MATH 101, SPRING 2020

WITH PETER KRONHEIMER

Contents

Preliminaries

These notes were taken during the spring semester of 2020 in Harvard's Math 101: *Sets, Groups, and Topology*. The course was taught by Dr. Peter Kronheimer and transcribed by Julian Asilis. The notes have not been carefully proofread and are sure to contain errors, for which Julian takes responsibility. Corrections are welcome at `asilis@college.harvard.edu`.

1. Abstract Groups

Today we'll talk about the symmetries of a square - if I label the vertices of a square by A, B, C, D, then I can describe its symmetries by describing where each vertex goes under a given symmetry. For instance, the identity transformation sends each vertex to itself and – if I wrote the labels in counter-clockwise order – then counter-clockwise rotation by $\frac{\pi}{2}$ is defined by $A \mapsto B$, $B \mapsto C$, $C \mapsto D$, and $D \mapsto A$.

We can similarly consider reflections across vertical, horizontal, and diagonal axes of symmetry. They're a bit hard to think about without drawing a picture, but they also correspond to bijections on the set of the square's vertices.

If we claim that these transformations have the structure of a group – and we do – then we need to check that it's closed under composition and inverses (we've already shown that the identity is an element of our set of transformations). One way to do this is to write down the 8×8 multiplication table of the (purported) group, which stores the product of each column entry with each row entry. Today's worksheet shows that each entry of the multiplication table indeed lies in our original collection of transformations, so it remains only to show that it's closed under inverses. That's the content of exercise 1.

Moving forward, we're going to think about groups slightly more abstractly, as a set of elements with a way of combining elements which satisfies the rules from our old setting of invertible functions (i.e. having a "do nothing" element called the identity, being closed under combinations, and having inverses). In order to make this formal, we first need a definition.

Definition 1.1. A *binary operation* on a set G is a function $G \times G \to G$. That is, it's a rule which assigns to each pair (g_1, g_2) for $g_i \in G$ a unique element of G, written $g_1 \cdot g_2$ or simply $g_1 g_2$.

Example 1.2. The following are binary operations on sets:

- Composition of symmetries of a square
- $G = \mathbb{Z}$ with addition or multiplication
- $G = \mathbb{N}$ with addition or multiplication

Definition 1.3. A *group* is a set G equipped with a binary operation, written $a \cdot b$, satisfying

- (Assocative law): $\forall a, b, c \in G$, $(a \cdot b) \cdot c = a \cdot (b \cdot c)$
- (Existence of identity): $\exists e \in G$ such that $\forall a \in G$, $e \cdot a = a \cdot e = a$
- (Existence of inverses): $\forall g \in G$, $\exists$ an element $g^{-1} \in G$ with $g^{-1} \cdot g = g \cdot g^{-1} = e$.

Note that we do not require that $a \cdot b = b \cdot a$ for all $a, b \in G$. When this property holds – usually called *commmutativity* of the group operation – G is referred to as an *abelian* group.

Example 1.4. $\mathbb{Z}$ with addition is an abelian group: $(a + b) + c = a + (b + c)$ and $a + b = b + a$. Furthermore, the identity element 0 satisfies $a + 0 = a$ and the inverse of $a \in \mathbb{Z}$ is $-a$, as $a + (-a) = 0$.

Example 1.5. $\mathbb{R} \setminus \{0\}$ is an abelian group under multiplication. $(xy)z = x(yz)$ and the identity element 1 indeed satisfies $1x = x = x1$. Finally, any non-zero $x \in \mathbb{R}$ has inverse $\frac{1}{x}$, as $x\frac{1}{x} = 0$.

Note that $\mathbb{R} \setminus \{0\}$ is not a group under addition. In fact, addition is not even a binary operation on this set - the sum of two non-zero real numbers in general need not be non-zero. Similarly $\mathbb{R}_{>0}$, the set of positive real numbers, fails to be a group under addition. Addition is a binary operation on this set, but there's no identity and no additive inverses.

Example 1.6. Let $M_2(\mathbb{R})$ denote the set of 2×2 matrices over $\mathbb{R}$ (i.e. with real entries) equipped with the binary operation of matrix multiplication. Matrix multiplication associates, i.e. $A(BC) = (AB)C$, and there's an identity matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. But arbitrary 2×2 matrices over $\mathbb{R}$ need not have inverses, so this doesn't form a group. $GL_2(\mathbb{R})$, the set of invertible 2×2 matrices, equipped with matrix multiplication does form a group, however. Proving this requires showing, among other things, that the identity matrix is invertible and that invertible matrices are closed under multiplication (i.e. the product of invertible matrices is itself invertible).

Now we'll try to deduce results from our new, more abstract theory.

Proposition 1.7 (Uniqueness of inverses)**.** *Let G be a group and $g \in G$. Then g has only one inverse in G.*

Proof. Suppose g', g'' are both inverses of g. We'll show that $g = g''$. Note that, because g' and g'' have inverses, we have $g'g = gg' = e$ and $g''g = gg'' = e$. Then we have

$$\begin{aligned} g'' &= g''e \\ &= g''(gg') \\ &= (g''g)g' \\ &= eg' \\ &= g' \end{aligned}$$

□

2. Group properties, Equivalence relations

Today we'll spend more time talking about groups and considering examples, and we'll also touch upon equivalence relations and equivalence classes, which relate to some of the exercises that appeared on the last problem set.

Last time we proved that inverses are unique in groups. Something similar holds, which is that the identity element of a group is unique. What we're saying is that – even though we didn't demand this explicitly in our group axioms – any group one could possibly think of has one and only one identity. In symbols, if G is a group and e, e' are elements of G such that $eg = ge = g \ \forall g \in G$ and $e'g = ge' = g \ \forall g \in G$, then it must be that $e = e'$. Here's another result:

Proposition 2.1. *Let G be a group. If $a, b, c \in G$ and $ab = ab$, then $b = c$. Similarly. if $ba = ca$, then $b = c$.*

Proof. On the next homework! □

In these early stages of learning group theory, it's a good idea to clearly denote all the properties you're making use of when writing a proof (e.g. even associativity). Here's another result that we *will* prove.

Proposition 2.2. *If $a \in G$, then $(a^{-1})^{-1} = a$.*

Proof. For any x, x^{-1} is the element of G with $xx^{-1} = x^{-1}x = e$. Take $x = a^{-1}$. Then x^{-1} satisfies $a^{-1}x^{-1} = x^{-1}a^{-1} = e$. But $a^{-1}a = aa^{-1} = e$, so taking $x^{-1} = a$ in the previous expressions indeed produces e. □

More practice with groups:

Proposition 2.3. $(ab)^{-1} = b^{-1}a^1$

Proof. This amounts to checking that $(ab)(b^{-1}a^{-1}) = e$ and $(b^{-1}a^{-1})(ab) = e$. And indeed we have

$$\begin{aligned} (ab)(b^{-1}a^{-1}) &= (a(bb^{-1}))a^{-1} \\ &= (ae)a^{-1} \\ &= aa^{-1} \\ &= e \end{aligned}$$

And similarly

$$\begin{aligned} (b^{-1}a^{-1})(ab) &= (b^{-1}(a^{-1}a))b \\ &= (b^{-1}e)b \\ &= b^{-1}b \\ &= e \end{aligned}$$

□

Note that associativity of triples of elements means that multiplication on larger tuples of elements is well-defined, i.e. any parenthesization of *abcd* or *abcde* gives rise to the same element.

As with any set[1], $|G|$ will denote the size of G (i.e. the number of distinct elements it has). For groups, however, size usually takes the name *order*, so that $|G|$ denote the order of the group G. Another matter of notation: g^n denotes $\underbrace{gg\dots g}_{n \text{ times}}$ for $n \in \mathbb{N}$.

Similarly, g^{-n} denotes the repeated product of g^{-1} (n times). You can check that g^{-n} is the inverse of g^n by repeatedly applying the fact that g^{-1} is the inverse of g. Exponent laws that hold for group elements – by proof, not by definition – are that $g^n g^m = g^{n+m}$ and $(g^n)^m = g^{nm}$.

We've already seen the group of symmetries of a square (a regular 4-gon). More generally, D_n denotes the group of symmetries (i.e. rotations, reflections) of a regular n-gon. In general, $|D_n| = 2n$, and for this reason some authors use D_{2n} to mean what we mean by D_n.

[1] Groups are sets with additional structure (a binary operation satisfying some properties), so in particular they're sets.

Example 2.4. Let C_n denote the rotational symmetries of the regular n-gon. Its elements are $\{e, r_1, \ldots, r_{n-1}\}$, where r_k denotes rotations by $\frac{2\pi k}{n}$. This is referred to as the *cyclic group of order n.*

Onto equivalence relations. We've seen that relations on a set S can be witnessed as subsets of $S \times S$. Equivalence relations are especially nice relations. In particular,

Definition 2.5. The relation $\sim$ on a set S is an *equivalence relation if it is*

(i) Reflexive: $x \sim x \; \forall x \in S$
(ii) Symmetric: if $x \sim y$*, then* $y \sim x$, $\forall x, y \in S$
(iii) Transitive: if $x \sim y$ *and* $y \sim z$*, then* $x \sim z$, $\forall x, y, z \in S$.

For instance, if S is the set of students in the class, then the relation defined by $x \sim y$ if students x and y have birthdays in the same month is an equivalence relation. The relation defined by $x \sim y$ if students x and y have birthdays within 30 days is not an equivalence relation, because it fails transitivity (can you see why?).

In the 'same month' example, we could have used our relation to partition the set into groups of students which are related to each other. That would have looked like grouping all the students with a January birthday, all those with a February birthday, and so on. Everyone must be a member at at least one such set, and they're disjoint – meaning they share no elements – so they form a partition. This procedure generalized to equivalence relations, via equivalence classes.

Definition 2.6. If $\sim$ is an equivalence relation on S, and $a \in S$, the *equivalence class* of a is $[a] = \{x \in S | x \sim a\}$.

Proposition 2.7. *If S is a set endowed with an equivalence relation* $\sim$*, then the equivalence classes of* $\sim$ *partition S (i.e. every element of S belong to an equivalence class and different equivalence classes have empty intersection.)*

Proof. For any $x \in S$, the equivalence class $[x] \subseteq S$ contains x, because $x \sim x$ by reflexivity. Now suppose that equivalence classes $[a], [b]$ have non-empty intersections - we'll show that $[a] = [b]$. If $[a], [b]$ have non-empty intersection, then there exists $z \in S$ with $z \in [a]$ and $z \in [b]$. That means $z \sim a$ and $z \sim b$. By symmetry, $a \sim z$. Then, by transitivity, wee have $a \sim b$.

To show $[a] = [b]$, we show $[a] \subseteq [b]$ and $[b] \subseteq [a]$. To see that $[a] \subseteq [b]$, note that if $x \in [a]$, then $x \sim a$. Since $a \sim b$, transitivity gives us $x \sim b$ and thus $x \in [b]$. The other containment is proven similarly. □

3. $\mathbb{Z}_n$, Subgroups

Last time we talked about equivalence relations.

Example 3.1. Fix a natural number n and write $a \sim b$ for $a, b \in \mathbb{Z}$ if $a - b$ is a multiple of n. This is an equivalence relation:

- $a \sim a$ because $a - a$ is a (zero) multiple of n
- $a \sim b \implies b \sim a$ because the negative of a multiple of n is itself a multiple of n
- $a \sim b, b \sim c \implies a \sim c$ because $(a - c) = (a - b) + (b - c)$ and the sum of multiples of n is itself a multiple of n.

Each integer belongs to one of the following equivalence classes: $[0], [1], \ldots, [n-1]$. Note that $[i]$ consists exactly of the integers with remainder i under division by n.

The above example also ties into our discussion of groups, because the equivalence classes attain the structure of a group with $[a] + [b] = [a + b] = [(a + b) \mod n]$. We'll write this group as $\mathbb{Z}_n$, and call it "the integers mod n". An equivalent definition for $\mathbb{Z}_n$ is that it is the set $\{0, 1, \ldots, n-1\}$ with addition given by adding the elements as integers and then taking the remainder on division by n.

Example 3.2. In $\mathbb{Z}_12$, $11 + 3 = 2$, precisely because $11 + 3 \equiv 2 \mod 12$ (and $2 \in \{0, 1, \ldots, 11\}$.

Definition 3.3. A subset $A \subseteq \mathbb{N}$ is called *sum-free* if $(A + A) \cap A = \emptyset$, i.e. there is no solution to $a_1 + a_2 = a_3$ for $a_1, a_2, a_3 \in A$.

What are examples of sum-free subsets in $\{1, 2, \ldots, 100\}$? One is $A = \{51, 52, \ldots, 100\}$, since the sum of any two elements of A is too big to live in A. Another is $A' = \{1, 3, 5, \ldots, 99\}$, since the sum of odd numbers is even.

Back to $\mathbb{Z}_n$ – by convention, one usually defines $\mathbb{Z}_0$ to be $\mathbb{Z}$ with the usual operation of addition. Note that $\mathbb{Z}_1$ is the group with one element, i.e. $\{0\}$, since any integer has remainder 0 under division by 1. Now we turn to subgroups, which are subsets of groups which have the structure of groups in their own right.

Definition 3.4. Let $(G, \cdot)$ be a group with identity element e. A *subgroup* of G is a subset $H \subseteq G$ such that:

(i) If $h_1, h_2 \in H$, then $h_1 \cdot h_2 \in H$
(ii) $e \in H$
(iii) If $h \in H$, then $h^{-1} \in H$.

Note then that a subgroup of G is a group in its own right: $(\cdot)$ is a binary operation on H because of (i), it has an identity because of (ii), and it has inverses because of (iii). Note that associativity is inherited from the associativity of the group operation on G, which we reuse for H.

Example 3.5. In any group G with identity element e, the set $\{e\}$ is a (trivial) subgroup.

Example 3.6. Let $G = (\mathbb{Z}, +)$ and $H = \{5n | n \in \mathbb{Z}\} = 5\mathbb{Z}$. This is closed under multiplication, has the identity $0 = 5 \cdot 0$ and has inverses $-(5n) = 5(-n)$, so it's a subgroup. This would have worked with any integer in place of 5.

Example 3.7. $\mathbb{N} \subseteq \mathbb{Z}$ isn't a subgroup under addition because it doesn't have inverses or the identity.

Definition 3.8. Let G be a group and $a \in G$. Define $\langle a \rangle = \{a^n | n \in \mathbb{Z}\} \subseteq G$.

Proposition 3.9. *$\langle a \rangle$ is a subgroup of G.*

Proof. It's closed because if $a^n, a^m \in \langle a \rangle$, then $a^n a^m = a^{n+m} \in \langle a \rangle$. The identity is a^0 by definition and the inverse of a^n is a^{-n}. □

Definition 3.10. If $G = \langle a \rangle$ for some $a \in G$, we say G is *cylic* and that a is its generator.

Proposition 3.11. *If G is cyclic, then G is abelian.*

Proof. Suppose $G = \langle a \rangle$. If $g_1, g_2 \in G$, then $g_1 = a^{n_1}$ and $g_2 = a^{n_2}$. So

$$g_1 g_2 = a^{n_1} a^{n_2} = a^{n_1+n_2} = a^{n_2+n_1} = a^{n_2} a^{n_1} = g_2 g_1$$

□

Corollary 3.12. *The operation on Rubik's cubes – which gives it the structure of a group – is not abelian, which is easy to see by playing with some examples. That shows that the group of transformations isn't cyclic, which is not so easy to see at all.*

4. Cyclic groups

Monday's midterm will cover the content from Homework 1 to Homework 5, and wasn't designed to be so long as to press people for time. Recall that it will be in Science Center C, rather than our usual room. The roadmap for today is:

- More on cyclic groups
- Multiplicative group of "units modulo n"
- Worksheet on symmetries of the cube

Recall that a cyclic group H takes the form $H = \langle a \rangle = \{a^n | n \in \mathbb{Z}\}$ for some $a \in H$. Consider first the elements

$$a^0 (= e), a, a^2, a^3, \ldots$$

There are two cases: 1) either the powers of a wrap back to the identity as powers increase, or 2) they never do. In the first case, there's a least d with $a^d = e$, which we investigate now.

Proposition 4.1. *In case 1, with d as above, we have $\forall n \in \mathbb{Z}$ that $a^n = e$ if and only if $d \mid n$.*

Proof. We first prove that $d \mid n$ implies $a^n = e$. Since $d \mid n$, we have $n = q \cdot d$ for some $q \in \mathbb{Z}$. Then

$$\begin{aligned} a^n &= a^{qd} \\ &= (a^d)^q \\ &= e^q \\ &= e \end{aligned}$$

For the other direction, suppose $a^n = e$. Division with remainder permits us to write $n = q \cdot d + r$ for $0 \leq r < d$. We have

$$\begin{aligned} a^n &= e \\ a^{qd+r} &= e \\ a^{qd}a^r &= e \\ ea^r &= e \\ a^r &= e \end{aligned}$$

Since $0 \leq r < d$, $a^r = e$ demands $r = 0$, meaning n was a multiple of d. □

Proposition 4.2. *In case 1, with d as above, the elements of $\langle a \rangle$ are exactly $\{e = a^0, a, a^2, \ldots, a^{d-1}\}$, all of which are distinct.*

Proof. If $x \in H$, then $x = a^n$ for some n. Writing $n = q \cdot d + r$ for $0 \leq r < d$, we have

$$\begin{aligned} a^n &= a^{qd}a^r \\ &= ea^r \\ &= a^r \end{aligned}$$

So $x = a^r$. Now suppose two elements in the set are equal. That means $a^i = a^j$ for some $i < j$ and, using cancellation, $a^{j-i} = e$. Since $0 < j - i < d$, that produces contradiction with minimality of d. □

Definition 4.3. In case 1, d is called the *order* of the element a. By the above proposition, this equals the number of elements in $\langle a \rangle$. For G an arbitrary group, the *order* of G is its cardinality, i.e. $|G|$. If G is cyclic, this coincides with the order of its generators.

In the second case, in which $a^n \neq e$ for any $n \geq 1$, all values of a^i must be distinct (by an argument like that in the above proof). Cyclic groups which fall into the second case are referred to as *infinite cyclic groups*.

Example 4.4. $(\mathbb{Z}_d, +)$ is a cyclic group of order d, generated by 1.

Example 4.5. $(\mathbb{Z}, +)$ is an infinite cylic group generated by 1.

Proposition 4.6. *If $H = \langle a \rangle$ and a has order d, the element $b = a^k$ has order $\frac{d}{gcf(d,k)}$.*

Proof. Suppose $m \geq 1$ and $b^m = e$. That means $a^{mk} = e$, as $b = a^k$. By a previous result, this occurs precisely when $d \mid mk$, i.e. when mk is a multiple of the LCM of d and k. The smallest m for which this occurs is then m with $mk = \text{lcm}(k, d)$, i.e. $m = \frac{d}{\text{gcf}(k,d)}$. □

Onto the worksheet!

6. SYMMETRIC GROUPS

First day of Zoom! Exciting stuff. Two Wednesdays ago, we started talking about permutation groups. The setup was the following: for X a set, a *permutation* of X is a bijection $\sigma : X \to X$. It turns out that for any set X, the collection of permutation of X forms a group, where the group operation is function composition. Most of the time, though, we'll be interested in sets X which take the form $\{1, 2, \ldots, n\}$ for $n \in \mathbb{N}$.

The standard notation for the group of permutations of such an X is S_n. To be clear, an element of S_n is a bijective map from $\{1, 2, \ldots, n\}$ to $\{1, 2, \ldots, n\}$. It turns out that $|S_n| = n!$ – recall that $n! = n(n-1) \cdots 2 \cdot 1$. Informally, if $f \in S_n$ then there are n choices for $f(1)$ – any element in $\{1, 2, \ldots, n\}$ – then $(n-1)$ choices for $f(2)$ – since f injects, $f(2)$ cannot equal $f(1)$ – and so on. S_n is conventionally referred to as the *symmetric group on n letters*.

A useful way of depicting the elements of S_n is cycle notation. For $\sigma \in S_n$, a cycle of σ is a collection of elements in $\{1, 2, \ldots, n\}$ which are mapped to each other by repeated application of σ. For instance, let $\sigma \in S_n$ be the function defined by:

$$1 \mapsto 5,\ 2 \mapsto 4,\ 3 \mapsto 2,\ 4 \mapsto 1,\ 5 \mapsto 3,\ 6 \mapsto 6,\ 7 \mapsto 8,\ 8 \mapsto 7$$

which is a bit of a pain to write and hard to make sense of. The cycles of σ are $1 \mapsto 5 \mapsto 3 \mapsto 2 \mapsto 4 \mapsto 1$ and $7 \mapsto 8 \mapsto 7$ and $6 \mapsto 6$. Note that cycles start and end with the same number. The cycle notation of σ then consists of listing its cycles in order - by convention, elements which map to themselves (i.e. cycles of length 1) are excluded in the notation. So the cycle notation of σ comes out to

$$\sigma = (1\,5\,3\,2\,4)(7\,8)$$

so σ is written using a 5-*cycle* and a 2-*cycle*.

Definition 6.1. An n-cycle of σ is a cycle with n elements in it, i.e. $(x_1\ x_2\ \ldots\ x_n)$ where $\sigma(x_i) = x_{i+1}$ and $\sigma(x_n) = x_1$.

There turns out to be a notion of parity for permutations, i.e. of even and odd permutations. Developing this idea requires a definition.

Definition 6.2. A *transposition* is a 2-cycle, i.e. a permutation of the form $\tau = (a\ b)$ with $a \neq b$.

An important fact is that every permutation can be written as a composition of transpositions. Let's try to convince ourselves of this: fix a k-cycle $\sigma = (a_1\ a_2\ \ldots\ a_k)$. If we hope to write it as a composition of transpositions, a first step is to use the transposition which sends a_1 to a_2, i.e. $(a_1\ a_2)$. Now $a_2 \mapsto a_1$, which isn't what we want. In order to make a_2 go to a_3, we can compose with the transposition that sends a_1 to a_3, i.e. $(a_1\ a_3)$. So our partial solution is

$$(a_1\ a_3) \circ (a_1\ a_2)$$

Continuing in this way, we arrive at a solution of

$$(a_1\ a_k) \circ \cdots \circ (a_1\ a_4) \circ (a_1\ a_3) \circ (a_1\ a_2)$$

It's a good exercise to convince yourself that this recreates the k-cycle we began with.

Theorem 6.3. *In the symmetric group S_n, every permutation σ can be written either as a product of an even number of transposition or an odd number of transposition, but not both.*

We set the stage for the proof the theorem by making use of a lemma:

Lemma 6.4. *The identity $e \in S_n$ cannot be written as a composition of an odd number of transpositions.*

Proof. Suppose $e = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_N$, where each τ_i is a transposition. We hope to show that N is even by inducting on N. Say $\tau_N = (a\ b)$. If $T_{N-1} = (a\ b)$ as well, then $e = \tau_1 \circ \cdots \circ T_{N-2}$. By the inductive hypothesis, $N - 2$ is even so N is even. Unfortunately we don't have time to work through the other cases for T_{N-1} and wrap this up formally. □

Definition 6.5. A permutation σ is *even* if it is a product of an even number of transpositions and *odd* if it can be written as a product of an odd number of transpositions.

By our previous claim that every permutation can be written as some product of permutations, every permutation is either even or odd. By Theorem 5.3, it can't be both – making the definition of parity for permutations a sensible one. The video of today's class, along with a scan of the notes and a proof of Lemma 6.4, will be posted on Canvas soon.

7. Alternating groups, Isomorphisms

Last time we saw that permutations in symmetric groups as S_n can always be witnessed as compositions of transpositions. We saw that this gives rise to a well-defined notion of parity for permutations, and today we'll start things off today by explicitly studying S_n for small n.

When $n = 2$, we have that $S_2 = \{e, (1,2)\}$, since the only bijections from $\{1,2\}$ to itself are the identity and the map that swaps 1 and 2. When $n = 3$, things get slightly more complicated. We can enumerate the elements of S_n by conditioning on lengths of cycles. In particular, the 2-cycles in S_n are (1 2), (2 3), and (3 1). Recall that (1 2) and (2 1) mean the same thing. Thee 3-cycles are (1 2 3) and (1 3 2). By similar reasoning, we have that

$$(1\ 2\ 3) = (2\ 3\ 1) = (3\ 1\ 2)$$

Along with the identity, we see that S_n has $3 + 2 + 1 = 6$ elements, agreeing with our earlier claim that $|S_n| = (n!)$.

When $n = 4$, we won't list everything out but it's worth noting that there are additional possibilities than a permutation consisting of a single k-cycle for some k. In particular, a permutation in S_4 can be written in cycle notation as the identity e (really, as an empty symbol), a 2-cycle $(a\ b)$, a 3-cycle $(a\ b\ c)$, a 4-cycle $(a\ b\ c\ d)$, or two 2-cycles $(a\ b)(c\ d)$. In general, an element of S_n consists of cycles of lengths $\ell_1, \ldots, \ell_k$ with $\ell_1 + \cdots + \ell_k = n$, and the convention that 1-cycles are swept under the rug.

What are the parities of these permutations? We saw last time that k-cycle is even if k odd and odd if k is even (since we witnessed a k-cycle as a product of $k - 1$ transpositions).

So the 2$-$ and 4$-$cycles in S_4 are odd and the 3-cycles are even. The double 2-cycles are even, precisely because $(a\ b)(c\ d) = (a\ b) \circ (c\ d)$.

Definition 7.1. $A_n \subseteq S_n$ is the set of even permutations. It is referred to as the *alternating group* on n letters.

Proposition 7.2. *A_n is a subgroup of S_n. More explicitly,*

(i) $e \in A_n$
(ii) If $\sigma, \pi \in A_n$ *then* $\sigma\pi \in A_n$
(iii) If $\sigma \in A_n$ *then* $\sigma^{-1} \in A_n$

Example 7.3. $A_2 = \{e\}$, the trivial group with one element.
$A_3 = \{e, (1\ 2\ 3), (1\ 3\ 2)\}$. More abstractly,

$$A_4 = \{\text{identity, all 3-cycles, all double 2-cyles}\}$$

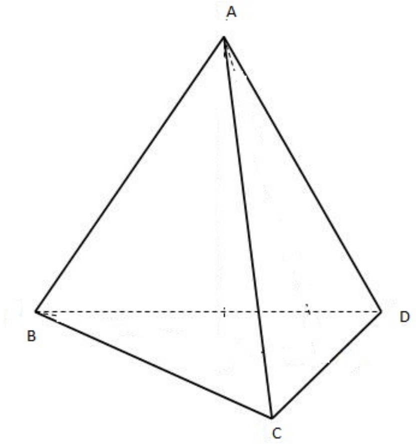
Proposition 7.4. $|A_n| = \frac{1}{2}n!$ *if* $n \geq 2$.

Proof. Let B_n be the set of odd permutations, so that $B_n \subseteq S_n$. Then $S_n = A_n \cup B_n$. Since $|S_n| = n!$, and A_n and B_n are disjoint, it suffices to show that $|B_n| = |A_n|$. We do this by exhibiting a bijection $T : A_n \to B_n$. Fixing a transposition $\tau \in S_n$, the map is $T(\sigma) = \sigma\tau$. If $T(\sigma_1) = T(\sigma_2)$, then $\sigma_1\tau = \sigma_2\tau$ so $\sigma_1 = \sigma_2$, by cancellation. So T injects. To see that T surjects, fix $p \in B_n$. Then $T(p\tau^{-1}) = p\tau^{-1}\tau = p$, so p is in the image of T and T surjects. We conclude $|B_n| = |A_n|$, as desired. □

Definition 7.5. A pair of groups G and H are *isomorphic* if there is a bijection $T : G \to H$ between them which 'respects' group multiplication. Formally, $\forall a, b \in G$, T has the property that $T(ab) = T(a)T(b)$.

Informally, for G and H to be isomorphic means that G and H are identical 'up to relabeling', including the structure of their multiplication tables. In other words, G can be 'relabeled' by T in such a way that the group operation of G applied to its elements (with new labels) coincides with the group operation of H.

Recall the groups of symmetries of regular polygons. For the moment, let G denote the group of rotational symmetries of the tetrahedron, i.e. of



Now some geometric results that we won't prove formally but that one can gain intuition for by playing with examples.

Theorem 7.6. *G is isomorphic to A_4, denoted $G \cong A_4$.*

Theorem 7.7. *The rotation group of the cube is isomorphic to S_4.*

Theorem 7.8. *The (rotational) symmetric group of the regular icosahedron (or of the regular dodecahedron) is isomorphic to A_5, a group of size $\frac{1}{2}5! = 60$.*

More generally, it is a result that finite groups of rotations of 3-space are isomorphic to one of C_n, D_n, A_4, S_4, or A_5.

This week's homework will be drawn from the fifth chapter of Judson's text. He makes use of the phrase 'permutation group', by which he means a subgroup of S_n for some n (e.g. S_n itself, A_n, etc.).

8. Cosets

Today's material comes from Chapter 6 of Judson and is the last lecture which will be covered on next week's midterm.

The setting for today is that G is a group and H a subgroup of G.

Definition 8.1. For any $g \in G$, the *left coset* of H with 'representative' g is

$$gH = \{gh | h \in H\}$$

Likewise, the *right coset* of H with representative g is

$$Hg = \{hg | h \in H\}$$

Example 8.2. Take $G = S_3 = \{e, (123), (132), (12), (13), (23)\}$ and $H = \{e, (12)\}$. Let $g = (23)$. Then

$$\begin{aligned} gH &= \{gh | h \in H\} \\ &= \{(23)e, (23)(12)\} \\ &= \{(23), (132)\} \end{aligned}$$

By examining all the cosets of $\{e, (12)\}$ in S_3, we can observe that the cosets partition S_3. This turns out not to be an coincidence but a result about cosets in a group. In particular, cosets of H form equivalence classes of G under the following relation:

Proposition 8.3. *Define a relation $\sim$ on G by $g_1 \sim g_2$ if $g_1^{-1} g_2 \in H$. This forms an equivalence relation.*

Proof. $g \sim g$ because $g^{-1} g = e \in H$. To see that $\sim$ is symmetric, suppose $g_1 \sim g_2$. Then $g_1^{-1} g_2 \in H$, so $(g_1^{-1} g_1)^2 \in H$, as H is a subgroup. That comes out to $g_2^{-1} g_1 \in H$, meaning $g_2 \sim g_1$.

Lastly, suppose $g_1 \sim g_2$ and $g_2 \sim g_3$. Then $g_1^{-1} g_2, g_2^{-1} g_3 \in H$. Multiplying them together - and recalling that H is closed under multiplication - we have

$$(g_1^{-1} g_2)(g_2^{-1} g_3) = g_1^{-1} g_3 \in H$$

so $g_1 \sim g_3$. □

To see that the equivalence classes of this relation are the cosets of H, note that

$$\begin{aligned} g_1 \sim g_2 &\iff g_1^{-1} g_2 \in H \\ &\iff g_1^{-1} g_2 = h && \text{(for some } h \in H) \\ &\iff g_2 = g_1 h \\ &\iff g_2 \in g_1 H \end{aligned}$$

A takeaway is that the representative of a coset is not unique. In particular, any element in a coset serves as a representative for it.

Example 8.4. Set $G = S_3$, $H = \{e, (12)\}$ and $g = (13)$. Then

$$\begin{aligned} Hg &= \{hg | h \in H\} \\ &= \{(13), (12)(13)\} \\ &= \{(13), (132)\} \end{aligned}$$

Notably, this is not a left coset of H. In general, the left and right cosets of H can differ.

When G is abelian, however, its left and right cosets coincide. When G is abelian, it is common to denote its operation additively, i.e. $(G, +)$ with a left coset (equivalently, right coset) taking the form $g + H$. The abelian condition is sufficient, but not necessary, in order for the left and right cosets to coincide. On Wednesday, we'll talk more about this property.

Proposition 8.5. *For G a group and H a finite subgroup, say $|H| = n$. Then every coset, left or right, of H also has exactly n elements.*

Proof. We exhibit a bijection between H and gH. The map is $h \mapsto gh$. This surjects by definition and injects because of the cancellation law in groups. □

Theorem 8.6 (Lagrange)**.** *If G is a finite group, then the order of any subgroup H in G divides the order of G.*

Proof. The cosets of H partition G and all have size $|H|$, so $|H|k = |G|$, k the number of cosets. □

Definition 8.7. $[G : H]$, called the *index* of H in G, denotes the number of cosets of H in G. When G is finite, $[G : H] = \frac{|G|}{|H|}$.

Example 8.8. $[\mathbb{Z} : 3\mathbb{Z}] = 3$, since there are three $3\mathbb{Z}$ cosets in $\mathbb{Z}$, with representative 0, 1, and 2. More generally, $[\mathbb{Z} : n\mathbb{Z}] = n$ with representatives $0, 1, \ldots, n-1$.

A consequence of Lagrange's theorem is that that for G a finite group, the order of g divides the order of G. To see why, note that the order of g is also the order of the cyclic subgroup $H = \{e, g, g^2, \ldots, \}$ that it generates. Since $|H|$ divides $|G|$, the order of g does as well.

9. Normal subgroups

Today's class will conclude group theory – after that, we'll pivot to topology and analysis. Last time we talked about left and right cosets, which we learned are not always the same. We saw that for $g \in G$ and $H \subseteq G$ a subgroup, then the left coset of H containing g is

$$gH = \{gh | h \in H\}$$

and the right coset containing g is

$$Hg = \{hg | h \in H\}$$

We may be motivated to ask: In general, when are the left coset containing g and the right coset containing g the same? For these to be the same, it must be the case that $\forall h \in H$, $gh = h'g$ for some $h' \in H$ (i.e. $gH \subseteq Hg$). This is equivalent to saying that for all h, $ghg^{-1} = h'$ for some $h' \in H$. In order for it to be the case that $gH = Hg$, it must furthermore be the case that $Hg \subseteq gH$. By similar reasoning, that condition amounts to $g^{-1}hg \in H$.

Motivated by this characterization of the subgroups H which induce equivalent left and right cosets, we introduce the following definition.

Definition 9.1. A subgroup H of G is a *normal subgroup* if $\forall g \in G, \forall h \in H$, $ghg^{-1} \in H$. Equivalently, for every $g \in G$, $gH = Hg$.

Example 9.2. $H = \{e, (1\ 2)\}$ is not a normal subgroup of S_3. Taking $g = (1\ 2\ 3)$, you can check $gH \neq Hg$.

We will use $H \leq G$ to denote that H is a subgroup of G. An important observation is that $H \leq G$ is obligated to be normal when G is abelian, as

$$gH = \{gh | h \in H\} = \{hg | h \in H\} = Hg$$

$H \lhd G$ is the standard notation to denote when H is a normal subgroup of G. Crucially, the reason people care about normal subgroups is that they give rise to a well-behaved notion of multiplication for cosets.

For A, B subsets of G, AB is defined as $\{ab | a \in A, b \in B\}$, i.e. the element-wise product.

Proposition 9.3. *If $H \leq G$, then $HH = H$.*

Proof. To see $HH \subseteq H$, consider an arbitrary element in HH. It takes the form $h_1 h_2$ for $h_i \in H$. Since H is a subgroup, it is closed under multiplication and thus $h_1 h_2 \in H$. Alternatively, consider $h \in H$. To see $h \in H$, note $h = he \in HH$. We made use of the fact that $H \leq G$ when concluding that $he \in H$; in particular, we used the fact that $e \in H$. □

We consider this notion of element-wise multiplication applied to cosets.

Proposition 9.4. *Let $H \lhd G$, and let $g_1 H$, $g_2 H$ be two cosets of H in G. Then $(g_1 H)(g_2 H) = (g_1 g_2)H$.*

Proof.

$$\begin{aligned}
(g_1 H)(g_2 H) &= g_1 (H g_2) H \\
&= g_1 (g_2 H) H && \text{(normality of } H\text{)} \\
&= (g_1 g_2)(HH) \\
&= (g_1 g_2) H && \text{(previous proposition)}
\end{aligned}$$

□

It turns out that, equipped with the above multiplication, the cosets of a normal subgroup themselves inherit the structure of a group.

Example 9.5. Consider $\mathbb{Z}_p$ for p prime. Its only subgroups are $\{e\}$ and $\mathbb{Z}_p$ itself. This is a consequence of Lagrange's theorem, and there's a term for this.

Definition 9.6. A group G is *simple* if its only normal subgroups are $\{e\}$ and G.

In addition to $\mathbb{Z}_p$, it turns out that A_5 – and indeed A_n for all $n \geq 5$ – is simple.

10. Real numbers

Today, we depart from group theory and begin the next topic in the course: analysis. This begins with an investigation of the real numbers, $\mathbb{R}$, and, in time, will lead us to such topics as sequences and their limits.

Let us first discuss the algebraic properties of $\mathbb{R}$:

(i) $\mathbb{R}$ is equipped with the binary operation of addition, written $a + b$

(ii) It is also equipped with the binary operation of multiplication, written $\times$ or $\cdot$ (or even simply with concatenation, as in ab)

Crucially, these operations obey certain structure.

(i) $(\mathbb{R}, +)$ is an abelian group with identity 0

(ii) $(\mathbb{R}, \times)$ is *not* a group, as 0 has no multiplicative inverse (i.e. no r such that $0 \cdot r = 1$). However, $(R \setminus \{0\}, \times)$ forms an abelian group with identity 1.

(iii) Moreover, addition and multiplication play nicely with one another. In particular, multiplication distributes over addition:

$$a(b+c) = ab + ac$$

Note that, since multiplication and addition both commute, multiplication distributes from the right hand side as well:

$$(b+c)a = ba + ca$$

The term for a set equipped with such properties is that of a *field*, so $(\mathbb{R}, +, \times)$ is a field. In addition to such algebraic structure, $\mathbb{R}$ has order properties as well. Making this formal calls for a definition.

Definition 10.1. A relation $\lhd$ on a set S forms a *total order* if

(i) for every $a, b \in S$, exactly one of the following hold: $a \lhd b$, $b \lhd a$, $a = b$. (Trichotomy)

(ii) For every $a, b, c \in S$, if $a \lhd b$ and $b \lhd c$, then $a \lhd c$. (Transitivity)

$\mathbb{R}$ equipped with the usual strict inequality $<$ gains the structure of a totally ordered set. Furthermore, this order is well-behaved with respect to addition and multiplication. In particular,

(i) for all $x, y, z \in \mathbb{R}$, if $x < y$ then $x + z < y + z$

(ii) If $0 < x$ and $0 < y$, then $0 < xy$.

We can now refer to $\mathbb{R}$ as an *ordered field*, i.e. a field equipped with a total order such that the preceding conditions are satisfied. This endows $\mathbb{R}$ with a lot of structure but does not distinguish it from, for instance, $\mathbb{Q}$. In the following, we introduce exhaustive axioms will totally determine $\mathbb{R}$.

Definition 10.2. The real numbers $(\mathbb{R}, +, \times, <)$ are the data obeying the following conditions.

(1) $(\mathbb{R}, +)$ is an abelian group with identity element 0

(2) $(\mathbb{R} \setminus \{0\}, \times)$ is an abelian group with identity element 1

(3) $x(y+z) = xy + xz$

(4) $<$ forms a total order with

  (i) if $x < y$ then $x + z < y + z$

  (ii) if $0 < x$ and $0 < y$, then $0 < xy$

(5) The completeness axiom, whose definition we postpone for now

Note that $\mathbb{Q} \subseteq \mathbb{R}$ satisfies conditions (1) through (4). In order for the preceding conditions to determine $\mathbb{R}$, and if $\mathbb{R}$ is distinct from $\mathbb{Q}$, it must be that $\mathbb{Q}$ fails to be 'complete', whatever that means. We will see that indeed this is the case.

As a demonstration of the fact that $\mathbb{R}$'s essential structure is captured in its axioms, we prove results on $\mathbb{R}$ by way of its axioms.

Example 10.3. Use (1), (2), (3) to deduce that $0 \cdot x = 0$ for all $x \in \mathbb{R}$. We have

$$0 \cdot x = (0+0) \cdot x = 0 \cdot x + 0 \cdot x$$

By cancellation in groups (or adding $-(0 \cdot x)$ to each side), we have $0 \cdot x = 0$, as desired.

Notation 10.4. Our notation is currently fairly restricted. In order to recover familiar notational tools such as $\leq$, we establish the following conventions:

- $a > b$ means $b < a$
- $a \leq b$ means ($a < b$ or $a = b$)
- $a - b$ means $a + (-b)$, for $-b$ the inverse of b in $(\mathbb{R}, +)$
- $\frac{a}{b}$ means $a \cdot (b^{-1})$, where b^{-1} is the inverse of b in $(\mathbb{R}, \cdot)$

We now turn to the completeness axiom of 10.2. This begins with consideration of subsets of R.

Example 10.5. Let $S \subseteq \mathbb{R}$. Examples include

- $S = [2,3] = \{x \in \mathbb{R} | 2 \leq x, x \leq 3\}$
- $S = [1, \infty) = \{x \in \mathbb{R} | x \geq 1\}$
- $S = \{\frac{1}{n} | n \in \mathbb{N}\} = \{\frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \dots\}$

Given a subset of $\mathbb{R}$, one natural task is to attempt to bound it.

Definition 10.6. An *upper bound* for $S \subseteq R$ is a number $b \in \mathbb{R}$ such that for all $s \in S, s \leq b$.

In words, an upper bound of S is a real number which is at least as big as all the elements of S.

Example 10.7. If $S = [2,3]$, then 100 is an upper bound of S. So is 50, or 3. But 2.7 is not an upper bound.

In the above example, several of the bounds provided seem excessive, or inefficient. An important task to formalize the idea of a sharpest upper bound.

Definition 10.8. A *least upper bound* for $S \subseteq \mathbb{R}$ is a number $b \in \mathbb{R}$ such that

(i) b is an upper bound for S
(ii) If b' is any upper bound for S, then $b \leq b'$

Example 10.9. The least upper bound of $S = [2,3]$ is 3. Indeed $3 \geq S$ and any $y < 3$ is exceeded by an element of S.

Example 10.10. Let $S = (2, 3)$. Then, once again, 3 is a least upper bound of S. Notably, least upper bounds of a set need not be elements of the set.

A synonym for the least upper bound for S is the *supremum* of S, written $\sup S$.

Proposition 10.11. *Least upper bounds, when they exist, are unique.*

Proof. Suppose b, b' are both supremums of S. Then, because b is a *least* upper bound and b' is some upper bound, it must be that $b \leq b'$. Similarly, $b' \leq b$. They imply $b = b'$. □

An arbitrary subset of $\mathbb{R}$ need not have a least upper bound. One obstruction is that the set not have any upper bounds at all, for instance $S = \mathbb{N} \subseteq \mathbb{R}$ or $S = (3, \infty)$. In principle, another obstruction is that there may upper bounds but there may not be a least one. This occurs, for instance, when $S = \emptyset$. In this case, every real number serves as an upper bound for S, and as such there is no least upper bound.

Definition 10.12 (Completeness Axioms for $\mathbb{R}$)**.** If a subset $S \subseteq \mathbb{R}$ is non-empty and has an upper bound, then it has a least upper bound.

In words, the above definition states that, in $\mathbb{R}$, the empty set is the only pathological case in which a set with an upper bound does not admit a least upper bound. In the coming lectures, we will further discuss what this definition captures and why it is an appropriate final axiom for $\mathbb{R}$.

11. Lemmas for $\mathbb{R}$

Today we'll keep moving with real numbers, looking to formally justify familiar and intuitive results.

Lemma 11.1. $(-1) \cdot (-1) = 1$

Proof. On the homework! □

Lemma 11.2. *For any $x \neq 0$, either $x > 0$ or $(-x) > 0$ (but not both).*

Proof. From the order axioms, if $x \neq 0$ then either $x > 0$ or $0 > x$. If $x > 0$, then we're done. If $0 > x$, then we have

$$\begin{aligned} 0 &> x \\ -x &> x + (-x) \\ -x &> 0 \end{aligned}$$

completing the proof. □

Lemma 11.3. $0 < 1$.

Proof. By contradiction - suppose not. Then by Lemma 2, $0 < (-1)$. The axioms of order said that $0 < x$ and $0 < y$ implies $0 < xy$. Taking $x = -1$ and $y = -1$, we deduce $0 < (-1)(-1)$. By our first lemma, that identically means $0 < 1$, producing contradiction. □

Lemma 11.4. *If $x > 0$, then $\frac{1}{x} > 0$.*

Proof. Assume, for contradiction, that $x > 0$ and $\frac{-1}{x} > 0$. Then $\frac{-1}{x}x = -1 > 0$, producing contradiction. □

Lemma 11.5. *If $x > 0$ and $y > 0$, then $x + y > 0$.*

Proof. On the HW! □

Now we introduce $2 = 1 + 1$. In particular, we are *defining* the symbol 2 to mean the output of the operation 1 + 1. Note that $2 > 0$, by the previous lemma. Thus, by induction, any natural number $n = \underbrace{1 + \cdots + 1}_{n \text{ times}}$ is positive. Making use of 11.4, we also have that $\frac{1}{n}$ is positive for any natural number n.

Lemma 11.6. *If $0 < x < y$, then $0 < \frac{1}{y} < \frac{1}{x}$.*

Proof. From $x < y$, get $0 < y - x$. From $0 < x, y$ we have $0 < xy$. So $0 < \frac{1}{xy}$. Then $0 < \frac{1}{xy}(y - x) = \frac{1}{x} - \frac{1}{y}$. □

It's sensible to ask whether there exists a minimal positive number or maximal negative number. This turns out not to be the case, which can be interpreted as the absence of a gap between 0 and the positive numbers to the right of it when drawing the conventional number line. This is made formal in the following lemma.

Lemma 11.7. *If $\epsilon > 0$, then $0 < \frac{\epsilon}{2} < \epsilon$.*

Proof. Since $0 < \epsilon$ and $0 < \frac{1}{2}$, we have $0 < \frac{\epsilon}{2}$. Adding $\frac{1}{2}\epsilon$ to each side, we get $\frac{1}{2}\epsilon < \epsilon$ as well. □

We haven't made use of the completeness axiom on $\mathbb{R}$ yet, only its order and field axioms. Recall that the completeness axiom states that any non-empty subset of $\mathbb{R}$ which admits an upper bound admits a (unique) least upper bound. This implies a dual result, which is that a non-empty subset of $\mathbb{R}$ which admits a low bound furthermore admits a (unique) greatest lower bound, referred to as the infimum.

Remark 11.8. For bounded $S \subseteq \mathbb{R}$, $\sup(S)$ and $\inf(S)$ may or may not belong to S.

Example 11.9. Let $S = [0, 1]$. Then $\sup(S) = 1 \in S$, and $\inf(S) = 0 \in S$. If $S = (0, 1)$, then $\sup(S) = 1 \notin S$ and $\inf(S) = 0 \notin S$.

When $\sup(S)$ exists and is known to be an element of S, it is referred to as the *maximum* of S. Likewise, when $\inf(S)$ exists and is known to belong to S, it is referred to as the *minimum* of S.

Recall that the completeness axiom is the only axiom which distinguishes $\mathbb{R}$ from $\mathbb{Q}$, as $\mathbb{Q}$ satisfies the field and order axioms.

Theorem 11.10. *For every real number c, there exists a natural number $n \in N$ which exceeds c, i.e. $c < n$.*

Proof. In pursuit of contradiction, suppose $c \in \mathbb{R}$ and $n \leq c$ for all $n \in \mathbb{N}$. Then c is an upper bound for $\mathbb{N}$, and there then exists a least upper bound b for $\mathbb{N}$. Since for any natural number n, $(n+1)$ is also a natural number, it also the case that $(n+1) \leq b$ for all $n \in \mathbb{N}$. Subtracting 1 from both sides gives $n \leq (b-1)$ for all $n \in \mathbb{N}$. So $b-1$ is an upper bound for $\mathbb{N}$. But $b-1 < b$, contradicting minimality of b among upper bounds. □

Now we'll do something which we can't do in $\mathbb{Q}$, meaning we'll need to make use of the completeness axiom.

Theorem 11.11. *There exists $\alpha \in \mathbb{R}$ such that $\alpha^2 = 2$.*

Proof. Set $S = \{x \in \mathbb{R} | x^2 \leq 2\}$. S is not empty because it contains 0, and it is bounded above by 2. So it has a supremum α. I claim $\alpha^2 = 2$. Suppose not, meaning $\alpha^2 > 2$ or $\alpha^2 < 2$. We don't have time left, but the idea is that if $\alpha^2 < 2$, then it can be increased and stll square to less than 2, in which case α is not an upper bound for S. If $\alpha^2 > 2$, then it can be reduced slightly and still have its square exceed 2, in which case it is not a minimal upper bound, likewise producing contradiction. □

12. Completeness

Until now, our discussion of the completeness axiom has centered around the order on $\mathbb{R}$, i.e. extracting least upper bounds from upper bounds. There tuns out to be an equivalent definition of completeness for $\mathbb{R}$ which makes use of the notion of distance, rather than order. This reformulation of completeness generalizes easily to settings in which one has the structure of a distance but not an order (such as $\mathbb{C}$).

Define the closed interval $[a,b] \subseteq \mathbb{R}$ to be to be the set $\{x | a \leq x \leq b\}$. Imagine now a sequence of intervals $[a_n, b_n]$, $n \in \mathbb{N}$ which are nested, meaning $[a_n, b_n] \sup [a_{n+1}, b_{n+1}]$.

Theorem 12.1. *If $[a_n, b_n]$ is a sequence of nested, closed intervals in the above sense, then $\bigcap_{n=1}^{\infty} [a_n, b_n] \neq \emptyset$. That is, there exists a real number y with $y \in [a_n, b_n]$ $\forall n \in \mathbb{N}$.*

Proof. We make use of the axiom of completeness. Consider $A = \{a_n : n \in \mathbb{N}\}$. The set is non-empty and has b_1 as an upper bound, as $b_1 > a_1 \geq a_2 \geq a_3 \geq \dots$. Consider then its least upper bound α. Because it is an upper bound, $\alpha \geq a_i$ for all $i \in \mathbb{N}$. Because it is the least upper bound, and b_1 is an upper bound, $\alpha \leq b_1$. Since b_1 is minimal amongst the b_i, then furthermore $\alpha \leq b_i$ $\forall i \in \mathbb{N}$. So $\alpha \in \bigcap [a_n, b_n]$. □

This idea can be generalized to settings without order by using, for instance, closed disks rather than closed intervals. Completeness of, say, $\mathbb{R}^2$ can be seen as corresponding to the fact that intersections of nested closed disks are non-empty. Note that the above claim fails if the intervals are open. In particular, set $a_n = 0$ and $b_n = \frac{1}{n}$. Then $\cap(0, \frac{1}{n}) = \emptyset$, since every $y > 0$ is less than $\frac{1}{n}$ for sufficiently large y.

How does distance enter the picture? In $\mathbb{R}$, the metric $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is defined $d(x,y) = |x-y|$, i.e. the distance between points x and y is defined as the absolute value of their difference. In higher dimensions, distances include the Euclidean distance - for which points $p = (p_1, \dots, p_n)$ and $q = (q_1, \dots, q_n)$ in $\mathbb{R}^n$ are distance $\sqrt{\sum |p_i - q_i|^2}$ apart - as well as the taxicab distance or Manhattan metric, in which p and q are distance $\sum |p_i - q_i|$ apart.

In nearly any setting there are a wide variety of choices for a formal notion of distance, but they are obligated to satisfy the following axioms in order to indeed qualify as distances, or metrics.

Definition 12.2. A *metric*, or *distance function*, on a set X is a function $d : X \times X \to \mathbb{R}$ such that $\forall p, q, r \in X$:

(i) $d(p,p) = 0$
(ii) $d(p,q) > 0$ if $p \neq q$
(iii) $d(p,q) = d(q,p)$
(iv) $d(p,r) \leq d(p,q) + d(q,r)$, known as the triangle inequality

Example 12.3. Set $X = \{\text{students in Math 101}\}$, and endow it with the function

$$d(p,q) = \begin{cases} 0 & p = q \\ 1 & p \neq q \end{cases}$$

You can check that this indeed satisfies the conditions of Definition 12.2, though it may not be very interesting. In particular, metric need not resemble anything like the Euclidean distance or Manhattan metric of $\mathbb{R}^n$.

Definition 12.4. Give a set X equipped with a metric d, set $p \in X$ and $r \in \mathbb{R}_{\geq 0}$. The open ball in X with center P and radius r is defined to be the set

$$B(p,r) = \{x \in X | d(x,p) < r\}.$$

Open balls under the Euclidean metric look like usual spheres, while open balls under the taxicab distance look like squares. Under the metric of Example 12.3, $B(p,r)$ is $\{p\}$ if $r \leq 1$ and all of X if $r > 1$.

Definition 12.5. A *metric space* is a set equipped with a distance function.

13. Sequences

Today we'll touch on sequences and their limits, ideas you may have encountered previously in a calculus course. Our intention is to bring out the mathematical points that lie behind familiar calculations, such as $\frac{\log x}{x} \to 0$ as $x \to \infty$, rather than performing more routine calculations.

A *sequence* of real numbers $(a_n)_{n\in\mathbb{N}}$ consists of a collection of real numbers indexed by the natural numbers $\mathbb{N}$. Formally, a sequence in R is given by a function $f : \mathbb{N} \to \mathbb{R}$, so that $a_n = f(n)$.

Example 13.1. There is a sequence $(a_n)_{n\in\mathbb{N}}$ with $a_n := \frac{n+1}{n}$. Alternatively, it can be thought of as a function $f : \mathbb{N} \to \mathbb{R}$ with $f(n) = \frac{n+1}{n}$. In either case, the sequence takes the form

$$(\frac{2}{1}, \frac{3}{2}, \frac{4}{3}, \dots)$$

This definition admits generalization.

Definition 13.2. Let X be a set. A *sequence* of elements in X is a collection $(a_n)_{n\in\mathbb{N}}$ of elements $a_n \in X$ indexed by $\mathbb{N}$. More formally, it is a function $f : \mathbb{N} \to X$.

Crucially, sequences need not have 'rules' determining their elements, just as functions need not have rules determining their outputs. Any arbitrary $\mathbb{N}$-indexed collection of elements of X, in which adjacent terms may appear to bear no relation to one another, defines a sequence. For the moment, we fix $X = \mathbb{R}$ and restrict focus to real sequences.

Likely the primary notion associated to the sequence is that of a limit. Informally, a sequence's limit - when it exists - can be described as the point that the sequence converges to or stabilizes at as its indices increase.

Definition 13.3. The sequence $(a_n)_{n\in\mathbb{N}}$ of real numbers converges to the limit $L \in \mathbb{R}$ if for any $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that $|a_{n'} - L| < \epsilon$ when $n' \geq N$. In this case, we write $\lim_{n\to\infty} a_n = L$.

In words, for any error tolerance ϵ, there exists an index N such that a_n stays within ϵ of the limit after N. Or, perhaps more simply, a sequence with limit L eventually stays arbitrarily close to L.

Example 13.4. We claim that $\lim_{n\to\infty} \frac{n+1}{n} = 1$. Fix $\epsilon > 0$. We need an N such that $|a_n - 1| < \epsilon$ for $n > N$.

> A bit of scratch work: as a first step, let's unpack $|a_n - 1|$. It takes the value $|\frac{n+1}{n} - 1| = \frac{1}{n}$. So we need N such that $\frac{1}{n} < \epsilon$ whenever $n \geq N$. This amounts to demanding that $N > \frac{1}{\epsilon}$. Back to the proof.

Given ϵ, we select N so that it strictly exceed $\frac{1}{\epsilon}$. Now suppose $n' \geq N$. Then

$$\begin{aligned}|a_{n'} - 1| &= \frac{1}{n'} \\ &\leq \frac{1}{N} \\ &< \epsilon\end{aligned}$$

as desired.

Definition 13.5. A sequence with a limit is said to *converge*. A sequence which fails to converge *diverges*.

It is worth noting that there are various definitions for divergence, though the above is the definition adopted for this class.

Definition 13.6. A *subsequence* of a sequence $(a_1, a_2, a_3, a_4, \dots)$ is a sequence with terms drawn from $(a_n)_{n\in\mathbb{N}}$ in order and without repetition. For instance $(a_1, a_3, a_5, a_7, \dots)$ determines a subsequence of $(a_n)_{n\in\mathbb{N}}$.

Alternatively, a subsequence of $(a_n)_{n\in\mathbb{N}}$ is determined by a collection of natural numbers $n_1 < n_2 < n_3 < \dots$, which give rise to the sequence $(a_{n_1}, a_{n_2}, a_{n_3}, \dots) = (a_{n_k})_{k\in\mathbb{N}}$.

Example 13.7. Take $a_n = (-1)^n$. It is not convergent, but it does have the convergent subsequence $(1, 1, 1, \ldots)$.

So divergent sequences may have convergent subsequences. Next time, we'll discuss the Bolzano-Weierstrass theorem, a central result in the convergence of real (sub)sequences. For the moment, here's the statement.

Theorem 13.8 (Bolzano-Weierstrass). *If $(a_n)_{n\in\mathbb{N}}$ is a bounded sequence of real numbers, then the sequence has a convergent subsequence.*

14. Bolzano-Weierstrass

Last time, we began discussing the Bolzano-Weierstrass theorem, which states that any bounded real sequence admits a convergent subsequence.

Let's make a first pass at the big idea - by appropriately scaling our sequence, which does not perturb whether it has a convergent subsequence, we may assume the interval $(a_n)_{n\in\mathbb{N}}$ lies in the interval [0,1]. The first step is to split up the interval into closed halves, i.e. into $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$. Then, because a_n has infinitely many terms in $[0, 1]$, it has infinitely many terms in at least one of $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$.

Then pick one of the interval which contains infinitely many of the a_i. Subdivide this interval into closed halves again, and select one with infinitely many of the a_i. Proceeding in this way, we obtain a sequence of closed, nested intervals. The elements of the subsequence can then be selected from successive intervals, and it can be shown that their limit is the intersection of the nested intervals.

Proof of Balzano-Weierstrass. We show that there exist intervals I_1 sup I_2 sup I_3 sup $\ldots$, each of which is closed with length $\frac{1}{2^k}$ and which contains infinitely many of the a_i. This is by induction, $I_0 = [0, 1]$. Given $I_k = [s, t]$, we write $I_k = [s, \frac{s+t}{2}] \cup [\frac{s+t}{2}, t]$. Because I_k has infinitely many of the a_i, at least one of its closed sub-intervals does as well. We take it to be I_{k+1}.

We now construct the elements a_{n_k} of the subsequence. In particular, select a_{n_k} so that $a_{n_k} \in I_k$ and $n_1 < n_2 < \ldots$. This too can be done inductively. Having selected n_k, the demand that $n_{k+1} > n_k$ eliminates finitely many terms in I_{k+1} from contention, thereby permitting selection of an n_{k+1}.

The final claim is that the subsequence a_{n_k} converges to a limit L. We take $L \in \cap I_j$, which exists by the nested interval property of $\mathbb{R}$. To show $a_{n_k} \to L$, fix $\epsilon > 0$. Choose k so that $2^{-k} < \epsilon$. Then if $n' > k$, $a_{n'} \in I_k$. The distance of any two elements of I_k is at most 2^{-k}, so in particular the distance between $a_{n'}$ and L is at most $2^{-k} < \epsilon$. □

One of the attractive features of this proof is that it generalizes fairly easily to, for instance, $\mathbb{R}^n$. The only modification needed to our original definition is the replacement of absolute value with Euclidean distance.

Definition 14.1. A sequence $\{a_n\}_{n\in\mathbb{N}}$ in $\mathbb{R}^2$ converges to $L \in \mathbb{R}^2$ if $\forall \epsilon > 0$, there exists $N \in \mathbb{N}$ such that $|a_n - L| < \epsilon$ for all $n \geq N$.

The proof of Bolzano-Weierstrass in 2 dimensions proceeds as in 1. A bounded sequence lies in a square which can be repeatedly divided into squares with half width and height which contain infinitely many points in the sequence. A nested square property over $\mathbb{R}^2$ then permits the construction of a convergent subsequence.

Let's return to $\mathbb{R}$, and review some of its essential properties:

(i) Every subset which is non-empty and admits an upper bound furthermore admits a least upper bound

(ii) The nested interval property

(iii) The Bolzano-Weierstrass theorem

Another important property of $\mathbb{R}$ concerns its Cauchy sequences: in particular, real Cauchy sequences converge.

Definition 14.2. A sequence of real numbers $(a_n)_{n\in\mathbb{N}}$ is a *Cauchy sequence* if for every $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $|a_n - a_m| < \epsilon$ when $n, m \geq N$.

One advantage to the Cauchy condition for convergence is that it does not make reference to a sequence's limit; the convergence can instead be detected using only relationships between the sequence's elements.

15. Cauchy sequences, Open/closed subsets

Today, we'll continue our discussion of Cauchy sequences and touch upon 'open' and 'closed' subsets in $\mathbb{R}$, $\mathbb{R}^d$, and indeed any metric space. Recall from last time that a Cauchy sequence in $\mathbb{R}$ is a sequence (a_n) so that for any $\epsilon > 0$, there exists an N with $|a_n - a_m| < \epsilon$ for all $n, m \geq N$. In words, Cauchy sequences are those whose terms eventually stay arbitrarily close to each other.

Proposition 15.1. *If a sequence (a_n) of real numbers is Cauchy, then it is convergent: there exists $L \in \mathbb{R}$ such that $a_n \to L$.*

Proof. First, note that (a_n) is bounded as a consequence of being Cauchy. Taking $\epsilon = 1$, we have an N so that $|a_N - a_m| < 1$ for $m > N$. Then $\sup_{i\in\mathbb{N}} |a_i| = \max\{|a_1|, \ldots, |a_{N-1}|, |a_N + 1|\}$.

Next, we use the consequence of completeness of $\mathbb{R}$: bounded sequences have convergent sub-sequences. In particular, there's a subsequence $(a_{n_k})_{k\in\mathbb{N}}$ with $a_{n_k} \to L$. We'll now prove that $a_n \to L$. Fix $\epsilon > 0$. Since $a_{n_k} \to L$, there's $K \in \mathbb{N}$ such that $|a_{n_k} - L| < \frac{\epsilon}{2}$ when $k > K$. By the Cauchy property of (a_n), there exists K' such that $|a_n - a_m| < \frac{\epsilon}{2}$ when $n, m > K'$.

Now set $N = \max\{n_K, K'\}$. Then if $n > N$, we have

$$\begin{aligned}
|a_n - L| &\leq |a_n - a_{n_K}| + |a_{n_K} + L| \\
&\leq \epsilon/2 + \epsilon/2 \\
&= \epsilon
\end{aligned}$$

□

The Cauchy property can be generalized to arbitrary metric spaces (X, d) by replacing $|a_n - a_m|$ with $d(a_n, a_m)$. This gives rise to a more general definition of completeness.

Definition 15.2. A metric space (X, d) is *complete* if every Cauchy sequence in X converges in X.

Example 15.3. $X = \mathbb{R} \setminus 0$ with the restriction of the usual Euclidean metric is not complete. The sequence $a_n = 2^{-n}$ is Cauchy but does not converge in X.

We now turn our attention to formalizing the notions of open and closed sets which we have made some use of informally (e.g. in referring to $[a, b]$ as a closed interval and (a, b) as an open one).

Definition 15.4. A subset $B \subseteq \mathbb{R}^d$ is *closed* if a convergent sequence in $\mathbb{R}^d$ with elements in B necessarily has limit in B.

Example 15.5. $[a, b]$ is a closed subset of $\mathbb{R}$, for any $a < b$.

Definition 15.6. A subset $A \subseteq \mathbb{R}^d$ is *open* if it is the complement in $\mathbb{R}^d$ of a closed set.

Example 15.7. (a, b) is a closed subset of $\mathbb{R}$, for any $a < b$.

An important observation is that subsets of $\mathbb{R}^d$ need not be open or closed. For instance, $[a, b)$ is neither open nor closed.

16. Open/closed sets, Countability

Today, we'll continue on the topic of open and closed subsets of metric spaces, observe applications of the completeness property of $\mathbb{R}$, and prove that the real numbers are not 'countable.'

Let (X, d) be a metric space, e.g. $X = \mathbb{R}^k$ and $d =$ Euclidean distance. We saw last time $P \subseteq X$ is closed if it contains its limit points and open if its complement is closed. In particular, P is closed if whenever a sequence (a_n) in L with each $a_n \in P$ converges to $L \in X$, it must be that $L \in P$.

There is an alternative, perhaps more intuitive, definition of open sets.

Definition 16.1. A subset $O \subseteq X$ of a metric space X is *open* if for every $x \in O$, there exists $\epsilon > 0$ such that $B(x, \epsilon) \subseteq O$.

In words, a set is open if all of its points admit some 'wiggle room' which is still in the set.

Example 16.2. $(0, 1) \subseteq \mathbb{R}$ is open. Given $x \in (0, 1)$, set $\epsilon = \min(x, 1 - x)$. $[0, 1] \subseteq \mathbb{R}$ is not open. Set $x = 0$. Then for any $\epsilon > 0$, $B(x, \epsilon)$ contains $-\frac{\epsilon}{2}$, which is not in $[0, 1]$.

For the moment, we adopt Definition 16.1 for open sets, and leave it to a proposition to show that it coincides with the previous definition as the complement of a closed set.

Proposition 16.3. *Suppose $X = O \cup P$ and $O \cap P = \emptyset$. Then O is open if and only P is closed.*

Proof. Suppose that O is open. To see that P is closed, let (a_n) be a sequence in P with limit $x \in X$. For a contradiction, suppose $x \notin P$, meaning $x \in O$. Because O is open, there exists $\epsilon > 0$ such that $B(x, \epsilon) \subseteq O$. By convergence of (a_n) to x, there exists N such that $a_{n'} \in B(x, \epsilon)$ when $n' > N$. Then such terms lie in O. Since O and P are disjoint, this produces contradiction with $a_n \in P$.

Now suppose P is closed, and that O is not open. Then there exists $x \in O$ such that for every $\epsilon > 0$, $B(x, \epsilon)$ is not contained in O. In particular, for every $n \in \mathbb{N}$, there exists $a_n \in B(x, \frac{1}{n}) \cap P$. This defines a sequence in P which converges to $x \notin P$, contradicting closedness of P. □

Remark 16.4. $\emptyset$ is always open, as conditions on the elements of a set hold vacuously for the empty set. Likewise, for any metric space (X, d), the whole space X is open; open balls $B(x, \epsilon)$ are obligated to remain in X. There's nowhere else to go!

Proposition 16.5. *Arbitrary unions of open sets are open. Finite intersections of open sets are open.*

Opens and closed sets are the beginnings of a mathematical discussions which leads to such concepts as continuity of functions. If things had turned out differently, these ideas may have formed the next topic in the course, known as *topology*. Consider Math 131 to learn more!

We'll conclude, however, with a discussion of countability of sets.

Definition 16.6. A set S is *countable* if there exists a map $f : \mathbb{N} \to S$ which surjects. Equivalently, if there exists a sequence $(a_n)_{n \in \mathbb{N}}$ in S for which each $s \in S$ equals a_n for some n.

Of course, $\mathbb{N}$ is countable by taking f to be the identity, or considering the sequence $a_n = n$. Perhaps surprisingly, it turns out that $\mathbb{Z}$ and even $\mathbb{N} \times \mathbb{N}$ are countable as well.

Theorem 16.7. $\mathbb{R}$ *is uncountable.*

Proof. Let (a_n) be a sequence of real numbers. We'll show there exists $c \in \mathbb{R}$ which is not equal to any a_n. We construct $I_n = [a, b]$ so that $I_1 \supset I_2 \supset \ldots$ and $a_n \notin I_n$. As the initial step, take I_1 a closed interval not containing a_1, such as $I_1 = [a_1 + 1, a_1 + 2]$. Given $I_n = [a, b]$ not containing a_n, select I_{n+1} to be a closed interval in I_n which does not contain a_{n+1}, such as $[a, a + \frac{b-a}{3}]$ or $[a + 2\frac{b-a}{3}, b]$. Now consider $\cap I_n$. By the nested interval property, the intersection is non-empty. Now fix $c \in \cap I_n$. By design, $c \neq a_n$ for all n, as desired. □

This is one of the first signs of there being different kinds of infinities. In particular, $\mathbb{R}$ is a larger infinity than $\mathbb{N}$, in the sense than it cannot be enumerated using the elements of $\mathbb{N}$.

That's all - thank you for joining us, and best of luck on the final!