I

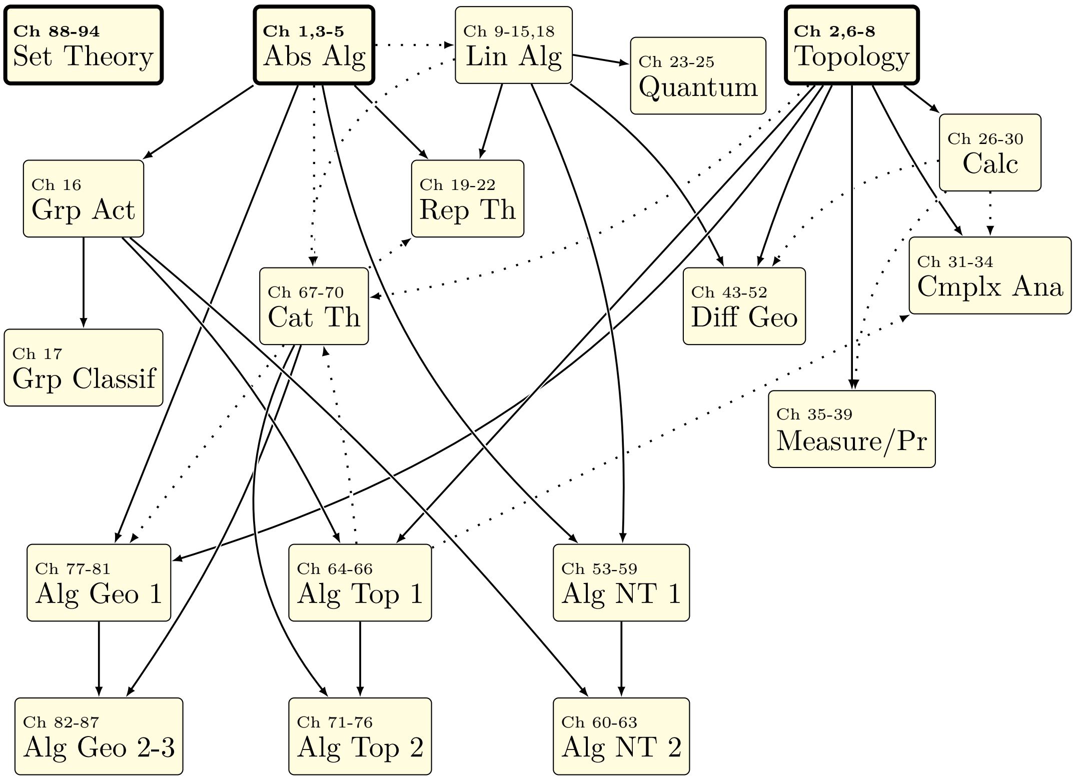
Starting Out

Part I: Contents

0 Sales pitches

This chapter contains a pitch for each part, to help you decide what you want to read and to elaborate more on how they are interconnected.

For convenience, here is again the dependency plot that appeared in the frontmatter.



§0.1 The basics

I. Starting Out.

I made a design decision that the first part should have a little bit of both algebra and topology: so this first chapter begins by defining a **group**, while the second chapter begins by defining a **metric space**. The intention is so that newcomers get to see two different examples of "sets with additional structure" in somewhat different contexts, and to have a minimal amount of literacy as these sorts of definitions appear over and over.[1]

II. Basic Abstract Algebra.

The algebraically inclined can then delve into further types of algebraic structures: some more details of **groups**, and then **rings** and **fields** — which will let you generalize $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$. So you'll learn to become familiar with all sorts of other nouns that appear in algebra, unlocking a whole host of objects that one couldn't talk about before.

[1] In particular, I think it's easier to learn what a homeomorphism is after seeing group isomorphism, and what a homomorphism is after seeing continuous map.

We'll also come to **ideals**, which generalize the GCD in $\mathbb{Z}$ that you might know of. For example, you know in $\mathbb{Z}$ that any integer can be written in the form $3a + 5b$ for $a, b \in \mathbb{Z}$, since $\gcd(3, 5) = 1$. We'll see that this statement is really a statement of ideals: "$(3, 5) = 1$ in $\mathbb{Z}$", and thus we'll understand in what situations it can be generalized, e.g. to polynomials.

III. **Basic Topology.**

The more analytically inclined can instead move into topology, learning more about spaces. We'll find out that "metric spaces" are actually too specific, and that it's better to work with **topological spaces**, which are based on the so-called **open sets**. You'll then get to see the buddings of some geometrical ideals, ending with the really great notion of **compactness**, a powerful notion that makes real analysis tick.

One example of an application of compactness to tempt you now: a continuous function $f : [0, 1] \to \mathbb{R}$ always achieves a *maximum* value. (In contrast, $f : (0, 1) \to \mathbb{R}$ by $x \mapsto 1/x$ does not.) We'll see the reason is that $[0, 1]$ is compact.

§0.2 Abstract algebra

IV. **Linear Algebra.**

In high school, linear algebra is often really unsatisfying. You are given these arrays of numbers, and they're manipulated in some ways that don't really make sense. For example, the determinant is defined as this funny-looking sum with a bunch of products that seems to come out of thin air. Where does it come from? Why does $\det(AB) = \det A \det B$ with such a bizarre formula?

Well, it turns out that you *can* explain all of these things! The trick is to not think of linear algebra as the study of matrices, but instead as the study of *linear maps*. In earlier chapters we saw that we got great generalizations by speaking of "sets with enriched structure" and "maps between them". This time, our sets are **vector spaces** and our maps are **linear maps**. We'll find out that a matrix is actually just a way of writing down a linear map as an array of numbers, but using the "intrinsic" definitions we'll de-mystify all the strange formulas from high school and show you where they all come from.

In particular, we'll see *easy* proofs that column rank equals row rank, determinant is multiplicative, trace is the sum of the diagonal entries. We'll see how the dot product works, and learn all the words starting with "eigen-". We'll even have a bonus chapter for Fourier analysis showing that you can also explain all the big buzz-words by just being comfortable with vector spaces.

V. **More on Groups.**

Some of you might be interested in more about groups, and this chapter will give you a way to play further. It starts with an exploration of **group actions**, then goes into a bit on **Sylow theorems**, which are the tools that let us try to *classify all groups*.

VI. **Representation Theory.**

If G is a group, we can try to understand it by implementing it as a *matrix*, i.e. considering embeddings $G \hookrightarrow \mathrm{GL}_n(\mathbb{C})$. These are called **representations** of G; it turns out that they can be decomposed into **irreducible** ones. Astonishingly we will find that we can *basically characterize all of them*: the results turn out to be short and completely unexpected.

For example, we will find out that there are finitely many irreducible representations of a given finite group G; if we label them V_1, V_2, $\dots$, V_r, then we will find that r is the number of conjugacy classes of G, and moreover that

$$|G| = (\dim V_1)^2 + \dots + (\dim V_r)^2$$

which comes out of nowhere!

The last chapter of this part will show you some unexpected corollaries. Here is one of them: let G be a finite group and create variables x_g for each $g \in G$. A $|G| \times |G|$ matrix M is defined by setting the (g,h)th entry to be the variable $x_{g \cdot h}$. Then this determinant will turn out to *factor*, and the factors will correspond to the V_i we described above: there will be an irreducible factor of degree $\dim V_i$ appearing $\dim V_i$ times. This result, called the **Frobenius determinant**, is said to have given birth to representation theory.

VII. Quantum Algorithms.

If you ever wondered what **Shor's algorithm** is, this chapter will use the built-up linear algebra to tell you!

§0.3 Real and complex analysis

VIII. Calculus 101.

In this part, we'll use our built-up knowledge of metric and topological spaces to give short, rigorous definitions and theorems typical of high school calculus. That is, we'll really define and prove most everything you've seen about **limits**, **series**, **derivatives**, and **integrals**.

Although this might seem intimidating, it turns out that actually, by the time we start this chapter, *the hard work has already been done*: the notion of limits, open sets, and compactness will make short work of what was swept under the rug in AP calculus. Most of the proofs will thus actually be quite short. We sit back and watch all the pieces slowly come together as a reward for our careful study of topology beforehand.

That said, if you are willing to suspend belief, you can actually read most of the other parts without knowing the exact details of all the calculus here, so in some sense this part is "optional".

IX. Complex Analysis.

It turns out that **holomorphic functions** (complex-differentiable functions) are close to the nicest things ever: they turn out to be given by a Taylor series (i.e. are basically polynomials). This means we'll be able to prove unreasonably nice results about holomorphic functions $\mathbb{C} \to \mathbb{C}$, like

- they are determined by just a few inputs,
- their contour integrals are all zero,
- they can't be bounded unless they are constant,
- $\dots$.

We then introduce **meromorphic functions**, which are like quotients of holomorphic functions, and find that we can detect their zeros by simply drawing loops in the plane and integrating over them: the famous **residue theorem** appears. (In the practice problems, you will see this even gives us a way to evaluate real integrals that can't be evaluated otherwise.)

X. **Measure Theory.**

Measure theory is the upgraded version of integration. The Riemann integration is for a lot of purposes not really sufficient; for example, if f is the function equals 1 at rational numbers but 0 at irrational numbers, we would hope that $\int_0^1 f(x)\,dx = 0$, but the Riemann integral is not capable of handling this function f.

The **Lebesgue integral** will handle these mistakes by assigning a *measure* to a generic space Ω, making it into a **measure space**. This will let us develop a richer theory of integration where the above integral *does* work out to zero because the "rational numbers have measure zero". Even the development of the measure will be an achievement, because it means we've developed a rigorous, complete way of talking about what notions like area and volume mean — on any space, not just $\mathbb{R}^n$! So for example the Lebesgue integral will let us integrate functions over any **measure space**.

XI. **Probability (TO DO).**

Using the tools of measure theory, we'll be able to start giving rigorous definitions of **probability**, too. We'll see that a **random variable** is actually a function from a measure space of worlds to $\mathbb{R}$, giving us a rigorous way to talk about its probabilities. We can then start actually stating results like the **law of large numbers** and **central limit theorem** in ways that make them both easy to state and straightforward to prove.

XII. **Differential Geometry.**

Multivariable calculus is often confusing because of all the partial derivatives. But we'll find out that, armed with our good understanding of linear algebra, that we're really looking at a **total derivative**: at every point of a function $f \colon \mathbb{R}^n \to \mathbb{R}$ we can associate a *linear map* Df which captures in one object the notion of partial derivatives. Set up this way, we'll get to see versions of **differential forms** and **Stokes' theorem**, and we finally will know what the notation dx really means. In the end, we'll say a little bit about manifolds in general.

§0.4 Algebraic number theory

XIV. **Algebraic NT I: Rings of Integers.**

Why is $3+\sqrt{5}$ the conjugate of $3-\sqrt{5}$? How come the norm $\left\lVert a+b\sqrt{5} \right\rVert = a^2-5b^2$ used in Pell's equations just happens to be multiplicative? Why is it we can do factoring into primes in $\mathbb{Z}[i]$ but not in $\mathbb{Z}[\sqrt{-5}]$? All these questions and more will be answered in this part, when we learn about **number fields**, a generalization of $\mathbb{Q}$ and $\mathbb{Z}$ to things like $\mathbb{Q}(\sqrt{5})$ and $\mathbb{Z}[\sqrt{5}]$. We'll find out that we have unique factorization into prime ideals, that there is a real *multiplicative norm* in play here, and so on. We'll also see that Pell's equation falls out of this theory.

XV. **Algebraic NT II: Galois and Ramification Theory.**

All the big buzz-words come out now: **Galois groups**, the **Frobenius**, and friends. We'll see quadratic reciprocity is just a shadow of the behavior of the Frobenius element, and meet the **Chebotarev density theorem**, which generalizes greatly the Dirichlet theorem on the infinitude of primes which are $a \pmod n$. Towards the end, we'll also state **Artin reciprocity**, one of the great results of **class field theory**, and how it generalizes quadratic reciprocity and cubic reciprocity.

§0.5 Algebraic topology

XVI. Algebraic Topology I: Homotopy.

What's the difference between an annulus and disk? Well, one of them has a "hole" in it, but if we are just given intrinsic topological spaces it's hard to make this notion precise. The **fundamental group** $\pi_1(X)$ and more general **homotopy group** will make this precise — we'll find a way to define an abelian group $\pi_1(X)$ for every topological space X which captures the idea there is a hole in the space, by throwing lassos into the space and seeing if we can reel them in.

Amazingly, the fundamental group $\pi_1(X)$ will, under mild conditions, tell you about ways to cover X with a so-called **covering projection**. One picture is that one can wrap a real line $\mathbb{R}$ into a helix shape and then project it down into the circle S^1. This will turn out to correspond to the fact that $\pi_1(S^1) = \mathbb{Z}$ which has only one subgroup. More generally the subgroups of $\pi_1(X)$ will be in bijection with ways to cover the space X!

XVII. Category Theory.

What do fields, groups, manifolds, metric spaces, measure spaces, modules, representations, rings, topological spaces, vector spaces, all have in common? Answer: they are all "objects with additional structure", with maps between them.

The notion of **category** will appropriately generalize all of them. We'll see that all sorts of constructions and ideas can be abstracted into the framework of a category, in which we *only* think about objects and arrows between them, without probing too hard into the details of what those objects are. This results in drawing many **commutative diagrams**.

For example, any way of taking an object in one category and getting another one (for example π_1 as above, from the category of spaces into the category of groups) will probably be a **functor**. We'll unify $G \times H$, $X \times Y$, $R \times S$, and anything with the × symbol into the notion of a product, and then even more generally into a **limit**. Towards the end, we talk about **abelian categories** and talk about the famous **snake lemma**, **five lemma**, and so on.

XVIII. Algebraic Topology II: Homology.

Using the language of category theory, we then resume our adventures in algebraic topology, in which we define the **homology groups** which give a different way of noticing holes in a space, in a way that is longer to define but easier to compute in practice. We'll then reverse the construction to get so-called **cohomology rings** instead, which give us an even finer invariant for telling spaces apart.

§0.6 Algebraic geometry

XIX. Algebraic Geometry I: Classical Varieties.

We begin with a classical study of classical **complex varieties**: the study of intersections of polynomial equations over $\mathbb{C}$. This will naturally lead us into the geometry of rings, giving ways to draw pictures of ideals, and motivating **Hilbert's nullstellensatz**. The **Zariski topology** will show its face, and then we'll play with **projective varieties** and **quasi-projective varieties**, with a bonus detour into **Bézout's theorem**. All this prepares us for our journey into schemes.

XX. **Algebraic Geometry II: Affine Schemes.**

We now get serious and delve into Grothendieck's definition of an **affine scheme**: a generalization of our classical varieties that allows us to start with any ring A and construct a space $\operatorname{Spec} A$ on it. We'll equip it with its own Zariski topology and then a sheaf of functions on it, making it into a **locally ringed space**; we will find that the sheaf can be understood effectively in terms of **localization** on it. We'll find that the language of commutative algebra provides elegant generalizations of what's going on geometrically: prime ideals correspond to irreducible closed subsets, radical ideals correspond to closed subsets, maximal ideals correspond to closed points, and so on. We'll draw lots of pictures of spaces and examples to accompany this.

§0.7 Set theory

XXI. **Set Theory I: ZFC, Ordinals, and Cardinals.**

Why is **Russell's paradox** such a big deal and how is it resolved? What is this **Zorn's lemma** that everyone keeps talking about? In this part we'll learn the answers to these questions by giving a real description of the **Zermelo-Frankel** axioms, and the **axiom of choice**, delving into the details of how math is built axiomatically at the very bottom foundations. We'll meet the **ordinal numbers** and **cardinal numbers** and learn how to do **transfinite induction** with them.

XXII. **Set Theory II: Model Theory and Forcing.**

The **continuum hypothesis** states that there are no cardinalities between the size of the natural numbers and the size of the real numbers. It was shown to be *independent* of the axioms — one cannot prove or disprove it. How could a result like that possibly be proved? Using our understanding of the ZF axioms, we'll develop a bit of **model theory** and then use **forcing** in order to show how to construct entire models of the universe in which the continuum hypothesis is true or false.

1 Groups

A group is one of the most basic structures in higher mathematics. In this chapter I will tell you only the bare minimum: what a group is, and when two groups are the same.

§1.1 Definition and examples of groups

Prototypical example for this section: The additive group of integers $(\mathbb{Z}, +)$ and the cyclic group $\mathbb{Z}/m\mathbb{Z}$. Just don't let yourself forget that most groups are non-commutative.

A group consists of two pieces of data: a set G, and an associative binary operation $\star$ with some properties. Before I write down the definition of a group, let me give two examples.

Example 1.1.1 (Additive integers)
The pair $(\mathbb{Z}, +)$ is a group: $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ is the set and the associative operation is *addition*. Note that

- The element $0 \in \mathbb{Z}$ is an *identity*: $a + 0 = 0 + a = a$ for any a.
- Every element $a \in \mathbb{Z}$ has an additive *inverse*: $a + (-a) = (-a) + a = 0$.

We call this group $\mathbb{Z}$.

Example 1.1.2 (Nonzero rationals)
Let $\mathbb{Q}^\times$ be the set of *nonzero rational numbers*. The pair $(\mathbb{Q}^\times, \cdot)$ is a group: the set is $\mathbb{Q}^\times$ and the associative operation is *multiplication*.

Again we see the same two nice properties.

- The element $1 \in \mathbb{Q}^\times$ is an *identity*: for any rational number, $a \cdot 1 = 1 \cdot a = a$.
- For any rational number $x \in \mathbb{Q}^\times$, we have an inverse x^{-1}, such that
$$x \cdot x^{-1} = x^{-1} \cdot x = 1.$$

From this you might already have a guess what the definition of a group is.

Definition 1.1.3. A **group** is a pair $G = (G, \star)$ consisting of a set of elements G, and a binary operation $\star$ on G, such that:

- G has an **identity element**, usually denoted 1_G or just 1, with the property that
$$1_G \star g = g \star 1_G = g \text{ for all } g \in G.$$
- The operation is **associative**, meaning $(a \star b) \star c = a \star (b \star c)$ for any $a, b, c \in G$. Consequently we generally don't write the parentheses.
- Each element $g \in G$ has an **inverse**, that is, an element $h \in G$ such that
$$g \star h = h \star g = 1_G.$$

Remark 1.1.4 (Unimportant pedantic point) — Some authors like to add a "closure" axiom, i.e. to say explicitly that $g \star h \in G$. This is implied already by the fact that $\star$ is a binary operation on G, but is worth keeping in mind for the examples below.

Remark 1.1.5 — It is not required that $\star$ is commutative ($a \star b = b \star a$). So we say that a group is **abelian** if the operation is commutative and **non-abelian** otherwise.

Example 1.1.6 (Non-Examples of groups)

- The pair $(\mathbb{Q}, \cdot)$ is NOT a group. (Here $\mathbb{Q}$ is rational numbers.) While there is an identity element, the element $0 \in \mathbb{Q}$ does not have an inverse.
- The pair $(\mathbb{Z}, \cdot)$ is also NOT a group. (Why?)
- Let $\operatorname{Mat}_{2\times 2}(\mathbb{R})$ be the set of 2×2 real matrices. Then $(\operatorname{Mat}_{2\times 2}(\mathbb{R}), \cdot)$ (where $\cdot$ is matrix multiplication) is NOT a group. Indeed, even though we have an identity matrix
$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
we still run into the same issue as before: the zero matrix does not have a multiplicative inverse.

  (Even if we delete the zero matrix from the set, the resulting structure is still not a group: those of you that know some linear algebra might recall that any matrix with determinant zero cannot have an inverse.)

Let's resume writing down examples. Here are some more **abelian examples** of groups:

Example 1.1.7 (Complex unit circle)
Let S^1 denote the set of complex numbers z with absolute value one; that is

$$S^1 := \{z \in \mathbb{C} \mid |z| = 1\}.$$

Then $(S^1, \times)$ is a group because

- The complex number $1 \in S^1$ serves as the identity, and
- Each complex number $z \in S^1$ has an inverse $\frac{1}{z}$ which is also in S^1, since $|z^{-1}| = |z|^{-1} = 1$.

There is one thing I ought to also check: that $z_1 \times z_2$ is actually still in S^1. But this follows from the fact that $|z_1 z_2| = |z_1| |z_2| = 1$.

Example 1.1.8 (Addition mod n)
Here is an example from number theory: Let $n > 1$ be an integer, and consider the residues (remainders) modulo n. These form a group under addition. We call this the **cyclic group of order** n, and denote it as $\mathbb{Z}/n\mathbb{Z}$, with elements $\overline{0}, \overline{1}, \dots$. The

identity is $\overline{0}$.

Example 1.1.9 (Multiplication mod p)

Let p be a prime. Consider the *nonzero residues modulo* p, which we denote by $(\mathbb{Z}/p\mathbb{Z})^\times$. Then $((\mathbb{Z}/p\mathbb{Z})^\times, \times)$ is a group.

Question 1.1.10. Why do we need the fact that p is prime?

(Digression: the notation $\mathbb{Z}/n\mathbb{Z}$ and $(\mathbb{Z}/p\mathbb{Z})^\times$ may seem strange but will make sense when we talk about rings and ideals. Set aside your worry for now.)

Here are some **non-abelian examples**:

Example 1.1.11 (General linear group)

Let n be a positive integer. Then $\mathrm{GL}_n(\mathbb{R})$ is defined as the set of $n \times n$ real matrices which have nonzero determinant. It turns out that with this condition, every matrix does indeed have an inverse, so $(\mathrm{GL}_n(\mathbb{R}), \times)$ is a group, called the **general linear group**.

(The fact that $\mathrm{GL}_n(\mathbb{R})$ is closed under × follows from the linear algebra fact that $\det(AB) = \det A \det B$, proved in later chapters.)

Example 1.1.12 (Special linear group)

Following the example above, let $\mathrm{SL}_n(\mathbb{R})$ denote the set of $n \times n$ matrices whose determinant is actually 1. Again, for linear algebra reasons it turns out that $(\mathrm{SL}_n(\mathbb{R}), \times)$ is also a group, called the **special linear group**.

Example 1.1.13 (Symmetric groups)

Let S_n be the set of permutations of $\{1, \dots, n\}$. By viewing these permutations as functions from $\{1, \dots, n\}$ to itself, we can consider *compositions* of permutations. Then the pair $(S_n, \circ)$ (here $\circ$ is function composition) is also a group, because

- There is an identity permutation, and
- Each permutation has an inverse.

The group S_n is called the **symmetric group** on n elements.

Example 1.1.14 (Dihedral group)

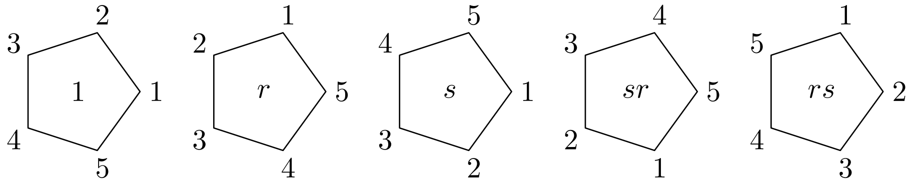
The **dihedral group of order $2n$**, denoted D_{2n}, is the group of symmetries of a regular n-gon $A_1 A_2 \dots A_n$, which includes rotations and reflections. It consists of the $2n$ elements

$$\left\{1, r, r^2, \dots, r^{n-1}, s, sr, sr^2, \dots, sr^{n-1}\right\}.$$

The element r corresponds to rotating the n-gon by $\frac{2\pi}{n}$, while s corresponds to reflecting it across the line OA_1 (here O is the center of the polygon). So rs means

"reflect then rotate" (like with function composition, we read from right to left).
In particular, $r^n = s^2 = 1$. You can also see that $r^k s = sr^{-k}$.

Here is a picture of some elements of D_{10}.



Trivia: the dihedral group D_{12} is my favorite example of a non-abelian group, and is the first group I try for any exam question of the form "find an example. . . ".

More examples:

Example 1.1.15 (Products of groups)

Let $(G, \star)$ and $(H, *)$ be groups. We can define a **product group** $(G \times H, \cdot)$, as follows. The elements of the group will be ordered pairs $(g, h) \in G \times H$. Then

$$(g_1, h_1) \cdot (g_2, h_2) = (g_1 \star g_2, h_1 * h_2) \in G \times H$$

is the group operation.

Question 1.1.16. What are the identity and inverses of the product group?

Example 1.1.17 (Trivial group)

The **trivial group**, often denoted 0 or 1, is the group with only an identity element. I will use the notation $\{1\}$.

Exercise 1.1.18. Which of these are groups?

(a) Rational numbers with odd denominators (in simplest form), where the operation is addition. (This includes integers, written as $n/1$, and $0 = 0/1$).

(b) The set of rational numbers with denominator at most 2, where the operation is addition.

(c) The set of rational numbers with denominator at most 2, where the operation is multiplication.

(d) The set of nonnegative integers, where the operation is addition.

§1.2 Properties of groups

Prototypical example for this section: $(\mathbb{Z}/p\mathbb{Z})^\times$ *is possibly best.*

Abuse of Notation 1.2.1. From now on, we'll often refer to a group $(G, \star)$ by just G. Moreover, we'll abbreviate $a \star b$ to just ab. Also, because the operation $\star$ is associative, we will omit unnecessary parentheses: $(ab)c = a(bc) = abc$.

Abuse of Notation 1.2.2. From now on, for any $g \in G$ and $n \in \mathbb{N}$ we abbreviate

$$g^n = \underbrace{g \star \cdots \star g}_{n \text{ times}}.$$

Moreover, we let g^{-1} denote the inverse of g, and $g^{-n} = (g^{-1})^n$.

In mathematics, a common theme is to require that objects satisfy certain minimalistic properties, with certain examples in mind, but then ignore the examples on paper and try to deduce as much as you can just from the properties alone. (Math olympiad veterans are likely familiar with "functional equations" in which knowing a single property about a function is enough to determine the entire function.) Let's try to do this here, and see what we can conclude just from knowing Definition 1.1.3.

It is a law in Guam and 37 other states that I now state the following proposition.

Fact 1.2.3. Let G be a group.

(a) The identity of a group is unique.

(b) The inverse of any element is unique.

(c) For any $g \in G$, $(g^{-1})^{-1} = g$.

Proof. This is mostly just some formal manipulations, and you needn't feel bad skipping it on a first read.

(a) If 1 and $1'$ are identities, then $1 = 1 \star 1' = 1'$.

(b) If h and h' are inverses to g, then $1_G = g \star h \implies h' = (h' \star g) \star h = 1_G \star h = h$.

(c) Trivial; omitted. □

Now we state a slightly more useful proposition.

Proposition 1.2.4 (Inverse of products)
Let G be a group, and $a, b \in G$. Then $(ab)^{-1} = b^{-1}a^{-1}$.

Proof. Direct computation. We have

$$(ab)(b^{-1}a^{-1}) = a(bb^{-1})a^{-1} = aa^{-1} = 1_G.$$

Similarly, $(b^{-1}a^{-1})(ab) = 1_G$ as well. Hence $(ab)^{-1} = b^{-1}a^{-1}$. □

Finally, we state a very important lemma about groups, which highlights why having an inverse is so valuable.

Lemma 1.2.5 (Left multiplication is a bijection)
Let G be a group, and pick a $g \in G$. Then the map $G \to G$ given by $x \mapsto gx$ is a bijection.

> **Exercise 1.2.6.** Check this by showing injectivity and surjectivity directly. (If you don't know what these words mean, consult Appendix E.)

Example 1.2.7

Let $G = (\mathbb{Z}/7\mathbb{Z})^\times$ (as in Example 1.1.9) and pick $g = 3$. The above lemma states that the map $x \mapsto 3 \cdot x$ is a bijection, and we can see this explicitly:

$$\begin{aligned} 1 &\xmapsto{\times 3} 3 \pmod 7 \\ 2 &\xmapsto{\times 3} 6 \pmod 7 \\ 3 &\xmapsto{\times 3} 2 \pmod 7 \\ 4 &\xmapsto{\times 3} 5 \pmod 7 \\ 5 &\xmapsto{\times 3} 1 \pmod 7 \\ 6 &\xmapsto{\times 3} 4 \pmod 7. \end{aligned}$$

The fact that the map is injective is often called the **cancellation law**. (Why do you think so?)

Abuse of Notation 1.2.8 (Later on, sometimes the identity is denoted 0 instead of 1)**.** You don't need to worry about this for a few chapters, but I'll bring it up now anyways. In most of our examples up until now the operation $\star$ was thought of like multiplication of some sort, which is why $1 = 1_G$ was a natural notation for the identity element.

But there are groups like $\mathbb{Z} = (\mathbb{Z}, +)$ where the operation $\star$ is thought of as addition, in which case the notation $0 = 0_G$ might make more sense instead. (In general, whenever an operation is denoted $+$, the operation is almost certainly commutative.) We will eventually start doing so too when we discuss rings and linear algebra.

§1.3 Isomorphisms

Prototypical example for this section: $\mathbb{Z} \cong 10\mathbb{Z}$.

First, let me talk about what it means for groups to be isomorphic. Consider the two groups

- $\mathbb{Z} = (\{\dots, -2, -1, 0, 1, 2, \dots\}, +)$.
- $10\mathbb{Z} = (\{\dots, -20, -10, 0, 10, 20, \dots\}, +)$.

These groups are "different", but only superficially so – you might even say they only differ in the names of the elements. Think about what this might mean formally for a moment.

Specifically the map

$$\phi\colon \mathbb{Z} \to 10\mathbb{Z} \text{ by } x \mapsto 10x$$

is a bijection of the underlying sets which respects the group operation. In symbols,

$$\phi(x + y) = \phi(x) + \phi(y).$$

In other words, ϕ is a way of re-assigning names of the elements without changing the structure of the group. That's all just formalism for capturing the obvious fact that $(\mathbb{Z}, +)$ and $(10\mathbb{Z}, +)$ are the same thing.

Now, let's do the general definition.

Definition 1.3.1. Let $G = (G, \star)$ and $H = (H, *)$ be groups. A bijection $\phi\colon G \to H$ is called an **isomorphism** if

$$\phi(g_1 \star g_2) = \phi(g_1) * \phi(g_2) \quad \text{for all } g_1, g_2 \in G.$$

If there exists an isomorphism from G to H, then we say G and H are **isomorphic** and write $G \cong H$.

Note that in this definition, the left-hand side $\phi(g_1 \star g_2)$ uses the operation of G while the right-hand side $\phi(g_1) * \phi(g_2)$ uses the operation of H.

Example 1.3.2 (Examples of isomorphisms)

Let G and H be groups. We have the following isomorphisms.

(a) $\mathbb{Z} \cong 10\mathbb{Z}$, as above.

(b) There is an isomorphism

$$G \times H \cong H \times G$$

by the map $(g, h) \mapsto (h, g)$.

(c) The identity map $\mathrm{id}\colon G \to G$ is an isomorphism, hence $G \cong G$.

(d) There is another isomorphism of $\mathbb{Z}$ to itself: send every x to $-x$.

Example 1.3.3 (Primitive roots modulo 7)

As a nontrivial example, we claim that $\mathbb{Z}/6\mathbb{Z} \cong (\mathbb{Z}/7\mathbb{Z})^\times$. The bijection is

$$\phi(a \bmod 6) = 3^a \bmod 7.$$

- This map is a bijection by explicit calculation:

  $$(3^0, 3^1, 3^2, 3^3, 3^4, 3^5) \equiv (1, 3, 2, 6, 4, 5) \pmod 7.$$

  (Technically, I should more properly write $3^{0 \bmod 6} = 1$ and so on to be pedantic.)

- Finally, we need to verify that this map respects the group action. In other words, we want to see that $\phi(a + b) = \phi(a)\phi(b)$ since the operation of $\mathbb{Z}/6\mathbb{Z}$ is addition while the operation of $(\mathbb{Z}/7\mathbb{Z})^\times$ is multiplication. That's just saying that $3^{a+b \bmod 6} \equiv 3^{a \bmod 6} 3^{b \bmod 6} \pmod 7$, which is true.

Example 1.3.4 (Primitive roots)

More generally, for any prime p, there exists an element $g \in (\mathbb{Z}/p\mathbb{Z})^\times$ called a **primitive root** modulo p such that $1, g, g^2, \dots, g^{p-2}$ are all different modulo p. One can show by copying the above proof that

$$\mathbb{Z}/(p-1)\mathbb{Z} \cong (\mathbb{Z}/p\mathbb{Z})^\times \text{ for all primes } p.$$

The example above was the special case $p = 7$ and $g = 3$.

Exercise 1.3.5. Assuming the existence of primitive roots, establish the isomorphism $\mathbb{Z}/(p-1)\mathbb{Z} \cong (\mathbb{Z}/p\mathbb{Z})^\times$ as above.

It's not hard to see that $\cong$ is an equivalence relation (why?). Moreover, because we really only care about the structure of groups, we'll usually consider two groups to be the same when they are isomorphic. So phrases such as "find all groups" really mean "find all groups up to isomorphism".

§1.4 Orders of groups, and Lagrange's theorem

Prototypical example for this section: $(\mathbb{Z}/p\mathbb{Z})^\times$.

As is typical in math, we use the word "order" for way too many things. In groups, there are two notions of order.

Definition 1.4.1. The **order of a group** G is the number of elements of G. We denote this by $|G|$. Note that the order may not be finite, as in $\mathbb{Z}$. We say G is a **finite group** just to mean that $|G|$ is finite.

Example 1.4.2 (Orders of groups)

For a prime p, $|(\mathbb{Z}/p\mathbb{Z})^\times| = p-1$. In other words, the order of $(\mathbb{Z}/p\mathbb{Z})^\times$ is $p-1$. As another example, the order of the symmetric group S_n is $n!$ and the order of the dihedral group D_{2n} is $2n$.

Definition 1.4.3. The **order of an element** $g \in G$ is the smallest positive integer n such that $g^n = 1_G$, or ∞ if no such n exists. We denote this by $\operatorname{ord} g$.

Example 1.4.4 (Examples of orders)

The order of -1 in $\mathbb{Q}^\times$ is 2, while the order of 1 in $\mathbb{Z}$ is infinite.

Question 1.4.5. Find the order of each of the six elements of $\mathbb{Z}/6\mathbb{Z}$, the cyclic group on six elements. (See Example 1.1.8 if you've forgotten what $\mathbb{Z}/6\mathbb{Z}$ means.)

Example 1.4.6 (Primitive roots)

If you know olympiad number theory, this coincides with the definition of an order of a residue mod p. That's why we use the term "order" there as well. In particular, a primitive root is precisely an element $g \in (\mathbb{Z}/p\mathbb{Z})^\times$ such that $\operatorname{ord} g = p-1$.

You might also know that if $x^n \equiv 1 \pmod{p}$, then the order of $x \pmod{p}$ must divide n. The same is true in a general group for exactly the same reason.

Fact 1.4.7. If $g^n = 1_G$ then $\operatorname{ord} g$ divides n.

Also, you can show that any element of a finite group has a finite order. The proof is just an olympiad-style pigeonhole argument. Consider the infinite sequence $1_G, g, g^2, \dots$, and find two elements that are the same.

Fact 1.4.8. Let G be a finite group. For any $g \in G$, $\operatorname{ord} g$ is finite.

What's the last property of $(\mathbb{Z}/p\mathbb{Z})^\times$ that you know from olympiad math? We have Fermat's little theorem: for any $a \in (\mathbb{Z}/p\mathbb{Z})^\times$, we have $a^{p-1} \equiv 1 \pmod p$. This is no coincidence: exactly the same thing is true in a more general setting.

Theorem 1.4.9 (Lagrange's theorem for orders)

Let G be any finite group. Then $x^{|G|} = 1_G$ for any $x \in G$.

Keep this result in mind! We'll prove it later in the generality of Theorem 3.4.1.

§1.5 Subgroups

Prototypical example for this section: $\mathrm{SL}_n(\mathbb{R})$ *is a subgroup of* $\mathrm{GL}_n(\mathbb{R})$.

Earlier we saw that $\mathrm{GL}_n(\mathbb{R})$, the $n \times n$ matrices with nonzero determinant, formed a group under matrix multiplication. But we also saw that a subset of $\mathrm{GL}_n(\mathbb{R})$, namely $\mathrm{SL}_n(\mathbb{R})$, also formed a group with the same operation. For that reason we say that $\mathrm{SL}_n(\mathbb{R})$ is a subgroup of $\mathrm{GL}_n(\mathbb{R})$. And this definition generalizes in exactly the way you expect.

Definition 1.5.1. Let $G = (G, \star)$ be a group. A **subgroup** of G is exactly what you would expect it to be: a group $H = (H, \star)$ where H is a subset of G. It's a **proper subgroup** if $H \neq G$.

Remark 1.5.2 — To specify a group G, I needed to tell you both what the set G was and the operation $\star$ was. But to specify a subgroup H of a given group G, I only need to tell you who its elements are: the operation of H is just inherited from the operation of G.

Example 1.5.3 (Examples of subgroups)

(a) $2\mathbb{Z}$ is a subgroup of $\mathbb{Z}$, which is isomorphic to $\mathbb{Z}$ itself!

(b) Consider again S_n, the symmetric group on n elements. Let T be the set of permutations $\tau\colon \{1, \dots, n\} \to \{1, \dots, n\}$ for which $\tau(n) = n$. Then T is a subgroup of S_n; in fact, it is isomorphic to S_{n-1}.

(c) Consider the group $G \times H$ (Example 1.1.15) and the elements $\{(g, 1_H) \mid g \in G\}$. This is a subgroup of $G \times H$ (why?). In fact, it is isomorphic to G by the isomorphism $(g, 1_H) \mapsto g$.

Example 1.5.4 (Stupid examples of subgroups)

For any group G, the trivial group $\{1_G\}$ and the entire group G are subgroups of G.

Next is an especially important example that we'll talk about more in later chapters.

Example 1.5.5 (Subgroup generated by an element)

Let x be an element of a group G. Consider the set

$$\langle x \rangle = \left\{ \dots, x^{-2}, x^{-1}, 1, x, x^2, \dots \right\}.$$

This is also a subgroup of G, called the subgroup generated by x.

Exercise 1.5.6. If $\operatorname{ord} x = 2015$, what is the above subgroup equal to? What if $\operatorname{ord} x = \infty$?

Finally, we present some non-examples of subgroups.

Example 1.5.7 (Non-examples of subgroups)

Consider the group $\mathbb{Z} = (\mathbb{Z}, +)$.

(a) The set $\{0, 1, 2, \dots\}$ is not a subgroup of $\mathbb{Z}$ because it does not contain inverses.

(b) The set $\{n^3 \mid n \in \mathbb{Z}\} = \{\dots, -8, -1, 0, 1, 8, \dots\}$ is not a subgroup because it is not closed under addition; the sum of two cubes is not in general a cube.

(c) The empty set $\varnothing$ is not a subgroup of $\mathbb{Z}$ because it lacks an identity element.

§1.6 Groups of small orders

Just for fun, here is a list of all groups of order less than or equal to ten (up to isomorphism, of course).

1. The only group of order 1 is the trivial group.

2. The only group of order 2 is $\mathbb{Z}/2\mathbb{Z}$.

3. The only group of order 3 is $\mathbb{Z}/3\mathbb{Z}$.

4. The only groups of order 4 are
   - $\mathbb{Z}/4\mathbb{Z}$, the cyclic group on four elements,
   - $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, called the Klein Four Group.

5. The only group of order 5 is $\mathbb{Z}/5\mathbb{Z}$.

6. The groups of order six are
   - $\mathbb{Z}/6\mathbb{Z}$, the cyclic group on six elements.
   - S_3, the permutation group of three elements. This is the first non-abelian group.

   Some of you might wonder where $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ is. All I have to say is: Chinese remainder theorem!

   You might wonder where D_6 is in this list. It's actually isomorphic to S_3.

7. The only group of order 7 is $\mathbb{Z}/7\mathbb{Z}$.

8. The groups of order eight are more numerous.

- $\mathbb{Z}/8\mathbb{Z}$, the cyclic group on eight elements.
- $\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.
- $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.
- D_8, the dihedral group with eight elements, which is not abelian.
- A non-abelian group Q_8, called the *quaternion group*. It consists of eight elements ± 1, $\pm i$, $\pm j$, $\pm k$ with $i^2 = j^2 = k^2 = ijk = -1$.

9. The groups of order nine are
   - $\mathbb{Z}/9\mathbb{Z}$, the cyclic group on nine elements.
   - $\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$.

10. The groups of order 10 are
    - $\mathbb{Z}/10\mathbb{Z} \cong \mathbb{Z}/5\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ (again Chinese remainder theorem).
    - D_{10}, the dihedral group with 10 elements. This group is non-abelian.

§1.7 Unimportant long digression

A common question is: why these axioms? For example, why associative but not commutative? This answer will likely not make sense until later, but here are some comments that may help.

One general heuristic is: Whenever you define a new type of general object, there's always a balancing act going on. On the one hand, you want to include enough constraints that your objects are "nice". On the other hand, if you include too many constraints, then your definition applies to too few objects.

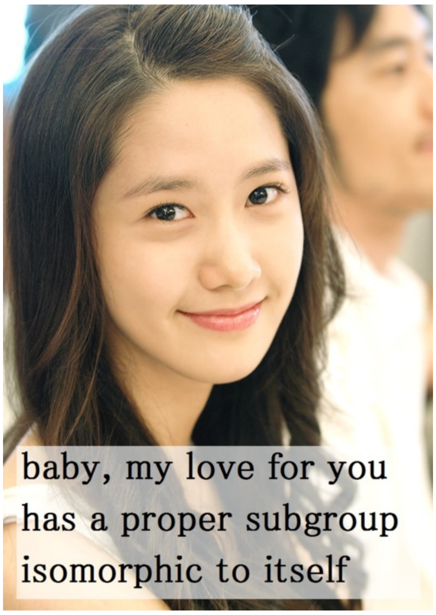
So, for example, we include "associative" because that makes our lives easier and most operations we run into are associative. In particular, associativity is required for the inverse of an element to necessarily be unique. However we don't include "commutative", because examples below show that there are lots of non-abelian groups we care about. (But we introduce another name "abelian" because we still want to keep track of it.)

Another comment: a good motivation for the inverse axioms is that you get a large amount of *symmetry*. The set of positive integers with addition is not a group, for example, because you can't subtract 6 from 3: some elements are "larger" than others. By requiring an inverse element to exist, you get rid of this issue. (You also need identity for this; it's hard to define inverses without it.)

Even more abstruse comment: Problem 1F† shows that groups are actually shadows of the so-called symmetric groups (defined later, also called permutation groups). This makes rigorous the notion that "groups are very symmetric".

§1.8 A few harder problems to think about

Problem 1A. What is the joke in the following figure? (Source: [Ge].)



Problem 1B. Prove Lagrange's theorem for orders in the special case that G is a finite abelian group.

Problem 1C. Show that $D_6 \cong S_3$ but $D_{24} \ncong S_4$.

Problem 1D*. Let p be a prime. Show that if G is a group of order p then $G \cong \mathbb{Z}/p\mathbb{Z}$.

Problem 1E (A hint for Cayley's theorem)**.** Find a subgroup H of S_8 which is isomorphic to D_8, and write the isomorphism explicitly.

Problem 1F†. Let G be a finite group.[1] Show that there exists a positive integer n such that

(a) (Cayley's theorem) G is isomorphic to some subgroup of the symmetric group S_n.

(b) (Representation Theory) G is isomorphic to some subgroup of the general linear group $\mathrm{GL}_n(\mathbb{R})$. (This is the group of invertible $n \times n$ matrices.)

Problem 1G. Find the smallest integer n such that the symmetric group S_n has a subgroup isomorphic to the dihedral group D_{2018} of order 2018.

Problem 1H (IMO SL 2005 C5)**.** There are n markers, each with one side white and the other side black. In the beginning, these n markers are aligned in a row so that their white sides are all up. In each step, if possible, we choose a marker whose white side is up (but not one of the outermost markers), remove it, and reverse the closest marker to the left of it and also reverse the closest marker to the right of it.

Prove that if $n \equiv 1 \pmod 3$ it's impossible to reach a state with only two markers remaining. (In fact the converse is true as well.)

Problem 1I. Let p be a prime and $F_1 = F_2 = 1$, $F_{n+2} = F_{n+1} + F_n$ be the Fibonacci sequence. Show that $F_{2p(p^2-1)}$ is divisible by p.

[1] In other words, permutation groups can be arbitrarily weird. I remember being highly unsettled by this theorem when I first heard of it, but in hindsight it is not so surprising.

2 Metric spaces

At the time of writing, I'm convinced that metric topology is the morally correct way to motivate point-set topology as well as to generalize normal calculus.[1] So here is my best attempt.

The concept of a metric space is very "concrete", and lends itself easily to visualization. Hence throughout this chapter you should draw lots of pictures as you learn about new objects, like convergent sequences, open sets, closed sets, and so on.

§2.1 Definition and examples of metric spaces

Prototypical example for this section: $\mathbb{R}^2$, *with the Euclidean metric.*

Definition 2.1.1. A **metric space** is a pair (M, d) consisting of a set of points M and a **metric** $d \colon M \times M \to \mathbb{R}_{\ge 0}$. The distance function must obey:

- For any $x, y \in M$, we have $d(x, y) = d(y, x)$; i.e. d is symmetric.
- The function d must be **positive definite** which means that $d(x, y) \ge 0$ with equality if and only if $x = y$.
- The function d should satisfy the **triangle inequality**: for all $x, y, z \in M$,
$$d(x, z) + d(z, y) \ge d(x, y).$$

Abuse of Notation 2.1.2. Just like with groups, we will abbreviate (M, d) as just M.

Example 2.1.3 (Metric spaces of $\mathbb{R}$)

(a) The real line $\mathbb{R}$ is a metric space under the metric $d(x, y) = |x - y|$.

(b) The interval $[0, 1]$ is also a metric space with the same distance function.

(c) In fact, any subset S of $\mathbb{R}$ can be made into a metric space in this way.

Example 2.1.4 (Metric spaces of $\mathbb{R}^2$)

(a) We can make $\mathbb{R}^2$ into a metric space by imposing the Euclidean distance function
$$d\left((x_1, y_1), (x_2, y_2)\right) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.$$

(b) Just like with the first example, any subset of $\mathbb{R}^2$ also becomes a metric space after we inherit it. The unit disk, unit circle, and the unit square $[0, 1]^2$ are special cases.

[1] Also, "metric" is a fun word to say.

Example 2.1.5 (Taxicab on $\mathbb{R}^2$)

It is also possible to place the **taxicab distance** on $\mathbb{R}^2$:

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

For now, we will use the more natural Euclidean metric.

Example 2.1.6 (Metric spaces of $\mathbb{R}^n$)

We can generalize the above examples easily. Let n be a positive integer.

(a) We let $\mathbb{R}^n$ be the metric space whose points are points in n-dimensional Euclidean space, and whose metric is the Euclidean metric

$$d\left((a_1,\dots,a_n),(b_1,\dots,b_n)\right) = \sqrt{(a_1-b_1)^2+\dots+(a_n-b_n)^2}.$$

This is the n-dimensional **Euclidean space**.

(b) The open **unit ball** B^n is the subset of $\mathbb{R}^n$ consisting of those points $(x_1,\dots,x_n)$ such that $x_1^2+\dots+x_n^2<1$.

(c) The **unit sphere** S^{n-1} is the subset of $\mathbb{R}^n$ consisting of those points $(x_1,\dots,x_n)$ such that $x_1^2+\dots+x_n^2=1$, with the inherited metric. (The superscript $n-1$ indicates that S^{n-1} is an $n-1$ dimensional space, even though it lives in n-dimensional space.) For example, $S^1 \subseteq \mathbb{R}^2$ is the unit circle, whose distance between two points is the length of the chord joining them. You can also think of it as the "boundary" of the unit ball B^n.

Example 2.1.7 (Function space)

We can let M be the space of continuous functions $f\colon [0,1]\to\mathbb{R}$ and define the metric by $d(f,g)=\int_0^1 |f-g| \, dx$. (It admittedly takes some work to check $d(f,g)=0$ implies $f=g$, but we won't worry about that yet.)

Here is a slightly more pathological example.

Example 2.1.8 (Discrete space)

Let S be any set of points (either finite or infinite). We can make S into a **discrete space** by declaring

$$d(x,y) = \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{if } x = y. \end{cases}$$

If $|S|=4$ you might think of this space as the vertices of a regular tetrahedron, living in $\mathbb{R}^3$. But for larger S it's not so easy to visualize...

Example 2.1.9 (Graphs are metric spaces)

Any connected simple graph G can be made into a metric space by defining the distance between two vertices to be the graph-theoretic distance between them. (The discrete metric is the special case when G is the complete graph on S.)

Question 2.1.10. Check the conditions of a metric space for the metrics on the discrete space and for the connected graph.

Abuse of Notation 2.1.11. From now on, we will refer to $\mathbb{R}^n$ with the Euclidean metric by just $\mathbb{R}^n$. Moreover, if we wish to take the metric space for a subset $S \subseteq \mathbb{R}^n$ with the inherited metric, we will just write S.

§2.2 Convergence in metric spaces

Prototypical example for this section: The sequence $\frac{1}{n}$ (for $n = 1, 2, \dots$) in $\mathbb{R}$.

Since we can talk about the distance between two points, we can talk about what it means for a sequence of points to converge. This is the same as the typical epsilon-delta definition, with absolute values replaced by the distance function.

Definition 2.2.1. Let $(x_n)_{n\ge1}$ be a sequence of points in a metric space M. We say that x_n **converges** to x if the following condition holds: for all $\varepsilon > 0$, there is an integer N (depending on ε) such that $d(x_n, x) < \varepsilon$ for each $n \ge N$. This is written

$$x_n \to x$$

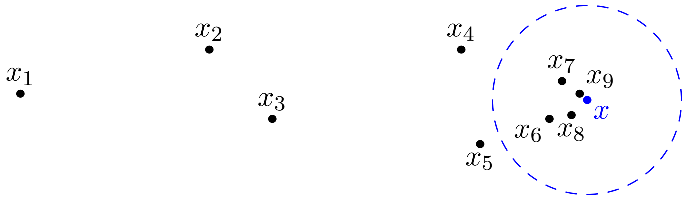
or more verbosely as

$$\lim_{n\to\infty} x_n = x.$$

We say that a sequence converges in M if it converges to a point in M.

You should check that this definition coincides with your intuitive notion of "converges".

Abuse of Notation 2.2.2. If the parent space M is understood, we will allow ourselves to abbreviate "converges in M" to just "converges". However, keep in mind that convergence is defined relative to the parent space; the "limit" of the space must actually be a point in M for a sequence to converge.



Example 2.2.3

Consider the sequence $x_1 = 1$, $x_2 = 1.4$, $x_3 = 1.41$, $x_4 = 1.414$, $\dots$.

(a) If we view this as a sequence in $\mathbb{R}$, it converges to $\sqrt{2}$.

(b) However, even though each x_i is in $\mathbb{Q}$, this sequence does NOT converge when we view it as a sequence in $\mathbb{Q}$!

Question 2.2.4. What are the convergent sequences in a discrete metric space?

§2.3 Continuous maps

In calculus you were also told (or have at least heard) of what it means for a function to be continuous. Probably something like

> A function $f\colon \mathbb{R} \to \mathbb{R}$ is continuous at a point $p \in \mathbb{R}$ if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $|x - p| < \delta \implies |f(x) - f(p)| < \varepsilon$.

Question 2.3.1. Can you guess what the corresponding definition for metric spaces is?

All we have do is replace the absolute values with the more general distance functions: this gives us a definition of continuity for any function $M \to N$.

Definition 2.3.2. Let $M = (M, d_M)$ and $N = (N, d_N)$ be metric spaces. A function $f\colon M \to N$ is **continuous** at a point $p \in M$ if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$d_M(x, p) < \delta \implies d_N(f(x), f(p)) < \varepsilon.$$

Moreover, the entire function f is continuous if it is continuous at every point $p \in M$.

Notice that, just like in our definition of an isomorphism of a group, we use the metric of M for one condition and the metric of N for the other condition.

This generalization is nice because it tells us immediately how we could carry over continuity arguments in $\mathbb{R}$ to more general spaces like $\mathbb{C}$. Nonetheless, this definition is kind of cumbersome to work with, because it makes extensive use of the real numbers (epsilons and deltas). Here is an equivalent condition.

Theorem 2.3.3 (Sequential continuity)

A function $f\colon M \to N$ of metric spaces is continuous at a point $p \in M$ if and only if the following property holds: if $x_1, x_2, \dots$ is a sequence in M converging to p, then the sequence $f(x_1), f(x_2), \dots$ in N converges to $f(p)$.

Proof. One direction is not too hard:

Exercise 2.3.4. Show that ε-δ continuity implies sequential continuity at each point.

Conversely, we will prove if f is not ε-δ continuous at p then it does not preserve convergence.

If f is not continuous at p, then there is a "bad" $\varepsilon > 0$, which we now consider fixed. So for each choice of $\delta = 1/n$, there should be some point x_n which is within δ of p, but which is mapped more than ε away from $f(p)$. But then the sequence x_n converges to p, and $f(x_n)$ is always at least ε away from $f(p)$, contradiction. □

Example application showcasing the niceness of sequential continuity:

Proposition 2.3.5 (Composition of continuous functions is continuous)

Let $f\colon M \to N$ and $g\colon N \to L$ be continuous maps of metric spaces. Then their composition $g \circ f$ is continuous.

Proof. Dead simple with sequences: Let $p \in M$ be arbitrary and let $x_n \to p$ in M. Then $f(x_n) \to f(p)$ in N and $g(f(x_n)) \to g(f(p))$ in L, QED. □

> **Question 2.3.6.** Let M be any metric space and D a discrete space. When is a map $f\colon D \to M$ continuous?

§2.4 Homeomorphisms

Prototypical example for this section: The unit circle S^1 is homeomorphic to the boundary of the unit square.

When do we consider two groups to be the same? Answer: if there's a structure-preserving map between them which is also a bijection. For metric spaces, we do exactly the same thing, but replace "structure-preserving" with "continuous".

Definition 2.4.1. Let M and N be metric spaces. A function $f\colon M \to N$ is a **homeomorphism** if it is a bijection, and both $f\colon M \to N$ and its inverse $f^{-1}\colon N \to M$ are continuous. We say M and N are **homeomorphic**.

Needless to say, homeomorphism is an equivalence relation.

You might be surprised that we require f^{-1} to also be continuous. Here's the reason: you can show that if ϕ is an isomorphism of groups, then ϕ^{-1} also preserves the group operation, hence ϕ^{-1} is itself an isomorphism. The same is not true for continuous bijections, which is why we need the new condition.

Example 2.4.2 (Homeomorphism $\neq$ continuous bijection)

(a) There is a continuous bijection from $[0, 1)$ to the circle, but it has no continuous inverse.

(b) Let M be a discrete space with size $|\mathbb{R}|$. Then there is a continuous function $M \to \mathbb{R}$ which certainly has no continuous inverse.

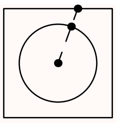
Note that this is the topologist's definition of "same" – homeomorphisms are "continuous deformations". Here are some examples.

Example 2.4.3 (Examples of homeomorphisms)

(a) Any space M is homeomorphic to itself through the identity map.

(b) The old saying: a doughnut (torus) is homeomorphic to a coffee cup. (Look this up if you haven't heard of it.)

(c) The unit circle S^1 is homeomorphic to the boundary of the unit square. Here's one bijection between them, after an appropriate scaling:



Example 2.4.4 (Metrics on the unit circle)

It may have seemed strange that our metric function on S^1 was the one inherited from $\mathbb{R}^2$, meaning the distance between two points on the circle was defined to be the length of the chord. Wouldn't it have made more sense to use the circumference of the smaller arc joining the two points?

In fact, it doesn't matter: if we consider S^1 with the "chord" metric and the "arc" metric, we get two homeomorphic spaces, as the map between them is continuous.

The same goes for S^{n-1} for general n.

Example 2.4.5 (Homeomorphisms really don't preserve size)

Surprisingly, the open interval $(-1,1)$ is homeomorphic to the real line $\mathbb{R}$! One bijection is given by

$$x \mapsto \tan(x\pi/2)$$

with the inverse being given by $t \mapsto \frac{2}{\pi}\arctan(t)$.

This might come as a surprise, since $(-1,1)$ doesn't look that much like $\mathbb{R}$; the former is "bounded" while the latter is "unbounded".

§2.5 Extended example/definition: product metric

Prototypical example for this section: $\mathbb{R} \times \mathbb{R}$ *is* $\mathbb{R}^2$.

Here is an extended example which will occur later on. Let $M = (M, d_M)$ and $N = (N, d_N)$ be metric spaces (say, $M = N = \mathbb{R}$). Our goal is to define a metric space on $M \times N$.

Let $p_i = (x_i, y_i) \in M \times N$ for $i = 1, 2$. Consider the following metrics on the set of points $M \times N$:

$$\begin{aligned} d_{\max}(p_1,p_2) &:= \max\left\{d_M(x_1,x_2), d_N(y_1,y_2)\right\} \\ d_{\text{Euclid}}(p_1,p_2) &:= \sqrt{d_M(x_1,x_2)^2 + d_N(y_1,y_2)^2} \\ d_{\text{taxicab}}(p_1,p_2) &:= d_M(x_1,x_2) + d_N(y_1,y_2). \end{aligned}$$

All of these are good candidates. We are about to see it doesn't matter which one we use:

Exercise 2.5.1. Verify that

$$d_{\max}(p_1,p_2) \le d_{\text{Euclid}}(p_1,p_2) \le d_{\text{taxicab}}(p_1,p_2) \le 2d_{\max}(p_1,p_2).$$

Use this to show that the metric spaces we obtain by imposing any of the three metrics are homeomorphic, with the homeomorphism being just the identity map.

Definition 2.5.2. Hence we will usually simply refer to *the* metric on $M \times N$, called the **product metric**. It will not be important which of the three metrics we select.

Example 2.5.3 ($\mathbb{R}^2$)
If $M = N = \mathbb{R}$, we get $\mathbb{R}^2$, the Euclidean plane. The metric d_{Euclid} is the one we started with, but using either of the other two metric works fine as well.

The product metric plays well with convergence of sequences.

Proposition 2.5.4 (Convergence in the product metric is by component)
We have $(x_n, y_n) \to (x, y)$ if and only if $x_n \to x$ and $y_n \to y$.

Proof. We have $d_{\max}((x, y), (x_n, y_n)) = \max\{d_M(x, x_n), d_N(y, y_n)\}$ and the latter approaches zero as $n \to \infty$ if and only if $d_M(x, x_n) \to 0$ and $d_N(y, y_n) \to 0$. □

Let's see an application of this:

Proposition 2.5.5 (Addition and multiplication are continuous)
The addition and multiplication maps are continuous maps $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$.

Proof. For multiplication: for any n we have

$$\begin{aligned} x_n y_n &= (x + (x_n - x))(y + (y_n - y)) \\ &= xy + y(x_n - x) + x(y_n - y) + (x_n - x)(y_n - y) \\ \implies |x_n y_n - xy| &\le |y| |x_n - x| + |x| |y_n - y| + |x_n - x| |y_n - y| . \end{aligned}$$

As $n \to \infty$, all three terms on the right-hand side tend to zero. The proof that $+ \colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is continuous is similar (and easier): one notes for any n that

$$|(x_n + y_n) - (x + y)| \le |x_n - x| + |y_n - y|$$

and both terms on the right-hand side tend to zero as $n \to \infty$. □

Problem 2C covers the other two operations, subtraction and division. The upshot of this is that, since compositions are also continuous, most of your naturally arising real-valued functions will automatically be continuous as well. For example, the function $\frac{3x}{x^2+1}$ will be a continuous function from $\mathbb{R} \to \mathbb{R}$, since it can be obtained by composing $+$, $\times$, $\div$.

§2.6 Open sets

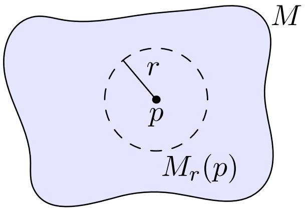
Prototypical example for this section: The open disk $x^2 + y^2 < r^2$ in $\mathbb{R}^2$.

Continuity is really about what happens "locally": how a function behaves "close to a certain point p". One way to capture this notion of "closeness" is to use metrics as we've done above. In this way we can define an r-neighborhood of a point.

Definition 2.6.1. Let M be a metric space. For each real number $r > 0$ and point $p \in M$, we define

$$M_r(p) := \{x \in M : d(x, p) < r\} .$$

The set $M_r(p)$ is called an ***r*-neighborhood** of p.



We can rephrase convergence more succinctly in terms of r-neighborhoods. Specifically, a sequence (x_n) converges to x if for every r-neighborhood of x, all terms of x_n eventually stay within that r-neighborhood.

Let's try to do the same with functions.

> **Question 2.6.2.** In terms of r-neighborhoods, what does it mean for a function $f \colon M \to N$ to be continuous at a point $p \in M$?

Essentially, we require that the pre-image of every ε-neighborhood has the property that some δ-neighborhood exists inside it. This motivates:

Definition 2.6.3. A set $U \subseteq M$ is **open** in M if for each $p \in U$, some r-neighborhood of p is contained inside U. In other words, there exists $r > 0$ such that $M_r(p) \subseteq U$.

Abuse of Notation 2.6.4. Note that a set being open is defined *relative to* the parent space M. However, if M is understood we can abbreviate "open in M" to just "open".

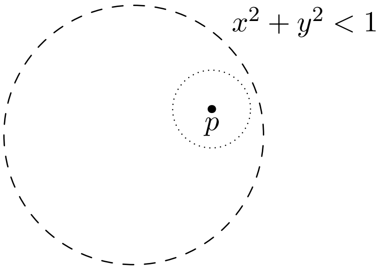


Figure 2.1: The set of points $x^2 + y^2 < 1$ in $\mathbb{R}^2$ is open in $\mathbb{R}^2$.

Example 2.6.5 (Examples of open sets)

(a) Any r-neighborhood of a point is open.

(b) Open intervals of $\mathbb{R}$ are open in $\mathbb{R}$, hence the name! This is the prototypical example to keep in mind.

(c) The open unit ball B^n is open in $\mathbb{R}^n$ for the same reason.

(d) In particular, the open interval $(0,1)$ is open in $\mathbb{R}$. However, if we embed it in $\mathbb{R}^2$, it is no longer open!

(e) The empty set $\varnothing$ and the whole set of points M are open in M.

Example 2.6.6 (Non-examples of open sets)

(a) The closed interval $[0,1]$ is not open in $\mathbb{R}$. There is no ε-neighborhood of the point 0 which is contained in $[0,1]$.

(b) The unit circle S^1 is not open in $\mathbb{R}^2$.

Question 2.6.7. What are the open sets of the discrete space?

Here are two quite important properties of open sets.

Proposition 2.6.8 (Intersections and unions of open sets)

(a) The intersection of finitely many open sets is open.

(b) The union of open sets is open, even if there are infinitely many.

Question 2.6.9. Convince yourself this is true.

Exercise 2.6.10. Exhibit an infinite collection of open sets in $\mathbb{R}$ whose intersection is the set $\{0\}$. This implies that infinite intersections of open sets are not necessarily open.

The whole upshot of this is:

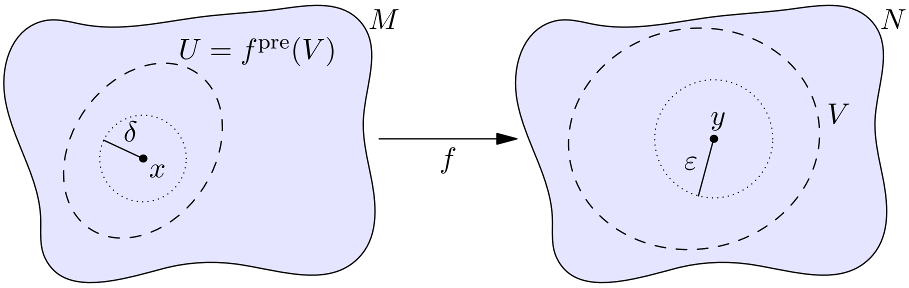
Theorem 2.6.11 (Open set condition)

A function $f\colon M \to N$ of metric spaces is continuous if and only if the pre-image of every open set in N is open in M.

Proof. I'll just do one direction...

Exercise 2.6.12. Show that δ-ε continuity follows from the open set condition.

Now assume f is continuous. First, suppose V is an open subset of the metric space N; let $U = f^{\text{pre}}(V)$. Pick $x \in U$, so $y = f(x) \in V$; we want an open neighborhood of x inside U.



As V is open, there is some small ε-neighborhood around y which is contained inside V. By continuity of f, we can find a δ such that the δ-neighborhood of x gets mapped by f into the ε-neighborhood in N, which in particular lives inside V. Thus the δ-neighborhood lives in U, as desired. □

§2.7 Closed sets

Prototypical example for this section: The closed unit disk $x^2 + y^2 \le r^2$ in $\mathbb{R}^2$.

It would be criminal for me to talk about open sets without talking about closed sets. The name "closed" comes from the definition in a metric space.

Definition 2.7.1. Let M be a metric space. A subset $S \subseteq M$ is **closed** in M if the following property holds: let $x_1, x_2, \dots$ be a sequence of points in S and suppose that x_n converges to x in M. Then $x \in S$ as well.

Abuse of Notation 2.7.2. Same caveat: we abbreviate "closed in M" to just "closed" if the parent space M is understood.

Here's another way to phrase it. The **limit points** of a subset $S \subseteq M$ are defined by

$$\lim S := \{p \in M : \exists (x_n) \in S \text{ such that } x_n \to p\}.$$

Thus S is closed if and only if $S = \lim S$.

> **Exercise 2.7.3.** Prove that $\lim S$ is closed even if S isn't closed. (Draw a picture.)

For this reason, $\lim S$ is also called the **closure** of S in M, and denoted $\overline{S}$. It is simply the smallest closed set which contains S.

Example 2.7.4 (Examples of closed sets)

(a) The empty set $\varnothing$ is closed in M for vacuous reasons: there are no sequences of points with elements in $\varnothing$.

(b) The entire space M is closed in M for tautological reasons. (Verify this!)

(c) The closed interval $[0,1]$ in $\mathbb{R}$ is closed in $\mathbb{R}$, hence the name. Like with open sets, this is the prototypical example of a closed set to keep in mind!

(d) In fact, the closed interval $[0,1]$ is even closed in $\mathbb{R}^2$.

Example 2.7.5 (Non-examples of closed sets)

Let $S = (0,1)$ denote the open interval. Then S is not closed in $\mathbb{R}$ because the sequence of points

$$\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots$$

converges to $0 \in \mathbb{R}$, but $0 \notin (0,1)$.

I should now warn you about a confusing part of this terminology. Firstly, **"most" sets are neither open nor closed**.

Example 2.7.6 (A set neither open nor closed)

The half-open interval $[0,1)$ is neither open nor closed in $\mathbb{R}$.

Secondly, it's **also possible for a set to be both open and closed**; this will be discussed in Chapter 7.

The reason for the opposing terms is the following theorem:

Theorem 2.7.7 (Closed sets are complements of open sets)
Let M be a metric space, and $S \subseteq M$ any subset. Then the following are equivalent:

- The set S is closed in M.
- The complement $M \setminus S$ is open in M.

> **Exercise 2.7.8** (Great)**.** Prove this theorem! You'll want to draw a picture to make it clear what's happening: for example, you might take $M = \mathbb{R}^2$ and S to be the closed unit disk.

§2.8 A few harder problems to think about

Problem 2A. Let $M = (M, d)$ be a metric space. Show that

$$d \colon M \times M \to \mathbb{R}$$

is itself a continuous function (where $M \times M$ is equipped with the product metric).

Problem 2B. Are $\mathbb{Q}$ and $\mathbb{N}$ homeomorphic subspaces of $\mathbb{R}$?

Problem 2C (Continuity of arithmetic continued)**.** Show that subtraction is a continuous map $- \colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, and division is a continuous map $\div \colon \mathbb{R} \times \mathbb{R}_{>0} \to \mathbb{R}$.

Problem 2D. Exhibit a function $f \colon \mathbb{R} \to \mathbb{R}$ such that f is continuous at $x \in \mathbb{R}$ if and only if $x = 0$.

🌶️🌶️ **Problem 2E.** Prove that a function $f \colon \mathbb{R} \to \mathbb{R}$ which is strictly increasing must be continuous at some point.

Problem 2F. Someone on the Internet posted the question "is $1/x$ a continuous function?", leading to great controversy on Twitter. How should you respond?