IB Groups, Rings and Modules

Martin von Hodenberg (mjv43@cam.ac.uk)

Last updated: April 12, 2022

These are my notes for the IB course 'Groups, Rings and Modules', which was lectured in Lent 2022 at Cambridge by Dr R.Zhou. These notes are written in LaTeX for revision purposes[1]. Any suggestions or feedback is welcome.

[1]These notes are posted online on my website.

Contents

0 Introduction

This is a second course in abstract algebra that will extend our knowledge of groups and introduce other fundamental objects.

The course is divided into several sections:

1. Groups; this will be a continuation from IA, focusing on simple groups, p-groups, and p-subgroups. The main result in this part of the course will be the Sylow theorems.

2. Rings; these are sets where you can add, subtract and multiply (e.g $\mathbb{Z}$ or $\mathbb{C}[X]$). We will study rings of integers such as $\mathbb{Z}[i], \mathbb{Z}[\sqrt{2}]$. These also generalise to polynomial rings. We will also study fields, which are rings where you can divide (e.g $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ or $\mathbb{Z}/p\mathbb{Z}$ for p prime).

3. Modules; these are an analogue of vector spaces where the scalars belong to a ring instead of a field. We will classify modules over certain "nice" rings. This allows us to prove Jordan Normal Form, and classify finite abelian groups.

1 Groups

1.1 Recall of IA Groups

This first subsection will just recap the results seen in IA Groups; it can be skipped by anyone with a sufficient knowledge of the course.

Definition (Group)
A **group** is a pair $(G, \cdot)$ where G is a set and $\cdot : G \times G \to G$ is a binary operation satisfying:

1. $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (associativity)
2. $\exists e \in G$ such that $e \cdot g = g \cdot e = g$ for all $g \in G$ (identity)
3. $\forall g \in G$, $\exists g^{-1} \in G$ such that $g \cdot g^{-1} = g^{-1} \cdot g = e$ (inverses)

Remarks.
- In practice, one often needs to check closure in order to check that $\cdot$ is well-defined.
- If using additive (respectively multiplicative) relations, we will often write 0 (or 1) for the identity.
- We write $|G|$ for the number of elements in G.

Definition (Subgroup)
A subset $H \subseteq G$ is a **subgroup** (written $H \leq G$) if H is closed under $\cdot$ and $(H, \cdot)$ is a group.

Remark. A non-empty subset H of G is a subgroup if $a, b \in H \implies a \cdot b^{-1} \in H$ (see IA Groups for the proof).

Example (Examples of groups)
Groups we have already seen include:

- Additive groups $(\mathbb{Z}, +) \leq (\mathbb{Q}, +) \leq (\mathbb{R}, +)$.
- Cyclic and dihedral groups C_n and D_{2n}.
- Abelian groups: those groups G such that $a \cdot b = b \cdot a$ for all $a, b \in G$.
- Symmetric and alternating groups $S_n =$ group of all permutations of $\{1, \ldots, n\}$ and $A_n \leq S_n$, the group of all even permutations.
- Quaternion group $Q_8 = \{\pm 1, \pm i, \pm j, \pm k\}$ where i, j, k are quaternions.
- General and special linear groups $GL_n(\mathbb{R}) = n \times n$ matrices on $\mathbb{R}$ with $\det \neq 0$, where the group operation is matrix multiplication. This contains the subgroup $SL_n(\mathbb{R}) \leq GL_n(\mathbb{R})$, which is the subgroup of matrices with $\det = 1$.

Definition (Direct product)

The **direct product** of groups G and H is the set $G \times H$ with operation

$$(g_1, h_1) \cdot (g_2, h_2) = (g_1 g_2, h_1 h_2).$$

Theorem 1.1 (Lagrange's theorem)
Let $H \leq G$. Then the left cosets of H in G are the sets $gH = \{gh : h \in H\}$ for $g \in G$. These partition G, and each has the same cardinality as H. From this we can deduce Lagrange's theorem:

If G is a finite group and $H \leq G$, then $|G| = |H|[G : H]$ where $[G : H]$ is the number of left cosets of H in G (the **index** of H in G).

Remark. Can also carry this out with right cosets. A corollary of Lagrange's theorem is thus that the number of left cosets = number of right cosets.

Definition (Order of an element)
Let $g \in G$. If $\exists n \geq 1$ such that $g^n = 1$, then the least such n is the **order** of g in G. If no such n exists, g has infinite order.

Remark. If g has order d, then

- $g^n = 1 \implies d | n$.
- $\{1, g, \ldots, g^{d-1}\} \leq G$ and so if G is finite, then $d \big| |G|$ (Lagrange).

Definition (Normal subgroup)
A subgroup $H \leq G$ is **normal** if $g^{-1}Hg = H$ for all $g \in G$. We write $H \trianglelefteq G$.

Proposition 1.2
If $H \trianglelefteq G$ then the set G/H of left cosets of H in G is a group (called the **quotient group**) with operation $g_1H \cdot g_2H = g_1g_2H$.

Proof. Check $\cdot$ is well-defined:

Suppose $g_1H = g_1'H$ and $g_2H = g_2'H$ for some $g_1, g_1', g_2, g_2' \in G$. Then $g_1' = g_1h_1$ and $g_2' = g_2h_2$ for some $h_1, h_2 \in H$. Therefore

$$\begin{aligned} g_1'g_2' &= g_1h_1g_2h_2 \\ &= g_1g_2 \underbrace{({g_2}^{-1}h_1g_2)}_{\in H} \underbrace{h_2}_{\in H} \end{aligned}$$

Therefore $g_1'g_2'H = g_1g_2H$. Associativity is inherited from G, the identity is $H = eH$, and the inverse of gH is $g^{-1}H$. $\square$

Definition (Homomorphism)
If G, H are groups, then a function $\phi : G \to H$ is a group **homomorphism** if

$\phi(g_1g_2) = \phi(g_1g_2) = \phi(g_1)\phi(g_2)$. It has kernel

$$\ker\phi = \{g \in G : \ \phi(g) = e\} \leq G.$$

and image

$$\operatorname{Im}\phi = \{\phi(g) : \ g \in G\} \leq H.$$

Remark. If $a \in \ker\phi$ and $g \in G$, then

$$\begin{aligned}\phi(g^{-1}ag) &= \phi(g^{-1})\phi(a)\phi(g) \\ &= \phi(g^{-1})\phi(g) \\ &= \phi(g^{-1}g) = \phi(e) = e.\end{aligned}$$

So $g^{-1}ag \in \ker\phi$ and hence $\ker\phi$ is a normal subgroup of G.

Definition (Isomorphism)
An **isomorphism** of groups is a group homomorphism that is also a bijection. We say G and H are **isomorphic** and write $G \cong H$ if there exists an isomorphism $\phi : G \to H$. (Note it follows from the definition that ϕ^{-1} is also a group homomorphism)

Theorem 1.3 (First Isomorphism Theorem)
Let $\phi : G \to H$ be a group homomorphism. Then $\ker\phi \trianglelefteq G$ and

$$G/\ker\phi \cong \operatorname{Im}\phi.$$

Proof. Let $K = \ker\phi$. We have already checked K is normal. Now we define $\Phi : G/K \to \operatorname{Im}\phi$ by

$$gK \to \phi(g).$$

To show Φ is well defined and injective:

$$\begin{aligned}g_1K = g_2K &\iff {g_2}^{-1}g_1 \in K \\ &\iff \phi({g_2}^{-1}g_1) = e \\ &\iff \phi(g_1) = \phi(g_2).\end{aligned}$$

To show Φ is a group hom.:

$$\begin{aligned}\Phi(g_1Kg_2K) &= \Phi(g_1g_2K) \\ &= \phi(g_1g_2) = \phi(g_1)\phi(g_2) \\ &= \Phi(g_1K)\Phi(g_2K)\end{aligned}$$

Showing Φ is surjective:

Let $x \in \operatorname{Im}\phi$, say $x = \phi(g)$ for some $g \in G$. Then $x = \phi(gR)$. $\square$

Example (Application of First Isomorphism Theorem)

Let $\phi : \mathbb{C} \to \mathbb{C}^x = \{x \in C : x \neq 0\}$ given by $z \mapsto e^z$.

Since $e^{z+w} = e^z e^w$, this is a group homomorphism from $(\mathbb{C}, +) \to (\mathbb{C}^x, \times)$. We have that

$$\begin{aligned} \ker\phi &= \{z \in \mathbb{C} :\ e^x = 1\} = 2\pi i\mathbb{Z} \\ \operatorname{Im}\phi &= \mathbb{C}^x \text{ by existence of log} \end{aligned}$$

Hence $\mathbb{C}/2\pi i\mathbb{Z} \cong \mathbb{C}^x$.

Theorem 1.4 (Second Isomorphism Theorem)
Let $H \leq G$, and $K \trianglelefteq G$. Then $HK = \{hk : h \in H, k \in K\} \leq G$ and $H \cap K \trianglelefteq H$. Moreover,

$$HK/K \cong H/(H \cap K).$$

Proof. Let $h_1k_1, h_2k_2 \in HK$ (so $h_1h_2 \in H,\ k_1k_2 \in K$). Now

$$h_1k_1(h_2k_2)^{-1} = \underbrace{h_1{h_2}^{-1}}_{\in H}(\underbrace{h_2k_1{k_2}^{-1}{h_2}^{-1}}_{\in K}) \in HK.$$

Thus $HK \leq G$ (by our previous remark). Let $\phi : H \to G/K$ be given by $h \to hK$. This is the composite of $H \to G$ and the quotient map $G \to G/K$; hence ϕ is a group homomorphism. Thus

$$\begin{aligned} &\ker\phi = \{h \in H : hK = K\} = H \cap K \trianglelefteq H \\ &\operatorname{Im}\phi = \{hK :\ h \in H\} = HK/K \end{aligned}$$

Now by the First Isomorphism Theorem $HK/K \cong H/(H \cap K)$. $\square$

Lemma 1.5
Suppose $K \trianglelefteq G$. There is a bijection

$$\{\text{subgroups of } G/K\} \leftrightarrow \{\text{subgroups of } G \text{ containing } K\},$$

where $X \mapsto \{g \in G : gK \in X\}$ and $H/K \leftarrow H$. This further restricts to a bijection

$$\{\text{normal subgroups of } G/K\} \leftrightarrow \{\text{normal subgroups of } G \text{ containing } K\}.$$

Theorem 1.6 (Third Isomorphism Theorem)
Let $K \leq H \leq G$ be normal subgroups of G. Then

$$\frac{G/K}{H/K} \cong G/H.$$

Proof. Let $\phi : G/K \to G/K$ be defined by $gK \mapsto gH$. If $g_1K = g_2K$, then ${g_2}^{-1}g_1 \in K \leq H \implies g_1H = g_2H$. Thus ϕ is well-defined.

Thus ϕ is a surjective homomorphim with kernel H/K. Now just apply the First Isomorphism Theorem. □

1.2 Simple groups

If $K \trianglelefteq G$, then studying the groups K and G/K gives some information about G. However, this approach is not always available. This is the case when a group is simple.

Definition (Simple group)
A group G is simple if $\{e\}$ and G are its only normal subgroups.

Remark. It is convention to not consider the trivial group a simple group.

Lemma 1.7
Let G be an abelian group. G is simple iff $G \cong C_p$ for some prime p.

Proof. $\Longleftarrow$ **:** Let $H \leq C_p$. Lagrange's theorem says that $|H|\ ||C_p| = p$. Since p is prime, $|H| = 1$ or p. So H is the trivial group or C_p.

$\Longrightarrow$ **:** Let $g \in G$ where $g \neq e$. Consider the subgroup generated by g:

$$\langle g \rangle = \left\{ \ldots, g^{-2}, g^{-1}, e, g, g^2, \ldots \right\}.$$

This is normal in G since G is abelian. Since G is simple, $\langle g \rangle = G$. If G is infinite, $G \cong (\mathbb{Z}, +)$ and $2\mathbb{Z} \leq \mathbb{Z}$ which gives a contradiction.

Otherwise, we now know $G \cong C_n$ for some n. Let g be a generator. If $m|n$ then $g^{n/m}$ generates a subgroup of order m and so G simple $\implies$ the only factors of n are 1 and n. Therefore n is prime. □

Lemma 1.8
If G is a finite group, then G has a composition series

$$e = G_0 \trianglelefteq G_1 \trianglelefteq \ldots \trianglelefteq G_m = G,$$

with each quotient G_i/G_{i-1} simple.

Proof. We induct on $|G|$. If $|G| = 1$ it's obvious. If $|G| > 1$, let G_{m-1} be a normal subgroup of largest possible order that isn't G itself. Lemma 1.5 implies G/G_{m-1} is simple. Then apply the induction hypothesis to G_{m-1}. □

1.3 Group actions

Definition (Permutation group)
For X any set, let $\mathrm{Sym}(X)$ be the group of all bijections $X \to X$ under composition. This clearly forms a group with $e = \mathrm{Id}_X$.

A group G is a permutation group of degree n if $G \leq \mathrm{Sym}(X)$ with $|X| = n$.

Example
Examples of permutation groups are:

- $S_n = \mathrm{Sym}(\{1, 2, \dots, n\})$ is a permutation group of degree n, as is $A_n \leq S_n$.
- $D_{2n} =$ (symmetries of a regular n-gon) is a subgroup of $\mathrm{Sym}(\{\text{vertices of n-gon}\})$.

Definition (Group action)
An action of a group G on a set X is a function $* : G \times X \to X$ satisfying

(i) $e * x = x$ for all $x \in X$

(ii) $(g_1 g_2) * x = g_1 * (g_2 * x)$ for all $g_1, g_2 \in G$, $x \in X$.

Proposition 1.9
An action of a group G on a set X is equivalent to specifying a group homomorphism $\phi : G \to \mathrm{Sym}(X)$.

Proof. For each $g \in G$, let $\phi_g : X \to X$ send $x \mapsto g * x$.

We have $\phi_{g_1 g_2}(x) = (g_1 g_2) * x = g_1 * (g_2 * x) = \phi_{g_1} \circ \phi_{g_2}(x)$. (†)

In particular, $\phi_g \circ \phi_{g^{-1}} = \phi_{g^{-1}} \circ \phi_g = \phi_e = \mathrm{Id}_X$. Thus $\phi_g \in \mathrm{Sym}(X)$. Then the map $\phi : G \to \mathrm{Sym}(X)$ given by $g \mapsto \phi_g$ is a group homomorphism by (†).

Conversely, let $\phi : G \to \mathrm{Sym}(X)$ be a group homomorphism. Define $* : G \times X \to X$ given by $(g, x) \mapsto \phi(g)(x)$. Then

(i) $e * x = \phi(e)(x) = \mathrm{Id}_X(x) = x$.

(ii) $(g_1 g_2) * x = \phi(g_1 g_2)(x) = \phi(g_1)(\phi(g_2)(x)) = g_1 * (g_2 * x)$.

□

Definition
We say $\phi : G \to \mathrm{Sym}(X)$ is a **permutation representation** of G.

Definition (Orbit and stabiliser)
Let G act on a set X.

(i) The **orbit** of $x \in X$ is $\mathrm{orb}_G(x) = \{g * x : g \in G\} \subset X$

(ii) The **stabiliser** of $x \in X$ is

$$G_x = \{g \in G : \ g * x = x\} \leq G.$$

Recall the Orbit-Stabiliser Theorem from IA Groups: There is a bijection $\operatorname{orb}_G(x) \leftrightarrow$ the set of left cosets of G_x in G. In particular if G is finite, then

$$|G| = |\operatorname{orb}_G(x)||G_x|.$$

This has lots of useful applications:

Example (Example of Orbit-Stabiliser)
Let G be the group of all symmetries of a cube, acting on the set of veretices X. We can reach any vertex from any other one, so $|\operatorname{orb}_G(x)| = 8$. Some basic geometry gives $|G_x| = 6$. Therefore $|G| = 48$.

Remark.

- $\ker \phi = \bigcap_{x\in X} G_x$ is called the kernel of the group action.
- The orbits partition X. We say the action is transitive if there is only one orbit.
- $G_{g*x} = gG_xg^{-1}$, so if $x, y \in X$ belong to the same orbit, then their stabilisers are conjugate.

Later on a lot of the proofs will involve picking a nice group action. So let's look at some examples of group actions.

(i) Let G act on itself by left multiplication, i.e $g * x = gx$. The kernel of this action is

$$\{g \in G : gx = x \quad \forall x \in G\} = e.$$

Thus G is injective into $\operatorname{Sym}(G)$. This proves Cayley's theorem:

Theorem 1.10 (Cayley's theorem)
Any finite group G is isomorphic to a subgroup of the symmetric group S_n for some n. (Take $n = |G|$.)

Proof. As above in (i). □

(ii) Let $H \leq G$; then G acts on G/H by left multiplication, i.e $g * xH = gxH$. This action is transitive (since $(x_2{x_1}^{-1})x_1H = x_2H$) with

$$\begin{aligned} G_{xH} &= \{g \in G : gxH = xH\} \\ &= \left\{g \in G : x^{-1}gx \in H\right\} \\ &= xHx^{-1} \end{aligned}$$

Thus $\ker(\phi) = \bigcap_{x\in G} xHx^{-1}$. This is the largest normal subgroup of G that is contained in H.

Theorem 1.11
Let G be a non-abelian simple group, and $H \leq G$ a subgroup of index $n > 1$. Then $n \geq 5$ and G is isomorphic to a subgroup of A_n.

Proof. Let G act on $X = G/H$ by left multiplication, and let $\phi : G \to \mathrm{Sym}(X)$ be the associated permutation representation. As G is simple, $\ker(\phi) = e$ or G. If $\ker(\phi) = G$, then $\mathrm{Im}(\phi) = e$. This is a contradiction since G acts transitively on X and $|X| > 1$. Thus $\ker(\phi) = e$ and $G \cong \mathrm{Im}(\phi) \leq S_n$.

Since $G \leq S_n$ and $A_n \trianglelefteq S_n$, the second isomorphism theorem gives $G \cap A_n \trianglelefteq G$ and $G/(G \cap A_n) \cong GA_n/A_n \leq S_n/A_n \cong C_2$. Since G is simple, $G \cap A_n = e$ (this is impossible as $G \leq C_2$ but G isn't abelian) or G. Thus $G \leq A_n$. Finally, if $n \leq 4$, then A_n has no non-abelian simple subgroups. $\square$

(iii) Let G act on itself by conjugation, i.e $g * x = gxg^{-1}$. We define the conjugacy class of $x \in G$ to be

$$\mathrm{ccl}_G(x) = \mathrm{orb}_G(x) = \left\{gxg^{-1} \in G : \quad g \in G\right\}.$$

We also define the centraliser of x by

$$C_G(x) = G_x = \{g \in G : \quad gx = xg\} \leq G.$$

We define the centre of G by

$$Z(G) = \mathrm{Ker}(\phi) = \{g \in G : \quad gx = xg \ \forall x \in G\}.$$

Note that the $\phi(g) : G \to G$ given by $h \mapsto ghg^{-1}$ satisfies

$$\phi(g)(h_1h_2) = gh_1h_2g^{-1} = gh_1g^{-1}gh_2g^{-1} = \phi(g)(h_1)\phi(g)(h_2).$$

Thus $\phi(g)$ is a group homomorphism, and also a bijection i.e $\phi(g)$ is an isomorphism.

Definition (Automorphism)
$\mathrm{Aut}(G) = \{$ group isomorphisms $\zeta : G \to G\}$. Then $\mathrm{Aut}(G) \leq \mathrm{Sym}(G)$ and $\phi : G \to \mathrm{Sym}(G)$ has image in $\mathrm{Aut}(G)$.

(iv) Let X be the set of all subgroups of G. Then G acts on X by conjugation, i.e $g * H = gHg^{-1}$. The stabiliser of H is

$$\left\{g \in G : \quad gHg^{-1} = H\right\} = N_G(H).$$

This is also called the normaliser of H in G, and is the largest subgroup of G containing H as a normal subgroup. In particular,

$$H \trianglelefteq G \iff N_G(H) = G.$$

1.4 Alternating groups

From IA Groups, we know that elements in S_n are conjugate iff they have the same cycle type. For example, in S_5, we have the following:

Cycle type	Number of elements	Sign
id	1	+1
$(**)$	10	-1
$(**)(**)$	15	+1
$(***)$	20	+1
$(**)(***)$	20	-1
$(****)$	30	-1
$(*****)$	24	+1
Total:	$120{=}5!{=}\lvert S_5\rvert$	

Let $g \in A_n$. Then $C_{A_n}(g) = C_{S_n}(g) \cap A_n$. We effectively have two cases:

- If there exists an odd permutation commuting with g, then $|C_{A_n}(g)| = \frac{1}{2}|C_{S_n}(g)|$ and by Orbit-Stabiliser, $|\operatorname{ccl}_{A_n}(g)| = |\operatorname{ccl}_{S_n}(g)|$.
- Otherwise, $|C_{A_n}(g)| = |C_{S_n}(g)|$ and by Orbit-Stabiliser, $|\operatorname{ccl}_{A_n}(g)| = \frac{1}{2}|\operatorname{ccl}_{S_n}(g)|$.

Example (Conjugacy classes of A_5)
If we take $n = 5$, then first consider the element $(12)(34)$, which commutes with (12). Also, (123) commutes with (45).

But if we take $g = (12345)$, then $h \in C_{S_5}(g)$ means

$$\begin{aligned}(12345) &= h(12345)h^{-1}\\ &= (h(1)h(2)h(3)h(4)h(5)) \implies h \in \langle g\rangle \leq A_5.\end{aligned}$$

In this case, the conjugacy class does split.

Thus A_5 has conjugacy classes of sizes 1,15,20,12,12.

Proposition 1.12
A_5 is simple.

Proof. If $H \trianglelefteq A_5$, then H is a union of conjugacy classes. Therefore

$$|H| = 1 + 15a + 20b + 12c \quad \text{for some } a, b \in \{0,1\} \text{ and } c \in \{0,1,2\}\,.$$

Since $H|60$, this implies $H = 1$ or 60, i.e A_5 is simple. □

Now we move on to a more general statement about A_n being simple. Before we can do that, we will need some lemmas for the proof.

Lemma 1.13
A_n is generated by 3-cycles.

Proof. Each $\sigma \in A_n$ is a product of an even number of transpositions. Thus it suffices to write the product of any two transpositions as a product of 3-cycles. For

a, b, c, d distinct, we can have

$$\begin{cases} (ab)(bc) = (abc) \\ (ab)(cd) = (acb)(acd). \end{cases}$$

□

Lemma 1.14
If $n \geq 5$, then all 3-cycles in A_n are conjugate.

Proof. We claim that every 3-cycle is conjugate to (123). Indeed, if (abc) is a 3-cycle, then $(abc) = \sigma(abc)\sigma^{-1}$ for some $\sigma \in S_n$. If $\sigma \notin A_n$, then replace σ by $\sigma(45)$ (using the fact that $n \geq 5$). □

Theorem 1.15
A_n is simple for all $n \geq 5$.

Proof. Let $e \neq N \trianglelefteq A_n$. Suffices to show that N contains a 3-cycle, since by 1.13 and 1.14 we then have $N = A_n$.

Take $e \neq \sigma \in N$ and write σ in its disjoint cycle decomposition. Consider the cases:

1. σ contains a cycle of length $r \geq 4$. WLOG $\sigma = (123\dots r)\tau$, where τ is some product of cycles that fixes $1, 2, \dots, r$.

   Let $\delta = (123)$. Then consider the element

   $$\underbrace{\sigma^{-1}}_{\in N}\underbrace{\delta^{-1}\sigma\delta}_{\in N} = (r\dots 21)(132)(12\dots r)(123) = (23r) \in N.$$

   Note τ gets cancelled as it fixes 1 to r. Therefore N contains a 3-cycle.

2. σ contains two 3-cycles. WLOG $\sigma = (123)(456)\tau$. Let $b = (124)$. Then

   $$\sigma^{-1}\delta^{-1}\sigma\delta = (132)(465)(142)(123)(456)(124) = (12436) \in N.$$

   Then we are back to case 1, so N contains a 3-cycle.

3. σ contains two 2-cycles. WLOG $\sigma = (12)(34)\tau$. Let $\delta = (123)$ and consider

   $$\sigma^{-1}\delta^{-1}\sigma\delta = (12)(34)(132)(12)(34)(123) = (14)(23) = \pi \in N.$$

   Let $\varepsilon = (235)$. Then

   $$\pi^{-1}\varepsilon^{-1}\pi\varepsilon = (14)(23)(253)(14)(23)(235) = (253).$$

   Thus N contains a 3-cycle.

We now consider the remaining cases:

1. Cycle type $(*\ *) \implies \sigma \notin A_n$.
2. Cycle type $(***) \implies \sigma$ is a 3-cycle.
3. Cycle type $(**)(***) \implies \sigma \notin A_n$.

This concludes the proof. $\square$

1.5 p-groups and p-subgroups

Definition (p-group)
Let p be a prime. A finite group G is a p-group if $|G| = p^n$, $n \geq 1$.

Theorem 1.16
If G is a p-group, then $Z(G) \neq 1$.

Proof. For $g \in G$, we have $|\operatorname{ccl}_G(g)||C_G(g)| = |G| = p^n$. So each conjugacy class must have size that is a power of p. Since G is a dijoint union of conjugacy classes,

$$\begin{aligned} |G| \equiv (\text{number of conjugacy classes of size 1}) &\mod p \\ \implies 0 \equiv |Z(G)| &\mod p \\ \implies Z(G) \neq 1. \end{aligned}$$

We have used the fact that the conjugacy classes of size 1 are precisely the elements of $Z(G)$:

$$g \in Z(G) \iff x^{-1}gx = g \ \forall x \in G \iff \operatorname{ccl}_G(g) = \{g\}.$$

□

Corollary 1.17
The only simple p-group is C_p.

Proof. Let G be a simple p-group. Since $Z(G) \trianglelefteq G$, we have $Z(G) = 1$ or G. By 1.16, we must have $Z(G) = G$. Therefore G is abelian. Conclude by Lemma 1.7. □

Corollary 1.18
Let G be a p-group of order p^n. Then G has a subgroup of order p^r for $0 \leq r \leq n$.

Proof. By Lemma 1.8, G has a composition series

$$1 = G_0 \trianglelefteq G_1 \trianglelefteq \ldots \trianglelefteq G_{m-1} \trianglelefteq G_m = G.$$

with each G_i/G_{i-1} simple. Since G is a p-group, all of the G_i/G_{i-1} must be p-groups. Therefore have by Proposition 1.17 that $G_i/G_{p-1} \cong C_p$.

Thus $|G_i| = p^i$ for all $0 \leq i \leq m$ and $m = n$. □

Lemma 1.19
For G a group, if $G/Z(G)$ is cyclic, then G is abelian (so in fact $G/Z(G) = 1$).

Proof. Let $gZ(G)$ be a generator for $G/Z(G)$. Then each coset is of the form $g^r Z(G)$ for some $r \in \mathbb{Z}$. Thus

$$G = \{g^r z : \ r \in \mathbb{Z}, z \in Z(G)\} .$$

We now check that elements in this group always commute:

$$\begin{aligned} g^{r_1} z_1 g^{r_2} z_2 &= g^{r_1+r_2} z_1 z_2 \qquad \text{since } z \in Z(G) \\ &= g^{r_1+r_2} z_2 z_1 \\ &= g^{r_2} z_2 g^{r_1} z_1 . \end{aligned}$$

Therefore G is abelian. □

Corollary 1.20

If $|G| = p^2$ then G is abelian.

Proof. We know that $|Z(G)| \in \{1, p, p^2\}$. We can't have 1 by 1.16. If we have p, then $|G/Z(G)| = p$ and therefore is cyclic. Now applying 1.19 we have that G is abelian. If we have p^2 then $Z(G) = G$ so G is abelian. □

1.6 The Sylow theorems

Theorem 1.21 (Sylow theorems)
Let G be a finite group of order $p^a m$ where p is a prime with $p \nmid m$. Then

(i) The set $\mathrm{Syl}_p(G) = \{P \leq G : \ |P| = p^a\}$ of Sylow p-subgroups is non-empty.

(ii) All elements of $\mathrm{Syl}_p(G)$ are conjugate.

(iii) We define $n_p = |\mathrm{Syl}_p(G)|$. This satisfies $n_p \equiv 1 \mod p$ and $n_p| \mid |G|$ (and so $n_p|m$)

Proof. (i) Let Ω be the set of all **subsets** of G of order p^a. We know that

$$|\Omega| = \binom{p^a m}{p^a} = \left(\frac{p^a m}{p^a}\right)\left(\frac{p^a m - 1}{p^a - 1}\right) \cdots \left(\frac{p^a m - p^a + 1}{1}\right).$$

For $0 \leq k \leq p^a$, the number $p^a m - k$ and $p^a - k$ are divisible by the same power of p.

Therefore $|\Omega|$ is coprime to p. $(\dagger)$

Let G act on Ω by left-multiplication, i.e for $g \in G$ and $x \in \Omega$ we have

$$g * X = \{gx : \ x \in X\} \in \Omega.$$

For any $X \in \Omega$ we have

$$|G_x| \ |\mathrm{orb}_G(X)| = |G| = p^a m.$$

By $(\dagger)$, there exists X such that $|\mathrm{orb}_G(X)|$ is coprime to p. This is because the orbits give a partition of Ω, and so they can't all divide p. Thus

$$p^a| \ |G_x \tag{1}$$

On the other hand, if $g \in G$ and $x \in X$, then $g \in (gx^{-1}) * X$ and hence

$$\begin{aligned} G = \bigcup_{g \in G} g * X &= \bigcup_{y \in \mathrm{orb}_G(X)} Y \\ \implies |G| &\leq |\mathrm{orb}_G| \cdot |X| \quad \text{since } |Y| = |X| \\ \implies |G_x| &= \frac{|G|}{|\mathrm{orb}_G(X)| = |X| = p^a}. \end{aligned} \tag{2}$$

By 1 and 2, $|G_X| = p^a$, i.e $G_x \in \mathrm{Syl}_p(G)$.

(ii) We prove a stronger result. We claim that if $P \in \mathrm{Syl}_p(G)$ and $Q \leq G$ is a p-subgroup, then $Q \leq gP^{-1}g^{-1}$ for some $g \in G$.

The proof is as follows: let Q act on the set of left cosets (not a group!) G/P by left multiplication, i.e $q * gP = qgP$. By Orbit-Stabiliser, each orbit has size $|Q|$, so either 1 or a multiple of p. Since $|G/P| = m$ by definition, and m is coprime to p, there must exist some orbit of size 1, i.e

$$\exists g \in G : \quad qgP = gP \quad \forall q \in Q$$

$$\implies g^{-1}qg \in P \quad \forall q \in Q$$
$$\implies Q \leq gPg^{-1}.$$

So we are done.

(iii) Let G act on $\mathrm{Syl}_p(G)$ by conjugation. Sylow (ii) tells us this action is transitive. Thus orbit-stabiliser implies

$$n_p = |\,\mathrm{Syl}_p(G)| \mid |G|.$$

Now to show that $n_p \equiv 1 \mod p$, let $P \in \mathrm{Syl}_p(G)$ act on $\mathrm{Syl}_p(G)$ by conjugation. The orbits have size dividing $|P| = p^a$, so either 1 or a multiple of p. To show $n_p \equiv 1 \mod p$, it suffices to show that $\{P\}$ is the unique orbit of size 1.

If $\{Q\}$ is another orbit of size 1, then P normalises Q, i.e $P \leq N_G(Q)$. Now P, Q are both Sylow p-subgroups of $N_G(Q)$ since $|N_G(Q)| \leq p^a$. Thus by (ii), P and Q are conjugate in $N_G(Q)$ - but $Q \trianglelefteq N_G(Q)$, thus $P = Q$. This completes the proof.

□

Now let's look at an application of these theorems.

Corollary 1.22
If $n_p = 1$, then the unique Sylow p-subgroup is normal.

Proof. Let $g \in G$ and $P \in \mathrm{Syl}_p(G)$. Then $gPg^{-1} \in \mathrm{Syl}_p(G)$, and so $gPg^{-1} = P$. Thus $P \trianglelefteq G$. □

This is very useful to show groups of certain orders can't be simple.

Example
Let $|G| = 1000 = 2^3 \cdot 5^3$. Then $n_5 = 1 \mod 5$ and $n_5 | 8$ so $n_5 = 1$. Thus the unique Sylow 5-subgroup is normal and hence G is not simple.

Example
Let $|G| = 132 = 2^2 \cdot 3 \cdot 11$. We have that $n_{11} = 1 \mod 11$ and $n_{11}|12$. So $n_{11} = 1$ or 12. Suppose G is simple. Then $n_{11} \neq 1$ (otherwise the Sylow 11-subgroup is normal). Hence $n_{11} = 12$.

Now $n_3 \equiv 1 \mod 3$ and $n_3|44$. Thus $n_3 \in \{1, 4, 22\}$. But the case $n_3 = 1$ can't occur as before. Now suppose $n_3 = 4$. Then letting G act on $\mathrm{Syl}_3(G)$ by conjugation gives a group homomorphism $\phi : G \to S_4$. But then $\operatorname{Ker} \phi \trianglelefteq G \implies \operatorname{Ker} \phi = 1$ or G. But $\operatorname{Ker} \phi = 1$ would mean that G injects into S_4, which is a contradiction as $|G| = 132 > 24 = |S_4|$, and $\operatorname{Ker} \phi = G$ would be a contradiction to Sylow (ii).

Thus $n_3 = 22$ and $n_{11} = 12$. Thus G has $22 \cdot (3 - 1) = 44$ elements of order 3 and $120 = 12 \cdot (11 - 1)$ elements of order 11. But $44 + 120 > 132 = |G|$ which is a contradiction. Hence G is not simple.

1.7 Matrix groups

Matrix groups provide a wealth of examples of finite groups, and are crucial in the classification of finite simple groups. First we will recap a few groups we've seen before in IA Groups.

For a field F, let $GL_n(F)$ be the set of $n \times n$ invertible matrices over F.

This contains the subgroup $SL_n = \mathrm{Ker}(GL_n(F) \xrightarrow{\det} F^\times)$. Here $F^\times = F \setminus \{0\}$.

Let $Z \trianglelefteq GL_n(F)$ be the normal subgroup of scalar multiples of I. This is in fact the centre of $GL_n(F)$, but we won't prove this in the course since the proof is pretty involved.

Definition

We define the projective general linear group by

$$PGL_n(F) = GL_n(F)/Z,$$

and the projective special linear group by

$$PGL_n(F) = \frac{SL_n(F)}{Z \cap SL_n(F)} \cong \frac{Z \cdot SL_n(F)}{Z} \leq PGL_n(F) \quad \text{by 2nd isom. theorem.}$$

Example

Consider $G = GL_n(\mathbb{Z}/p\mathbb{Z})$. A list of n vectors in $(\mathbb{Z}/p\mathbb{Z})^n$ are the columns of some $A \in G$ iff they are linearly independent. Thus

$$\begin{aligned} |G| &= \underbrace{(p^n - 1)}_{\text{1st col.}} \underbrace{(p^n - p)}_{\text{2nd col.}} \underbrace{(p^n - p^2)}_{\text{3rd col.}} \cdots \underbrace{(p^n - p^{n-1})}_{\text{last col.}} \\ &= p^{1+2+\ldots+n-1}(p^{n-1} - 1)(p^n - 1)\ldots(p-1) \\ &= p^{n(n-1)/2} \prod_{i=1}^{n} (p^i - 1). \end{aligned}$$

So the Sylow p-subgroups have size $p^{n(n-1)/2}$. Let

$$U = \left\{ \begin{pmatrix} 1 & * & * & * \\ & 1 & * & * \\ & & \ddots & * \\ & & & 1 \end{pmatrix} \right\} \leq G$$

be the set of upper triangular matrices with 1's on the diagonal. Then $U \in \mathrm{Syl}_p(G)$, since it has $n(n-1)/2$ entries and each can take p values.

Just as $PGL_2(\mathbb{C})$ acts on $\mathbb{C} \cup \{\infty\}$ via Möbius transformations, $PGL_2(\mathbb{Z}/p\mathbb{Z})$ acts on $\mathbb{Z}/p\mathbb{Z} \cup \{\infty\}$ via

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \frac{az + b}{cz + d}.$$

Since scalar matrices act trivially, we obtain an action of $PGL_2(\mathbb{Z}/p\mathbb{Z})$.

Lemma 1.23
The permutation representation $PGL_2(\mathbb{Z}/p\mathbb{Z}) \to S_{p+1}$ is injective (in fact an isomorphism if $p = 2$ or 3).

Proof. Suppose $\frac{az+b}{cz+d} = z$ for all $z \in \mathbb{Z}/p\mathbb{Z} \cup \{\infty\}$.

- Setting $z = 0$ gives $b = 0$.
- Setting $z = \infty$ gives $c = 0$.
- Setting $z = 1$ gives $a = d$.

So $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is a scalar matrix, hence it is trivial in $PGL_2(\mathbb{Z}/p\mathbb{Z})$. □

Lemma 1.24
If p is an odd prime, then

$$|PSL_2(\mathbb{Z}/p\mathbb{Z})| = p(p-1)(p+1)/2.$$

Proof. By the previous example, $|GL_2(\mathbb{Z}/p\mathbb{Z})| = p(p^2-1)(p-1)$. The homomorphism $GL_2(\mathbb{Z}/p\mathbb{Z}) \xrightarrow{\det} (\mathbb{Z}/p\mathbb{Z})^\times$ is surjective.

Thus $|SL_2(\mathbb{Z}/p\mathbb{Z})| = |GL_2(\mathbb{Z}/p\mathbb{Z})|/(p-1) = p(p-1)(p+1)$. If $\begin{pmatrix} \lambda & 0 \\ \lambda & 0 \end{pmatrix} \in SL_2(\mathbb{Z}/p\mathbb{Z})$, then $\lambda^2 \equiv 1 \mod p \implies \lambda \equiv \pm 1 \mod p$ (since p is prime).

Thus $Z \cap SL_2(\mathbb{Z}/p\mathbb{Z}) = \{\pm I\}$ which are distinct since $p > 2$. Therefore

$$|PSL_2(\mathbb{Z}/p\mathbb{Z})| = \frac{1}{2}|SL_2(\mathbb{Z}/p\mathbb{Z})| = p(p-1)(p+1)/2.$$

□

Example
Let $G = PSL_2(\mathbb{Z}/5\mathbb{Z})$. Then $|G| = \frac{4\cdot 5\cdot 6}{2} = 60$.

Let G act on $\mathbb{Z}/5\mathbb{Z} \cup \{\infty\}$ via Mobius transformations. By Lemma 1.23, the permutation representation

$$\phi : G \to \mathrm{Sym}(\{0,1,2,3,4\} \cup \infty) \cong S_6 \quad \text{is injective..}$$

Claim. $\mathrm{Im}(\phi) \leq A_6$, i.e $\psi : G \to S_6 \xrightarrow{\mathrm{sgn}} \{\pm 1\}$ is trivial.

Proof. Let $h \in G$ have order $2^n m$, m odd. If $\psi(h^m) = 1$, then $\psi(h)^m = 1 \implies \psi(h) = 1$. So suffices to show $\psi(g) = 1$ for all $g \in G$ with order a power of 2. But we know that every such g belongs to a Sylow 2-subgroup. It then suffices to show $\psi(H) = 1$, for H a Sylow 2-subgroup (since $\operatorname{Ker}\psi$ is normal and all Sylow 2-subgroups are conjugate). Take

$$H = \left\langle \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} (\pm I), \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} (\pm I) \right\rangle.$$

We compute that $\phi \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} = (14)(23)$, since it acts as $z \mapsto -z$. Also, $\phi \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = (0\infty)(14)$, since it acts as $z \mapsto -\frac{1}{z}$. Thus $\psi(H) = 1$. This proves the claim. $\square$

See Example Sheet 1 Q14 for a similar result: If $G \leq A_6$ and $|G| = 60$, then $G \cong A_5$.

Remarks. (Not proved in this course)

- $PSL_n(\mathbb{Z}/p\mathbb{Z})$ is a simple group for all $n \geq 2$ and p a prime, e.g $(n, p) = (2, 2), (2, 3)$ (finite groups of Lie type).
- The smallest non-abelian simple groups are $A_5 \cong PSL_2(\mathbb{Z}/5\mathbb{Z})$ of order 60, and $PSL_2(\mathbb{Z}/7\mathbb{Z})$ of order 168.

1.8 Finite abelian groups

We now investigate finite abelian groups, which we can actually characterise very effectively.

Lemma 1.25
If $m, n \in \mathbb{N}$ are coprime, then $C_m \times C_n \cong C_{mn}$.

Proof. Let g and h be generators of C_n and C_m. Then $(g, h) \in C_m \times C_n$ and $(g, h)^r = (g^r, h^r)$. Hence

$$\begin{aligned}(g,h)^r = 1 &\iff m|r \text{ and } n|r\\ &\iff mn|r \quad \text{as } m, n \text{ coprime.}\end{aligned}$$

Thus (g, h) has order $mn = |C_m \times C_n|$. So

$$C_m \times C_n \cong C_{mn} \cong \langle (g, h) \rangle.$$

□

Corollary 1.26
Let G be a finite abelian group. Then

$$G \cong C_{n_1} \times C_{n_2} \times \ldots \times C_{n_k},$$

where n_i are prime powers.

Proof. If $n = p_1^{a_1} \ldots p_r^{a_r}$, where $p_1, \ldots, p_r$ are distinct primes, then we just apply our previous lemma inductively to get

$$C_n \cong C_{p_1^{a_1}} \times C_{p_2^{a_2}} \times \ldots \times C_{p_r^{a_r}}.$$

□

Theorem 1.27
Every finite abelian group G is isomorphic to a product of cyclic groups.

Proof. Immediate by applying Corollary 1.26. □

Remark. Note such an isomorphism is not unique.

Theorem 1.28
Let G be a finite abelian group. Then $G \cong C_{d_1} \times C_{d_2} \times \ldots C_{d_t}$, for some $d_1|d_2|\ldots|d_t$ (they are successively divisible).

Proof. Omitted for now (see Modules section). □

This almost immediately shows that finite abelian groups are pretty easy to work with. Let's use our results to compute what the abelian groups of various orders are in an example.

Example (i) The abelian groups of order 8 are

$$C_8, C_2 \times C_4, C_2 \times C_2 \times C2.$$

(ii) For the abelian groups of order 12, using 1.27 we get that they are

$$C_2 \times C_2 \times C_3, C_4 \times C_3.$$

Using 1.28 we get

$$C_2 \times C_6, C_{12}.$$

This isn't a problem as these are pairwise isomorphic.

Definition (Exponent of a group)
The **exponent** of a group G is the least integer $n \geq 1$ such that $g^n = 1$ for all $g \in G$, i.e the LCM of the orders of the elements of G. For example, A_4 has exponent 6.

Corollary 1.29
Every finite abelian group contains an element whose order is the exponent of the group.

Proof. If $G \cong C_{d_1} \times C_{d_2} \times \ldots C_{d_t}$ with $d_1|d_2|\ldots|d_t$, then every $g \in G$ has order dividing d_t, and $h \in C_{d_t}$ of order d_t, then $(1, 1, \ldots, h) \in G$ has order d_t. Thus G has exponent d_t. $\square$

2 Rings

2.1 Definitions and examples

Definition (Ring)
A **ring** is a triple $(R, +, \cdot)$ consisting of a set R and two binary operations $+ : R \times R \to R$, $\cdot : R \times R \to R$ satisfying:

(i) Addition: $(R, +)$ is an abelian group, with identity element 0.

(ii) Multiplication: the operation $\cdot$ is associative, and has an identity 1.

(iii) Distributivity: $x \cdot (y + z) = x \cdot y + x \cdot z$ and $(y + z) \cdot x = y \cdot x + z \cdot x$ for all $x, y, z \in R$.

We say R is a **commutative ring** if $x \cdot y = y \cdot x$ for all $x, y \in R$. In this course, we will only consider commutative rings.

Remarks.

(i) As in the case of groups, we should check closure!

(ii) For any $x \in R$, write $-x$ for the inverse of x under $+$, and abbreviate $x + (-y)$ as $x - y$.

(iii) $0 \cdot x = (0 + 0) \cdot x = 0 = 0 \cdot x + 0 \cdot x$. Therefore $0 \cdot x = 0$ for all $x \in R$.

(iv) $0 = 0 \cdot x = (1 - 1) \cdot x = 1 \cdot x + (-1) \cdot x = x + (-1) \cdot x$. So $(-1) \cdot x = -x$ for all $x \in R$.

Definition (Subring)
A subset $S \subset R$ is a subring (written $S \leq R$) if it is a ring under $+$ and $\cdot$, with the same identity elements 0 and 1.

Example

(i) $\mathbb{Z} \leq \mathbb{Q} \leq \mathbb{R} \leq \mathbb{C}$ are all rings.

(ii) $\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\} \leq \mathbb{C}$ is a ring. This is called the ring of Gaussian integers.

(iii) $\mathbb{Q}[\sqrt{2}] = \left\{a + b\sqrt{2} : a, b \in Q\right\} \leq \mathbb{R}$.

(iv) $\mathbb{Z}/n\mathbb{Z} =$ (integers modulo n)

(v) If R, S are rings, we define the **product ring** to be the set $R \times S$ via

$$\begin{aligned} (r_1, s_1) + (r_2, s_2) &= (r_1 + r_2, s_1 + s_2) \\ (r_1, s_1) \cdot (r_2, s_2) &= (r_1 \cdot r_2, s_1 \cdot s_2) \\ 0_{R \times S} = (0_R, 0_S) \text{ and } 1_{R \times S} &= (1_R, 1_S). \end{aligned}$$

(vi) For R a ring, a **polynomial** f over R is an expression

$$f = a_0 + a_1 X + a_2 X^2 + a_n X^n, \quad a_i \in R.$$

"X" is just a formal symbol (i.e our definition of a polynomial just means some finite sequence in R). The degree of f is the largest $n \in \mathbb{N}$ s.t $a_n \neq 0$. We write $R[X]$ for the set of all polynomials over R.

Remark. We say f is a **monic** polynomial if $a_n = 1_R$.

If $g = b_0 + b_1 X + \ldots + b_m X^m$ is another polynomial, set

$$f + g = \sum_i (a_i + b_i) X^i$$

$$f \cdot g = \sum_i (\sum_{j=0}^{i} a_j b_{i-j}) X^i$$

Then $R[X]$ is a ring with identities 0_R and 1_R, which are constant polynomials. We identify R with the subring of constant polynomials (i.e $a_i = 0$ for all $i > 0$)

Definition (Unit)
An element $r \in R$ is a **unit** if it has an inverse under multiplication, i.e $\exists s \in R : s \cdot r = 1$.

The units in R form an abelian group $(R^\times, \cdot)$ under multiplication. For example, $\mathbb{Z}^\times = \{\pm 1\}$ and $\mathbb{Q}^\times = \mathbb{Q} \backslash 0$.

Definition (Field)
A **field** is a ring with $0 \neq 1$, such that every non-zero element is a unit. It's "a ring where you can divide". Examples of rings are $\mathbb{Q}$ and $\mathbb{Z}/p\mathbb{Z}$ for p prime.

Remark. If R is a ring where $0 = 1$, then for all $x \in R$ we have

$$x = 1 \cdot x = 0 \cdot x = 0.$$

So $R = \{0\}$ is the trivial ring. This is why we stipulate that $0 \neq 1$ for a ring to be a field.

Proposition 2.1 (Euclidean algorithm for rings)
Let $f, g \in R[X]$. Suppose the leading coefficient of g is a unit. Then there exist $q, r \in R[X]$ s.t.

$$f(X) = q(X) g(X) + r(X) \quad \text{where } \deg(r) < \deg(g).$$

Proof. Induction on $n = \deg(f)$. Write

$$f(X) = a_n X^n + a_{n-1}X^{n-1} + \ldots + a_0, \quad a_n \neq 0$$
$$g(X) = b_n X^n + b_{n-1}X^{n-1} + \ldots + b_0, \quad b_n \neq 0$$

If $n < m$, then let $q = 0$ and $r = f$. Done. Otherwise, we have $n \geq m$ and we set

$$f_1(X) = f(X) - a_n {b_m}^{-1} g(X) X^{n-m}.$$

The coefficient of X^n is $a_n - a_n {b_m}^{-1} b_m = 0$. Thus $\deg(f_1) < n$.

By the inductive hypothesis, $\exists q_1, r \in R[X]$ such that $f_1(X) = q_1(X)g(X) + r(X)$ where $\deg(r) < \deg(g)$. Therefore

$$f(X) = \underbrace{q_1(X) + a_n {b_m}^{-1} X^{n-m}}_{q(X)} + r(X).$$

So we are done. □

Example

Let's look at further examples of rings.

(i) If R is a ring and X is a set, then the set of all functions $X \to R$ is a ring under pointwise operations:

$$(f + g)(x) = f(x) + g(x)$$
$$(f \cdot g)(x) = f(x) \cdot g(x)$$

Further interesting examples appear as subrings, for example the ring of all C^1 functions from $\mathbb{R} \to \mathbb{R}$.

(ii) Power series ring $R[[X]] = a_0 + a_1 X + a_2 X^2 + \ldots$, where the $a_i \in R$. We use the same operations $+$ and $\cdot$ as the polynomial ring. We could view this as an "infinite version" of the polynomial ring.

(iii) Laurent polynomials:

$$R[X, X^{-1}] = \left\{ \sum_{i \in \mathbb{Z}} a_i X^i : a_i \in R, \text{ and } a_i \text{ is non-zero for finitely many } i \right\}.$$

2.2 Homomorphisms, ideals and quotients

We now define a homomorphism for rings, which similarly to the case of groups preserves the structure of a ring.

Definition (Ring homomorphism)
Let R and S be rings. a function $\phi : R \to S$ is a ring homomorphism if

(i) $\phi(r_1 + r_2) = \phi(r_1) + \phi(r_2)$ for all $r_1, r_2 \in R$

(ii) $\phi(r_1 \cdot_R r_2) = \phi(r_1) \cdot_S \phi(r2)$ for all $r_1, r_2 \in R$

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

A ring homomorphism that is also a bijection is called an isomorphism.

The kernel of ϕ is $\text{Ker}(\phi) = \{r \in R : \phi(r) = 0\}$.

Lemma 2.2
A ring homomorphism $\phi : R \to S$ is injective iff $\text{Ker}(\phi) = 0_R$.

Proof. $\phi : (R, +) \to (S, +)$ is a group homomorphism. □

Definition (Ideal)
A subset $I \subset R$ is an **ideal**, written $I \trianglelefteq R$ if

(i) $(I, +)$ is a subgroup of $(R, +)$

(ii) If $r \in R$ and $x \in I$, then $rx \in I$.

We say I is proper if $I \neq R$.

The point of ideals is that they should arise as the kernels of ring homomorphisms; they are analogous to normal subgroups.

Lemma 2.3
If $\phi : R \to S$ is a ring homomorphism, then $\text{Ker}(\phi)$ is an ideal of R.

Proof. $\phi : (R, +) \to (S, +)$ is a group homomorphism, so $\text{Ker}\,\phi$ is a subgroup of $(R, +)$. If $r \in R$ and $x \in \text{Ker}\,\phi$, then

$$\phi(rx) = \phi(r)\phi(x) = \phi(r)0_S = 0_S \implies rx \in \text{Ker}\,\phi.$$

□

Remark. If I contains a unit u, then $1_R \in I$ (by multiplying by $-u$). Hence $I = R$ (multiplying any element in R by 1_R). Thus if I is a proper ideal, $1_R \notin I$, so I is not a subring of R.

Lemma 2.4
The ideals in $\mathbb{Z}$ are $n\mathbb{Z} = \{kn : \ k \in \mathbb{Z}\}$ for $n = 0, 1, \ldots$.

Proof. Certainly $n\mathbb{Z} \trianglelefteq \mathbb{Z}$. Let $I \trianglelefteq \mathbb{Z}$ be a non-zero ideal and n be the smallest positive integer in I. Then $n\mathbb{Z} \subseteq I$. If $m \in I$, then write $m = qn + r$ with $q, r \in \mathbb{Z}$ and $0 \le r \le n$. Then $r = m - qn \in I$. This contradicts the fact that n is the smallest positive integer in I unless $r = 0$. But then $m \in n\mathbb{Z}$, i.e $I = n\mathbb{Z}$. □

Definition (Generated ideal)
For $a \in R$, write $(a) = \{ra : r \in R\} \trianglelefteq R$. This is the **ideal generated by** a. More generally, if $a_1, \dots, a_n \in R$, we write

$$(a_1, \dots, a_n) = \{r_1 a_1 + r_2 a_2 + \dots r_n a_n : r_i \in R\} \trianglelefteq R.$$

Definition (Principal ideal)
Let $I \trianglelefteq R$. We say I is a **principal ideal** if $I = (u)$ for some $a \in R$.

Now we show the converse of Lemma 2.3.

Theorem 2.5
If $I \trianglelefteq R$, then the set R/I of cosets of I in $(R, +)$ forms a ring (called the **quotient ring**) with the operations

$$\begin{aligned}(r_1 + I) + (r_2 + I) &= r_1 + r_2 + I \\ (r_1 + I) \cdot (r_2 + I) &= r_1 \cdot r_2 + I\end{aligned}$$

and $0_{R/I} = 0_R + I$, $1_{R/I} = 1_R + I$. Moreover, the map $R \to \frac{R}{I}$ with $r \mapsto r + I$ is a ring homomorphism (called the quotient map) with kernel I.

Proof. Already know $(R/I, +)$ is a group by our definition of the quotient group.

If $r_1 + I = r_1' + I$ and $r_2 + I = r_2' + I$, this means we can write $r_1' = r_1 + a_1$, $r_2' = r_2 + a_2$ for some $a_1, a_2 \in I$. Then

$$\begin{aligned}r_1' r_2' &= (r_1 + a_1)(r_2 + a_2) \\ &= r_1 r_2 + \underbrace{r_1 a_2 + r_2 a_1 + a_1 a_2}_{\in I}.\end{aligned}$$

Thus $r_1 r_2 + I = r_1' + r_2' + I$. The remaining properties for R/I follow from those properties for R. □

Example 2.6 (Examples of quotient rings)

1. $n\mathbb{Z} \trianglelefteq \mathbb{Z}$ is an ideal, with quotient ring $\mathbb{Z}/n\mathbb{Z}$. The cosets are $0 + n\mathbb{Z}, 1 + n\mathbb{Z}, \dots, (n-1) + n\mathbb{Z}$. Addition and multiplication are carried out modulo n.
2. Consider $(X) \trianglelefteq \mathbb{C}[X]$ be the ideal of polynomials with constant term 0.

If $f(X) = \underbrace{a_n X^n + \ldots + a_1 X}_{\in (X)} + a_0$ for $a_i \in \mathbb{C}$, then $f(X) + (X) = a_0 + (X)$.

There is a bijection $\mathbb{C}[X]/(X) \to \mathbb{C}$ given by

$$f(X) + (X) \mapsto f(0), \quad a + (X) \leftarrow a.$$

These maps are ring homomorphisms. Thus $\mathbb{C}[X]/(X) \cong \mathbb{C}$.

3. Consider $(X^2+1) \trianglelefteq \mathbb{R}[X]$. We want to understand

$$\mathbb{R}[X]/(X^2+1) = \left\{ f(X) + (X^2+1) : \ f(X) \in \mathbb{R}[X] \right\}.$$

By Proposition 2.1, $f(X) = g(X)(X^2+1) + r(X)$ with $\deg r < 2$, i.e. $r(X) = a + bX$ for $a, b \in \mathbb{R}$. Thus $\mathbb{R}[X]/(X^2+1) = \left\{ a + bX + (X^2+1) : \ a, b \in \mathbb{R} \right\}$. Now let's investigate if this representation is unique.

If $a+bX+(X^2+1) = a'+b'X+(X^2+1)$, then $(a-a')+(b-b')X = g(X)(X^2+1)$ for some $g \in R[X]$. Comparing degrees, we see $g(X) = 0$. Thus $a = a'$ and $b = b'$.

Consider the bijection $\mathbb{R}[X]/(X^2+1) \xrightarrow{\phi} \mathbb{C}$ defined by $a + bX + (X^2+1) \mapsto a + bi$. We show ϕ is a ring homomorphism. It preserves addition and maps $1 + (X^2+1)$ to 1.

$$\begin{aligned}
&\phi((a+bX) + (X^2+1)) \left(c + dX + \left(X^2+1\right)\right) \\
&= \phi\left((a+bX)(c+dX)\left(X^2+1\right)\right) \\
&= \phi\left(ac + (ad+bc)X + bd\left(X^2+1\right) - bd + \left(X^2+1\right)\right) \\
&= ac - bd + (ad-bc)\,i \\
&= (a+bi)(c+di) \\
&= \phi\left((a+bX) + (X^2+1)\right)\phi\left(c + dX + \left(X^2+1\right)\right).
\end{aligned}$$

Thus $\mathbb{R}[X]/\left(X^2+1\right) \cong \mathbb{C}$.

We move on to the isomorphism theorems for rings; they are closely related to those for groups.

Theorem 2.7 (First isomorphism theorem for rings)
Let $\phi : R \to S$ be a ring homomorphism, then $\operatorname{Ker}\phi \trianglelefteq R$, $\operatorname{Im}\phi \leq S$ and there exists an isomorphism $R/\operatorname{Ker}\phi \cong \operatorname{Im}\phi$.

Proof. We know that by Lemma 2.3 $\operatorname{Ker}\phi \trianglelefteq R$, and $\operatorname{Im}\phi$ is a subgroup of $(S, +)$. To check that the image is closed:

$$\phi(r_1 r_2) = \phi(r_1)\,\phi(r_2) \in \operatorname{Im}\phi.$$

We also note that $1_S = \phi(1_R)$. Thus $\operatorname{Im}\phi$ is a subring of S.

Let $K = \operatorname{Ker}\phi$. Define

$$\Phi : R/K \to \operatorname{Im}\phi, \quad r + K \mapsto \phi(r).$$

By the first isomorphism theorem for groups, this is well-defined, a bijection, and a group homomorphism under $+$. So it remains to prove that Φ preserves multiplication and that it maps the identity to the identity. We have $\Phi(1_R + K) = \phi(1_R) = I_S$, and

$$\begin{aligned}\Phi\left((r_1+K)(r_2+K)\right) &= \Phi(r_1r_2+K)\\ &= \phi(r_1r_2) = \phi(r_1)\phi(r_2)\\ &= \Phi(r_1+K)\Phi(r_2+K).\end{aligned}$$

Thus Φ is a ring isomorphism. ☐

Theorem 2.8 (Second isomorphism theorem for rings)
Let $R \leq S$ and $J \trianglelefteq S$. Then $R \cap J \trianglelefteq R$ and $R + J \equiv \{r + j : \ r \in R, j \in J\} \leq S$ and

$$R/(R\cap J) \cong (R+J)/J \leq S/J.$$

Proof. By the second isomorphism theorem for groups, $R+J$ is a subgroup of $(S,+)$ and we have $1_S = \underset{1_S}{\in R} + \underset{0}{J} \in R+J$. If $r_1, r_2 \in R$ and $j_1, j_2 \in J$:

$$(r_1+j_1)(r_2+j_2) = \underset{r_1r_2}{\in R} + \underset{r_1j_2+r_2j_1+r_2j_2}{\in J} \in R+J.$$

Therefore $R + J \leq S$. Define

$$\phi : R \to S/J, \quad r \mapsto r + J.$$

This is the composite of inclusion $R \subseteq$, and the quotient $S \mapsto S/J$, hence ϕ is a ring homomorphism. We can calculate

$$\begin{aligned}\operatorname{Ker}\phi &= \{r \in R : r + J = J\} = R \cap J \trianglelefteq R\\ \operatorname{Im}\phi &= \{r + J : r \in J\} = (R+J)/J \leq S/J.\end{aligned}$$

Now apply the first isomorphism theorem for rings, which concludes the proof. ☐

Remark. Let $I \trianglelefteq R$. There exists a bijection

$$\{\text{ Ideals in } R/I\} \leftrightarrow \{\text{ Ideals of } R \text{ containing } I\},$$

given by

$$K \mapsto \{r \in R : r + I \in K\}, \quad J/I \leftarrow J.$$

Theorem 2.9 (Third isomorphism theorem for rings)
Let $I \trianglelefteq R$, $J \trianglelefteq R$ with $I \subset S$. Then $J/I \trianglelefteq R/I$, and

$$(R/I)/(I/J) \cong R/J.$$

Proof. Consider $\phi : R/I \to R/J$ with $r + I \mapsto r + J$. This is well-defined surjective ring homomorphism by the third isomorphism theorem for groups (Exercise). Furthermore,

$$\operatorname{Ker}\phi = \{r + I : r \in J\} = J/I \trianglelefteq R/I.$$

Apply the first isomorphism theorem to finish the proof. □

Now we go back to Example 2.6, and reprove the last part much faster using the isomorphism theorem.

Example

There is a surjective ring homomorphism

$$\phi : \mathbb{R}[X] \to \mathbb{C}, \quad f(X) = \sum a_n X^n \mapsto f(i) = \sum a_n i^n.$$

Proposition 2.1 implies $\operatorname{Ker}\phi = \left(X^2+1\right)$. By the first isomorphism theorem,

$$\mathbb{R}[X]/\left(X^2+1\right) \cong \mathbb{C}.$$

If R is a ring, there exists a unique ring homomorphism $i : \mathbb{Z} \to R$, given by

$$\begin{aligned} 0 &\mapsto 0_R \\ 1 &\mapsto 1_R \\ n &\mapsto \underbrace{1_R + \ldots + 1_R}_{n \text{ times}}. \end{aligned}$$

Since $\operatorname{Ker}(i) \trianglelefteq \mathbb{Z}$, have $\operatorname{Ker}(i) = n\mathbb{Z}$ for some $n \in \{0,1,2,\ldots\}$. By the first isomorphism theorem,

$$\mathbb{Z}/n\mathbb{Z} \cong \operatorname{Im}(i) \leq R.$$

This motivates our next definition:

Definition (Characteristic)

We call n the **characteristic** of R. For example,

- $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ have characteristic 0.
- $\mathbb{Z}/p\mathbb{Z}, \mathbb{Z}/p\mathbb{Z}[X]$ has characteristic p.

2.3 Integral domains, maximal ideals and prime ideals

Definition (Integral domain/Zero-divisor)
An **integral domain** is a ring with $0 \neq 1$ such that for $a, b \in R$, $ab = 0 \implies a = 0$ or $b = 0$.

A **zero-divisor** in a ring R is a non-zero element $a \in R$ such that $ab = 0$ for some $0 \neq b \in R$. So an integral domain is a ring with no zero-divisors.

Example
1. All fields are integral domains (if $ab = 0$ with $b = 0$, multiply by b^{-1} to get $a = 0$).
2. Any subring of an integral domain is an integral domain, e.g $\mathbb{Z} \leq \mathbb{Q}$, $\mathbb{Z}[i] \leq \mathbb{C}$.
3. $\mathbb{Z} \times \mathbb{Z}$ is not an integral domain since $(1, 0) \cdot (0, 1) = (0, 0)$.

Lemma 2.10
If R is an integral domain, then $R[X]$ is an integral domain.

Proof. Write

$$\begin{aligned} f(X) &= a_m X^m + \ldots + a_1 X + a_0, \quad a_m \neq 0 \\ g(X) &= b_n X^m + \ldots + b_1 X + b_0, \quad b_n \neq 0. \end{aligned}$$

Then $f(X)\, g(X) = \underbrace{a_m b_n}_{\neq 0 \text{ as } R \text{ integral domain}} + \ldots$

and $\deg(fg) = m + n = \deg f + \deg g$ and $f \cdot g \neq 0$. □

Lemma 2.11
Let R be an integral domain and $0 \neq f \in R[X]$. Let the set of roots of f be

$$\mathcal{R} = \{a \in R : \ f(a) = 0\}.$$

Then $|\mathcal{R}| \leq \deg(f)$.

Proof. Uses Proposition 2.1; see Example Sheet 1. □

Theorem 2.12 (Subgroups of a field are cyclic)
Let F be a field. Then any finite subgroup $G \leq (F^\times, \cdot)$ is cyclic.

Proof. G is a finite abelian group. If G is not cyclic, then by Theorem , there exists $H \leq G$ such that $H \cong C_{d_1} \times C_{d_1}$ for some $d_1 \geq 2$. But then the polynomial

$$f(X) = X^{d_1} - 1 \in F[X]$$

has degree d_1, and $\geq d_1^2$ roots which contradicts our previous lemma. □

Example
This theorem tells us that the field $(\mathbb{Z}/p\mathbb{Z})^\times$ is cyclic.

Proposition 2.13
Any finite integral domain is a field.

Proof. Let R be a finite integral domain. Let $a \in R$, $a \neq 0$.

Consider the map $\phi : R \to R$: $x \mapsto ax$.

If $\phi(x) = \phi(y)$ then $a(x - y) = 0$. Since R is an integral domain and $a \neq 0$, this gives $x - y = 0$. So $x = y$. Thus ϕ is injective, and hence surjective since R is finite. Therefore

$$\exists b \in R : a \cdot b = 1 \implies a \text{ is a unit. } .$$

Thus R is a field. ☐

The next theorem tells us about the converse of this statement.

Theorem 2.14 (Field of fractions)
Let R be an integral domain. Then there exists a field F such that:

(i) $R \leq F$.

(ii) Every element of F can be written in the form $a \cdot b^{-1}$ where $a, b \in R$ and $b \neq 0$.

Remark. Condition (ii) guarantees our field is the "smallest" field with $R \leq F$. F is called the **field of fractions** of R.

Proof. Consider the set $S = \{(a, b) \in R : \ b \neq 0\}$, and the equivalence relation $\sim$ on S given by

$$(a, b) \sim (c, d) \iff ad - bc = 0.$$

Clearly this is reflexive and symmetric. For transitivity, if $(a, b) \sim (c, d) \sim (e, f)$, then

$$\begin{aligned} (ad)f = (bc)f = b(cf) = b(de) \\ \implies d(af - be) = 0 \end{aligned}$$

Since R is integral and $d \neq 0$, this gives $af - be = 0$, i.e $(a, b) \sim (e, f)$.

Let $F = S/\sim$ be the set of equivalence classes and (somewhat suggestively) write $\frac{a}{b}$ for $[(a, b)]$. Define

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

It can easily be checked that these operations are well-defined and make F into a ring with $0_F = \frac{0_R}{1_R}$ and $1_F = \frac{1_R}{1_R}$. If $\frac{a}{b} \neq 0_F$, then $a \neq 0_R$, and $\frac{a}{b} \cdot \frac{b}{a} = \frac{1_R}{1_R} = 1_F$. So F is a field. Now let's check the conditions we gave.

(i) Identify R with the subring $\left\{\frac{r}{1_R : r \in R}\right\} \leq F$.

(ii) $\frac{a}{b} = a \cdot b^{-1}$.

This concludes the proof. □

Remark. In this proof, we have essentially just mimicked the construction of the rationals. This leads on to our first example.

Example 1. $\mathbb{Z}$ is an integral domain with field of fractions $\mathbb{Q}$.

2. $\mathbb{C}[X]$ has field of fractions

$$\mathbb{C}(X) \equiv \text{ field of rational functions in } X.$$

Now we move on to discussing maximal and prime ideals, which we will use extensively in the next section.

Definition (Maximal ideal)
An ideal $I \trianglelefteq R$ is **maximal** if $I \neq R$ and if $I \leq J \trianglelefteq R$, then $J = I$ or R. This is the biggest *proper* ideal in R.

Lemma 2.15
A (non-zero) ring R is a field iff its only ideals are $\{0\}$ and R.

Proof. If R is a field and $I \trianglelefteq R$ is nontrivial, then I contains a unit and hence $I = R$.

Conversely, suppose the only ideals in R are $\{0\}$ and R. Take $x \in R$, $x \neq 0$, then the ideal (x) is non-zero, hence $(x) = R$. So $\exists y \in R : x \cdot y = 1$. So x is a unit. □

Proposition 2.16
Let $I \trianglelefteq R$ be an ideal. I is maximal iff R/I is a field.

Proof. R/I is a field $\iff$ I/I and R/I are the only ideals in R/I

$\iff$ I and R are the only ideals in R which contain I

$\iff$ $I \trianglelefteq R$ is maximal. □

Definition (Prime ideal)
An ideal $I \trianglelefteq R$ is prime if $I \neq R$ and whenever $a, b \in R$ with $ab \in I$, we have $a \in I$ or $b \in I$.

Example
The ideal $n\mathbb{Z} \trianglelefteq \mathbb{Z}$ is a prime ideal iff $n = 0$ or $n = p$ is prime.

Checking: If $ab \in p\mathbb{Z}$, then $p|ab$, so $p|a$ or $p|b$, i.e $a \in p\mathbb{Z}$ or $b \in p\mathbb{Z}$.

Conversely, if $n = uv$ with $u, v > 1$, then $uv \in n\mathbb{Z}$, but $u \notin n\mathbb{Z}$, but $u \notin n\mathbb{Z}$, $v \notin n\mathbb{Z}$.

Remark. The motivation of a prime ideal is to generalise the notion of prime numbers to arbitrary rings.

Proposition 2.17

Let $I \trianglelefteq R$ be an ideal. Then I is prime iff R/I is an integral domain.

Proof. I is prime

$$\begin{aligned} &\iff \text{Whenever } a, b \in R \text{ with } ab \in I, \text{ we have } a \in I \text{ or } b \in I. \\ &\iff \text{Whenever } a + I, b + I \in R/I \text{ with } (a + I)(b + I) = 0 + I, \\ &\qquad \text{we have } a + I = 0 + I \text{ or } b + I = 0 + I. \\ &\iff R/I \text{ is an integral domain.} \end{aligned}$$

□

Remark. Proposition 2.16 and Proposition 2.17 show that if I is a maximal ideal, it is also a prime ideal (since all fields are integral domains).

Remark. If $\text{char}(R) = n$, then $\mathbb{Z}/n\mathbb{Z} \leq R$. So if R is an integral domain, then $\mathbb{Z}/n\mathbb{Z}$ is an integral domain. This implies that $n\mathbb{Z} \trianglelefteq \mathbb{Z}$ is a prime ideal, so $n = 0$ or $n = p$ prime.

In particular, a field has characteristic 0 (and so contains $\mathbb{Q}$) or has characteristic p, in which case it contains $\mathbb{Z}/p\mathbb{Z} = \mathbb{F}_p$.

2.4 Factorisation in integral domains

In this section we investigate the following property: we know that in $\mathbb{N}$ we can factorise any number as a product of prime factors, and we want to know to which extent this is true in integral domains.

Remark. In this section we will always take R to be an integral domain.

Definition

We will quickly establish some key definitions that will help us work with the objects of this chapter.

(i) Recall that $a \in R$ is a unit if there exists $b \in R$ with $ab = 1$ (equivalently $(a) = R$).

(ii) $a \in R$ **divides** $b \in R$ (written $a|b$) if there exists $c \in R$ such that $b = ac$ (equivalently $(b) = (a)$).

(iii) $a, b \in R$ are **associates** if $a = bc$ for some unit $c \in R$.

(iv) $r \in R$ is **irreducible** if $r \neq 0$, r is not a unit, and $r = ab \implies a$ or b is a unit.

(v) $r \in R$ is **prime** if $r \neq 0$, r is not a unit and $r|ab \implies r|a$ or $r|b$.

Remark. These properties of an element depend on the ambient ring R: for example 2 is prime and irreducible in $\mathbb{Z}$, but not in $\mathbb{Q}$.

Another example is that $2X$ is irreducible in $\mathbb{Q}[X]$, but not in $\mathbb{Z}[X]$ (since then we can factor it as $2 \cdot X$, and 2 is a unit in $\mathbb{Z}$).

Now we want to chase some definitions to make sure our definitions of prime element and prime ideal agree.

Lemma 2.18

Consider $(r) \trianglelefteq R$. This is a prime ideal iff $r = 0$ or r is a prime element.

Proof. Suppose (r) is prime and $r \neq 0$. Since prime ideals are proper, $(r) \neq R$. Therefore r is not a unit. If $r|ab$, then $ab \in (r)$, so $a \in (r)$ or $b \in (r)$. Thus $r|a$ or $r|b$. So r is prime.

Conversely, $\{0\} \trianglelefteq R$ is a prime ideal since R is an integral domain. Let $r \in R$ be a prime. $(r) \neq R$ since $r \in R^\times$. If $ab \in (r)$ then $r|ab$.

$$\begin{aligned} &\implies r|a \text{ or } r|b \\ &\implies r \in (a) \text{ or } r \in (b), \text{ i.e } r \text{ is a prime ideal.} \end{aligned}$$

□

Lemma 2.19

If $r \in R$ is prime, then it is irreducible.

Proof. Since r is a prime, $r \neq 0$ and $r \notin R^\times$. Suppose $r = ab$. Then $r|ab$ so $r|a$ or

$r|b$. WLOG assume $r|a$, so $a = rc$ for some element $c \in R$.

Then $r = ab = rbc \implies r(1 - bc) = 0 \implies bc = 1$. This uses the fact that R is an integral domain and $r \neq 0$. So b is a unit. $\square$

The converse of this lemma does not hold in general.

Example 2.20
Let
$$R = \mathbb{Z}[\sqrt{-5}] = \left\{a + b\sqrt{-5} : \ a, b \in \mathbb{Z}\right\} \leq C, \quad R \cong \mathbb{Z}[X]/(X^2 + 5).$$

R is a subring of a field, so it's an integral domain. Define a function (called the **norm**)
$$N : R \to \mathbb{Z}_{\geq 0}, \quad a + b\sqrt{-5} \mapsto a^2 + 5b^2.$$

Note that $N(z_1 z_2) = N(z_1)N(z_2)$.

Claim. $R^\times = \{\pm 1\}$.

Proof. If $r \in R^\times$, i.e $rs = 1$ for some $s \in R$, then
$$N(r)N(s) = N(1) = 1 \implies N(r) = 1 \text{ WLOG.}$$
But the only integer solutions to $a^2 + 5b^2 = 1$ are $(a, b) = (1, 0)$ or $(-1, 0)$. $\square$

Claim. $2 \in R$ is irreducible.

Proof. Suppose $2 = rs$ where $r, s \in R$. We want to show one of these elements is a unit. Then
$$4 = N(2) = N(r)N(s).$$
Since $a^2 + 5b^2 = 2$ has no integer solutions, R has no elements of norm 2. Thus $N(r) = 1$ and $N(s) = 4$ WLOG. But if $N(r) = 1$, this implies that r is a unit. $\square$

Similarly 3, $1+\sqrt{-5}$, $1 - \sqrt{-5}$ are all irreducible (as there are no elements of norm 3). Now $(1 + \sqrt{-5})(1 - \sqrt{-5}) = 6 = 2 \cdot 3$. Thus $2|(1 + \sqrt{-5})(1 - \sqrt{-5})$, but $2 \nmid 1 + \sqrt{-5}$ or $1 - \sqrt{-5}$ (check by taking norms).

So 2 is irreducible but not prime in R.

Remarks. We now know that being irreducible does not imply being prime. Importantly, in the above example $2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5})$ gives two distinct factorisations into irreducibles. Since $R^\times = \{\pm 1\}$, these irreducibles are not associates.

This example tells us if we want to have a good definition of factorisation in a ring, we should impose some extra properties.

Definition (Principal ideal domain)
An integral domain R is a **principal ideal domain** (PID) if any ideal $I \trianglelefteq R$ is principal. In other words, any ideal $I \trianglelefteq R$ is generated by some element $r \in R$.

For example, $\mathbb{Z}$ is a PID by Lemma 2.4.

Proposition 2.21
Let R be a PID. Then every irreducible element of R is prime.

Proof. Let $r \in R$ be irreducible. We already know that $r \neq 0$ and r is not a unit. Assume $r|ab$ and $r \nmid a$.

Since R is a PID, $(a, r) = (d)$ for some $d \in R$. In particular, $r = cd$ for some $c \in R$. Since r is irreducible, either c or d is a unit. If c is a unit, then $(a, r) = (r)$. So $r|a$. Contradiction. If d is a unit, then $(a, r) = R$.

So there exists $s, t \in R$ such that $sa + tr = 1$. Then $b = sab + trb$. So $r|b$ since $r|ab$. Thus r is prime. □

Lemma 2.22
Let R be a PID, and $0 \neq r \in R$. Then r is irreducible iff (r) is a maximal ideal.

Proof. If r is irreducible, $r \notin R^\times$ so $(r) \neq R$. Suppose $(r) \subset J \subset R$ where $J \trianglelefteq R$. Since R is a PID, $J = (a)$ for some $a \in R$. Therefore $r = ab$ for some $b \in R$. Since r is irreducible, either $a \in R^\times \implies J = R$ or $b \in R^\times \implies J = (r)$. Thus (r) is maximal.

Conversely, suppose (r) is maximal. So $r \notin R^\times$. Suppose $r = ab$. Then $(r) \subset (a) \subset R$. Since (r) is maximal, either $(a) = (r)$ or R. If $(a) = (r)$, then a and r are associates and so b is a unit. If $(a) = R$ then a is a unit. Thus r is irreducible.

Remark. This proof of the converse implication actually holds for a general integral domain R; it doesn't assume R is a PID.

□

Remark. Let R be a PID, and take a nonzero $r \in R$. Then (r) is maximal iff r is irreducible, which is equivalent to r being prime. This is again equivalent to (r) being prime. We can use this to deduce that there exists a bijection

$$\{\text{ non-zero prime ideals }\} \leftrightarrow \{\text{ non-zero maximal ideals }\}.$$

This is a nice property of PIDs. We can cook up some examples of PIDs using th following definition.

Definition (Euclidean domain)
An integral domain is a **Euclidean domain** (ED) if there is a function $\phi : R \setminus \{0\} \to \mathbb{Z}_{\geq 0}$ (a Euclidean function) such that

(i) If $a|b$, then $\phi(a) \leq \phi(b)$.

(ii) If $a, b \in R$ with $b \neq 0$, then there exist $q, r \in R$ with $a = bq + r$ and either $r = 0$, or $\phi(r) < \phi(b)$.

Example
$\mathbb{Z}$ is a Euclidean domain with Euclidean function $\phi(n) = |n|$.

Proposition 2.23

If R is a Euclidean domain, then it is a principal ideal domain.

Proof. Let R have Euclidean function ϕ. Let $I \trianglelefteq R$ be non-zero. Choose $b \in I \setminus \{0\}$ with $\phi(b)$ minimal. Then $(b) \in I$.

For $a \in I$, write $a = bq + r$ with $q, r \in R$ and either $r = 0$ or $\phi(r) < \phi(b)$. Since $r = bq - a \in I$, cannot have $\phi(r) < \phi(b)$ as we chose b to be minimal. So $r = 0$, and $a = bq$. Hence $I = (b)$. $\square$

Remark. We only used property (ii) of Euclidean domains in this proof. We include property (i) in the definition as it allows us to describe the units in R as

$$R^\times = \{u \in R \setminus \{0\} : \ \phi(a) = \phi(1)\}.$$

Example

1. Let F be a field. $F[X]$ is an ED with Euclidean function $\phi(f) = \deg f$ for $f \in F[X]$. ϕ is Euclidean by Proposition 2.1.
2. $R = \mathbb{Z}[i]$ is an ED with Euclidean function

$$\phi(a + ib) = N(a + ib) = |a + ib|^2 = a^2 + b^2.$$

Since N is multiplicative, this tells us that property (i) is satisfied.

For property (ii), let $z_1, z_2 \in \mathbb{Z}[i]$ with $z_2 \neq 0$. Consider $\frac{z_1}{z_2} \in \mathbb{C}$. This has distance less than 1 from the nearest element of $\mathbb{Z}[i]$, i.e there exists $q \in \mathbb{Z}[i]$ such that

$$\left| \frac{z_1}{z_2} - q \right| < 1. \tag{†}$$

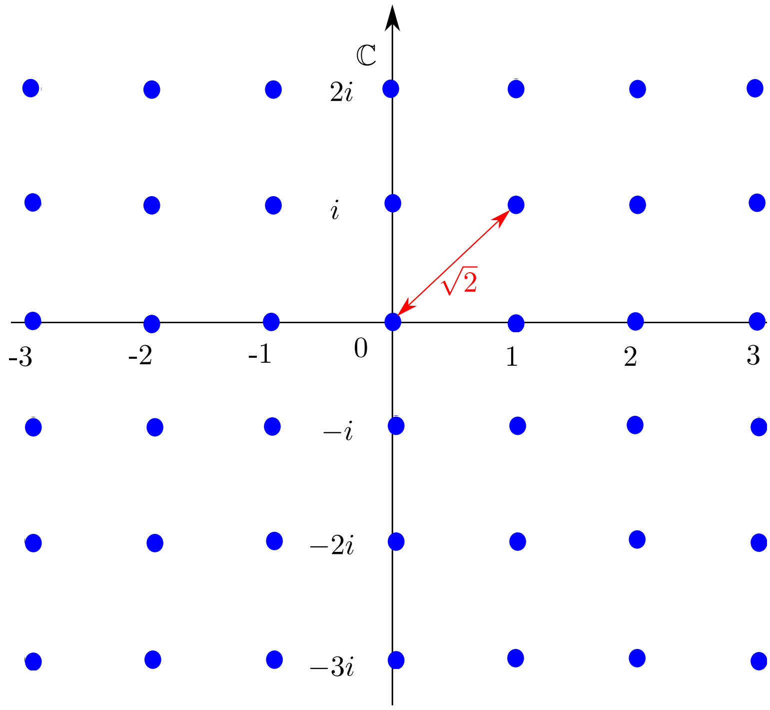

Figure 1: Every point in $\mathbb{C}$ is at most $1/\sqrt{2}$ away from a lattice point.

Set $r = z_1 - z_2 q \in \mathbb{Z}[i]$. Then $z_1 = qz_2 + r$ and

$$\phi(r) = |r|^2 = \underbrace{|z_1 - z_2 q|^2 < |z_2|^2}_{\text{by } (\dagger)} = \phi(z_2).$$

Thus Proposition 2.23 implies that both $\mathbb{Z}[i]$ and $F[X]$ for F a field are PIDs.

The first application we will see involves the construction of the minimal polynomial of a matrix.

Example
Let A be an $n \times n$ matrix over a field F. Let $I = \{f \in F[X] : f(A) = 0\}$.

Claim. I is an ideal in F.

Proof. If $f, g \in I$, then $(f - g)(A) = f(A) - g(A) = 0 \implies f - g \in I$. So $(I, +) \leq (F, +)$.

If $f \in F[X]$ and $g \in I$, then $(f \cdot g)(A) = f(A) \cdot g(A) = 0$. So $f \cdot g \in I$. □

So $I = (f)$ for some $f \in F[X]$ since $F[X]$ is a PID. Of course, this f isn't unique. We may assume f is monic by multiplying by a unit in F. Then for $g \in F[X]$,

$$g(A) = 0 \iff g \in I \iff g \in (f) \iff f | g.$$

So f is the minimal polynomial of A; we've derived it using rings (see IB Linear Algebra)

The next example will be of interest to those who want to study Part II Galois Theory.

Example (Field of order 8)
Let $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$. Now let $f(X) = X^3 + X + 1 \in \mathbb{F}_2[X]$. If $f(X) = g(X)h(X)$ with $g, h \in \mathbb{F}_2[X]$ and $\deg g, \deg h > 0$, then $\deg g + \deg h = 3$. Thus either $\deg g = 1$ or $\deg h = 1$, and so f has a root.

But $f(0) \neq 0$ and $f(1) \neq 0$ in $\mathbb{F}_2$. Thus f is irreducible. Since $\mathbb{F}_2$ is a PID, Lemma 2.22 implies $(f) \trianglelefteq \mathbb{F}_2[X]$ is a maximal ideal. Hence

$$\mathbb{F}_2[X]/(f) = \left\{aX^2 + bX + c : \ a, b, c \in \mathbb{F}_2\right\}$$

is a field of order 8.

Example 2.24
$\mathbb{Z}[X]$ is not a PID. Consider $I = (2, X)$. Then

$$\begin{aligned} I &= \{2f_1(X) + Xf_2(X) : \ f_1, f_2 \in \mathbb{Z}[X]\} \\ &= \{f \in \mathbb{Z}[X] : \ f(0) \text{ is even } \} \end{aligned}$$

Suppose $I = (f)$ for some $f \in \mathbb{Z}[X]$. Thus $2 = fg$ for some $g \in \mathbb{Z}[X]$. Thus $\deg f = \deg g = 0$, and $f \in \mathbb{Z}$. So $f = \pm 1$ or ± 2. Therefore $I = \mathbb{Z}[X]$ or $I = 2\mathbb{Z}$. Both these cases can't happen: if $I = \mathbb{Z}[X]$ we get a contradiction since $1 \notin I$, and in the other case we get a contradiction since $X \in I$.

Definition (Unique factorisation domain)
An integral domain is a **unique factorisation domain** (UFD) if

(i) Every non-zero, non-unit element is a product of irreducible elements.

(ii) If $p_1, \dots, p_m = q_1 q_2 \dots q_m$ where p_i, q_i are irreducible, then $m = n$ and we can reorder so that p_i is an associate of q_i for all $i = 1, \dots, n$.

Goal. We want to show that PIDs are UFDs, i.e that PIDs have a nice factorisation as we are familiar with from $\mathbb{N}$.

Proposition 2.25
Let R be an integral domain satisfying (i) in the definition of UFD. Then R is a UFD $\iff$ every irreducible is prime.

Proof. Suppose R satisfies (i). Suppose $p \in R$ is irreducible and $p | ab$. Then $ab = pc$ for some $c \in R$. Writing a, b, c as products of irreducibles, it follows from (i) that $p | a$ or $p | b$.

Conversely, suppose $p_1 \dots p_m = q_1 \dots q_n$ with each p_i and q_i irreducible. Since p_1 is prime and $p_1 | q_1 \dots q_n$, we have $p_1 | q_i$ for some i. Upon reordering, may assume that $p_1 | q_1$, i.e $q_1 = p_1 u$ for some $u \in R$. But q_1 is irreducible, and p_1 is irreducible, so u has to be a unit. Thus p_1 and q_1 are associates. So we can cancel p_1, which gives $p_2 \dots p_m = (uq_2) \dots q_n$. We now repeat the process and the result follows by induction. $\square$

Lemma 2.26

Let R be a PID and let $I_1 \subseteq I_2 \subseteq I_3 \subseteq \ldots$ be a nested sequence of ideals. Then there exists $N \in \mathbb{N}$ such that $I_n = I_{n+1}$ for all $n \geq N$.

Remark. Rings satisfying this "ascending chain condition" are called **Noetherian rings**. More on this later.

Proof. Let $I = \bigcup_{i=1}^{\infty} I_i$. This is an ideal in R (see Example Sheet 2). Since R is a PID, we have that $I = (a)$ for some $a \in R$. Then $a \in \bigcup_{i=1}^{\infty} I_i$, so $a \in I_N$ for some N. Consequently for any $n \geq N$ we have

$$(a) \subseteq I_N \subseteq I_n \subseteq I = (a).$$

So these inclusions are equalities and $I_n = I$. □

Theorem 2.27 (PIDs are UFDs)

If R is a principal ideal domain, then it is a unique factorisation domain.

Proof. Check (i) and (ii) in the definition of UFD.

(i) Let $x \in R$ be a nonzero unit. We proceed by contradiction: suppose x is not a product of irreducibles. Then x is not irreducible, so we can write $x = x_1 y_1$ where x_1, y_1 are not units. Then either x_1 or y_1 is not a product of irreducibles, say x_1 WLOG. We have $(x) \subseteq (x_1)$ and this inclusion is strict because y_1 is not a unit. Now we write $x_1 = x_2 y_2$ where x_2, y_2 are not units. Repeat this procedure to get

$$(x) \subset (x_1) \subset (x_2) \subset \ldots$$

with each inclusion strict. This contradicts Lemma 2.26.

(ii) By Proposition [todo], it suffices to show irreducibles are primes. We conclude by Proposition [todo].

□

Remark. We now know that for a general ring, the following chain of implications holds:

$$\text{ED} \implies \text{PID} \implies \text{UFD} \implies \text{Integral domain.}$$

Let's give some examples of rings that fall into different categories.

	ED	PID	UFD	ID
$\mathbb{Z}/4\mathbb{Z}$[1]	×	×	×	×
$\mathbb{Z}[\sqrt{-5}]$[2]	×	×	×	✓
$\mathbb{Z}[X]$[3]	×	×	✓	✓
$\mathbb{Z}[\frac{1+\sqrt{-19}}{2}]$[4]	×	✓	✓	✓
$\mathbb{Z}[i]$[5]	✓	✓	✓	✓

Definition (gcd/lcm)
Let R be an integral domain.

1. We say $d \in R$ is a **greatest common divisor** (gcd) of the elements $a_1, \dots, a_n \in R$ if $d | a_i$ for all i and if $d' | a_i$ for all i, then $d' | d$.
2. $m \in R$ is a **lowest common multiple** (lcm) of $a_1, \dots, a_n$ if $a_i | m$ for all i and if $a_i | m'$ for all i, then $m | m'$.

Both gcd and lcm (when they exist) are unique up to associates.

Proposition 2.28
In a UFD, both gcd's and lcm's exist.

Proof. Write $a_i = u_i \prod_j p_j^{n_{ij}}$ where $1 \leq i \leq n$, where u_i is a unit, the p_j are irreducible which are not associates of each other, and $n_{ij} \in Z_{\geq 0}$.

Claim. $d = \prod_i p_i^{m_i}$ where $m_i = \min_{1 \leq i \leq n} n_{ij}$ is the gcd of $a_1, \dots, a_n$.

Proof. Certainly $d | a_i$ for all i. If $d' | a_i$ for all i, then writing $d' = u \prod_j p_j^{t_j}$, we find that $t_j \leq n_{ij}$ for all i. So $t_i \leq m_j$. Therefore $d' | d$. □

The argument for lcm's is similar (just replacing min by max). □

[1] This isn't an integral domain as $2 \cdot 2 = 0$, i.e 2 is a zero-divisor.
[2] This isn't a UFD by Example 2.20.
[3] We showed that $\mathbb{Z}[X]$ isn't a PID in Example 2.24, and we will later show that it is a PID by Theorem 2.29.
[4] This result is beyond the scope of the course; see Part II Number Fields.
[5] We know this is an ED and all the other implications follow.

2.5 Factorisation in polynomial rings

Goal. Our first goal in this section is to prove the following theorem:

Theorem 2.29
If R is a UFD then $R[X]$ is also a UFD.

In this section, R is a UFD with field of fractions F. We have $R[X] \leq F[X]$. Moreover $F[X]$ is a ED, hence a PID and UFD. Note this isn't yet enough to show that $R[X]$ is a UFD- since it's just a subring of a UFD, which isn't necessarily a UFD. However, we will show that today.

First we will prove some preliminary results.

Definition (Content)
The **content** of $f = a_n X^n + \ldots + a_1 X + a_0 \in R[X]$ is

$$c(f) = \gcd(a_0, a_1, \ldots, a_n).$$

Note this is well-defined up to multiplication oby a unit. We say f is **primitive** if $c(f)$ is a unit.

Remark. It might be easier to view R as $\mathbb{Z}$if you are seeing this for the first time, in order to more intutively understand the definitions.

Lemma 2.30 (i) If $f, g \in R[X]$ are primitive, then fg is primitive.

(ii) If $f, g \in R[X]$, then $c(fg) = c(f)c(g)$ (up to multiplication by a unit).

There are quite a few different proofs of this lemma, but we will give a concrete one that involves directly computing the coefficients of fg.

Proof. (i) Let

$$f = a_n X^n + \ldots + a_1 X + a_0$$
$$g = b_m X^m + \ldots + b_1 X + b_0$$

If fg is not primitive, $c(fg)$ is not a unit, so there is some prime p such that $p|c(fg)$ (since it factors into irreducibles, but in a UFD irreducibles are primes). Since f, g primitive, $p \nmid c(f)$ and $p \nmid c(g)$.

Suppose $p|a_0, p|a_1 \ldots$ but $p \nmid a_k$ (it is the smallest such k). Let b_l be defined similarly. Then the coefficient of X^{k+l} in fg is

$$\sum_{i+j=k+l} a_i b_j = \underbrace{\ldots + a_{k-1}b_{l+1}}_{\text{div. by } p} + a_k b_l + \underbrace{a_{k+1}b_{l-1} + \ldots}_{\text{div. by } p}.$$

Thus $p|a_k b_l \implies p|a_k$ or $p|b_l$ since p is prime. Contradiction. So fg is primitive.

(ii) Write $f = c(f) \cdot f_0$ and $g = c(g) \cdot g_0$, where $f_0, g_0 \in R[x]$ are primitive. Then $fg = c(f)c(g)f_0 g_0$, where $f_0 g_0$ is primitive by (i). Taking the content, $c(fg) = c(f)c(g)$ up to units.

□

Corollary 2.31

Let $p \in R$ be prime. Then p is prime in $R[X]$.

Proof. $R[X]^\times = R^\times$ as R is an integral domain, so p is not a unit in $R[X]$. Let $f \in R[X]$. Then $p|f$ in $R[X]$ iff $p|c(f)$ in R. Thus if $p|gh$ in $R[X]$, then $p|c(gh) = c(g)c(h)$. Thus $p|c(g)$ or $p|c(h)$, so $p|g$ or $p|h$. So p is prime in $R[X]$. □

Lemma 2.32

Let $f, g \in R[X]$ with g primitive. If $g|f$ in $F[X]$, then $g|f$ in $R[X]$.

Proof. Let $f = gh$ where $h \in F[X]$. Let $a \in R, a \neq 0$ such that $ah \in R[X]$. (We are "cancelling the denominators".) Write $ah = c(ah)h_0$ with h_0 primitive. Then $af = c(ah) \underbrace{h_0 g}_{\text{primitive}}$. Taking contents, we find $a|c(ah)$. We can get rid of the a's since R is an integral domain, so $h|R[X]$ and $g|f$ in $R[X]$. □

Lemma 2.33 (Gauss' Lemma)

Let $f \in R[X]$ be primitive. Then f irreducible in $R[X] \implies f$ irreducible in $F[X]$.

Proof. Since $f \in R[X]$ is irreducible and primitive, we have $\deg(f) > 0$ (else it would be a unit in R), and so f is not a unit in $F[X]$.

Suppose that f is not irreducible in $F[X]$, say $f = gh$ where $g, h \in F[X]$ with $\deg(g), \deg(h) > 0$. Let $\lambda \in F^\times$ such that $\lambda^{-1} g \in R[X]$ is primitive. (For example, let $0 \neq b \in R$ such that $bg \in R[X]$, then $bg = c(bg) g_0$ with g_0 primitive, so $\lambda = c(bg)/b \in F^\times$.) Upon replacing g by $\lambda^{-1} g$ and h by λh, we may assume $g \in R[X]$ is primitive. Then Lemma 2.32 implies $h \in R[X]$ and so $f = gh$ in $R[X]$, with $\deg g, \deg h > 0$. This is a contradiction and we are done. $\square$

Remark. We'll see that the reverse implication also holds.

Lemma 2.34

Let $g \in R[X]$ be primitive. Then g prime in $F[X] \implies g$ is prime in $R[X]$.

Proof. Suppose $f_1, f_2 \in R[X]$ and $g | f_1 f_2$ in $R[X]$. g prime in $F[X]$ implies that $g | f_1$ or $g | f_2$ in $F[X]$. Lemma 2.32 implies $g | f_1$ or $g | f_2$ in $R[X]$. So g is prime in $R[X]$. $\square$

We now have all the tools necessary to prove our theorem.

Proof of Theorem 2.29. Let $f \in R[X]$. Write $f = c(f) f_0$, with $f_0 \in R[X]$ primitive. Since R is a UFD, $c(f)$ is a product of irreducibles in R (which are also irreducible in $R[X]$). If f_0 is not irreducible, say $f_0 = gh$, then $\deg g, \deg h > 0$ since f is primitive, and g, h are also primitive. By induction on degree, f_0 is a product of irreducibles in $R[X]$. This establishes (i) in the definition of a UFD.

By Proposition [todo], it suffices to show that if $f \in R[X]$ is irreducible, then f is prime. Write $f = c(f) f_0$ as before. Since f is irreducible, this means f is a constant or primitive.

Case. If f is constant, f is irreducible in $R[X]$ (hence also in R). So f is prime in R since R is a UFD. We showed in Corollary 2.31 that this implies f is prime in $R[X]$.

Case. If f is primitive, then f is irreducible in $R[X]$ and by Gauss' Lemma also in $F[X]$. Since $F[X]$ is a UFD, f is prime in $F[X]$. By Lemma 2.34, f is prime in $R[X]$.

This concludes the proof. $\square$

Remark. By Lemma [todo], the last three implications are equivalences.

Example

Let's give some applications of Theorem 2.29.

- $\mathbb{Z}[X]$ is a UFD.
- Let $R[X_1, \dots, X_n]$ be the polynomial ring in n variables $X_1, \dots, X_n$ with coefficients in R. We can also define it inductively by $R[X_1, \dots, X_n] = R[X_1, \dots, X_{n-1}][X_n]$. Then applying the theorem, this is a UFD.

Lemma 2.35 (Eisenstein's criterion)
Let R be a UFD and $f(X) = a_n X^n + \ldots + a_1 X + a_0 \in R[X]$ be primitive. Suppose there exists an irreducible element $p \in R$ (=prime) such that

(i) $p \nmid a_n$,

(ii) $p | a_i$ for all $0 \leq i \leq n-1$,

(iii) $p^2 \nmid a_0$.

Then f is irreducible in $R[X]$.

Proof. Suppose $f = gh$, with $g, h \in R[X]$ not units. Since f is primitive, $\deg(g), \deg(h) > 0$. Let

$$\begin{aligned} g &= r_k X^k + \ldots + r_1 X + r_0 \\ h &= s_l X^l + \ldots + s_1 X + s_0 \end{aligned}$$

with $k + l = n$. Then $p \nmid a_n = r_k s_l \implies p \nmid r_k$ and $p \nmid s_l$. $p | a_0 = r_0 s_0 \implies p | r_0$ or $p | s_0$. Let $p | r_0$ WLOG.

Then $\exists\ j \leq k$ such that $p | r_0, \ldots, p | r_{i-1}$ but $p \nmid r_j$. Consider

$$\underbrace{a_j}_{\text{div. by } p \text{ by (ii)}} = \underbrace{r_0 s_j + r_1 s_{j-1} + \ldots + r_{j-1} s_0}_{\text{div. by } p} + r_j s_0.$$

So $p | r_j s_0$. Thus $p | s_0$, so $p^2 | r_0 s_0 = a_0$. Contradiction. $\square$

Example (i) Consider $f(X) = X^3 + 2X + 5 \in \mathbb{Z}[X]$. If f is not irreducible in $\mathbb{Z}[X]$, then

$$f(X) = (X + a)(X^2 + bX + c) \quad \text{for } a, b, c \in \mathbb{Z}.$$

Thus $ac = 5$. But $\pm 1, \pm 5$ are not roots of f. Contradiction. By Gauss' Lemma, f is irreducible in $\mathbb{Q}[X]$. Thus $\mathbb{Q}[X]/(f)$ is a field.

(ii) Let $p \in \mathbb{Z}$ be prime. Apply Eisenstein's criterion to $X^n - p$, which satisfies all three conditions and is therefore irreducible in $\mathbb{Z}[X]$. By Gauss' Lemma, it is irreducible in $\mathbb{Q}[X]$.

(iii) Let $f(X) = X^{p-1} + X^{p-2} + \ldots + X + 1 \in \mathbb{Z}[X]$, where $p \in \mathbb{Z}$ is prime. Eisenstein's criterion does not apply directly to f. But note that $f(X) = \frac{X^p - 1}{X - 1}$, and substituting $Y = X - 1$ gives

$$f(Y+1) = \frac{(Y+1)^p - 1}{Y} = Y^{p-1} + \binom{p}{1} Y^{p-2} + \ldots + \binom{p}{p-2} Y + \binom{p}{p-1}.$$

Now $p | \binom{p}{i}$ for all $1 \leq i \leq p-1$ and $p^2 \nmid \binom{p}{p-1} = p$. Thus $f(Y+1)$ is irreducible in $\mathbb{Z}[Y]$, so $f(X)$ is irreducible in $\mathbb{Z}[X]$. This is because if we had $f(X) = g(X)h(X) \implies f(Y+1) = g(Y+1)h(Y+1)$.

2.6 Algebraic integers

Recall $\mathbb{Z}[i] = \{a + bi : \ a, b \in \mathbb{Z}\} \leq C$. This is the subring of Gaussian integers. We have equipped them with the norm function

$$N : \ \mathbb{Z}[i] \to \mathbb{N}_0, \quad a + bi \mapsto a^2 + b^2t.$$

This satisfies $N(z_1 z_2) = N(z_1)N(z_2)$, so is a Euclidean function. Thus $\mathbb{Z}[i]$ is an ED, hence a PID and UFD, and so *the primes in* $\mathbb{Z}[i]$ *are the irreducibles in* $\mathbb{Z}[i]$. The units in $\mathbb{Z}[i]$ are $\pm 1, \pm i$ (only elements of norm 1).

Goal. We would like to know the primes/irreducibles in $\mathbb{Z}[i]$.

Example

- $2 = (1+i)(1-i)$ and $5 = (2+i)(2-i)$ are not primes in $\mathbb{Z}[i]$ (even though they are primes in $\mathbb{Z}$).
- 3 is prime: we know $N(3) = 9$. So if $3 = ab$ in $\mathbb{Z}[i]$, $N(a)N(b) = 9$. But $\mathbb{Z}[i]$ has no elements of norm 3. Thus either a or b is a unit.

  Similarly 7 is prime in $\mathbb{Z}[i]$.

Proposition 2.36
Let $p \in \mathbb{Z}$ be a prime number. Then the following are equivalent:

(i) p is not prime in $\mathbb{Z}[i]$

(ii) $p = a^2 + b^2$ for some $a, b \in \mathbb{Z}$

(iii) $p = 2$, or $p \equiv 1 \mod 4$.

Proof of (i) $\implies$ (ii). Let $p = xy$, for $x, y \in \mathbb{Z}[i]$ not units. Then $p^2 = N(p) = N(x)N(y)$. Thus $N(x) = N(y) = p$. Writing $x = a + bi$ gives $p = N(x) = a^2 + b^2$. □

Proof of (ii) $\implies$ (iii). The squares mod 4 are 0 and 1. Thus if $p = a^2 + b^2$, then $p \not\equiv 3 \mod 4$. □

Proof of (iii) $\implies$ (i). Already saw 2 is not prime in $\mathbb{Z}[i]$. By Theorem [todo], $(\mathbb{Z}/p\mathbb{Z})^\times$ is cyclic of order $p - 1$. So if $p \equiv 1 \mod 4$, then $(\mathbb{Z}/p\mathbb{Z})$ contains an element of order 4, i.e $\exists\ x \in \mathbb{Z}$ with $x^4 \equiv 1 \mod p$ but $x^2 \not\equiv 1 \mod p$. So $x^2 \equiv -1 \mod p$.

Now $p | x^2 + 1 = (x - i)(x + i)$ but $p \nmid x + i$ and $p \nmid x - i$, so p is not prime in $\mathbb{Z}[i]$. □

Remark. In the course of this proof, we showed (iii) $\implies$ (ii): the statement that given a prime number p: $p = 2$, or $p \equiv 1 \mod 4$ is equivalent to p being expressible as the sum of two squares.

This is difficult to prove with elementary number theory, and shows the power of the techniques we have developed.

Theorem 2.37
The primes in $\mathbb{Z}[i]$ (up to associates) are

1. $a+bi$, where $a,b \in \mathbb{Z}$ and $a^2+b^2=p$ is a prime number with $p=2$, or $p \equiv 1 \mod 4$.
2. Prime numbers $p \in \mathbb{Z}$ with $p \equiv 3 \mod 4$.

Proof. First we check these are primes:

1. $N(a+bi)=p$. If $a+bi=uv$, then either $N(u)=1$ or $N(v)=1$. Thus $a+bi$ is irreducible and hence prime.
2. Immediate by the previous proposition.

Now let $z \in \mathbb{Z}[i]$ be a prime (irreducible). Then $\overline{z} \in \mathbb{Z}[i]$ is also irreducible and $N(z)=z\overline{z}$ is a factorisation into irreducibles. Let $p \in \mathbb{Z}$ be a prime number divinding $N(z)$ (exists as $N(z) \neq 1$).

If $p \equiv 3 \mod 4$, then p itself is prime in $\mathbb{Z}[i]$ by the first part of the proof. So $p|N(z)=z\overline{z}$. So $p|z$ or $p|\overline{z}$. Note that if $p|\overline{z}$, then $p|z$ by taking complex conjugates. So $p|z$. Since both p and z are irreducible, they must be equal up to associates.

Otherwise, we get $p=2$ or $p \equiv 1 \mod 4$. If $p \equiv 1 \mod 4$ then $p-1=4k$ for some $k \in \mathbb{Z}$. As $\mathbb{F}_p^\times \cong C_{p-1}=C_{4k}$, there is a unique element of order 2. This must be -1. Now let $a \in \mathbb{F}_p^\times$ be an element of order 4. Then a^2 has order 2, so $(a^2)=(-1)$. This is the same as saying we can find an a such that $p|a^2+1$. Thus $p|(a+i)(a-i)$. In the case where $p=2$, we know by checking directly that $2=(1+i)(1-i)$. In either case, we deduce that p is not prime (hence irreducible), since it clearly doesn't divide $a \pm i$. So we can write $p=z_1z_2$ for $z_1, z_2 \in \mathbb{Z}[i]$ not units. Now we get

$$p^2=N(p)=N(z_1)N(z_2).$$

As the z_i aren't units, we know $N(z_1)=N(z_2)=p$. By definition, this means $p=z_1\overline{z_1}=z_2\overline{z_2}$. But also $p=z_1z_2$. So we must have $\overline{z_1}=z_2$. Finally, we have $p=z_1\overline{z_1}|N(z)=z\overline{z}$. All these z, z_i are irreducible. So z must be an associate of z_1 or $\overline{z_1}$. So in particular $N(z)=p$. □

Remark. If $p=a^2+b^2$ (case 1.), then $a+bi$ and $a-bi$ are not associates unless $p=2$. $[(1+i)=(1-i)i]$.

Corollary 2.38
An integer $n \geq 1$ is the sum of two squares iff every prime factor p of n with $p \equiv 3 \mod 4$ divides n to an even power.

Proof.

$$\begin{aligned} n=a^2+b^2 &\iff n=N(x) \text{ for some } x \in \mathbb{Z}[i] \\ &\iff n \text{ is a product of norms of primes in } \mathbb{Z}[i]. \end{aligned}$$

Theorem 2.37 implies that the norms of primes in $\mathbb{Z}[i]$ are the primes $p \in \mathbb{Z}$ with $p \not\equiv 3 \mod 4$, and squares of primes $p \in \mathbb{Z}$ with $p \equiv 3 \mod 4$. □

Example

$65 = 5 \cdot 13$. We know by this corollary that it can be written as the sum of two squares. Factoring into primes into $\mathbb{Z}[i]$ gives

$$5 = (2+i)(2-i), \quad 13 = (2+3i)(2-3i).$$

This allows us to factor 65 into two complex conjugates:

$$65 = (2+3i)(2+i)(2-3i)(2+i) = N\left((2+3i)(2+i)\right) = N(1+8i) = 1^2 + 8^2.$$

But by factoring it as $65 = N\left((2+i)(2-3i)\right) = N(7-4i) = 7^2 + 4^2$.

Now we introduce a more general version of the Gaussian integers.

Definition (Algebraic number) (i) $\alpha \in \mathbb{C}$ is an **algebraic number** if there exists non-zero $f \in \mathbb{Q}[X]$ with $f(a) = 0$.

(ii) $\alpha \in \mathbb{C}$ is an **algebraic integer** if there exists *monic* non-zero $f \in \mathbb{Q}[X]$ with $f(\alpha) = 0$.

Notation. Let R be a subring of S, and $\alpha \in S$. We write $R[\alpha]$ for the smallest subring of S containing R and α.

We can view it as the image of the homomorphism from $R[X] \to S$ with $g(X) \mapsto g(\alpha)$.

Definition (Minimal polynomial)

Let α be an algebraic number, and let $\phi : \mathbb{Q}[X] \to \mathbb{C}$ be given by $g(X) \mapsto g(\alpha)$. We know $\mathbb{Q}[X]$ is a PID, which means $\operatorname{Ker}\phi = (f)$ for some $f \in \mathbb{Q}[X]$. Then $f \neq 0$ since α is an algebraic number. Upon multiplying f by a unit, we may assume f is monic WLOG.

Therefore f is characterised by α: we call it the **minimal polynomial** of α.

Remark. By isomorphism theorem,

$$\mathbb{Q}[X]/(f) \cong \mathbb{Q}[\alpha] \leq \mathbb{C}.$$

Thus $\mathbb{Q}[\alpha]$ is an integral domain, which implies f is irreducible in $\mathbb{Q}[X]$. So $\mathbb{Q}[\alpha]$ is a field.

Proposition 2.39

Let α be an algebraic integer, and $f \in \mathbb{Q}[X]$ its minimal polynomial. Then $f \in \mathbb{Z}[X]$ and $(f) = \operatorname{Ker}(\theta) \trianglelefteq \mathbb{Z}[X]$, where $\theta : \mathbb{Z}[X] \to \mathbb{C}$ is the map $g(X) \to g(\alpha)$.

Proof. Let $\lambda \in \mathbb{Q}$ such that $\lambda f \in \mathbb{Z}[X]$ is primitive. Then $\lambda f(\alpha) = 0$, so $\lambda f \in \operatorname{Ker}(\theta)$. Let $g \in \operatorname{Ker}(\theta)$. Then $g \in Ker(\phi)$ and hence $\lambda f | g$ in $\mathbb{Q}[X]$. Lemma [todo] tells us that $\lambda f | g$ in $\mathbb{Z}[X]$. Thus $\operatorname{Ker}(\theta) = (\lambda f)$. Now since α is an algebraic integer, $\exists\ g \in \operatorname{Ker}(\theta)$ which is monic. Then $\lambda f | g$ in $\mathbb{Z}[X]$, which implies $\lambda = \pm 1$. Hence $f \in \mathbb{Z}[X]$, and $(f) = \operatorname{Ker}\theta$. $\square$

Let α be an algebraic integer. Applying the isomorphism theorem to θ gives an isomorphism $\mathbb{Z}[X]/(f) \cong \mathbb{Z}[\alpha]$.

Example

$i, \sqrt{2}, \frac{-1+\sqrt{-3}}{2}, \sqrt[n]{p}$ have minimal polynomials $X^2+1, X^2-2, X^2+X+1, X^n-p$. So these are all algebraic integers. Thus

$$\mathbb{Z}[X]/(X^2+1) \cong Z[i], \quad \mathbb{Z}[X]/(X^2-2) \cong \mathbb{Z}[\sqrt{2}], \quad \text{etc.}$$

Corollary 2.40

If α is an algebraic integer and $\alpha \in \mathbb{Q}$, then $\alpha \in \mathbb{Z}$.

Proof. Let $0 \neq \alpha$ be an algebraic integer. Then Proposition 2.39 implies that the minimal polynomial has coefficients in $\mathbb{Z}$. Since $\alpha \in \mathbb{Q}$, the minimal polynomial is $X - \alpha$. Thus $\alpha \in \mathbb{Z}$. $\square$

2.7 Noetherian rings

We showed that any PID R satisfies the "ascending chain condition" (ACC): see Proposition 2.26. More generally, this gives us a characterisation of the rings that satisfy the ACC.

Lemma 2.41
Let R be a ring. Then R satisfies the ACC iff all ideals in R are finitely generated.

Proof. The reverse implication is a generalisation of our previous proof that we did for PIDs. Let $I_1 \subseteq I_2 \subseteq \dots$ be a chain of ideals and $I = \bigcup_{n \geq 1} I_n$, which is again an ideal.

By assumption $I = (a_1, \dots, a_m)$ for some $a_1, \dots, a_m \in R$. These elements belong to a nested union, so $\exists N \in \mathbb{N}$ such that $a_1, \dots, a_m \in I_N$. Then for $n \geq N$,

$$(a_1, \dots, a_m) \subseteq I_N \subseteq I_n \subseteq I = (a_1, \dots, a_n).$$

So $I_n = I_N = I$.

Conversely, assume that $J \trianglelefteq R$ is not finitely generated and R satisfies the ACC. Choose $a_1 \in J$, then $J \neq (a_1)$, so we can choose $a_2 \in J \backslash (a_1)$. Then $J \neq (a_1, a_2)$, so choose an element $a_3 \in J \backslash (a_1, a_2)$. Continuing this process, we obtain a chain of ideals

$$(a_1) \subset (a_1, a_2) \subset (a_1, a_2, a_3) \subset \dots$$

where the inclusions are strict by our choice of the a_i. This contradicts the ACC. $\square$

Definition (Noetherian ring)
A ring satisfying the ACC is called **Noetherian**.

Theorem 2.42 (Hilbert basis theorem)
If R is a Noetherian ring, then $R[X]$ is also a Noetherian ring.

Proof. Assume $J \trianglelefteq R[X]$ is not finitely generated. Choose $f_1 \in J$ of minimal degree. Then $(f_1) \neq J$. Choose $f_2 \in J \backslash (f_1)$ of minimal degree. Then $(f_1, f_2) \neq J$ and choose $f_3 \in J \backslash (f_1, f_2)$ of minimal degree. Continue this to obtain a sequence $f_1, f_2, f_3, \ldots \in R[X]$ with $\deg f_i \leq \deg f_{i+1}$ (since we have picked the f_i of minimal degree).

Set $a_i :=$ leading coefficient of f_i. We obtain a chain of ideals in R:

$$(a_1) \subseteq (a_1, a_2) \subseteq (a_1, a_2, a_3) \subseteq \ldots .$$

Since R is Noetherian, $\exists\ m \in \mathbb{N}$ such that $a_{m+1} \in (a_1, \ldots, a_m)$. Let $a_{m+1} = \sum_{i=1}^{m} \lambda_i a_i$, and set

$$g = \sum_{i=1}^{m} \lambda_i X^{\deg(f_{m+1} - f_i)} f_i .$$

Then $\deg f_{m+1} = \deg g$ and they have the same leading coefficient a_{m+1}. Then $f_{m+1} - g \in J$ and $\deg(f_{m+1} - g) < \deg f_{m+1}$. This implies $f_{m+1} - g \in (f_1, \ldots, f_m)$ by minimality of our choice of $\deg f_{m+1}$. $\implies$ $f_{m+1} \in (f_1, \ldots, f_m)$. Contradiction. So J is finitely generated and $R[X]$ is Noetherian. □

Corollary 2.43

We can use this to generate some examples of Noetherian rings:

1. $\mathbb{Z}[X_1, \ldots, X_n]$ is Noetherian.
2. $F[X_1, \ldots, X_n]$ is Noetherian for F a field.

Proof. Immediate. □

Example

Let $R = \mathbb{C}[X_1, \ldots, X_n]$. Let $V \subseteq \mathbb{C}^n$ be a subset of the form

$$\{(a_1, \ldots, a_n) \in \mathbb{C}^n : f(a_1, \ldots, a_n) = 0\ \forall f \in \mathcal{F}\},$$

where $\mathcal{F} \subseteq R$ is a possibly infinite set of polynomials. Let

$$I = \left\{ \sum_{i=1}^{m} \lambda_i f_i : \ m \in \mathbb{N}, \lambda_i \in R, f_i \in \mathcal{F} \right\}.$$

Then $I \trianglelefteq R$. Since R is Noetherian (by Hilber basis theorem), it is generated by a finite set of polynomials:

$$I = (g_1, \ldots, g_r), g_i \in I.$$

Thus $V = \{(a_1, \ldots, a_n) \in C^n : \ g(a_1, \ldots, a_n) = 0 \text{ and } i = 1, \ldots, r\}$.

Lemma 2.44

Let R be a Noetherian ring and $I \trianglelefteq R$. Then R/I is Noetherian.

Proof. Let $J_1' \subseteq J_2' \subseteq \ldots$ be a chain of ideals in R/I. By the ideal correspondence, we have $J_i' = J_i/I$ for some $J_1 \subseteq J_2 \subseteq \ldots$ a chain of ideals in R (containing I).

Since R is Noetherian, $\exists\ N \in \mathbb{N}$ such that $J_n = J_{n+1}$ for all $n \geq N$. This implies that $J_n' = J_{n+1}'$ for all $n \geq N$. Thus R/I is Noetherian. □

Example
- $\mathbb{Z}[i] = \mathbb{Z}[X]/(X^2 + 1)$ is Noetherian.
- If $R[X]$ is Noetherian, $R \cong R[X]/(X)$ is Noetherian. This is the converse to the Hilbert basis theorem.

We've basically been proving that every ring we have seen is Noetherian, so it would be nice to see some examples of rings that aren't Noetherian (to make sure everything we have been doing isn't trivial!).

Example (Examples of non-Noetherian rings)
Here are some examples.

1. (Polynomial ring in countably many variables). Let $R = \mathbb{Z}[X_1, X_2, \ldots] = \bigcup_{n\geq 1} \mathbb{Z}[X_1, \ldots, X_n]$. Consider the chain of ideals
$$(X_1) \subset (X_1, X_2) \subset (X_1, X_2, X_3) \subset \ldots.$$
This is an infinite ascending chain. So R isn't Noetherian.
2. Let $R = \{f \in \mathbb{Q}[X] : f(0) \in \mathbb{Z}\} \leq \mathbb{Q}[X]$. We can consider
$$(X) \subset (\frac{1}{2}X) \subset (\frac{1}{4}X) \subset (\frac{1}{8}X) \subset ....$$
These containments are strict since $2 \in R$ is not a unit. So R isn't Noetherian.

3 Modules

3.1 Definitions and examples

We define modules in essentially the same way as vector spaces, except for the fact that the scalars form only a ring, not a field.

Definition (Module)
Let R be a ring. A **module** over R is a triple $(M, +, \cdot)$ consisting of a set M and two operations

$$+ : M \times M \to M, \quad \cdot : R \times M \to M$$

such that

(i) $(M, +)$ is an abelian group, say with identity 0 $(= 0_M)$.

(ii) The operation $\cdot$ satisfies

$$\begin{aligned}
&(r_1 + r_2) \cdot m = r_1 \cdot m + r_2 \cdot m \quad \forall r_1, r_2 \in R, \quad \forall m \in M.\\
&r \cdot (m_1 + m_2) = r \cdot 1 + r \cdot m_2 \quad \forall r \in R, \quad \forall m_1, m_2 \in M.\\
&r_1 \cdot (r_2 \cdot m) = (r_1 \cdot r_2) \cdot m, \quad \forall r_1, r_2 \in R, \quad \forall m \in M.\\
&1_R \cdot m = m \quad \forall m \in M.
\end{aligned}$$

Remark. Don't forget to check closure when checking that $+, \cdot$ are well-defined.

Example (Examples of modules)

1. Let $R = F$ be a field. Then an F-module is precisely a vector space over F by definition.
2. Let $R = \mathbb{Z}$. A $\mathbb{Z}$-module is precisely the same as an abelian group A, where $\cdot : \mathbb{Z} \times A \to A$ is given by

$$(a, a) \mapsto \begin{cases} \underbrace{a + \ldots + a}_{n \text{ times}} & \text{if } n > 0;\\ 0 & \text{if } n = 0;\\ -(\underbrace{a + \ldots + a}_{|n| \text{ times}}) & \text{if } n < 0. \end{cases}$$

3. Let F be a field, and V be a vector space over F. Let $\alpha : V \to V$ be an endomorphism of V. We can make V into an $F[X]$-module via

$$\cdot : F[X] \times V \to V, \quad (f, v) \mapsto (f(\alpha))(v).$$

Note that different choices of α make V into different $F[X]$-modules. We sometimes write $V = V_\alpha$.

Example (General constructions)

1. For any ring R, R^n is an R-module via

$$r \cdot (r_1, \ldots, r_n) = (rr_1, \ldots, rr_n).$$

In particular, taking $n = 1$, we find that R is an R-module.

2. If $I \trianglelefteq R$, then I is an R-module (restricting multiplication on R to I, which can be done since I is an ideal). Moreover, R/I is an R-module via
$$r \cdot (s + I) = rs + I.$$

3. Let $\phi : R \to S$ be a ring homomorphism. Then any S-module M may be regarded as an R-module via
$$r \cdot m = \phi(r)m.$$
In particular, if $R \leq S$ then any S-module may be viewed as an R-module.

Definition (Submodule)
Let M be an R-module. A subset $N \subseteq M$ is an **R-submodule** (written $N \leq M$) if it is a subgroup of $(M, +)$ and $r \cdot n \in N \quad \forall r \in R, \quad \forall n \in N$.

Example
1. A subset of R is an R-submodule precisely when it is an ideal.
2. When $R = F$ is a field, a module is a vector space, and the submodules are vector subspaces.

Definition (Quotient module)
If $N \leq M$ is an R-submodule, the **quotient module** M/N is the quotient group M/N under $+$ with
$$r \cdot (m + N) = r \cdot m + N.$$
This is well-defined, and makes M/N an R-module.

Definition (Homomorphism)
Let M, N be R-modules. A function $f : M \to N$ is an **R-module homomorphism** if it is a homomorphism of abelian groups under $+$ and $f(r \cdot m) = r \cdot f(m)$ for all $r \in R,\ m \in M$.

Example
If $R = F$ is a field, an F-module homomorphism is just a linear map.

Theorem 3.1 (First isomorphism theorem for modules)
Let $f : M \to N$ be an R-module homomorphism. Then we define
$$\mathrm{Ker}(f) = \{m \in M : \ f(m) = 0\} \leq M$$
$$\mathrm{Im}(f) = \{f(m) \in N : m \in M\} \leq N$$
and we have the relation
$$M/\mathrm{Ker}(f) \cong \mathrm{Im}(f).$$

Proof. Similar to before. □

Theorem 3.2 (Second isomorphism theorem for modules)
Let $A, B \leq M$ be R-submodules. Define $A + B = \{a + b : \ a \in A, b \in B\}$. This is a submodule of M and so is $A \cap B$. Furthermore,

$$A/(A \cap B) \cong (A + B)/B.$$

Proof. Apply the first isomorphism theorem to the composite map $m \mapsto m + B$. □

Remark. For the third isomorphism theorem, note there exists a bijection

$$\{\text{submodules of } M/N\} \leftrightarrow \{\text{submodules of } M \text{ containing } N\}.$$

Theorem 3.3 (Third isomorphism theorem for modules)
If $N \leq L \leq M$ are R-submodules, then

$$\frac{M/N}{L/N} \cong M/L.$$

Proof. Similar to the case of groups. □

Remark. These isomorphism theorems apply to vector spaces (compare them with results from IB Linear Algebra). For example, the first isomorphism theorem tells us the rank-nullity theorem, and the second relates $\dim(A + B)$ to $\dim(A \cap B)$.

—-break—-

Let M be an R-module. If $m \in M$, write

$$Rm = \{rm \in M : r \in R\} = \text{submodule generated by } m.$$

Definition (Finitely generated module)
M is **finitely generated** if there exist $m_1, m_2, \ldots, m_n \in M$ such that $M = Rm_1 + Rm_2 + \ldots + Rm_n$.

Remark. This is the analogue for modules of a finite dimensional vector space.

Lemma 3.4
A module M is finitely generated $\iff$ $\exists$ a surjective R-module homomorphism $f : R^n \to M$ for some $n \in \mathbb{N}$.

Proof. If M is finitely generated, write $M = Rm_1 + \ldots + Rm_n$. Define $f : R^n \to M$ by $(r_1, \ldots, r_n) \mapsto \sum_{i=1}^n r_i m_i$. We can easily check this is an R-module homomorphism, and since M is generated by $m_1, \ldots, m_n$, f is surjective.

Conversely, if $\exists$ a surjective R-module homomorphism $f : R^n \to M$ for some $n \in \mathbb{N}$: Define e_i as the ith standard basis vector in R^n. Given our f, set $m_i = f(e_i)$. Then any $m \in M$ is of the form

$$m = f(r_1, \ldots, r_n) = f(\sum_i r_i e_i) = \sum_i r_i f(e_i) = \sum_i r_i m_i.$$

Thus $M = Rm_1 + \ldots + Rm_n$. □

Corollary 3.5
Let $N \leq M$ be an R-submodule. If M is finitely generated, then M/N is finitely generated.

Proof. This is easily deduced using the previous lemma. Let $f : R^n \to M$ be a surjective R-module homomorphism. Then the quotient map $R^n \to M \to M/N$ sending $m \mapsto m + N$ is a surjective R-module homomorphism. Thus M/N is finitely generated. □

Remark. A submodule of a finitely generated module need not be finitely generated. Consider the following counterexample:

Let R be a non-Noetherian ring and $I \trianglelefteq R$ be an ideal in R that isn't finitely generated. Then R is a finitely generated R-module (generated by 1) and I is a submodule which is not finitely generated.

In fact, a submodule of a finitely generated module over a Noetherian ring is finitely generated. (Example Sheet 4)

Definition (Torsion)
Let M be an R-module. Then

1. An element $m \in M$ is **torsion** if $\exists 0 \neq r \in R$ with $r \cdot m = 0$.
2. M is a **torsion module** if every $m \in M$ is torsion.
3. M is **torsion-free** if every non-zero $m \in M$ is not torsion.

Example
The torsion elements in a $\mathbb{Z}$-module (an abelian group) are the elements of finite order.

Any F-module (vector space) is torsion-free.

3.2 Direct sums and free modules

Definition (Direct sum)
Let $M_1, \ldots, M_n$ be R-modules. The **direct sum** $M_1 \oplus \ldots \oplus M_n$ is the set $M_1 \times \ldots \times M_n$ with operations

$$(m_1, \ldots, m_n) + (m'_1, \ldots, m'_n) = (m_1 + m'_1, \ldots, m_n + m'_n)$$
$$r \cdot (m_1, \ldots, m_n) = (rm_1, \ldots, rm_n)$$

The direct sum is an R-module.

Example
$R^n = \underbrace{R \oplus \ldots \oplus R}_{n \text{ times}}$.

Lemma 3.6
If $M = \oplus_{i=1}^n M_i$ and $N_i \leq M_i$ for all i, then setting $N = \oplus_{i=1}^n N_i$, we have

$$M/N \cong \oplus_{i=1}^n M_i/N_i.$$

Proof. Apply the 1st isomorphism theorem to the surjective R-module homomorphism from $M \to \oplus_{i=1}^n M_i/N_i$ with $(m_1, \ldots, m_n) \mapsto (m_1 + N_1, \ldots, m_n + N)$. This has kernel N, and the result follows. □

Definition (Independence)
Let $m_1, \ldots, m_n \in M$. The set $\{m_1, \ldots, m_n\}$ is **independent** if $\sum_{i=1}^n r_i m_i = 0 \implies r_1 = r_2 = \ldots = r_m = 0$.

Definition
A subset $S \subseteq M$ generates M freely if

1. S generates M, i.e. for all $m \in M$, $m = \sum r_i s_i$, $r_i \in R$, $s_i \in I$.
2. Any function $\psi : S \to N$ where N is an R-module extends to an R-module homomorphism $\theta : M \to N$.

   Remark. Such an extension θ is always unique, if it exists.

An R-module which is freely generated by some subset $S \subseteq M$ is called **free** and S is called a **free basis**.

Proposition 3.7
For a finite subset $S = \{m_1, \ldots, m_n\} \subseteq M$, the following are equivalent:

(i) S generates M freely.

(ii) S generates M and S is independent.

(iii) Every element of M can be written uniquely as $r_1 m_1 + \ldots + r_n m_n$ for some $r_1, \ldots, r_n \in R$.

(iv) The R-module homomorphism from $R^n \to M$ defined by $(r_1, \ldots, r_n) \mapsto \sum_i r_i m_i$ is an isomorphism.

Proof.
- (i) $\implies$ (ii): Let S generate M freely. If S is not independent, then $\exists r_1, \ldots, r_n \in R$ with $\sum_i r_i m_i = 0$ and some $r_j \neq 0$. Define $\psi : S \to R$ by
$$m_i \mapsto \begin{cases} 1 & \text{if } i = j; \\ 0 & \text{if } i \neq j. \end{cases}.$$
This extends to an R-module hom. $\theta : M \to R$ since S is a free basis. Then
$$0 = \theta(0) = \theta(\sum_i r_i m_i) = \sum_i r_i \theta(m_i) = r_j.$$
This gives our contradiction. Thus S is independent.

Other implications exercise TODO 3 imp 2 imp 1, then show 3 iff 4 □

Example
- A non-trivial finite abelian group is not a free $\mathbb{Z}$-module.
- The set $\{2, 3\} \subset \mathbb{Z}$ generates $\mathbb{Z}$ as a $\mathbb{Z}-$module, but they are not independent since $3 \cdot 2 + (-2) \cdot 3 = 0$. Furthermore, no subset of $\{2, 3\}$ is a free basis since $\{2\}, \{3\}$ do not generate $\mathbb{Z}$.

Proposition 3.8 (Invariance of dimension)
Let R be a non-zero ring. If $R^m \cong R^n$ as R-modules, then $m = n$. (The cardinality of any two free bases are the same.)

Proof. First, we introduce a general construction. Let $I \trianglelefteq R$ and M and R-module. Define $IM = \{\sum_i a_i m_i : a_i \in I, m_i \in M\} \leq M$. The quotient M/IM is an R/I-module via $(r + I) \cdot (m + IM) = rm + IM$. This is well defined since if $b \in I$, $b \cdot (m + IM) = bm + IM = 0 + IM$.

Suppose $R^m \cong R^n$. Choose $I \trianglelefteq R$ to be a maximal ideal[2]. By the above, get an isomorphism of R/I-modules
$$(R/I)^m \cong R^m/IR^m \cong R^n/IR^n \cong (R/I)^n.$$
But $I \trianglelefteq R$ is maximal, so R/I is a field. So $m = n$ by invariance of dimension for vector spaces. □

[2]The existence of a maximal ideal follows from Zorn's lemma and Example Sheet 2 Q4.

3.3 The structure theorem and applications

Remark. Until further notice, we will take R to be a Euclidean domain and $\phi : R\setminus\{0\} \to \mathbb{N}$ to be a Euclidean function.

Let A be an $m \times n$ matrix with entries in R.

Definition (Row operations)
The elementary row operations on A are as follows:

(ER1) Add λ multiples of the jth row to the ith row ($\lambda \in R, i \neq j$)

(ER2) Swap the ith and jth rows.

(ER3) Multiply the ith row by $u \in R^\times$. (We pick u a unit so that all of these are reversible operations.)

Each of these can be realised by left multiplication by an $m \times m$ invertible matrix (recall these from IB Linear Algebra). Similarly we can define the elementary column operations (EC1-3), realised by right multiplication by an $n \times n$ invertible matrix.

Definition
Two $m \times n$ matrices A and B are **equivalent** if $\exists$ a sequence of elementary row operations taking A to B.

Theorem 3.9 (Smith Normal Form)
An $m \times n$ matrix $A = (a_{ij})$ over a Euclidean domain R is equivalent to a diagonal matrix

$$\begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_t \end{pmatrix}$$

where the $d_i \neq 0$ (but there can be 0s on the diagonal), and $d_1|d_2|\dots|d_t$. The d_i are called **invariant factors** - we will show they are unique up to associates.

Proof. If $A = 0$, done. Otherwise upon swapping rows and columns, may assume $a_{11} \neq 0$. We will reduce $\phi(a_{11})$ as much as possible via the following algorithm, until it divides every other element in the matrix:

- *Step 1:* If $a_{11} \nmid a_{1j}$ for some $j \geq 2$, then write $a_{1j} = qa_{11} + r$ for $q, r \in R$ where $\phi(r) < \phi(a_{11})$. Subtracting q times column 1 from column j, and swapping these columns makes our top left entry r.
- *Step 2:* If $a_{11} \nmid a_{i1}$ for some $i \geq 2$, then repeat the above process with row operations.
- *Step 3:* Steps 1 and 2 decrease $\phi(a_{11})$, so can repeat them finitely many times until $a_{11}|a_{1j}$ $\forall j \geq 2$ and $a_{11}|a_{i1}$ $\forall i \geq 2$. Subtracting multiples of the first

row/column from the others gives us a matrix

$$A = \left(\begin{array}{c|c} a_{11} & 0 \dots 0 \\ \hline 0 & \\ \vdots & A' \\ 0 & \end{array}\right)$$

where A' is an $(m-1) \times (m-1)$ matrix.

- *Step 4:* If $a_{11} \nmid a_{ij}$ for some $i, j \geq 2$, then add ith row to the first row, and perform column operations as before to decrease $\phi(a_{11})$. Restart the algorithm. After finitely many steps obtain

$$A = \left(\begin{array}{c|c} a_{11} & 0 \dots 0 \\ \hline 0 & \\ \vdots & A' \\ 0 & \end{array}\right)$$

with $a_{11} = d_1$ say, so that $d_1 | a_{ij}$ $\forall i, j$.

Applying the same method to A' now gives the result. □

For uniqueness of the invariant factors, introduce the minors of A.

Definition
A $k \times k$ **minor** of A is the determinant of a $k \times k$ submatrix (i.e a matrix formed by deleting $n-k$ rows and $n-k$ columns).

Definition
The kth **Fitting ideal** $\mathrm{Fit}_k(A) \trianglelefteq R$ is the ideal generated by the $k \times k$ minors of A.

Lemma 3.10
If A and B are equivalent matrices, then $\mathrm{Fit}_k(A) = \mathrm{Fit}_k(B)$ for all k.

Proof. We show that (ER1-3) don't change $\mathrm{Fit}_k(A)$ (same proof works for (EC1-3)).

For (ER1), we are adding λ times the jth row to the ith row, so A becomes A'. Let C be a $k \times k$ submatrix of A and C' the corresponding submatrix of A'.

- If we did not choose the ith row, then $C = C'$ and hence $\det(C) = \det(C')$.
- If we chose both of the rows i and j, then C and C' differ by a row operation. Since each of the elementary matrices clearly have determinant 1, $\det(C) = \det(C')$.
- If we chose the ith row but not the jth row, then by expanding along the ith row using the definition of the determinant gives

$$\det(C') = \det(C) + \lambda \det(D)$$

where D is another $k \times k$ submatrix of A, choosing the jth row instead of the

ith row. Thus $\det(C') \in \mathrm{Fit}_k(A)$. Hence $\mathrm{Fit}_k(A') \subseteq \mathrm{Fit}_k(A)$. Since (ER1) is reversible, we apply the same argument in reverse to get $\mathrm{Fit}_k(A) \subseteq \mathrm{Fit}_k(A')$. Hence $\mathrm{Fit}_k(A') = \mathrm{Fit}_k(A)$

Checking (ER2) and (ER3) are similar but easier (we avoid the last case). □

Now if A has SNF

$$\begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_t \end{pmatrix} \quad \text{with } d_1|d_2|\dots|d_t,$$

then $\mathrm{Fit}_k(A) = (d_1 d_2 \dots d_k)$. Thus the products $d_1 \dots d_k$ (up to associates) depend only on A for each k. Cancelling out in the usual way for an integral domain shows that each d_i (up to an associate) depends only on A.

Example

Consider the matrix $A = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$ over $\mathbb{Z}$. To find the SNF of A:

$$\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \underbrace{\longrightarrow}_{C_1 \mapsto C_1 + C_2} \begin{pmatrix} 1 & -1 \\ 3 & 2 \end{pmatrix} \underbrace{\longrightarrow}_{C_2 \mapsto C_1 + C_2} \begin{pmatrix} 1 & 0 \\ 3 & 5 \end{pmatrix} \underbrace{\longrightarrow}_{R_2 \mapsto R_2 - 3R_1} \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}.$$

But also $(d_1) = (2, -1, 1, 2) = (1) \implies d_1 = \pm 1$; and $(d_1 d_2) = (\det(A)) = (5) \implies d_2 = \pm 5$. This gives us another way of calculating the SNF.

We will use SNF to prove the structure theorem. First, some preparation.

Lemma 3.11

Let R be an ED. Any submodule of R^m is generated by at most m elements.

Proof. Let $N \leq R^m$. Consider the ideal

$$I = \{r \in R : \ \exists r_2, \dots, r_m \in R \text{ s.t } (r, r_2, \dots, r_m) \in N\}.$$

(Here the brackets isn't an ideal generated, it's writing the elements in R^m by coordinates). Since EDs are PIDs, we have $I = (a)$ for some $a \in R$. Choose some $n = (a, a_2, \dots, a_m) \in N$. For $(r_1, \dots, r_m) \in N$, we have $r_1 = ra$ for some r, so $(r_1, r_2, \dots, r_m) - rn = (0, r_2 - ra_2, \dots, r_m - ra_m)$. This lies in

$$N' = (N \cap \{0\}) \times R^{m-1} \leq R^{m-1}.$$

Hence $N = Rn + N'$. By induction, N' is generated by $n_2, \dots, n_m$ and thus $\{n, n_2, \dots, n_m\}$ generate N. □

Theorem 3.12

Let R be an ED and $N \leq R^m$. There is a free basis $x_1, \dots, x_m$ for R^m such that N is generated by $d_1 x_1, \dots, d_t x_t$ for some $t \leq m$ and $d_1, d_2, \dots, d_t \in R$ with $d_1|d_2|\dots|d_t$.

Proof. By Lemma 3.11, we have $N = Ry_1 + \ldots + Ry_n$ for some $n \leq m$. Each y_i belongs to R^m, so we can form an $m \times n$ matrix given by

$$\left(\begin{array}{c|c|c|c} y_1 & y_2 & \cdots & y_n \end{array} \right).$$

By Smith Normal Form, A is equivalent to

$$A' = \begin{pmatrix} d_1 & & & & \\ & \ddots & & & \\ & & d_t & & \\ & & & 0 & \\ & & & & \ddots \end{pmatrix} \quad \text{with } d_1 | d_2 | \ldots | d_t.$$

A' is obtained from A by elementary row and column operations. Each row operation changes our choice of free basis for R^m. Each column operation changes our set of generators for N. Thus after changing the free basis of R^m to $x_1, \ldots, x_m$ (say) the submodule N is generated by $d_1x_1, d_2x_2, \ldots, d_tx_t$, as claimed. □

The following theorem is the main theorem of the modules part of the course.

Theorem 3.13 (Structure theorem)
Let R be an ED and M be a finitely generated R-module. Then

$$M \cong R/(d_1) \oplus R/(d_2) \oplus \ldots \oplus R/(d_t) \oplus \underbrace{R \oplus \ldots \oplus R}_{k \text{ copies}}$$

for some $0 \neq d_i \in R$ with $d_1 | d_2 | \ldots | d_t$ and $k \geq 0$. The d_i are called invariant factors.

Proof. Since M is finitely generated, there exists a surjective R-module homomorphism $\phi : R^m \to M$ for some m (Lemma 3.4). By the first isomorphism theore, $M \cong R^m / \operatorname{Ker}(\phi)$.

By the previous theorem, $\exists$ a free basis $x_1, \ldots, x_m$ of R^m such that $\operatorname{Ker}(\phi)$ is generated by $d_1x_1, \ldots, d_2x_2, \ldots, d_tx_t$ with $d_1 | d_2 | \ldots | d_t$. Then

$$\begin{aligned} M &\cong \frac{R \oplus R \oplus \ldots R \oplus R \oplus \ldots \oplus R}{d_1R \oplus d_2R \oplus \ldots \oplus d_tR \oplus 0 \oplus \ldots \oplus 0} \\ &\cong R/(d_1) \oplus R/(d_2) \oplus \ldots \oplus R/(d_t) \oplus \underbrace{R \oplus \ldots \oplus R}_{m-t \text{ copies}} \text{ by Lemma 3.6.} \end{aligned}$$

□

Remark. After deleting those d_i which are units, the module M uniquely determines them (up to associates). We omit the proof.

Corollary 3.14

Let R be an ED. Then any finitely generated torsion-free module is free.

Proof. If M is torsion-free, then there are no submodules of the form $R/(d)$ with $d \neq 0$ by the structure theorem. Thus $M \cong R^m$ for some m. □

Example

$R = \mathbb{Z}$. Consider the abelian group generated by a and b subject to the relations $2a + b = 0, \quad -a + 2b = 0$. Then $G \cong \mathbb{Z}^2/N$, where N is generated by $(2, 1)$ and $(-1, 2)$.

$A = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$ has SNF $\begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}$. This can change the basis for $\mathbb{Z}^2$ so that N is generated by $(1, 0)$ and $(0, 5)$. Thus

$$G \cong \frac{\mathbb{Z} + \mathbb{Z}}{\mathbb{Z} \oplus 5\mathbb{Z}} \cong \mathbb{Z}/5\mathbb{Z}.$$

We can expand on this result:

Theorem 3.15 (Structure theorem for finitely generated abelian groups)
Any finitely generated abelian group G is isomorphic to

$$\mathbb{Z}/d_1\mathbb{Z} \oplus \mathbb{Z}/d_2\mathbb{Z} \oplus \ldots \oplus \mathbb{Z}/d_t\mathbb{Z} \oplus \mathbb{Z}^r,$$

where $d_1|d_2|d_t$ and $r \geq 0$.

Proof. Take $R = \mathbb{Z}$ in the proof of the structure theorem. □

Remark. The special case where G is finite ($r = 0$) was quoted as Theorem 1.28.

In Section 1.8, we saw that any finite abelian group can be written as a product of C_p's where p is a prime number. To generalise this, we need the following lemma:

Lemma 3.16
Let R be a PID and $a, b \in R$ with $\gcd(a, b) = 1$. Then

$$R/(ab) \cong R/(a) \oplus R/(b) \qquad \text{as } R-\text{modules}.$$

(We already proved the case $R = \mathbb{Z}$ in Section 1.8.)

Proof. If R is a PID, then $(a, b) = (d)$ for some $d \in R$. But $\gcd(a, b) = 1 \implies d$ is a unit. So $\exists r, s \in R$ such that $ra + sb = 1$. (This is like an analogue of Bezout's theorem.) Define an R-module homomorphism

$$\phi : R \to R/(a) \oplus R/(b) \text{ given by } x \mapsto (x + (a), x + (b)).$$

Then $\phi(sb) = (1+(a), 0+(b))$ and $\phi(ra) = (0+(a), 1+(b))$. Thus $\phi(sbx+ray) = (x + (a), y + (b))$ for all $x, y \in R$, hence ϕ is surjective.

Clearly $(ab) \subseteq \mathrm{Ker}(\phi)$. Conversely, if $x \in \mathrm{Ker}(\phi)$, then $x \in (a) \cap (b)$ and $x = x(ra + sb) = r(ax) + s(xb) \in (ab)$. So $\mathrm{Ker}(\phi) = (ab)$ and the first isomorphism theorem concludes the proof. □

Theorem 3.17 (Primary decomposition theorem)
Let R be an ED and M a finitely generated R-module. Then

$$M \cong R/(p_1^{n_1}) \oplus \ldots \oplus R/(p_k^{n_k}) \oplus R^m$$

(as R-modules), where $p_1, \ldots, p_n$ are primes (not necessarily distinct), and $m \geq 0$.

Proof. By the structure theorem, we know that

$$M \cong R/(d_1) \oplus R/(d_2) \oplus \ldots \oplus R/(d_t) \oplus R^m.$$

So it suffices to consider $M = R/(d_i)$. We know $d_i = up_1^{a_1} \ldots p_r^{a_r}$ where u is a unit and $p_1, \ldots, p_r$ are distinct (non-associate) primes. Lemma 3.16 tells us

$$R/(d_i) = R/(p_1^{a_1}) \oplus \ldots \oplus R/(p_r^{a_r}).$$

The result follows. □

Let V be a vector space over a field F, and let α be an endomorphism on V. Let V_α denote the $F[X]$-module where $F[X] \times V \to V$ is given $(f(X), v) \mapsto f(\alpha)(v)$.

Lemma 3.18
If V is finite dimensional, then V_α is a finitely generated $F[X]$-module.

Proof. If $v_1, \ldots, v_n$ generate F as a vector space, then they generate V_α as an $F[X]$-module since $F \leq F[X]$. □

Example 1. Suppose $V_\alpha \cong F[X]/(X^n)$ as an $F[X]$-module. Then $1, X, X^2, \ldots, X^{n-1}$ is a basis for $F[X]/X^n$ as an F-vector space, and wrt this basis α has matrix

$$\mathcal{M} = \begin{pmatrix} 0 & 0 & \ldots & 0 \\ 1 & 0 & \ldots & 0 \\ 0 & 1 & & \\ \vdots & & \ddots & \\ 0 & & & 1 \end{pmatrix}$$

since α acts as "multiplication" by X.