-connected, homotopy.* A set that is both closed and open is called **clopen**. We say a space is **connected** if it has no non-empty clopen sets.

Most of the spaces we've seen so far are connected, but one exception is $\mathbb{Q}$ with the absolute distance metric. Consider the set of numbers in $\mathbb{Q}$ less than $\sqrt{2}$. Seeing that it's open shouldn't be too hard. Seeing that it's closed might be a bit harder, but remember that we're in the metric space $\mathbb{Q}$, and not $\mathbb{R}$, so for it to be closed, we only have to consider sequences that converge in $\mathbb{Q}$. Thus, this set is clopen, and $\mathbb{Q}$ is disconnected.

Exercise 3.10. Consider the absolute distance metric and the set of real numbers in $[0, 1] \cup [2, 3]$. This is a metric space. Convince yourself that $[0, 1]$ is a clopen set in this space. Thus, this space is disconnected.

A **path-connected** space is one that has a path joining any two points in it. The metric space $\mathbb{R}$ is path-connected, but not the metric space $[0, 1] \cup [2, 3]$, over the same metric.

Remark 3.11. All path-connected spaces are connected, but not vice-versa! My favorite example is to think about the comb space. We'll use the fact that if L is a subspace of M, which is a subspace of the closure of L, and L and its closure are both connectde, then so is M. We won't prove this, but think about why it feels true.

Consider the comb, which is a subspace of $\mathbb{R}^2$ with the usual distance metric. It consists of the line joining $(0, 0)$ to $(1, 0)$, which is its shaft, and a bunch of lines joining $\left(\frac{1}{n}, 0\right)$ to $\left(\frac{1}{n}, 1\right)$, and a line joining $(0, 0)$ to $(0, 1)$. This is connected. Now remove the last line; this is still connected, and its closure is the comb space. If we add back the point $(0, 1)$, it follows that it's still connected. But this last space isn't path-connected, as there's no path from $(0, 0)$ to $(0, 1)$.

A **homotopy** between two paths is a continuous deformation from one path to the other. The formal definition involves a continuous function from $[0, 1]$ to paths in the metric space, such that at 0 it's the first path and at 1 it's the last path. There's no good way of explaining this without a picture°.

A space is simply connected if it's path-connected, and, for any two points, all the paths joining them have homotopies between them. The space $\mathbb{R}^2$ is simply connected. If you take out a hole in the middle, it isn't, because there's no homotopy between two paths that go around the hole in different ways.

Remark 3.12. One of the lines of the song is “When we first met, we simply connected”. There's no math joke here, it's just a pun.

4 Set theory (TBD)

separable. *Also: countable.* Consider the statement “the intersection of any number of open sets is open.” We can try to prove this by induction on the number of sets. When you have $n = 1$ set, then it's open. Otherwise, you can take the intersection of the first $n - 1$ sets, which is open by inductive hypothesis, and you only need to show that the intersection of two open sets is open.

Exercise 4.1. Prove that the intersection of two open sets is open! Think of the “challenge” definition. Let's say I pick a point in the intersection of two open sets, and you need to give me an ϵ that works. How can you use the fact that the original sets were open, to find one that works?

Of course, this statement isn't actually true, because it's not true for an infinite number of sets! Indeed, the sets $\left(-\frac{1}{n}, \frac{1}{n}\right)$ are all open, but their intersection is just $\{0\}$, which isn't open.

This example shows that things get tricky when we jump from finite to infinite. A large part of set theory is about dealing with infinity, and what happens when you deal with really large things. And part of that is counting, and labeling the sizes of sets. We say a set is **countable** if it is finite, or if there is a bijection between it and the natural numbers.

As an example, the integers and the rational numbers are all countable, and so is “all numbers that can be described with a single sentence”. We'll explain why fully in class, but the basis of the proof is why the rationals are countable°. The real numbers are uncoutable, which we'll show later.

And completely unrelated to any of that, but because I have to define it anyway, a metric space is **separable** if it contains a countable, dense subset. The space $\mathbb{R}$ with the absolute distance metric is separable because it contains $\mathbb{Q}$. The discrete space is never separable, no matter how big you make it.

Remark 4.2. The reason it's called “separable” is because you can imagine this countable, dense subset, as “separating” the space. For example, in $\mathbb{R}$, any two real numbers are “separated” by a rational number.

class. *Also: power set.* As I said, a large part of set theory is about dealing with infinities. Part of the reason why that's tricky is because dealing with infinities can lead to weird questions about “what is a set?”

I promised a proof of why the real numbers aren't countable, and the reason is Cantor's paradox. Given a set S, let 2^S be its **power set**, or the set of its subsets. Cantor's diagonal argument says that there's never a surjective function $f : S \to 2^S$. If we did, what goes wrong?

The idea is to imagine a table, with the elements of S going down the rows and columns. Each row corresponds to applying f to that element, say x. It results in a subset of S, so across that row we write 1s or 0s, corresponding to whether the element in the column appears in $f(x)$ or not. Then we “invert” the diagonal to get a new subset, which can't be in this list by construction. A similar argument shows that there's no surjective function from $\mathbb{N}$ to $\mathbb{R}$, which shows that the real numbers aren't countable.

Why is this important? This means that we can't just define a set as "any collection of objects". Because if this was a set, which we can call V, then V would contain each of its subsets. But that would mean that there's a surjective function from V to 2^V, contradiction! That means V can't be a set. This is Cantor's paradox.

If V isn't a set, does it even exist at all? Well, we can clearly *define* V, so it has to, right? And if it's not a set, what is it? For convenience, we call it a **class**. It's the collection of sets which satisfy some property, in this case, being a set. It's not a set itself, as that leads to the size issues we talked about, but set theorists find the need to work with classes anyway.

axiom of choice. *Also: axiom, well-ordering theorem.* We just talked about how we can't just say sets are collections of objects. If so, then what *are* sets? In the twentieth century, mathematicians working on foundations gave the widely-agreed upon answer that we use today: a set is something that can be built from a certain list of rules, which we call **axioms**.

These include rules like "the empty set is a set" and "we can take unions of sets" and so on. The list of axioms most mathematicians agree to use is called ZFC, which stands for the Zermelo–Fraenkel with choice. As you can tell from the name alone, this "choice" axiom is apparently important enough to get its own letter in the acronym! We use ZF to refer to the ZFC axioms without choice.

The **axiom of choice** says that given a set of non-empty sets, you can pick an element out of each set, and make a new set. The typical example is to imagine lots of drawers with socks. The axiom of choice says that you can pick a sock from each drawer. If it sounds simple to you, then in a sense, it is—it only gets complicated when you deal with infinite things.

That's because the axiom of choice is equivalent to this axiom called the **well-ordering theorem**. A well-ordering of a set is a way to define "less than" on its elements, such that every subset of the set has a least element, according to this "less than". The well-ordering theorem says that every set has a well-ordering.

Remark 4.3. Here, equivalent means that you can prove one given the other. More precisely, if you take the axioms of ZF, you can prove that the well-ordering theorem implies the axiom of choice, and that the axiom of choice implies the well-ordering theorem.

The natural numbers, for example, have a well-ordering given by $<$. The integers aren't well-ordered by $<$, because the set of negative integers doesn't have a least element. On the other hand, they *are* well-ordered by an ordering like "the smallest absolute value, but in the event of a tie, the negative number is smaller". Those make sense—but what about something like the real numbers? The well-ordering theorem tells us that *this* is well-ordered too, which I find harder to believe.

Exercise 4.4. The axiom of choice and well-ordering theorem aren't as controversial on countable sets. Convince yourself that the rational numbers are well-ordered by coming up with a well-ordering. But don't try too hard to come up with one for the real numbers—it can be proven that there's no formula that describes one, even if one exists.

chain, upper bound. *Also: poset, maximal element, Zorn's lemma.*

The axiom of choice and well-ordering theorem are both equivalent to a third theorem, called Zorn's lemma. The statement is a bit complicated: "if every chain in a poset has an upper bound, then it has a maximal element." We'll explain what those words mean.

Remark 4.5. There's a classic joke about how the three are equivalent, which I've heard attributed to Jerry Bona: "The axiom of choice is obviously true, the well-ordering thoerem is obviously false, and who can tell about Zorn's lemma?"

A **poset**, short for partially-ordered set, is a set with a relation satisfying certain properties. The relation has to be reflexive, so (x, x) is always in the relation. It must be anti-symmetric, which means that if (x, y) and (y, x) are both in the relation, then (x, x) is too. And it must be transitive: if (x, y) and (y, z) are in the relation, so is (x, z). We typically write this relation with a $\leq$ sign, but note that it's not exactly the same as $\leq$ over numbers.

The typical example of a poset is the "is a factor of" relation on the positive integers, that we talked about way, way earlier. It's not necessary that, given two integers, one is a factor of the other—that's what puts the "partial" in "partially-ordered". Other examples are "is less than or equal to" over the real numbers, or the "is a subset of" relation, over the subsets of a given set. We'll talk about the "is a factor of" poset as an example for the rest of this section, but in class, we'll draw pictures of *Hasse diagrams*, which are a way to visualize smaller posets.

A **chain** is a subset of a poset, whose elements can all be compared with each other. For example, $\{2, 8, 24, 1, 72\}$ is a chain. Chains can be infinite, like $\{1, 2, 4, 8, \dots\}$. An **upper bound** of a chain is an element u, such that for each element c in the chain, (c, u) is in the relation. The second chain we gave doesn't have an upper bound, but the first chain has several, like 72 or 144.

A **maximal element** of a poset is an element m, such that (m, M) isn't in the relation for any M. Our "is a factor of" poset doesn't have any maximal elements. Finally, **Zorn's lemma** says that if every chain in a poset has an upper bound, then there has to be some maximal element.

Exercise 4.6. Two lines of the song are "You're the upper bound in the chains of my heart / You're my axiom of choice, you know it's true". What's the joke here? Do you find it funny?

5 Linear algebra (TBD)

tensor. *Also: vector, vector space.*

map, operator, rank.

complexification.

6 Differential geometry (TBD)

wedge, form.

principal bundle. *Also: bundle.*

stable equivalence.

7 Category theory (TBD)

directed system, faithful, forgetful, free, functor, limit.

8 Proof (TBD)

corollary, prove, proposition, QED, without loss of generality.

9 Conclusion (TBD)

10 Lyrics

The path° of love is never smooth°
But mine's continuous° for you
You're the upper bound° in the chains° of my heart
You're my axiom of choice,° you know it's true

But lately our relation's° not so well-defined°
And I just can't function° without you
I'll prove° my proposition° and I'm sure you'll find
We're a finite° simple° group° of order° two

I'm losing my identity°
I'm getting tensor° every day
And without loss of generality°
I will assume that you feel the same way

Since every time I see you, you just quotient° out
The faithful° image° that I map° into
But when we're one-to-one° you'll see what I'm about
Cause we're a finite simple group of order two

Our equivalence was stable°
A principal love bundle° sitting deep inside
But then you drove a wedge° between our two-forms°
Now everything is so complexified°

When we first met, we simply connected
My heart was open° but too dense°
Our system was already directed°
To have a finite limit°, in some sense

I'm living in the kernel° of a rank-one° map
From my domain°, its image looks so blue
Cause all I see are zeroes, it's a cruel trap
But we're a finite simple group of order two

I'm not the smoothest operator° in my class°
But we're a mirror pair°, me and you
So let's apply forgetful° functors° to the past
And be a finite simple group, a finite simple group

Let’s be a finite simple group of order two
(Why not three?)

I’ve proved my proposition now, as you can see
So let’s both be associative° and free°
And by corollary,° this shows you and I to be
Purely inseparable,° QED°