II

Basic Abstract Algebra

Part II: Contents

3 Homomorphisms and quotient groups

§3.1 Generators and group presentations

Prototypical example for this section: $D_{2n} = \langle r, s \mid r^n = s^2 = 1 \rangle$

Let G be a group. Recall that for some element $x \in G$, we could consider the subgroup

$$\left\{ \dots, x^{-2}, x^{-1}, 1, x, x^2, \dots \right\}$$

of G. Here's a more pictorial version of what we did: **put x in a box, seal it tightly, and shake vigorously**. Using just the element x, we get a pretty explosion that produces the subgroup above.

What happens if we put two elements x, y in the box? Among the elements that get produced are things like

$$xyxyx, \quad x^2y^9x^{-5}y^3, \quad y^{-2015}, \quad \dots$$

Essentially, I can create any finite product of x, y, x^{-1}, y^{-1}. This leads us to define:

Definition 3.1.1. Let S be a subset of G. The subgroup **generated** by S, denoted $\langle S \rangle$, is the set of elements which can be written as a finite product of elements in S (and their inverses). If $\langle S \rangle = G$ then we say S is a set of **generators** for G, as the elements of S together create all of G.

Exercise 3.1.2. Why is the condition "and their inverses" not necessary if G is a finite group? (As usual, assume Lagrange's theorem.)

Example 3.1.3 ($\mathbb{Z}$ is the infinite cyclic group)

Consider 1 as an element of $\mathbb{Z} = (\mathbb{Z}, +)$. We see $\langle 1 \rangle = \mathbb{Z}$, meaning $\{1\}$ generates $\mathbb{Z}$. It's important that -1, the inverse of 1 is also allowed: we need it to write all integers as the sum of 1 and -1.

This gives us an idea for a way to try and express groups compactly. Why not just write down a list of generators for the groups? For example, we could write

$$\mathbb{Z} \cong \langle a \rangle$$

meaning that $\mathbb{Z}$ is just the group generated by one element.

There's one issue: the generators usually satisfy certain properties. For example, consider $\mathbb{Z}/100\mathbb{Z}$. It's also generated by a single element x, but this x has the additional property that $x^{100} = 1$. This motivates us to write

$$\mathbb{Z}/100\mathbb{Z} = \left\langle x \mid x^{100} = 1 \right\rangle.$$

I'm sure you can see where this is going. All we have to do is specify a set of generators and **relations** between the generators, and say that two elements are equal if and only if you can get from one to the other using relations. Such an expression is appropriately called a **group presentation**.

Example 3.1.4 (Dihedral group)
The dihedral group of order $2n$ has a presentation

$$D_{2n} = \left\langle r, s \mid r^n = s^2 = 1, rs = sr^{-1} \right\rangle.$$

Thus each element of D_{2n} can be written uniquely in the form r^α or sr^α, where $\alpha = 0, 1, \dots, n-1$.

Example 3.1.5 (Klein four group)
The **Klein four group**, isomorphic to $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, is given by the presentation

$$\left\langle a, b \mid a^2 = b^2 = 1, ab = ba \right\rangle.$$

Example 3.1.6 (Free group)
The **free group on n elements** is the group whose presentation has n generators and no relations at all. It is denoted F_n, so

$$F_n = \langle x_1, x_2, \dots, x_n \rangle.$$

In other words, $F_2 = \langle a, b \rangle$ is the set of strings formed by appending finitely many copies of a, b, a^{-1}, b^{-1} together.

Question 3.1.7. Notice that $F_1 \cong \mathbb{Z}$.

Abuse of Notation 3.1.8. One might unfortunately notice that "subgroup generated by a and b" has exactly the same notation as the free group $\langle a, b \rangle$. We'll try to be clear based on context which one we mean.

Presentations are nice because they provide a compact way to write down groups. They do have some shortcomings, though.[1]

Example 3.1.9 (Presentations can look very different)
The same group can have very different presentations. For instance consider

$$D_{2n} = \left\langle x, y \mid x^2 = y^2 = 1, (xy)^n = 1 \right\rangle.$$

(To see why this is equivalent, set $x = s$, $y = rs$.)

§3.2 Homomorphisms

Prototypical example for this section: The "mod out by 100" map, $\mathbb{Z} \to \mathbb{Z}/100\mathbb{Z}$.

How can groups talk to each other?

[1] Actually, determining whether two elements of a presentation are equal is undecidable. In fact, it is undecidable to even determine if a group is finite from its presentation.

Two groups are "the same" if we can write an isomorphism between them. And as we saw, two metric spaces are "the same" if we can write a homeomorphism between them. But what's the group analogy of a continuous map? We simply drop the "bijection" condition.

Definition 3.2.1. Let $G = (G, \star)$ and $H = (H, *)$ be groups. A **group homomorphism** is a map $\phi \colon G \to H$ such that for any $g_1, g_2 \in G$ we have

$$\phi(g_1 \star g_2) = \phi(g_1) * \phi(g_2).$$

(Not to be confused with "homeomorphism" from last chapter: note the spelling.)

Example 3.2.2 (Examples of homomorphisms)
Let G and H be groups.

(a) Any isomorphism $G \to H$ is a homomorphism. In particular, the identity map $G \to G$ is a homomorphism.

(b) The **trivial homomorphism** $G \to H$ sends everything to 1_H.

(c) There is a homomorphism from $\mathbb{Z}$ to $\mathbb{Z}/100\mathbb{Z}$ by sending each integer to its residue modulo 100.

(d) There is a homomorphism from $\mathbb{Z}$ to itself by $x \mapsto 10x$ which is injective but not surjective.

(e) There is a homomorphism from S_n to S_{n+1} by "embedding": every permutation on $\{1, \dots, n\}$ can be thought of as a permutation on $\{1, \dots, n+1\}$ if we simply let $n+1$ be a fixed point.

(f) A homomorphism $\phi : D_{12} \to D_6$ is given by $s_{12} \mapsto s_6$ and $r_{12} \mapsto r_6$.

(g) Specifying a homomorphism $\mathbb{Z} \to G$ is the same as specifying just the image of the element $1 \in \mathbb{Z}$. Why?

The last two examples illustrate something: suppose we have a presentation of G. To specify a homomorphism $G \to H$, we only have to specify where each generator of G goes, in such a way that the relations are all satisfied.

Important remark: the right way to think about an isomorphism is as a "bijective homomorphism". To be explicit,

Exercise 3.2.3. Show that $G \cong H$ if and only if there exist homomorphisms $\phi \colon G \to H$ and $\psi \colon H \to G$ such that $\phi \circ \psi = \mathrm{id}_H$ and $\psi \circ \phi = \mathrm{id}_G$.

So the definitions of homeomorphism of metric spaces and isomorphism of groups are not too different.

Some obvious properties of homomorphisms follow.

Fact 3.2.4. Let $\phi \colon G \to H$ be a homomorphism. Then $\phi(1_G) = 1_H$ and $\phi(g^{-1}) = \phi(g)^{-1}$.

Proof. Boring, and I'm sure you could do it yourself if you wanted to. □

Now let me define a very important property of a homomorphism.

Definition 3.2.5. The **kernel** of a homomorphism $\phi\colon G \to H$ is defined by

$$\ker \phi \coloneqq \{g \in G : \phi(g) = 1_H\}.$$

It is a *subgroup* of G (in particular, $1_G \in \ker\phi$ for obvious reasons).

> **Question 3.2.6.** Verify that $\ker\phi$ is in fact a subgroup of G.

We also have the following important fact, which we also encourage the reader to verify.

Proposition 3.2.7 (Kernel determines injectivity)
The map ϕ is injective if and only if $\ker\phi = \{1_G\}$.

To make this concrete, let's compute the kernel of each of our examples.

Example 3.2.8 (Examples of kernels)

(a) The kernel of any isomorphism $G \to H$ is trivial, since an isomorphism is injective. In particular, the kernel of the identity map $G \to G$ is $\{1_G\}$.

(b) The kernel of the trivial homomorphism $G \to H$ (by $g \mapsto 1_H$) is all of G.

(c) The kernel of the homomorphism $\mathbb{Z} \to \mathbb{Z}/100\mathbb{Z}$ by $n \mapsto \overline{n}$ is precisely

$$100\mathbb{Z} = \{\dots, -200, -100, 0, 100, 200, \dots\}.$$

(d) The kernel of the map $\mathbb{Z} \to \mathbb{Z}$ by $x \mapsto 10x$ is trivial: $\{0\}$.

(e) There is a homomorphism from S_n to S_{n+1} by "embedding", but it also has trivial kernel because it is injective.

(f) A homomorphism $\phi\colon D_{12} \to D_6$ is given by $s_{12} \mapsto s_6$ and $r_{12} \mapsto r_6$. You can check that

$$\ker\phi = \left\{1, r_{12}^3\right\} \cong \mathbb{Z}/2\mathbb{Z}.$$

(g) Exercise below.

> **Exercise 3.2.9.** Fix any $g \in G$. Suppose we have a homomorphism $\mathbb{Z} \to G$ by $n \mapsto g^n$. What is the kernel?

> **Question 3.2.10.** Show that for any homomorphism $\phi : G \to H$, the image $\phi^{\text{img}}(G)$ is a subgroup of H. Hence, we'll be especially interested in the case where ϕ is surjective.

§3.3 Cosets and modding out

Prototypical example for this section: Modding out by n: $\mathbb{Z}/(n\cdot\mathbb{Z}) \cong \mathbb{Z}/n\mathbb{Z}$.

The next few sections are a bit dense. If this exposition doesn't work for you, try [Go11].

Let G and Q be groups, and suppose there exists a *surjective* homomorphism

$$\phi\colon G \twoheadrightarrow Q.$$

In other words, if ϕ is injective then $\phi\colon G \to Q$ is a bijection, and hence an isomorphism. But suppose we're not so lucky and $\ker\phi$ is bigger than just $\{1_G\}$. What is the correct interpretation of a more general homomorphism?

Let's look at the special case where $\phi\colon \mathbb{Z} \to \mathbb{Z}/100\mathbb{Z}$ is "modding out by 100". We already saw that the kernel of this map is

$$\ker\phi = 100\mathbb{Z} = \{\dots, -200, -100, 0, 100, 200, \dots\}.$$

Recall now that $\ker\phi$ is a subgroup of G. What this means is that ϕ **is indifferent to the subgroup** $100\mathbb{Z}$ **of** $\mathbb{Z}$:

$$\phi(15) = \phi(2000 + 15) = \phi(-300 + 15) = \phi(700 + 15) = \dots.$$

So $\mathbb{Z}/100\mathbb{Z}$ is what we get when we "mod out by 100". Cool.

In other words, let G be a group and $\phi\colon G \twoheadrightarrow Q$ be a surjective homomorphism with kernel $N \subseteq G$.

We claim that Q should be thought of as the quotient of G by N.

To formalize this, we will define a so-called **quotient group** G/N in terms of G and N only (without referencing Q) which will be naturally isomorphic to Q.

For motivation, let's give a concrete description of Q using just ϕ and G. Continuing our previous example, let $N = 100\mathbb{Z}$ be our subgroup of G. Consider the sets

$$\begin{aligned} N &= \{\dots, -200, -100, 0, 100, 200, \dots\} \\ 1 + N &= \{\dots, -199, -99, 1, 101, 201, \dots\} \\ 2 + N &= \{\dots, -198, -98, 2, 102, 202, \dots\} \\ &\vdots \\ 99 + N &= \{\dots, -101, -1, 99, 199, 299, \dots\}. \end{aligned}$$

The elements of each set all have the same image when we apply ϕ, and moreover any two elements in different sets have different images. Then the main idea is to notice that

We can think of Q as the group whose *elements* are the *sets* above.

Thus, given ϕ we define an equivalence relation $\sim_N$ on G by saying $x \sim_N y$ for $\phi(x) = \phi(y)$. This $\sim_N$ divides G into several equivalence classes in G which are in obvious bijection with Q, as above. Now we claim that we can write these equivalence classes very explicitly.

Exercise 3.3.1. Show that $x \sim_N y$ if and only if $x = yn$ for some $n \in N$ (in the mod 100 example, this means they "differ by some multiple of 100"). Thus for any $g \in G$, the equivalence class of $\sim_N$ which contains g is given explicitly by

$$gN \coloneqq \{gn \mid n \in N\}.$$

Here's the word that describes the types of sets we're running into now.

Definition 3.3.2. Let H be any subgroup of G (not necessarily the kernel of some homomorphism). A set of the form gH is called a **left coset** of H.

Remark 3.3.3 — Although the notation might not suggest it, keep in mind that g_1N is often equal to g_2N even if $g_1 \neq g_2$. In the "mod 100" example, $3 + N = 103 + N$. In other words, these cosets are *sets*.

This means that if I write "let gH be a coset" without telling you what g is, you can't figure out which g I chose from just the coset itself. If you don't believe me, here's an example of what I mean:

$$x + 100\mathbb{Z} = \{\dots, -97, 3, 103, 203, \dots\} \implies x = ?.$$

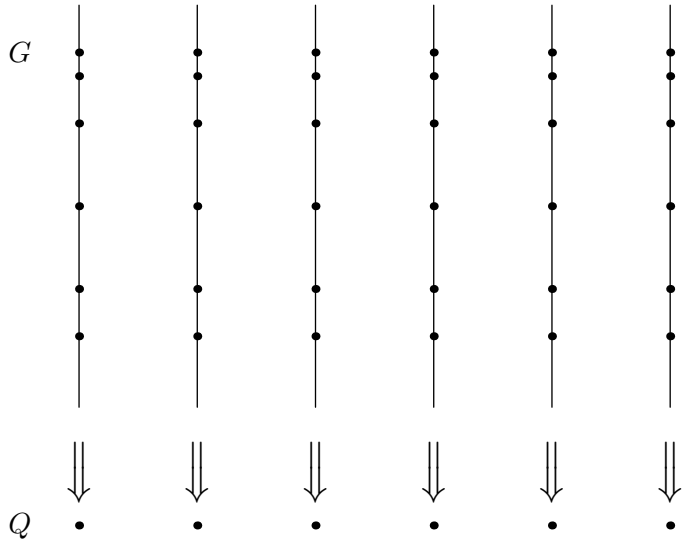
There's no reason to think I picked $x = 3$. (I actually picked $x = -13597$.)

Remark 3.3.4 — Given cosets g_1H and g_2H, you can check that the map $x \mapsto g_2g_1^{-1}x$ is a bijection between them. So actually, all cosets have the same cardinality.

So, long story short,

Elements of the group Q are naturally identified with left cosets of N.

In practice, people often still prefer to picture elements of Q as single points (for example it's easier to think of $\mathbb{Z}/2\mathbb{Z}$ as $\{0, 1\}$ rather than $\{\{\dots, -2, 0, 2, \dots\}, \{\dots, -1, 1, 3, \dots\}\}$). If you like this picture, then you might then draw G as a bunch of equally tall fibers (the cosets), which are then "collapsed" onto Q.



Now that we've done this, we can give an *intrinsic* definition for the quotient group we alluded to earlier.

Definition 3.3.5. A subgroup N of G is called **normal** if it is the kernel of some homomorphism. We write this as $N \trianglelefteq G$.

Definition 3.3.6. Let $N \trianglelefteq G$. Then the **quotient group**, denoted G/N (and read "G mod N"), is the group defined as follows.

- The elements of G/N will be the left cosets of N.

- We want to define the product of two cosets C_1 and C_2 in G/N. Recall that the cosets are in bijection with elements of Q. So let q_1 be the value associated to the coset C_1, and q_2 the one for C_2. Then we can take the product to be the coset corresponding to q_1q_2.

  Quite importantly, **we can also do this in terms of representatives of the cosets**. Let $g_1 \in C_1$ and $g_2 \in C_2$, so $C_1 = g_1N$ and $C_2 = g_2N$. Then $C_1 \cdot C_2$ should be the coset which contains g_1g_2. This is the same as the above definition since $\phi(g_1g_2) = \phi(g_1)\phi(g_2) = q_1q_2$; all we've done is define the product in terms of elements of G, rather than values in H.

  Using the gN notation, and with Remark 3.3.3 in mind, we can write this even more succinctly:
  $$(g_1N) \cdot (g_2N) \coloneqq (g_1g_2)N.$$

And now you know why the integers modulo n are often written $\mathbb{Z}/n\mathbb{Z}$!

Question 3.3.7. Take a moment to digest the above definition.

By the way we've built it, the resulting group G/N is isomorphic to Q. In a sense we think of G/N as "G modulo the condition that $n = 1$ for all $n \in N$".

§3.4 (Optional) Proof of Lagrange's theorem

As an aside, with the language of cosets we can now show Lagrange's theorem in the general case.

Theorem 3.4.1 (Lagrange's theorem)
Let G be a finite group, and let H be any subgroup. Then $|H|$ divides $|G|$.

The proof is very simple: note that the cosets of H all have the same size and form a partition of G (even when H is not necessarily normal). Hence if n is the number of cosets, then $n \cdot |H| = |G|$.

Question 3.4.2. Conclude that $x^{|G|} = 1$ by taking $H = \langle x \rangle \subseteq G$.

Remark 3.4.3 — It should be mentioned at this point that in general, if G is a finite group and N is normal, then $|G/N| = |G|/|N|$.

§3.5 Eliminating the homomorphism

Prototypical example for this section: Again $\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}/n\mathbb{Z}$.

Let's look at the last definition of G/N we provided. The short version is:

- The elements of G/N are cosets gN, which you can think of as equivalence classes of a relation $\sim_N$ (where $g_1 \sim_N g_2$ if $g_1 = g_2n$ for some $n \in N$).
- Given cosets g_1N and g_2N the group operation is
  $$g_1N \cdot g_2N \coloneqq (g_1g_2)N.$$

Question: where do we actually use the fact that N is normal? We don't talk about ϕ or Q anywhere in this definition.

The answer is in Remark 3.3.3. The group operation takes in two cosets, so it doesn't know what g_1 and g_2 are. But behind the scenes, **the normal condition guarantees that the group operation can pick any g_1 and g_2 it wants and still end up with the same coset.** If we didn't have this property, then it would be hard to define the product of two cosets C_1 and C_2 because it might make a difference which $g_1 \in C_1$ and $g_2 \in C_2$ we picked. The fact that N came from a homomorphism meant we could pick any representatives g_1 and g_2 of the cosets we wanted, because they all had the same ϕ-value.

We want some conditions which force this to be true without referencing ϕ at all. Suppose $\phi\colon G \to K$ is a homomorphism of groups with $H = \ker \phi$. Aside from the fact H is a group, we can get an "obvious" property:

Question 3.5.1. Show that if $h \in H$, $g \in G$, then $ghg^{-1} \in H$. (Check $\phi(ghg^{-1}) = 1_K$.)

Example 3.5.2 (Example of a non-normal subgroup)
Let $D_{12} = \langle r, s \mid r^6 = s^2 = 1, rs = sr^{-1} \rangle$. Consider the subgroup of order two $H = \{1, s\}$ and notice that

$$rsr^{-1} = r(sr^{-1}) = r(rs) = r^2 s \notin H.$$

Hence H is not normal, and cannot be the kernel of any homomorphism.

Well, duh – so what? Amazingly it turns out that this is the *sufficient* condition we want. Specifically, it makes the nice "coset multiplication" we wanted work out.

Remark 3.5.3 (For math contest enthusiasts) — This coincidence is really a lot like functional equations at the IMO. We all know that normal subgroups H satisfy $ghg^{-1} \in H$; the surprise is that from the latter seemingly weaker condition, we can deduce H is normal.

Thus we have a new criterion for "normal" subgroups which does not make any external references to ϕ.

Theorem 3.5.4 (Algebraic condition for normal subgroups)
Let H be a subgroup of G. Then the following are equivalent:

- $H \trianglelefteq G$.
- For every $g \in G$ and $h \in H$, $ghg^{-1} \in H$.

Proof. We already showed one direction.

For the other direction, we need to build a homomorphism with kernel H. So we simply *define* the group G/H as the cosets. To put a group operation, we need to verify:

Claim 3.5.5. If $g_1' \sim_H g_1$ and $g_2' \sim_H g_2$ then $g_1'g_2' \sim_H g_1g_2$.

Proof. Boring algebraic manipulation (again functional equation style). Let $g_1' = g_1h_1$ and $g_2' = g_2h_2$, so we want to show that $g_1h_1g_2h_2 \sim_H g_1g_2$. Since H has the property,

$g_2^{-1}h_1g_2$ is some element of H, say h_3. Thus $h_1g_2 = g_2h_3$, and the left-hand side becomes $g_1g_2(h_3h_2)$, which is fine since $h_3h_2 \in H$. ■

With that settled we can just *define* the product of two cosets (of normal subgroups) by

$$(g_1H) \cdot (g_2H) = (g_1g_2)H.$$

Thus the claim above shows that this multiplication is well-defined (this verification is the "content" of the theorem). So G/H is indeed a group! Moreover there is an obvious "projection" homomorphism $G \to G/H$ (with kernel H), by $g \mapsto gH$. □

Example 3.5.6 (Modding out in the product group)

Consider again the product group $G \times H$. Earlier we identified a subgroup

$$G' = \{(g, 1_H) \mid g \in G\} \cong G.$$

You can easily see that $G' \trianglelefteq G \times H$. (Easy calculation.)

Moreover, you can check that

$$(G \times H)/(G') \cong H.$$

Indeed, we have $(g, h) \sim_{G'} (1_G, h)$ for all $g \in G$ and $h \in H$.

Example 3.5.7 (Quotients and products don't necessarily cancel)

It is not necessarily true that $(G/H) \times H \cong G$. For example, consider $G = \mathbb{Z}/4\mathbb{Z}$ and the normal subgroup $H = \{0, 2\} \cong \mathbb{Z}/2\mathbb{Z}$. Then $G/H \cong \mathbb{Z}/2\mathbb{Z}$, but $\mathbb{Z}/4\mathbb{Z} \ncong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. (Footnote: the precise condition for this kind of "canceling" is called the Schur-Zassenhaus lemma.)

Example 3.5.8 (Another explicit computation)

Let $\phi : D_8 \to \mathbb{Z}/4\mathbb{Z}$ be defined by

$$r \mapsto \overline{2}, \quad s \mapsto \overline{2}.$$

The kernel of this map is $N = \{1, r^2, sr, sr^3\}$.

We can do a quick computation of all the elements of D_8 to get

$$\phi(1) = \phi(r^2) = \phi(sr) = \phi(sr^3) = \overline{0} \text{ and } \phi(r) = \phi(r^3) = \phi(s) = \phi(sr^2) = \overline{2}.$$

The two relevant fibers are

$$\phi^{\text{pre}}(\overline{0}) = 1N = r^2N = srN = sr^3N = \{1, r^2, sr, sr^3\}$$

and

$$\phi^{\text{pre}}(\overline{2}) = rN = r^3N = sN = sr^2N = \{r, r^3, s, sr^2\}.$$

So we see that $|D_8/N| = 2$ is a group of order two, or $\mathbb{Z}/2\mathbb{Z}$. Indeed, the image of ϕ is

$$\{\overline{0}, \overline{2}\} \cong \mathbb{Z}/2\mathbb{Z}.$$

Question 3.5.9. Suppose G is abelian. Why does it follow that any subgroup of G is normal?

Finally here's some food for thought: suppose one has a group presentation for a group G that uses n generators. Can you write it as a quotient of the form F_n/N, where N is a normal subgroup of F_n?

§3.6 (Digression) The first isomorphism theorem

One quick word about what other sources usually say.

Most textbooks actually *define* normal using the $ghg^{-1} \in H$ property. Then they define G/H for normal H in the way I did above, using the coset definition

$$(g_1H) \cdot (g_2H) = g_1g_2H.$$

Using purely algebraic manipulations (like I did) this is well-defined, and so now you have this group G/H or something. The underlying homomorphism isn't mentioned at all, or is just mentioned in passing.

I think this is incredibly dumb. The normal condition looks like it gets pulled out of thin air and no one has any clue what's going on, because no one has any clue what a normal subgroup actually should look like.

Other sources like to also write the so-called first isomorphism theorem.[2] It goes like this.

Theorem 3.6.1 (First isomorphism theorem)
Let $\phi\colon G \to H$ be a homomorphism. Then $G/\ker\phi$ is isomorphic to $\phi^{\text{img}}(G)$.

To me, this is just a clumsier way of stating the same idea.

About the only merit this claim has is that if ϕ is injective, then the image $\phi^{\text{img}}(G)$ is an *isomorphic copy* of G inside the group H. (Try to see this directly!) This is a pattern we'll often see in other branches of mathematics: whenever we have an *injective structure-preserving map*, often the image of this map will be some "copy" of G. (Here "structure" refers to the group multiplication, but we'll see some more other examples of "types of objects" later!)

In that sense an injective homomorphism $\phi\colon G \hookrightarrow H$ is an *embedding* of G into H.

§3.7 A few harder problems to think about

Problem 3A (18.701 at MIT)**.** Determine all groups G for which the map $\phi\colon G \to G$ defined by

$$\phi(g) = g^2$$

is a homomorphism.

[2]There is a second and third isomorphism theorem. But four years after learning about them, I *still* don't remember what they are. So I'm guessing they weren't very important.

Problem 3B. Consider the dihedral group $G = D_{10}$.

(a) Is $H = \langle r \rangle$ a normal subgroup of G? If so, compute G/H up to isomorphism.

(b) Is $H = \langle s \rangle$ a normal subgroup of G? If so, compute G/H up to isomorphism.

Problem 3C. Does S_4 have a normal subgroup of order 3?

Problem 3D. Let G and H be finite groups, where $|G| = 1000$ and $|H| = 999$. Show that a homomorphism $G \to H$ must be trivial.

Problem 3E. Let $\mathbb{C}^\times$ denote the nonzero complex numbers under multiplication. Show that there are five homomorphisms $\mathbb{Z}/5\mathbb{Z} \to \mathbb{C}^\times$ but only two homomorphisms $D_{10} \to \mathbb{C}^\times$, even though $\mathbb{Z}/5\mathbb{Z}$ is a subgroup of D_{10}.

Problem 3F. Find a non-abelian group G such that every subgroup of G is normal. (These groups are called **Hamiltonian**.)

Problem 3G (PRIMES entrance exam, 2018)**.** Let G be a group with presentation given by

$$G = \left\langle a, b, c \mid ab = c^2a^4,\ bc = ca^6,\ ac = ca^8,\ c^{2018} = b^{2019} \right\rangle.$$

Determine the order of G.

Problem 3H (Homophony group)**.** The homophony group (of English) is the group with 26 generators a, b, $\ldots$, z and one relation for every pair of English words which sound the same. For example $knight = night$ (and hence $k = 1$). Prove that the group is trivial.

4 Rings and ideals

§4.1 Some motivational metaphors about rings vs groups

In this chapter we'll introduce the notion of a **commutative ring** R. It is a larger structure than a group: it will have two operations addition and multiplication, rather than just one. We will then immediately define a **ring homomorphism** $R \to S$ between pairs of rings.

This time, instead of having normal subgroups $H \trianglelefteq G$, rings will instead have subsets $I \subseteq R$ called **ideals**, which are not themselves rings but satisfy some niceness conditions. We will then show you how to define R/I, in analogy to G/H as before. Finally, like with groups, we will talk a bit about how to generate ideals.

Here is a possibly helpful table of analogies to help you keep track:

	Group	Ring
Notation	G	R
Operations	$\cdot$	$+$, $\times$
Commutativity	only if abelian	for us, always
Sub-structure	subgroup	(not discussed)
Homomorphism	grp hom. $G \to H$	ring hom. $R \to S$
Kernel	normal subgroup	ideal
Quotient	G/H	R/I

§4.2 (Optional) Pedagogical notes on motivation

I wrote most of these examples with a number theoretic eye in mind; thus if you liked elementary number theory, a lot of your intuition will carry over. Basically, we'll try to generalize properties of the ring $\mathbb{Z}$ to any abelian structure in which we can also multiply. That's why, for example, you can talk about "irreducible polynomials in $\mathbb{Q}[x]$" in the same way you can talk about "primes in $\mathbb{Z}$", or about "factoring polynomials modulo p" in the same way we can talk "unique factorization in $\mathbb{Z}$". Even if you only care about $\mathbb{Z}$ (say, you're a number theorist), this has a lot of value: I assure you that trying to solve $x^n + y^n = z^n$ (for $n > 2$) requires going into a ring other than $\mathbb{Z}$!

Thus for all the sections that follow, keep $\mathbb{Z}$ in mind as your prototype.

I mention this here because commutative algebra is *also* closely tied to algebraic geometry. Lots of the ideas in commutative algebra have nice "geometric" interpretations that motivate the definitions, and these connections are explored in the corresponding part later. So, I want to admit outright that this is not the only good way (perhaps not even the most natural one) of motivating what is to follow.

§4.3 Definition and examples of rings

Prototypical example for this section: $\mathbb{Z}$ all the way! Also $R[x]$ and various fields (next section).

Well, I guess I'll define a ring[1].

[1]Or, according to some authors, a "ring with identity"; some authors don't require rings to have multiplicative identity. For us, "ring" always means "ring with 1".

Definition 4.3.1. A **ring** is a triple $(R, +, \times)$, the two operations usually called addition and multiplication, such that

(i) $(R, +)$ is an abelian group, with identity 0_R, or just 0.

(ii) $\times$ is an associative, binary operation on R with some identity, written 1_R or just 1.

(iii) Multiplication distributes over addition.

The ring R is **commutative** if $\times$ is commutative.

Abuse of Notation 4.3.2. As usual, we will abbreviate $(R, +, \times)$ to R.

Abuse of Notation 4.3.3. For simplicity, assume all rings are commutative for the rest of this chapter. We'll run into some noncommutative rings eventually, but for such rings we won't need the full theory of this chapter anyways.

These definitions are just here for completeness. The examples are much more important.

Example 4.3.4 (Typical rings)

(a) The sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ are all rings with the usual addition and multiplication.

(b) The integers modulo n are also a ring with the usual addition and multiplication. We also denote it by $\mathbb{Z}/n\mathbb{Z}$.

Here is also a trivial example.

Definition 4.3.5. The **zero ring** is the ring R with a single element. We denote the zero ring by 0. A ring is **nontrivial** if it is not the zero ring.

Exercise 4.3.6 (Comedic)**.** Show that a ring is nontrivial if and only if $0_R \neq 1_R$.

Since I've defined this structure, I may as well state the obligatory facts about it.

Fact 4.3.7. For any ring R and $r \in R$, $r \cdot 0_R = 0_R$. Moreover, $r \cdot (-1_R) = -r$.

Here are some more examples of rings.

Example 4.3.8 (Product ring)

Given two rings R and S the **product ring**, denoted $R \times S$, is defined as ordered pairs (r, s) with both operations done component-wise. For example, the Chinese remainder theorem says that

$$\mathbb{Z}/15\mathbb{Z} \cong \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}$$

with the isomorphism $n \bmod 15 \mapsto (n \bmod 3, n \bmod 5)$.

Remark 4.3.9 — Equivalently, we can define $R \times S$ as the abelian group $R \oplus S$, and endow it with the multiplication where $r \cdot s = 0$ for $r \in R$, $s \in S$.

Question 4.3.10. Which (r, s) is the identity element of the product ring $R \times S$?

Example 4.3.11 (Polynomial ring)

Given any ring R, the **polynomial ring** $R[x]$ is defined as the set of polynomials with coefficients in R:

$$R[x] = \left\{ a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0 \mid a_0, \ldots, a_n \in R \right\}.$$

This is pronounced "R adjoin x". Addition and multiplication are done exactly in the way you would expect.

Remark 4.3.12 (Digression on division) — Happily, polynomial division also does what we expect: if $p \in R[x]$ is a polynomial, and $p(a) = 0$, then $(x - a)q(x) = p(x)$ for some polynomial q. Proof: do polynomial long division.

With that, note the caveat that

$$x^2 - 1 \equiv (x - 1)(x + 1) \pmod 8$$

has *four* roots 1, 3, 5, 7 in $\mathbb{Z}/8\mathbb{Z}$.

The problem is that $2 \cdot 4 = 0$ even though 2 and 4 are not zero; we call 2 and 4 *zero divisors* for that reason. In an *integral domain* (a ring without zero divisors), this pathology goes away, and just about everything you know about polynomials carries over. (I'll say this all again next section.)

Example 4.3.13 (Multi-variable polynomial ring)

We can consider polynomials in n variables with coefficients in R, denoted $R[x_1, \ldots, x_n]$. (We can even adjoin infinitely many x's if we like!)

Example 4.3.14 (Gaussian integers are a ring)

The **Gaussian integers** are the set of complex numbers with integer real and imaginary parts, that is

$$\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\}.$$

Abuse of Notation 4.3.15 (Liberal use of adjoinment)**.** Careful readers will detect some abuse in notation here. $\mathbb{Z}[i]$ should officially be "integer-coefficient polynomials in a variable i". However, it is understood from context that $i^2 = -1$; and a polynomial in $i = \sqrt{-1}$ "is" a Gaussian integer.

Example 4.3.16 (Cube root of 2)

As another example (using the same abuse of notation):

$$\mathbb{Z}[\sqrt[3]{2}] = \left\{ a + b\sqrt[3]{2} + c\sqrt[3]{4} \mid a, b, c \in \mathbb{Z} \right\}.$$

§4.4 Fields

Prototypical example for this section: $\mathbb{Q}$ *is a field, but* $\mathbb{Z}$ *is not.*

Although we won't need to know what a field is until next chapter, they're so convenient for examples I will go ahead and introduce them now.

As you might already know, if the multiplication is invertible, then we call the ring a field. To be explicit, let me write the relevant definitions.

Definition 4.4.1. A **unit** of a ring R is an element $u \in R$ which is invertible: for some $x \in R$ we have $ux = 1_R$.

Example 4.4.2 (Examples of units)

(a) The units of $\mathbb{Z}$ are ± 1, because these are the only things which "divide 1" (which is the reason for the name "unit").

(b) On the other hand, in $\mathbb{Q}$ everything is a unit (except 0). For example, $\frac{3}{5}$ is a unit since $\frac{3}{5} \cdot \frac{5}{3} = 1$.

(c) The Gaussian integers $\mathbb{Z}[i]$ have four units: ± 1 and $\pm i$.

Definition 4.4.3. A nontrivial (commutative) ring is a **field** when all its nonzero elements are units.

Colloquially, we say that

A field is a structure where you can add, subtract, multiply, and divide.

Depending on context, they are often denoted either k, K, F.

Example 4.4.4 (First examples of fields)

(a) $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ are fields, since the notion $\frac{1}{c}$ makes sense in them.

(b) If p is a prime, then $\mathbb{Z}/p\mathbb{Z}$ is a field, which we usually denote by $\mathbb{F}_p$.

The trivial ring 0 is *not* considered a field, since we require fields to be nontrivial.

§4.5 Homomorphisms

Prototypical example for this section: $\mathbb{Z} \to \mathbb{Z}/5\mathbb{Z}$ *by modding out by* 5.

This section is going to go briskly – it's the obvious generalization of all the stuff we did with quotient groups.[2]

First, we define a homomorphism and isomorphism.

Definition 4.5.1. Let $R = (R, +_R, \times_R)$ and $S = (S, +_S, \times_S)$ be rings. A **ring homomorphism** is a map $\phi \colon R \to S$ such that

[2] I once found an abstract algebra textbook which teaches rings before groups. At the time I didn't understand why, but now I think I get it – modding out by things in commutative rings is far more natural, and you can start talking about all the various flavors of rings and fields. You also have (in my opinion) more vivid first examples for rings than for groups. I actually sympathize a lot with this approach — maybe I'll convert Napkin to follow it one day.

(i) $\phi(x +_R y) = \phi(x) +_S \phi(y)$ for each $x, y \in R$.

(ii) $\phi(x \times_R y) = \phi(x) \times_S \phi(y)$ for each $x, y \in R$.

(iii) $\phi(1_R) = 1_S$.

If ϕ is a bijection then ϕ is an **isomorphism** and we say that rings R and S are **isomorphic**.

Just what you would expect. The only surprise is that we also demand $\phi(1_R)$ to go to 1_S. This condition is not extraneous: consider the map $\mathbb{Z} \to \mathbb{Z}$ called "multiply by zero".

Example 4.5.2 (Examples of homomorphisms)

(a) The identity map, as always.

(b) The map $\mathbb{Z} \to \mathbb{Z}/5\mathbb{Z}$ modding out by 5.

(c) The map $\mathbb{R}[x] \to \mathbb{R}$ by $p(x) \mapsto p(0)$ by taking the constant term.

(d) For any ring R, there is a trivial ring homomorphism $R \to 0$.

Example 4.5.3 (Non-examples of homomorphisms)

Because we require 1_R to 1_S, some maps that you might have thought were homomorphisms will fail.

(a) The map $\mathbb{Z} \to \mathbb{Z}$ by $x \mapsto 2x$ is not a ring homomorphism. Aside from the fact it sends 1 to 2, it also does not preserve multiplication.

(b) If S is a nontrivial ring, the map $R \to S$ by $x \mapsto 0$ is not a ring homomorphism, even though it preserves multiplication.

(c) There is no ring homomorphism $\mathbb{Z}/2016\mathbb{Z} \to \mathbb{Z}$ at all.

In particular, whereas for groups G and H there was always a trivial group homomorphism sending everything in G to 1_H, this is not the case for rings.

§4.6 Ideals

Prototypical example for this section: The multiples of 5 are an ideal of $\mathbb{Z}$.

Now, just like we were able to mod out by groups, we'd also like to define quotient rings. So once again,

Definition 4.6.1. The **kernel** of a ring homomorphism $\phi\colon R \to S$, denoted $\ker \phi$, is the set of $r \in R$ such that $\phi(r) = 0$.

In group theory, we were able to characterize the "normal" subgroups by a few obviously necessary conditions (namely, $gHg^{-1} = H$). We can do the same thing for rings, and it's in fact easier because our operations are commutative.

First, note two obvious facts:

- If $\phi(x) = \phi(y) = 0$, then $\phi(x + y) = 0$ as well. So $\ker \phi$ should be closed under addition.

- If $\phi(x) = 0$, then for any $r \in R$ we have $\phi(rx) = \phi(r)\phi(x) = 0$ too. So for $x \in \ker \phi$ and *any* $r \in R$, we have $rx \in \ker \phi$.

A (nonempty) subset $I \subseteq R$ is called an ideal if it satisfies these properties. That is,

Definition 4.6.2. A nonempty subset $I \subseteq R$ is an **ideal** if it is closed under addition, and for each $x \in I$, $rx \in I$ for all $r \in R$. It is **proper** if $I \neq R$.

Note that in the second condition, r need not be in I! So this is stronger than merely saying I is closed under multiplication.

Remark 4.6.3 — If R is not commutative, we also need the condition $xr \in I$. That is, the ideal is *two-sided*: it absorbs multiplication from both the left and the right. But since rings in Napkin are commutative we needn't worry with this distinction.

Example 4.6.4 (Prototypical example of an ideal)
Consider the set $I = 5\mathbb{Z} = \{\dots, -10, -5, 0, 5, 10, \dots\}$ as an ideal in $\mathbb{Z}$. We indeed see I is the kernel of the "take mod 5" homomorphism:

$$\mathbb{Z} \twoheadrightarrow \mathbb{Z}/5\mathbb{Z}.$$

It's clearly closed under addition, but it absorbs multiplication from *all* elements of $\mathbb{Z}$: given $15 \in I$, $999 \in \mathbb{Z}$, we get $15 \cdot 999 \in I$.

Exercise 4.6.5 (Mandatory: fields have two ideals)**.** If K is a field, show that K has exactly two ideals. What are they?

Now we claim that these conditions are sufficient. More explicitly,

Theorem 4.6.6 (Ring analog of normal subgroups)
Let R be a ring and $I \subsetneq R$. Then I is the kernel of some homomorphism if and only if it's an ideal.

Proof. It's quite similar to the proof for the normal subgroup thing, and you might try it yourself as an exercise.

Obviously the conditions are necessary. To see they're sufficient, we *define* a ring by "cosets"

$$S = \{r + I \mid r \in R\}.$$

These are the equivalence classes under $r_1 \sim r_2$ if and only if $r_1 - r_2 \in I$ (think of this as taking "mod I"). To see that these form a ring, we have to check that the addition and multiplication we put on them is well-defined. Specifically, we want to check that if $r_1 \sim s_1$ and $r_2 \sim s_2$, then $r_1 + r_2 \sim s_1 + s_2$ and $r_1 r_2 \sim s_1 s_2$. We actually already did the first part – just think of R and S as abelian groups, forgetting for the moment that we can multiply. The multiplication is more interesting.

Exercise 4.6.7 (Recommended)**.** Show that if $r_1 \sim s_1$ and $r_2 \sim s_2$, then $r_1 r_2 \sim s_1 s_2$. You will need to use the fact that I absorbs multiplication from *any* elements of R, not just those in I.

Anyways, since this addition and multiplication is well-defined there is now a surjective homomorphism $R \to S$ with kernel exactly I. □

Definition 4.6.8. Given an ideal I, we define as above the **quotient ring**

$$R/I := \{r + I \mid r \in R\}.$$

It's the ring of these equivalence classes. This ring is pronounced "R mod I".

Example 4.6.9 ($\mathbb{Z}/5\mathbb{Z}$)
The integers modulo 5 formed by "modding out additively by 5" are the $\mathbb{Z}/5\mathbb{Z}$ we have already met.

But here's an important point: just as we don't actually think of $\mathbb{Z}/5\mathbb{Z}$ as consisting of $k + 5\mathbb{Z}$ for $k = 0, \dots, 4$, we also don't really want to think about R/I as elements $r + I$. The better way to think about it is

R/I is the result when we declare that elements of I are all zero; that is, we "mod out by elements of I".

For example, modding out by $5\mathbb{Z}$ means that we consider all elements in $\mathbb{Z}$ divisible by 5 to be zero. This gives you the usual modular arithmetic!

Exercise 4.6.10. Earlier, we wrote $\mathbb{Z}[i]$ for the Gaussian integers, which was a slight abuse of notation. Convince yourself that this ring could instead be written as $\mathbb{Z}[x]/(x^2+1)$, if we wanted to be perfectly formal. (We will stick with $\mathbb{Z}[i]$ though — it's more natural.) Here the shorthand $(x^2+1) := (x^2+1)\mathbb{Z}[x] = \{(x^2+1)f \mid f \in \mathbb{Z}[x]\}$ denotes the ideal of multiples of x^2+1 within $\mathbb{Z}[x]$.
Figure out the analogous formalization of $\mathbb{Z}[\sqrt[3]{2}]$.

§4.7 Generating ideals

Prototypical example for this section: In $\mathbb{Z}$, the ideals are all of the form (n).

Let's give you some practice with ideals.

An important piece of intuition is that once an ideal contains a unit, it contains 1, and thus must contain the entire ring. That's why the notion of "proper ideal" is useful language. To expand on that:

Proposition 4.7.1 (Proper ideal $\iff$ no units)
Let R be a ring and $I \subseteq R$ an ideal. Then I is proper (i.e. $I \neq R$) if and only if it contains no units of R.

Proof. Suppose I contains a unit u, i.e. an element u with an inverse u^{-1}. Then it contains $u \cdot u^{-1} = 1$, and thus $I = R$. Conversely, if I contains no units, it is obviously proper. □

As a consequence, if K is a field, then its only ideals are (0) and K (this was Exercise 4.6.5). So for our practice purposes, we'll be working with rings that aren't fields.

First practice: $\mathbb{Z}$.

Exercise 4.7.2. Show that the only ideals of $\mathbb{Z}$ are precisely those sets of the form $n\mathbb{Z}$, where n is a nonnegative integer.

Thus, while ideals of fields are not terribly interesting, ideals of $\mathbb{Z}$ look eerily like elements of $\mathbb{Z}$. Let's make this more precise.

Definition 4.7.3. Let R be a ring. The **ideal generated** by a set of elements $x_1, \dots, x_n \in R$ is denoted by $I = (x_1, x_2, \dots, x_n)$ and given by

$$I = \{r_1 x_1 + \cdots + r_n x_n \mid r_i \in R\}.$$

One can think of this as "the smallest ideal containing all the x_i".

The analogy of putting the $\{x_i\}$ in a sealed box and shaking vigorously kind of works here too.

Remark 4.7.4 (Linear algebra digression) — If you know linear algebra, you can summarize this as: an ideal is an R-module. The ideal $(x_1, \dots, x_n)$ is the submodule spanned by $x_1, \dots, x_n$.

In particular, if $I = (x)$ then I consists of exactly the "multiples of x", i.e. numbers of the form rx for $r \in R$.

Remark 4.7.5 — We can also apply this definition to infinite generating sets, as long as only finitely many of the r_i are not zero (since infinite sums don't make sense in general).

Example 4.7.6 (Examples of generated ideals)

(a) As $(n) = n\mathbb{Z}$ for all $n \in \mathbb{Z}$, every ideal in $\mathbb{Z}$ is of the form (n).

(b) In $\mathbb{Z}[i]$, we have $(5) = \{5a + 5bi \mid a, b \in \mathbb{Z}\}$.

(c) In $\mathbb{Z}[x]$, the ideal (x) consists of polynomials with zero constant terms.

(d) In $\mathbb{Z}[x, y]$, the ideal (x, y) again consists of polynomials with zero constant terms.

(e) In $\mathbb{Z}[x]$, the ideal $(x, 5)$ consists of polynomials whose constant term is divisible by 5.

Question 4.7.7. Please check that the set $I = \{r_1 x_1 + \cdots + r_n x_n \mid r_i \in R\}$ is indeed always an ideal (closed under addition, and absorbs multiplication).

Now suppose $I = (x_1, \dots, x_n)$. What does R/I look like? According to what I said at the end of the last section, it's what happens when we "mod out" by each of the elements x_i. For example...

Example 4.7.8 (Modding out by generated ideals)

(a) Let $R = \mathbb{Z}$ and $I = (5)$. Then R/I is literally $\mathbb{Z}/5\mathbb{Z}$, or the "integers modulo 5": it is the result of declaring $5 = 0$.

(b) Let $R = \mathbb{Z}[x]$ and $I = (x)$. Then R/I means we send x to zero; hence $R/I \cong \mathbb{Z}$ as given any polynomial $p(x) \in R$, we simply get its constant term.

(c) Let $R = \mathbb{Z}[x]$ again and now let $I = (x-3)$. Then R/I should be thought of as the quotient when $x - 3 \equiv 0$, that is, $x \equiv 3$. So given a polynomial $p(x)$ its image after we mod out should be thought of as $p(3)$. Again $R/I \cong \mathbb{Z}$, but in a different way.

(d) Finally, let $I = (x-3, 5)$. Then R/I not only sends x to three, but also 5 to zero. So given $p \in R$, we get $p(3) \pmod 5$. Then $R/I \cong \mathbb{Z}/5\mathbb{Z}$.

Remark 4.7.9 (Mod notation) — By the way, given an ideal I of a ring R, it's totally legit to write

$$x \equiv y \pmod I$$

to mean that $x - y \in I$. Everything you learned about modular arithmetic carries over.

§4.8 Principal ideal domains

Prototypical example for this section: $\mathbb{Z}$ is a PID, $\mathbb{Z}[x]$ is not. $\mathbb{C}[x]$ is a PID, $\mathbb{C}[x,y]$ is not.

What happens if we put multiple generators in an ideal, like $(10, 15) \subseteq \mathbb{Z}$? Well, we have by definition that $(10, 15)$ is given as a set by

$$(10, 15) \coloneqq \{10x + 15y \mid x, y \in \mathbb{Z}\}.$$

If you're good at number theory you'll instantly recognize this as $5\mathbb{Z} = (5)$. Surprise! In $\mathbb{Z}$, the ideal (a, b) is exactly $\gcd(a, b)\mathbb{Z}$. And that's exactly the reason you often see the GCD of two numbers denoted (a, b).

We call such an ideal (one generated by a single element) a **principal ideal**. So, in $\mathbb{Z}$, every ideal is principal. But the same is not true in more general rings.

Example 4.8.1 (A non-principal ideal)

In $\mathbb{Z}[x]$, $I = (x, 2015)$ is *not* a principal ideal.

For if $I = (f)$ for some polynomial $f \in I$ then f divides x and 2015. This can only occur if $f = \pm 1$, but then I contains ± 1, which it does not.

A ring with the property that all its ideals are principal is called a **principal ideal ring**. We like this property because they effectively let us take the "greatest common factor" in a similar way as the GCD in $\mathbb{Z}$.

In practice, we actually usually care about so-called **principal ideal domains (PID's)**. But we haven't defined what a domain is yet. Nonetheless, all the examples below are actually PID's, so we will go ahead and use this word for now, and tell you what the additional condition is in the next chapter.

Example 4.8.2 (Examples of PID's)

To reiterate, for now you should just verify that these are principal ideal rings, even though we are using the word PID.

(a) As we saw, $\mathbb{Z}$ is a PID.

(b) As we also saw, $\mathbb{Z}[x]$ is not a PID, since $I = (x, 2015)$ for example is not principal.

(c) It turns out that for a field k the ring $k[x]$ is always a PID. For example, $\mathbb{Q}[x]$, $\mathbb{R}[x]$, $\mathbb{C}[x]$ are PID's.

If you want to try and prove this, first prove an analog of Bézout's lemma, which implies the result.

(d) $\mathbb{C}[x, y]$ is not a PID, because (x, y) is not principal.

§4.9 Noetherian rings

Prototypical example for this section: $\mathbb{Z}[x_1, x_2, \dots]$ *is not Noetherian, but most reasonable rings are. In particular polynomial rings are. (Equivalently, only weirdos care about non-Noetherian rings).*

If it's too much to ask that an ideal is generated by *one* element, perhaps we can at least ask that our ideals are generated by *finitely many* elements. Unfortunately, in certain weird rings this is also not the case.

Example 4.9.1 (Non-Noetherian ring)
Consider the ring $R = \mathbb{Z}[x_1, x_2, x_3, \dots]$ which has *infinitely* many free variables. Then the ideal $I = (x_1, x_2, \dots) \subseteq R$ cannot be written with a finite generating set.

Nonetheless, most "sane" rings we work in *do* have the property that their ideals are finitely generated. We now name such rings and give two equivalent definitions:

Proposition 4.9.2 (The equivalent definitions of a Noetherian ring)
For a ring R, the following are equivalent:

(a) Every ideal I of R is finitely generated (i.e. can be written with a finite generating set).

(b) There does *not* exist an infinite ascending chain of ideals

$$I_1 \subsetneq I_2 \subsetneq I_3 \subsetneq \dots.$$

The absence of such chains is often called the **ascending chain condition**.

Such rings are called **Noetherian**.

Example 4.9.3 (Non-Noetherian ring breaks ACC)
In the ring $R = \mathbb{Z}[x_1, x_2, x_3, \dots]$ we have an infinite ascending chain

$$(x_1) \subsetneq (x_1, x_2) \subsetneq (x_1, x_2, x_3) \subsetneq \dots.$$

From the example, you can kind of see why the proposition is true: from an infinitely generated ideal you can extract an ascending chain by throwing elements in one at a

time. I'll leave the proof to you if you want to do it.[3]

Question 4.9.4. Why are fields Noetherian? Why are PID's (such as $\mathbb{Z}$) Noetherian?

This leaves the question: is our prototypical non-example of a PID, $\mathbb{Z}[x]$, a Noetherian ring? The answer is a glorious yes, according to the celebrated Hilbert basis theorem.

Theorem 4.9.5 (Hilbert basis theorem)
Given a Noetherian ring R, the ring $R[x]$ is also Noetherian. Thus by induction, $R[x_1, x_2, \ldots, x_n]$ is Noetherian for any integer n.

The proof of this theorem is really olympiad flavored, so I couldn't possibly spoil it – I've left it as a problem at the end of this chapter.

Noetherian rings really shine in algebraic geometry, and it's a bit hard for me to motivate them right now, other than to say "most rings you'll encounter are Noetherian". Please bear with me!

§4.10 A few harder problems to think about

Problem 4A. The ring $R = \mathbb{R}[x]/(x^2+1)$ is one that you've seen before. What is its name?

Problem 4B. Show that $\mathbb{C}[x]/(x^2 - x) \cong \mathbb{C} \times \mathbb{C}$.

Problem 4C. In the ring $\mathbb{Z}$, let $I = (2016)$ and $J = (30)$. Show that $I \cap J$ is an ideal of $\mathbb{Z}$ and compute its elements.

Problem 4D*. Let R be a ring and I an ideal. Find an inclusion-preserving bijection between

- ideals of R/I, and
- ideals of R which contain I.

Problem 4E. Let R be a ring.

(a) Prove that there is exactly one ring homomorphism $\mathbb{Z} \to R$.

(b) Prove that the number of ring homomorphisms $\mathbb{Z}[x] \to R$ is equal to the number of elements of R.

Problem 4F. Prove the Hilbert basis theorem, Theorem 4.9.5.

Problem 4G (USA Team Selection Test 2016)**.** Let $\mathbb{F}_p$ denote the integers modulo a fixed prime number p. Define $\Psi \colon \mathbb{F}_p[x] \to \mathbb{F}_p[x]$ by

$$\Psi\left(\sum_{i=0}^{n} a_i x^i\right) = \sum_{i=0}^{n} a_i x^{p^i}.$$

Let S denote the image of Ψ.

(a) Show that S is a ring with addition given by polynomial addition, and multiplication given by *function composition*.

(b) Prove that $\Psi \colon \mathbb{F}_p[x] \to S$ is then a ring isomorphism.

[3] On the other hand, every undergraduate class in this topic I've seen makes you do it as homework. Admittedly I haven't gone to that many such classes.

5 Flavors of rings

We continue our exploration of rings by considering some nice-ness properties that rings or ideals can satisfy, which will be valuable later on. As before, number theory is interlaced as motivation. I guess I can tell you at the outset what the completed table is going to look like, so you know what to expect.

Ring noun	Ideal adjective	Relation
PID	principal	R is a PID $\iff$ R is an integral domain, and every I is principal
Noetherian ring	finitely generated	R is Noetherian $\iff$ every I is fin. gen.
field	maximal	R/I is a field $\iff$ I is maximal
integral domain	prime	R/I is an integral domain $\iff$ I is prime

§5.1 Fields

Prototypical example for this section: $\mathbb{Q}$ *is a field, but* $\mathbb{Z}$ *is not.*

We already saw this definition last chapter: a field K is a nontrivial ring for which every nonzero element is a unit.

In particular, there are only two ideals in a field: the ideal (0), which is maximal, and the entire field K.

§5.2 Integral domains

Prototypical example for this section: $\mathbb{Z}$ *is an integral domain.*

In practice, we are often not so lucky that we have a full-fledged field. Now it would be nice if we could still conclude the zero product property: if $ab = 0$ then either $a = 0$ or $b = 0$. If our ring is a field, this is true: if $b \neq 0$, then we can multiply by b^{-1} to get $a = 0$. But many other rings we consider like $\mathbb{Z}$ and $\mathbb{Z}[x]$ also have this property, despite not having division.

Not all rings though: in $\mathbb{Z}/15\mathbb{Z}$,

$$3 \cdot 5 \equiv 0 \pmod{15}.$$

If $a, b \neq 0$ but $ab = 0$ then we say a and b are **zero divisors** of the ring R. So we give a name to such rings.

Definition 5.2.1. A nontrivial ring with no zero divisors is called an **integral domain**.[1]

> **Question 5.2.2.** Show that a field is an integral domain.

> **Exercise 5.2.3** (Cancellation in integral domains)**.** Suppose $ac = bc$ in an integral domain, and $c \neq 0$. Show that $a = b$. (There is no c^{-1} to multiply by, so you have to use the definition.)

[1] Some authors abbreviate this to "domain", notably Artin.

Example 5.2.4 (Examples of integral domains)
Every field is an integral domain, so all the previous examples apply. In addition:

(a) $\mathbb{Z}$ is an integral domain, but it is not a field.

(b) $\mathbb{R}[x]$ is not a field, since there is no polynomial $P(x)$ with $xP(x) = 1$. However, $\mathbb{R}[x]$ is an integral domain, because if $P(x)Q(x) = 0$ then one of P or Q is zero.

(c) $\mathbb{Z}[x]$ is also an example of an integral domain. In fact, $R[x]$ is an integral domain for any integral domain R (why?).

(d) $\mathbb{Z}/n\mathbb{Z}$ is a field (hence integral domain) exactly when n is prime. When n is not prime, it is a ring but not an integral domain.

The trivial ring 0 is *not* considered an integral domain.

At this point, we go ahead and say:

Definition 5.2.5. An integral domain where all ideals are principal is called a **principal ideal domain (PID)**.

Recall that the ideal (a, b) is the ring-analog of the *gcd* operation, so essentially what this definition is saying is that: If any family of elements $\{a_i\}$ is taken, then the ideal generated by all of the a_i is in fact generated by a single element a.

In other words,

In a PID, you can take the *gcd* of any collection of elements.

The ring $\mathbb{Z}/6\mathbb{Z}$ is an example of a ring which is a principal ideal ring, but not an integral domain. As we alluded to earlier, we will never really use "principal ideal ring" in any real way: we typically will want to strengthen it to PID.

§5.3 Prime ideals

Prototypical example for this section: (5) *is a prime ideal of* $\mathbb{Z}$.

We know that every integer can be factored (up to sign) as a unique product of primes; for example $15 = 3 \cdot 5$ and $-10 = -2 \cdot 5$. You might remember the proof involves the so-called Bézout's lemma, which essentially says that $(a, b) = (\gcd(a, b))$; in other words we've carefully used the fact that $\mathbb{Z}$ is a PID.

It turns out that for general rings, the situation is not as nice as factoring elements because most rings are not PID's. The classic example of something going wrong is

$$6 = 2 \cdot 3 = \left(1 - \sqrt{-5}\right)\left(1 + \sqrt{-5}\right)$$

in $\mathbb{Z}[\sqrt{-5}]$. Nonetheless, we can sidestep the issue and talk about factoring *ideals*: somehow the example $10 = 2 \cdot 5$ should be $(10) = (2) \cdot (5)$, which says "every multiple of 10 is the product of a multiple of 2 and a multiple of 5". I'd have to tell you then how to multiply two ideals, which I do in the chapter on unique factorization.

Let's at least figure out what primes are. In $\mathbb{Z}$, we have that $p \neq 1$ is prime if whenever $p \mid xy$, either $p \mid x$ or $p \mid y$. We port over this definition to our world of ideals.

Definition 5.3.1. A *proper* ideal $I \subsetneq R$ is a **prime ideal** if whenever $xy \in I$, either $x \in I$ or $y \in I$.

The condition that I is proper is analogous to the fact that we don't consider 1 to be a prime number.

Example 5.3.2 (Examples and non-examples of prime ideals)

(a) The ideal (7) of $\mathbb{Z}$ is prime.

(b) The ideal (8) of $\mathbb{Z}$ is not prime, since $2 \cdot 4 = 8$.

(c) The ideal (x) of $\mathbb{Z}[x]$ is prime.

(d) The ideal (x^2) of $\mathbb{Z}[x]$ is not prime, since $x \cdot x = x^2$.

(e) The ideal $(3, x)$ of $\mathbb{Z}[x]$ is prime. This is actually easiest to see using Theorem 5.3.5 below.

(f) The ideal $(5) = 5\mathbb{Z} + 5i\mathbb{Z}$ of $\mathbb{Z}[i]$ is not prime, since the elements $3 + i$ and $3 - i$ have product $10 \in (5)$, yet neither is itself in (5).

Remark 5.3.3 — Ideals have the nice property that they get rid of "sign issues". For example, in $\mathbb{Z}$, do we consider -3 to be a prime? When phrased with ideals, this annoyance goes away: $(-3) = (3)$. More generally, for a ring R, talking about ideals lets us ignore multiplication by a unit. (Note that -1 is a unit in $\mathbb{Z}$.)

Exercise 5.3.4. What do you call a ring R for which the zero ideal (0) is prime?

We also have:

Theorem 5.3.5 (Prime ideal $\iff$ quotient is integral domain)

An ideal I is prime if and only if R/I is an integral domain.

Exercise 5.3.6 (Mandatory)**.** Convince yourself the theorem is true; it is just definition chasing. (A possible start is to consider $R = \mathbb{Z}$ and $I = (15)$.)

I now must regrettably inform you that unique factorization is still not true even with the notion of a "prime" ideal (though again I haven't told you how to multiply two ideals yet). But it will become true with some additional assumptions that will arise in algebraic number theory (relevant buzzword: Dedekind domain).

§5.4 Maximal ideals

Prototypical example for this section: The ideal $(x, 5)$ is maximal in $\mathbb{Z}[x]$, by quotient-ing.

Here's another flavor of an ideal.

Definition 5.4.1. A proper ideal I of a ring R is **maximal** if it is not contained in any other proper ideal.

Example 5.4.2 (Examples of maximal ideals)

(a) The ideal $I = (7)$ of $\mathbb{Z}$ is maximal, because if an ideal J contains 7 and an element n not in I it must contain $\gcd(7, n) = 1$, and hence $J = \mathbb{Z}$.

(b) The ideal (x) is *not* maximal in $\mathbb{Z}[x]$, because it's contained in $(x, 5)$ (among others).

(c) On the other hand, $(x, 5)$ is indeed maximal in $\mathbb{Z}[x]$. This is actually easiest to verify using Theorem 5.4.4 below.

(d) Also, (x) is maximal in $\mathbb{C}[x]$, again appealing to Theorem 5.4.4 below.

Exercise 5.4.3. What do you call a ring R for which the zero ideal (0) is maximal?

There's an analogous theorem to the one for prime ideals.

Theorem 5.4.4 (I maximal $\iff$ R/I field)
An ideal I is maximal if and only if R/I is a field.

Proof. A ring is a field if and only if (0) is the only maximal ideal. So this follows by Problem 4D*. □

Corollary 5.4.5 (Maximal ideals are prime)
If I is a maximal ideal of a ring R, then I is prime.

Proof. If I is maximal, then R/I is a field, hence an integral domain, so I is prime. □

In practice, because modding out by generated ideals is pretty convenient, this is a very efficient way to check whether an ideal is maximal.

Example 5.4.6 (Modding out in $\mathbb{Z}[x]$)

(a) This instantly implies that $(x, 5)$ is a maximal ideal in $\mathbb{Z}[x]$, because if we mod out by x and 5 in $\mathbb{Z}[x]$, we just get $\mathbb{F}_5$, which is a field.

(b) On the other hand, modding out by just x gives $\mathbb{Z}$, which is an integral domain but not a field; that's why (x) is prime but not maximal.

As we saw, any maximal ideal is prime. But now note that $\mathbb{Z}$ has the special property that all of its nonzero prime ideals are also maximal. It's with this condition and a few other minor conditions that you get a so-called *Dedekind domain* where prime factorization of ideals *does* work. More on that later.

§5.5 Field of fractions

Prototypical example for this section: $\operatorname{Frac}(\mathbb{Z}) = \mathbb{Q}$.

As long as we are here, we take the time to introduce a useful construction that turns any integral domain into a field.

Definition 5.5.1. Given an integral domain R, we define its **field of fractions** or **fraction field** $\text{Frac}(R)$ as follows: it consists of elements a/b, where $a, b \in R$ and $b \neq 0$. We set $a/b \sim c/d$ if and only if $bc = ad$. Addition and multiplication is defined by

$$\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}$$
$$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}.$$

In fact everything you know about $\mathbb{Q}$ basically carries over by analogy. You can prove if you want that this indeed a field, but considering how comfortable we are that $\mathbb{Q}$ is well-defined, I wouldn't worry about it...

Definition 5.5.2. Let k be a field. We define $k(x) = \text{Frac}(k[x])$ (read "k of x"), and call it the **field of rational functions**.

Example 5.5.3 (Examples of fraction fields)

(a) By *definition*, $\text{Frac}(\mathbb{Z}) = \mathbb{Q}$.

(b) The field $\mathbb{R}(x)$ consists of rational functions in x:

$$\mathbb{R}(x) = \left\{ \frac{f(x)}{g(x)} \mid f, g \in \mathbb{R}[x] \right\}.$$

For example, $\frac{2x}{x^2-3}$ might be a typical element.

Example 5.5.4 (Gaussian rationals)

Just like we defined $\mathbb{Z}[i]$ by abusing notation, we can also write $\mathbb{Q}(i) = \text{Frac}(\mathbb{Z}[i])$. Officially, it should consist of

$$\mathbb{Q}(i) = \left\{ \frac{f(i)}{g(i)} \mid g(i) \neq 0 \right\}$$

for polynomials f and g with rational coefficients. But since $i^2 = -1$ this just leads to

$$\mathbb{Q}(i) = \left\{ \frac{a+bi}{c+di} \mid a, b, c, d \in \mathbb{Q}, (c,d) \neq (0,0) \right\}.$$

And since $\frac{1}{c+di} = \frac{c-di}{c^2+d^2}$ we end up with

$$\mathbb{Q}(i) = \{a + bi \mid a, b \in \mathbb{Q}\}.$$

§5.6 Unique factorization domains (UFD's)

Prototypical example for this section: $\mathbb{Z}$ and polynomial rings in general.

Here is one stray definition that will be important for those with a number-theoretic inclination. Over the positive integers, we have a fundamental theorem of arithmetic, stating that every integer is uniquely the product of prime numbers.

We can even make an analogous statement in $\mathbb{Z}$ or $\mathbb{Z}[i]$, if we allow representations like $6 = (-2)(-3)$ and so on. The trick is that we only consider everything *up to units*; so $6 = (-2)(-3) = 2 \cdot 3$ are considered the same.

The general definition goes as follows.

Definition 5.6.1. A nonzero non-unit of an integral domain R is **irreducible** if it cannot be written as the product of two non-units.

An integral domain R is a **unique factorization domain** if every nonzero non-unit of R can be written as the product of irreducible elements, which is unique up to multiplication by units.

Question 5.6.2. Verify that $\mathbb{Z}$ is a UFD.

Example 5.6.3 (Examples of UFD's)

(a) Fields are a "degenerate" example of UFD's: every nonzero element is a unit, so there is nothing to check.

(b) $\mathbb{Z}$ is a UFD. The irreducible elements are p and $-p$, for example 5 or -17.

(c) $\mathbb{Q}[x]$ is a UFD: polynomials with rational coefficients can be uniquely factored, up to scaling by constants (as the units of $\mathbb{Q}[x]$ are just the rational numbers).

(d) $\mathbb{Z}[x]$ is a UFD.

(e) The Gaussian integers $\mathbb{Z}[i]$ turns out to be a UFD too (and this will be proved in the chapters on algebraic number theory).

(f) $\mathbb{Z}[\sqrt{-5}]$ is the classic non-example of a UFD: one may write

$$6 = 2 \cdot 3 = \left(1 - \sqrt{-5}\right)\left(1 + \sqrt{-5}\right)$$

but each of 2, 3, $1 \pm \sqrt{-5}$ is irreducible. (It turns out the right way to fix this is by considering prime *ideals* instead, and this is one big motivation for Part XIV.)

(g) Theorem we won't prove: if R is a UFD, so is $R[x]$ (and hence by induction so is $R[x, y]$, $R[x, y, z]$, ...).

We have the following theorem:

Theorem 5.6.4

Let R be a PID. Then R is a UFD.

If we look at the non-example above:

$$6 = 2 \cdot 3 = \left(1 - \sqrt{-5}\right)\left(1 + \sqrt{-5}\right)$$

The failure of $\mathbb{Z}[\sqrt{-5}]$ to be a UFD here is reflected by the fact that we cannot decompose any factor into further irreducible factor, indeed, the ideal

$$\left(2, 1 + \sqrt{-5}\right)$$

is not principal — there is no element that is the gcd of 2 and $1 + \sqrt{-5}$.

In a similar manner, we can prove that a PID is an UFD, assuming a decomposition into irreducible factors exist.

§5.7 Extra: Euclidean domains

This chapter will not be used later on, but it is of historical interest.

Recall that a PID is a ring where you can take the gcd of any family of elements.

We all know that the most popular algorithm to compute the gcd of two elements in $\mathbb{Z}$ is the Euclidean algorithm:

- Start with two integers a and b.
- If either of a and b is 0, we're done. The gcd is the nonzero element.
- Otherwise, assume $|a| \geq |b|$, divide a by b to get some remainder r such that $|r| < |b|$, replace a with r, and continue the algorithm.

This algorithm is very efficient — it can be proven that the algorithm only takes logarithmically many steps in the size of a and b — for instance, if a and b are on the order of 10^{100}, at worst 500 steps are needed.

Naturally, the following questions are raised:

> On which rings can we perform the same algorithm?

We will see that we can in fact do it on several rings! For example, the ring of Gaussian integers $\mathbb{Z}[i]$, the Eisenstein integers $\mathbb{Z}[\omega]$, and so on.

If we look at the algorithm description above, what makes the algorithm work? It's the absolute value $|\cdot|$ which is used to compare the magnitude of two numbers, and this absolute value satisfies two conditions:

- It outputs nonnegative integer values — that way, the algorithm will eventually terminate.
- For any two ring elements a and b, where $b \neq 0$, there exist some q such that $r = a - qb$ has smaller absolute value than b.

So, naturally, for any ring with a similar integer-valued function, we can perform the algorithm. We call a function $\mathrm{N}\colon R \to \mathbb{Z}_{\geq 0}$ that satisfies the two conditions above an **Euclidean norm**, and an integral domain R that has a norm an **Euclidean domain**.

Example 5.7.1 (The ring of Gaussian integers is an Euclidean domain)

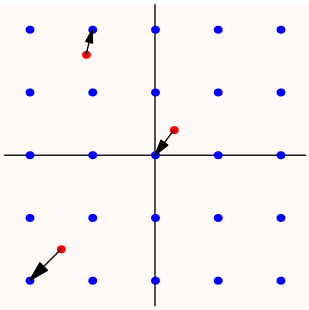
On $\mathbb{Z}[i]$, the usual norm

$$|a + bi| = a^2 + b^2$$

is an Euclidean norm.

Indeed, for any elements a and b with $b \neq 0$, we can compute the remainder r by dividing a by b, let q be the Gaussian integer that is closest to $\frac{a}{b}$ (that is, $|\frac{a}{b} - q|$ is minimized) and let $r = a - bq$, then it can be proven that $|r| < |b|$.

The proof is done by showing $|\frac{a}{b} - q| < 1$ — if we look at the lattice of points contained in $\mathbb{Z}[i]$ embedded in the complex plane, then for any value of $\frac{a}{b} \in \mathbb{C}$, rounding it to the nearest integer will move it by at most $\frac{\sqrt{2}}{2} < 1$.



Example 5.7.2 (The ring of Eisenstein integers is an Euclidean domain)

Similarly, let $\omega = \frac{1+\sqrt{3}i}{2}$ (that is $\omega^3 = -1$), then $\mathbb{Z}[\omega]$ is an Euclidean domain with the usual norm

$$|a+bi| = a^2+b^2$$

or equivalently

$$|a+b\omega| = a^2+ab+b^2.$$

Example 5.7.3 (The ring $\mathbb{Z}[\sqrt{11}]$ is an Euclidean domain)

As before. This time, the natural norm[a] will be:

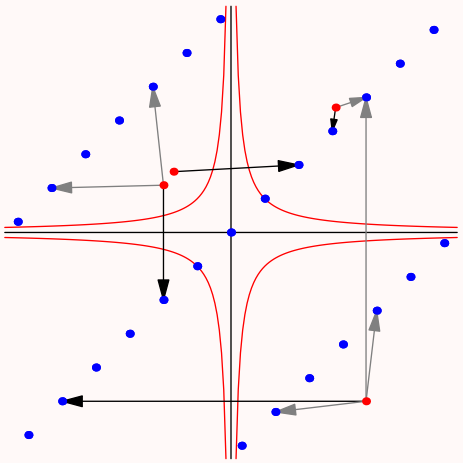
$$\mathrm{N}_{\mathbb{Q}(\sqrt{11})/\mathbb{Q}}(a+b\sqrt{11}) = (a+b\sqrt{11})(a-b\sqrt{11}) = a^2-11b^2.$$

Since we need an Euclidean norm, we will take $\mathrm{N}(a+b\sqrt{11}) = |a^2-11b^2|$.

Given two elements a and b in $\mathbb{Z}[\sqrt{11}]$ with $b \neq 0$, we will try to compute r such that $\mathrm{N}(r) < \mathrm{N}(b)$ as $r = a - qb$ as before.

This time around, we cannot draw $\mathbb{Z}[\sqrt{11}]$ as a lattice of points — it is dense in $\mathbb{R}$ — so, each point $a+b\sqrt{11}$ will be drawn at the coordinate $(a+b\sqrt{11}, a-b\sqrt{11})$.

The set of points with norm < 1 will be drawn below.



Instead of a ball (as in imaginary quadratic fields, that is, $\mathbb{Q}(\sqrt{-d})$ for integer d), the set of points with norm 1 forms a hyperbola.

As such, rounding to the nearest point is not always the best way — nevertheless, it can be proven (by exhaustive case checking, similar to the case of $\mathbb{Z}[i]$) that for

every value of $\frac{a}{b} \in \mathbb{Q}(\sqrt{11})$, there is some $q \in \mathbb{Z}[\sqrt{11}]$ such that $\mathrm{N}(\frac{a}{b} - q) < 1$. Thus N is an Euclidean norm.

[a] See Section 54.1 for the explanation why this norm is the natural one.

That having said, sometimes the natural norm of an Euclidean domain need not be Euclidean. $\mathbb{Z}[\frac{1+\sqrt{69}}{2}]$ is the first example.

Example 5.7.4 ($\mathbb{Q}[x]$ is an Euclidean domain)

Similarly, in $\mathbb{Q}[x]$ we can let the norm be the degree of a polynomial — the polynomial division with remainder algorithm will take care of computing the gcd.

Back to the topic of PID. In an Euclidean domain, you can compute the gcd of any two elements. What about an infinite family of elements?

Turns out the situation is very nice:

Proposition 5.7.5

An Euclidean domain is a PID.

Actually, we don't need to provide an explicit algorithm to compute the gcd of an infinite family of elements — of course any such algorithm cannot terminate in a finite amount of time! — but we only need to show the gcd exists, we can cheat our way out.

Note that in the Euclidean algorithm, the norm of the elements *keep decreasing* until one of the elements become 0. So, if we're given an arbitrary family of elements, we take the ideal generated by these elements — certainly the gcd is inside that ideal — and we *take the nonzero element with the smallest norm*. This is the gcd.

With that intuition in mind, we formalize our proof:

Proof. Let I be any ideal. We need to show I is principal.

Let a be a nonzero element with smallest norm in I — such an element exists because $\mathbb{Z}_{\geq 0}$ is well-ordered.

Question 5.7.6. Show that, for every other elements $b \in I$, then $a \mid b$.

Thus, $I = (a)$, we're done. □

Example 5.7.7 (The ring $\mathbb{Z}[\frac{1+\sqrt{-19}}{2}]$ is not an Euclidean domain)

Let $R = \mathbb{Z}[\frac{1+\sqrt{-19}}{2}]$.

With the above example in mind, what can we say about this ring?

This is in fact a principal ideal domain (we will not prove it here), but there is in fact no Euclidean norm on this ring.

We will prove the above claim. The general plan is:

- Show that the existence of an Euclidean norm implies the existence of something that we calls an **universal side divisor**.
- Show that R has no universal side divisor.

- Thus, R cannot have an Euclidean norm.

First, look at the examples above of $\mathbb{Z}[i]$ and $\mathbb{Z}[\omega]$. We see that the units are the elements with the smallest norm.

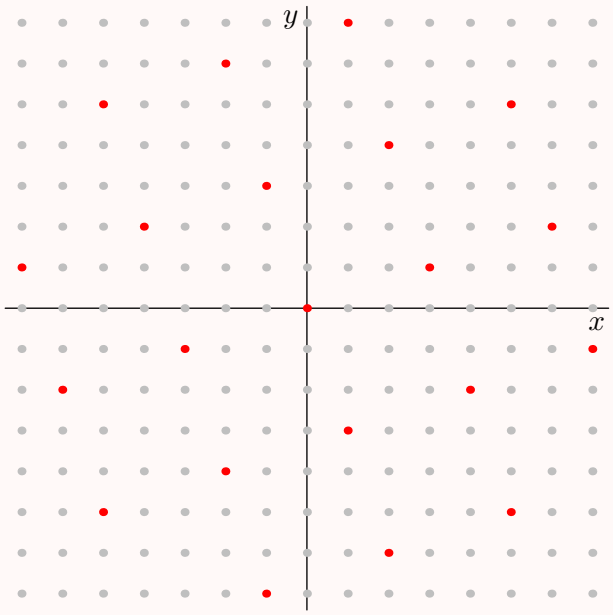
So far, nothing useful — every ring has the unit 1. Next, we look at the elements with the next-smallest norm:

- In $\mathbb{Z}[i]$, the element $1+i$ has norm 2, and is an element with smallest norm that is not 0 or an unit. (The other elements are $\pm 1 \pm i$.)
- In $\mathbb{Z}[\omega]$, the units are $\{\pm 1, \pm\omega, \pm(\omega-1)\}$ with norm 1. An elements with next-smallest norms are $1+\omega$ with norm 3.

Now, in order to proceed with the proof, we have to define a *side divisor*. Recall that b is a divisor of a if there is some q such that $a = bq$. We say b is a **side divisor** (read: "almost divisor") of a if there is some q such that the remainder $a - bq$ is either 0 or an unit.

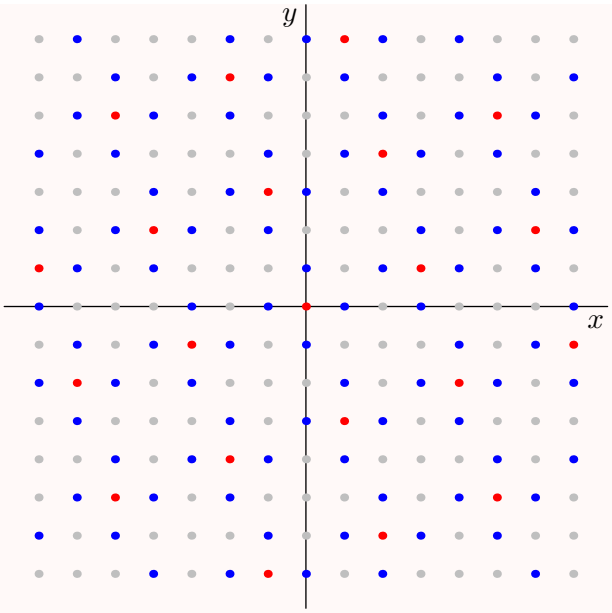
Example 5.7.8

In $\mathbb{Z}[i]$, consider $b = 3+i$. The set of numbers a for which $b \mid a$ is drawn in red below.



If we add an unit to these values of a, we get the set of numbers a for which b is a side divisor of a, thus give a picture to the concept of "almost divisor".

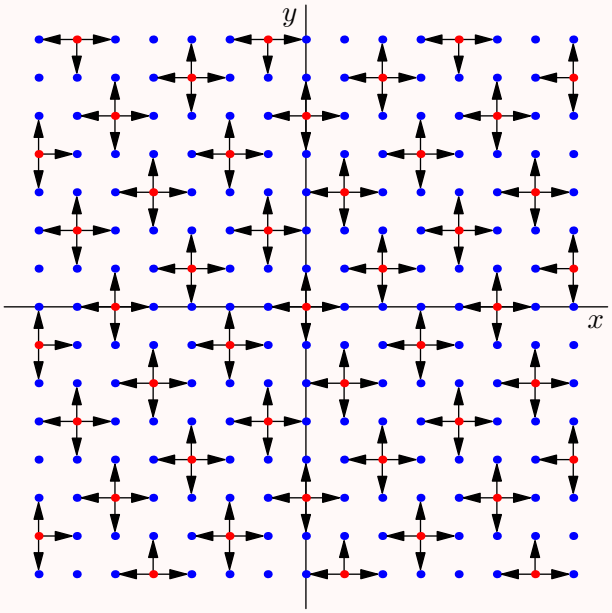
The points a where b is a side divisor of a is marked in red and blue below.



Finally, we define a **universal side divisor** to be a number b such that b is a side divisor of every element $a \in R$.

Example 5.7.9

Drawing a picture similar to the above, in $\mathbb{Z}[i]$, then $2+i$ and $1+i$ are side divisors.



Now, the connection between the two concepts considered above.

Lemma 5.7.10

In an Euclidean domain, the smallest-norm nonzero element b that is not an unit is a universal side divisor.

Proof. Just run the Euclidean algorithm between any number and b for one step, the remainder must be 0 or an unit. □

And finally,

Proposition 5.7.11

There is no universal side divisor in R.

fix

Proof is tedious, omitted.

§5.8 A few harder problems to think about

Not olympiad problems, but again the spirit is very close to what you might see in an olympiad.

Problem 5A. Consider the ring

$$\mathbb{Q}[\sqrt{2}] = \left\{ a + b\sqrt{2} \mid a, b \in \mathbb{Q} \right\}.$$

Is it a field?

Problem 5B (Homomorphisms from fields are injective)**.** Let K be a field and R a nontrivial ring. Prove that any homomorphism $\psi \colon K \to R$ is injective.[2]

Problem 5C* (Pre-image of prime ideals)**.** Suppose $\phi \colon R \to S$ is a ring homomorphism, and $I \subseteq S$ is a prime ideal. Prove that $\phi^{\text{pre}}(I)$ is prime as well.

Problem 5D*. Let R be an integral domain with finitely many elements. Prove that R is a field.

Problem 5E* (Krull's theorem)**.** Let R be a ring and J a proper ideal.

(a) Prove that if R is Noetherian, then J is contained in a maximal ideal I.

(b) Use Zorn's lemma (Chapter 88) to prove the result even if R isn't Noetherian.

Problem 5F (Spec $k[x]$)**.** Describe the prime ideals of $\mathbb{C}[x]$ and $\mathbb{R}[x]$.

Problem 5G†. How many prime ideals of $\mathbb{Z}[\sqrt{2017}]$ are *not* maximal ideals?

Problem 5H. Let R denote the set of rational numbers q such that, when q is written in lowest terms, the denominator is not a multiple of 5. Then R is a ring (under the usual addition and multiplication). Classify all the ideals of R. Which of these ideals are prime / maximal?

[2] Note that ψ cannot be the zero map for us, since we require $\psi(1_K) = 1_R$. You sometimes find different statements in the literature.