Groups, Rings and Modules

Cambridge University Mathematical Tripos: Part IB

12th June 2024

Contents

1 Review of IA Groups

This section contains material covered by IA Groups.

1.1 Definitions

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

- $a \cdot (b \cdot c) = (a \cdot b) \cdot c$;
- there exists $e \in G$ such that for all $g \in G$, we have $g \cdot e = e \cdot g = g$; and
- for all $g \in G$, there exists an inverse $h \in G$ such that $g \cdot h = h \cdot g = e$.

Remark. (i) Sometimes, such as in IA Groups, a closure axiom is also specified. However, this is implicit in the type definition of $\cdot$. In practice, this must normally be checked explicitly.

(ii) Additive and multiplicative notation will be used interchangeably. For additive notation, the inverse of g is denoted $-g$, and for multiplicative notation, the inverse is instead denoted g^{-1}. The identity element is sometimes denoted 0 in additive notation and 1 in multiplicative notation.

A subset $H \subseteq G$ is a *subgroup* of G, written $H \leq G$, if $h \cdot h' \in H$ for all $h, h' \in H$, and $(H, \cdot)$ is a group. The closure axiom must be checked, since we are restricting the definition of $\cdot$ to a smaller set.

Remark. A non-empty subset $H \subseteq G$ is a subgroup of G if and only if

$$a, b \in H \implies a \cdot b^{-1} \in H$$

An *abelian* group is a group such that $a \cdot b = b \cdot a$ for all a, b in the group. The *direct product* of two groups G, H, written $G \times H$, is the group over the Cartesian product $G \times H$ with operation $\cdot$ defined such that $(g_1, h_1) \cdot (g_2, h_2) = (g_1 \cdot_G g_2, h_1 \cdot_H h_2)$.

1.2 Cosets

Let $H \leq G$. Then, the *left cosets* of H in G are the sets gH for all $g \in G$. The set of left cosets partitions G. Each coset has the same cardinality as H. Lagrange's theorem states that if G is a finite group and $H \leq G$, we have $|G| = |H| \cdot [G : H]$, where $[G : H]$ is the number of left cosets of H in G. $[G : H]$ is known as the *index* of H in G. We can construct Lagrange's theorem analogously using right cosets. Hence, the index of a subgroup is independent of the choice of whether to use left or right cosets; the number of left cosets is equal to the number of right cosets.

1.3 Order

Let $g \in G$. If there exists $n \geq 1$ such that $g^n = 1$, then the least such n is the *order* of G. If no such n exists, we say that g has infinite order. If g has order d, then:

(i) $g^n = 1 \implies d \mid n$;

(ii) $\langle g \rangle = \{1, g, \ldots, g^{d-1}\} \leq G$, and by Lagrange's theorem (if G is finite) $d \mid |G|$.

1.4 Normality and quotients

A subgroup $H \leq G$ is *normal*, written $H \trianglelefteq G$, if $g^{-1}Hg = H$ for all $g \in G$. In other words, H is preserved under conjugation over G. If $H \trianglelefteq G$, then the set G/H of left cosets of H in G forms the *quotient group*. The group action is defined by $g_1H \cdot g_2H = (g_1 \cdot g_2)H$. This can be shown to be well-defined.

1.5 Homomorphisms

Let G, H be groups. A function $\phi : G \to H$ is a *group homomorphism* if $\phi(g_1 \cdot_G g_2) = \phi(g_1) \cdot_H \phi(g_2)$ for all $g_1, g_2 \in G$. The *kernel* of ϕ is defined to be $\ker\phi = \{g \in G : \phi(g) = 1\}$, and the *image* of ϕ is $\operatorname{Im}\phi = \{\phi(g) : g \in G\}$. The kernel is a normal subgroup of G, and the image is a subgroup of H.

1.6 Isomorphisms

An *isomorphism* is a homomorphism that is bijective. This yields an inverse function, which is of course also an isomorphism. If $\varphi : G \to H$ is an isomorphism, we say that G and H are isomorphic, written $G \cong H$. Isomorphism is an equivalence relation. The isomorphism theorems are

(i) if $\varphi : G \to H$, then $G/\ker\varphi \cong \operatorname{Im}\varphi$;

(ii) if $H \leq G$ and $N \trianglelefteq G$, then $H \cap N \trianglelefteq H$ and $H/H \cap N \cong HN/N$;

(iii) if $N \leq M \leq G$ such that $N \trianglelefteq G$ and $M \trianglelefteq G$, then $M/N \trianglelefteq G/N$, and $G/N/M/N = G/M$.

2 Simple groups

2.1 Introduction

If $K \trianglelefteq G$, then studying the groups K and G/K give information about G itself. This approach is available only if G has nontrivial normal subgroups. It therefore makes sense to study groups with no normal subgroups, since they cannot be decomposed into simpler structures in this way.

Definition. A group G is *simple* if $\{1\}$ and G are its only normal subgroups.

By convention, we do not consider the trivial group to be a simple group. This is analogous to the fact that we do not consider one to be a prime.

Lemma. Let G be an abelian group. G is simple if and only if $G \cong C_p$ for some prime p.

Proof. Certainly C_p is simple by Lagrange's theorem. Conversely, since G is abelian, all subgroups are normal. Let $1 \neq g \in G$. Then $\langle g \rangle \trianglelefteq G$. Hence $\langle g \rangle = G$ by simplicity. If G is infinite, then $G \cong \mathbb{Z}$, which is not a simple group; $2\mathbb{Z} \lhd \mathbb{Z}$. Hence G is finite, so $G \cong C_{o(g)}$. If $o(g) = mn$ for $m, n \neq 1, p$, then $\langle g^m \rangle \leq G$, contradicting simplicity. □

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

$$1 \cong G_0 \lhd G_1 \lhd \cdots \lhd G_n = G$$

where each quotient G_{i+1}/G_i is simple.

Remark. It is not the case that necessarily G_i be normal in G_{i+k} for $k \geq 2$.

Proof. We will consider an inductive step on $|G|$. If $|G| = 1$, then trivially $G = 1$. Conversely, if $|G| > 1$, let G_{n-1} be a normal subgroup of largest possible order not equal to $|G|$. Then, G/G_{n-1} exists, and is simple by the correspondence theorem. □

3 Group actions

3.1 Definitions

Definition. Let X be a set. Then $\mathrm{Sym}(X)$ is the group of permutations of X; that is, the group of all bijections of X to itself under composition. The identity can be written id or id_X.

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

Example. The symmetric group S_n is exactly equal to $\mathrm{Sym}(\{1, \ldots, n\})$, so is a permutation group of order n. A_n is also a permutation group of order n, as it is a subgroup of S_n. D_{2n} is a permutation group of order n.

Definition. A *group action* of a group G on a set X is a function $\alpha : G \times X \to X$ satisfying

$$\alpha(e, x) = x; \quad \alpha(g_1 \cdot g_2, x) = \alpha(g_1, \alpha(g_2, x))$$

for all $g_1, g_2 \in G, x \in X$. The group action may be written $*$, defined by $g * x \equiv \alpha(g, x)$.

Proposition. An action of a group G on a set X is uniquely characterised by a group homomorphism $\varphi : G \to \mathrm{Sym}(X)$.

Proof. For all $g \in G$, we can define $\varphi_g : X \to X$ by $x \mapsto g * x$. Then, for all $x \in X$,

$$\varphi_{g_1 g_2}(x) = (g_1 g_2) * x = g_1 * (g_2 * x) = \varphi_{g_1}(\varphi_{g_2}(x))$$

Thus $\varphi_{g_1 g_2} = \varphi_{g_1} \circ \varphi_{g_2}$. In particular, $\varphi_g \circ \varphi_{g^{-1}} = \varphi_e$. We now define

$$\varphi : G \to \mathrm{Sym}(X); \quad \varphi(g) = \varphi_g \implies \varphi(g)(x) = g * x$$

This is a homomorphism.

Conversely, any group homomorphism $\varphi : G \to \mathrm{Sym}(X)$ induces a group action $*$ by $g * x = \varphi(g)$. This yields $e * x = \varphi(e)(x) = \mathrm{id}\, x = x$ and $(g_1 g_2) * x = \varphi(g_1 g_2)x = \varphi(g_1)\varphi(g_2)x = g_1 * (g_2 * x)$ as required. □

Definition. The homomorphism $\varphi : G \to \mathrm{Sym}(X)$ defined in the above proof is called a *permutation representation* of G.

Definition. Let $G \curvearrowright X$. Then,
(i) the orbit of $x \in X$ is $\mathrm{Orb}_G(x) = \{g * x : g \in G\} \subseteq X$;
(ii) the stabiliser of $x \in X$ is $G_x = \{g \in G : g * x = x\} \leq G$.

Theorem (Orbit-stabiliser theorem)**.** The orbit $\mathrm{Orb}_G(x)$ bijects with the set ${}^{G}\!/_{G_x}$ of left cosets of G_x in G (which may not be a quotient group). In particular, if G is finite, we have

$$|G| = |\mathrm{Orb}(x)| \cdot |G_x|$$

Example. If G is the group of symmetries of a cube and we let X be the set of vertices in the cube, $G \curvearrowright X$. Here, for all $x \in X$, $|\mathrm{Orb}(x)| = 8$ and $|G_x| = 6$ (including reflections), hence $|G| = 48$.

Remark. Note that $\ker \varphi = \bigcap_{x \in X} G_x$. The kernel of the permutation representation φ is also referred to as the kernel of the group action itself. If the kernel is trivial the action is said to be *faithful*.

The orbits partition X. In particular, if there is exactly one orbit, the group action is said to be *transitive*.

Note that $G_{g*x} = gG_xg^{-1}$. Hence, if x, y lie in the same orbit, their stabilisers are conjugate.

Example. G acts on itself by left multiplication. This is known as the *left regular action*. The kernel is trivial, hence the action is faithful. The action is transitive, since for all $g_1, g_2 \in G$, the element $g_2g_1^{-1}$ maps g_1 to g_2.

3.2 Cayley's theorem

Theorem (Cayley's theorem)**.** Any finite group G is a permutation group of order $|G|$; it is isomorphic to a subgroup of $S_{|G|}$.

Example. Let $H \leq G$. Then $G \curvearrowright {}^{G}\!/_{H}$ by left multiplication, where ${}^{G}\!/_{H}$ is the set of left cosets of H in G. This is known as the *left coset action*. This action is transitive using the construction above for the left regular action. We have $\ker \varphi = \bigcap_{x \in G} xHx^{-1}$, which is the largest normal subgroup of G contained within H.

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

Proof. Let $G \curvearrowright X = {}^{G}\!/_{H}$ by left multiplication. Let $\varphi : G \to \mathrm{Sym}(X)$ be the permutation representation associated to this group action. Since G is simple, $\ker \varphi = 1$ or $\ker \varphi = G$. If $\ker \varphi = G$, then $\mathrm{Im}\, \varphi = \mathrm{id}$, which is a contradiction since G acts transitively on X, which has index greater than one. Thus $\ker \varphi = 1$, and $G \cong \mathrm{Im}\, \varphi \leq S_n$. Since $G \leq S_n$ and $A_n \lhd S_n$, the second isomorphism theorem

shows that $G \cap A_n \lhd G$, and

$$G/G \cap A_n \cong GA_n/A_n \leq S_n/A_n \cong C_2$$

Since G is simple, $G \cap A_n = 1$ or $G \cap A_n = G$. If $G \cap A_n = 1$, then G is isomorphic to a subgroup of C_2, but this is false, since G is non-abelian. Hence $G \cap A_n = G$ so $G \leq A_n$. Finally, if $n \leq 4$ we can check manually that A_n is not simple; A_n has no non-abelian simple subgroups. □

3.3 Conjugation actions

Example. Let $G \curvearrowright G$ by conjugation, so $g * x = gxg^{-1}$. This is known as the *conjugation action*.

Definition. The orbit of the conjugation action is called the *conjugacy class* of a given element $x \in G$, written $\mathrm{ccl}_G(x)$. The stabiliser of the conjugation action is the set C_x of elements which commute with a given element x, called the *centraliser* of x in G. The kernel of φ is the set $Z(G)$ of elements which commute with all elements in x, which is the *centre* of G. This is always a normal subgroup.

Remark. $\varphi : G \to G$ satisfies

$$\varphi(g)(h_1h_2) = gh_1h_2g^{-1} = hh_1g^{-1}gh_2g^{-1} = \varphi(g)(h_1)\varphi(g)(h_2)$$

Hence $\varphi(g)$ is a group homomorphism for all g. It is also a bijection, hence $\varphi(g)$ is an isomorphism from $G \to G$.

Definition. An isomorphism from a group to itself is known as an *automorphism*. We define $\mathrm{Aut}(G)$ to be the set of all group automorphisms of a given group. This set is a group. Note, $\mathrm{Aut}(G) \leq \mathrm{Sym}(G)$, and the $\varphi : G \to \mathrm{Sym}(G)$ above has image in $\mathrm{Aut}(G)$.

Example. Let X be the set of subgroups of G. Then $G \curvearrowright X$ by conjugation: $g * H = gHg^{-1}$. The stabiliser of a subgroup H is $\{g \in G : gHg^{-1} = H\} = N_G(H)$, called the *normaliser* of H in G. The normaliser of H is the largest subgroup of G that contains H as a normal subgroup. In particular, $H \lhd G$ if and only if $N_G(H) = G$.

4 Alternating groups

4.1 Conjugation in alternating groups

We know that elements in S_n are conjugate if and only if they have the same cycle type. However, elements of A_n that are conjugate in S_n are not necessarily conjugate in A_n. Let $g \in A_n$. Then $C_{A_n}(g) = C_{S_n}(g) \cap A_n$. There are two possible cases.

- If there exists an odd permutation that commutes with g, then $2|C_{A_n}|(g) = |C_{S_n}|(g)$. By the orbit-stabiliser theorem, $|\mathrm{ccl}_{A_n}(g)| = |\mathrm{ccl}_{S_n}(g)|$.
- If there is no odd permutation that commutes with g, we have $|C_{A_n}|(g) = |C_{S_n}|(g)$. Similarly, $2|\mathrm{ccl}_{A_n}(g)| = |\mathrm{ccl}_{S_n}(g)|$.

Example. For $n = 5$, the product (1 2)(3 4) commutes with (1 2), and (1 2 3) commutes with (4 5). Both of these elements are odd. So the conjugacy classes of the above inside S_5 and A_5 are the same. However, (1 2 3 4 5) does not commute with any odd permutation. Indeed, if that were true for some h, we would have

$$(1\ 2\ 3\ 4\ 5) = h(1\ 2\ 3\ 4\ 5)h^{-1} = (h(1)\ h(2)\ h(3)\ h(4)\ h(5))$$

Hence h must be a 5-cycle in the subgroup of A_5 generated by (1 2 3 4 5).

We can then show that A_5 has conjugacy classes of size $1, 15, 20, 12, 12$. If $H \trianglelefteq A_5$, $|H|$ must be a sum of the sizes of the above conjugacy classes. By Lagrange's theorem, $|H|$ must divide 60. We can check explicitly that this is not possible unless $|H| = 1$ or $|H| = 60$. Hence A_5 is simple.

4.2 Simplicity of alternating groups

Lemma. A_n is generated by 3-cycles.

Proof. All elements of A_n are generated by an even number of transpositions. It therefore suffices to show that a product of two transpositions can be written as a product of 3-cycles. Explicitly,

$$(a\ b)(c\ d) = (a\ c\ b)(a\ c\ d); \quad (a\ b)(b\ c) = (a\ b\ c)$$

□

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

Proof. We claim that every 3-cycle is conjugate to (1 2 3). If $(a\ b\ c)$ is a 3-cycle, we have $(a\ b\ c) = \sigma(1\ 2\ 3)\sigma^{-1}$ for some $\sigma \in S_n$. If $\sigma \in A_n$, then the proof is finished. Otherwise, $\sigma \mapsto \sigma(4\ 5) \in A_n$ suffices, since (4 5) commutes with (1 2 3). □

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

Proof. Suppose $1 \neq N \lhd A_n$. To disprove normality, it suffices to show that N contains a 3-cycle by the lemmas above, since the normality of N would imply N contains all 3-cycles and hence all elements of A_n.

Let $1 \neq \sigma \in N$, writing σ as a product of disjoint cycles.

(i) Suppose σ contains a cycle of length $r \geq 4$. Without loss of generality, let $\sigma = (1\ 2\ 3 \dots r)\tau$ where τ fixes $1, \dots, r$. Now, let $\delta = (1\ 2\ 3)$. We have

$$\underbrace{\sigma^{-1}}_{\in N}\ \underbrace{\delta^{-1}\sigma\delta}_{\in N} = (r \dots 2\ 1)(1\ 3\ 2)(1\ 2 \dots r) = (2\ 3\ r)$$

So N contains a 3-cycle.

(ii) Suppose σ contains two 3-cycles, which can be written without loss of generality as $(1\,2\,3)(4\,5\,6)\tau$. Let $\delta = (1\,2\,4)$, and then

$$\sigma^{-1}\delta^{-1}\sigma\delta = (1\,3\,2)(4\,6\,5)(1\,4\,2)(1\,2\,3)(4\,5\,6)(1\,2\,4) = (1\,2\,4\,3\,6)$$

Therefore, there exists an element of N which contains a cycle of length $5 \geq 4$. This reduces the problem to case (i).

(iii) Finally, suppose σ contains two 2-cycles, which will be written $(1\,2)(3\,4)\tau$. Then let $\delta = (1\,2\,3)$ and

$$\sigma^{-1}\delta^{-1}\sigma\delta = (1\,2)(3\,4)(1\,3\,2)(1\,2)(3\,4)(1\,2\,3) = (1\,4)(2\,3) = \pi$$

Let $\varepsilon = (2\,3\,5)$. Then

$$\underbrace{\pi^{-1}}_{\in N}\underbrace{\varepsilon^{-1}\pi\varepsilon}_{\in N} = (1\,4)(2\,3)(2\,5\,3)(1\,4)(2\,3)(2\,3\,5) = (2\,5\,3)$$

Thus N contains a 3-cycle.

There are now three remaining cases, where σ is a transposition, a 3-cycle, or a transposition composed with a 3-cycle. Note that the remaining cases containing transpositions cannot be elements of A_n. If σ is a 3-cycle, we already know A_n contains a 3-cycle, namely σ itself. □

5 p-groups

5.1 p-groups

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

Theorem. If G is a p-group, the centre $Z(G)$ is non-trivial.

Proof. For $g \in G$, due to the orbit-stabiliser theorem, $|\mathrm{ccl}(g)||C(g)| = p^n$. In particular, $|\mathrm{ccl}(g)|$ divides p^n, and they partition G. Since G is a disjoint union of conjugacy classes, modulo p we have

$$|G| \equiv \text{number of conjugacy classes of size } 1 \equiv 0 \implies |Z(G)| \equiv 0$$

Hence $Z(G)$ has order zero modulo p so it cannot be trivial. We can check this by noting that $g \in Z(G) \iff x^{-1}gx = g$ for all x, which is true if and only if $\mathrm{ccl}_G(g) = \{g\}$. □

Corollary. The only simple p-groups are the cyclic groups of order p.

Proof. Let G be a simple p-group. Since $Z(G)$ is a normal subgroup of G, we have $Z(G) = 1$ or $Z(G) = G$. But $Z(G)$ may not be trivial, so $Z(G) = G$. This implies G is abelian. The only abelian simple groups are cyclic of prime order, hence $G \cong C_p$. □

Corollary. Let G be a p-group of order p^n. Then G has a subgroup of order p^r for all $r \in \{0, \dots, n\}$.

Proof. Recall that any group G has a composition series $1 = G_1 \lhd \cdots \lhd G_N = G$ where each quotient G_{i+1}/G_i is simple. Since G is a p-group, G_{i+1}/G_i is also a p-group. Each successive quotient is an order p group by the previous corollary, so we have a composition series of nested subgroups of order p^r for all $r \in \{0, \dots, n\}$. □

Lemma. Let G be a group. If $G/Z(G)$ is cyclic, then G is abelian. This then implies that $Z(G) = G$, so in particular $G/Z(G) = 1$.

Proof. Let $gZ(G)$ be a generator for $G/Z(G)$. Then, each coset of $Z(G)$ in G is of the form $g^r Z(G)$ for some $r \in \mathbb{Z}$. Thus, $G = \{g^r z : r \in \mathbb{Z}, z \in Z(G)\}$. Now, we multiply two elements of this group and find

$$g^{r_1} z_1 g^{r_2} z_2 = g^{r_1+r_2} z_1 z_2 = g^{r_1+r_2} z_2 z_1 = z_2 z_1 g^{r_1+r_2} = g^{r_2} z_2 g^{r_1} z_1$$

So any two elements in G commute. □

Corollary. Any group of order p^2 is abelian.

Proof. Let G be a group of order p^2. Then $|Z(G)| \in \{1, p, p^2\}$. The centre cannot be trivial as proven above, since G is a p-group. If $|Z(G)| = p$, we have that $G/Z(G)$ is cyclic as it has order p. Applying the previous lemma, G is abelian. However, this is a contradiction since the centre of an abelian group is the group itself. If $|Z(G)| = p^2$ then $Z(G) = G$ and then G is clearly abelian. □

5.2 Sylow theorems

Theorem. Let G be a finite group of order $p^a m$ where p is a prime and p does not divide m. Then:

(i) The set $\mathrm{Syl}_p(G) = \{P \leq G : |P| = p^a\}$ of Sylow p-subgroups is non-empty.
(ii) All Sylow p-subgroups are conjugate.
(iii) The amount of Sylow p-subgroups $n_p = \left|\mathrm{Syl}_p(G)\right|$ satisfies

$$n_p \equiv 1 \mod p; \quad n_p \mid |G| \implies n_p \mid m$$

Proof. (i) Let Ω be the set of all subsets of G of order p^a. We can directly find

$$|\Omega| = \binom{p^a m}{p^a} = \frac{p^a m}{p^a} \cdot \frac{p^a m - 1}{p^a - 1} \cdots \frac{p^a m - p^a + 1}{1}$$

Note that for $0 \leq k < p^a$, the numbers $p^a m - k$ and $p^a - k$ are divisible by the same power of p. In particular, $|\Omega|$ is coprime to p.

Let $G \curvearrowright \Omega$ by left-multiplication, so $g * X = \{gx : x \in X\}$. For any $X \in \Omega$, the orbit-stabiliser theorem can be applied to show that

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

By the above, there must exist an orbit with size coprime to p, since orbits partition Ω. For such an X, $p^a \mid |G_X|$.

Conversely, note that if $g \in G$ and $x \in X$, then $g \in (gx^{-1}) * X$. Hence, we can consider

$$G = \bigcup_{g \in G} g * X = \bigcup_{Y \in \mathrm{orb}_G(X)} Y$$

Thus $|G| \leq |\mathrm{orb}_G(X)| \cdot |X|$, giving $|G_X| = \frac{|G|}{|\mathrm{orb}_G(X)|} \leq |X| = p^a$.

Combining with the above, we must have $|G_X| = p^a$. In other words, the stabiliser G_X is a Sylow p-subgroup of G.

(ii) We will prove a stronger result for this part of the proof. We claim that if P is a Sylow p-subgroup and $Q \leq G$ is a p-subgroup, then $Q \leq gPg^{-1}$ for some $g \in G$. Indeed, let Q act on the set of left cosets of P in G by left multiplication. By the orbit-stabiliser theorem, each orbit has size which divides $|Q| = p^k$ for some k. Hence each orbit has size p^r for some r.

Since G/P has size m, which is coprime to p, there must exist an orbit of size 1. Therefore there exists $g \in G$ such that $q * gP = gP$ for all $q \in Q$. Equivalently, $g^{-1}qg \in P$ for all $q \in Q$. This implies that $Q \leq gPg^{-1}$ as required. This then weakens to the second part of the Sylow theorems.

(iii) Let G act on $\mathrm{Syl}_p(G)$ by conjugation. Part (ii) of the Sylow theorems implies that this action is transitive. By the orbit-stabiliser theorem, $n_p = \left|\mathrm{Syl}_p(G)\right| \mid |G|$.

Let $P \in \mathrm{Syl}_p(G)$. Then let P act on $\mathrm{Syl}_p(G)$ by conjugation. Since P is a Sylow p-subgroup, the orbits of this action have size dividing $|P| = p^a$, so the size is some power of p. To show $n_p \equiv 1 \bmod p$, it suffices to show that $\{P\}$ is the unique orbit of size 1. Suppose $\{Q\}$ is another orbit of size 1, so Q is a Sylow p-subgroup which is preserved under conjugation by P. P normalises Q, so $P \leq N_G(Q)$. Notice that P and Q are both Sylow p-subgroups of $N_G(Q)$. By (ii), P and Q are conjugate inside $N_G(Q)$. Hence $P = Q$ since $Q \trianglelefteq N_G(Q)$. Thus, $|P|$ is the unique orbit of size 1, so $n_p \equiv 1 \bmod p$ as required.

□

Corollary. If $n_p = 1$, then there is only one Sylow p-subgroup, and it is normal.

Proof. Let $g \in G$ and $P \in \mathrm{Syl}_p(G)$. Then gPg^{-1} is a Sylow p-subgroup, hence $gPg^{-1} = P$. P is normal in G. □

Example. Let G be a group with $|G| = 1000 = 2^3 \cdot 5^3$. Here, $n_5 \equiv 1 \bmod 5$, and $n_5 \mid 8$, hence $n_5 = 1$. Thus the unique Sylow 5-subgroup is normal. Hence no group of order 1000 is simple.

Example. Let G be a group with $|G| = 132 = 2^2 \cdot 3 \cdot 11$. n_{11} satisfies $n_{11} \equiv 1 \bmod 11$ and $n_{11} \mid 12$, thus $n_{11} \in \{1, 12\}$. Suppose G is simple. Then $n_{11} = 12$. The amount of Sylow 3-subgroups satisfies $n_3 \equiv 1 \bmod 3$ and $n_3 \mid 44$ so $n_3 \in \{1, 4, 22\}$. Since G is simple, $n_3 \in \{4, 22\}$.

Suppose $n_3 = 4$. Then $G \curvearrowright \mathrm{Syl}_3(G)$ by conjugation, and this generates a group homomorphism $\varphi : G \to S_4$. But the kernel of this homomorphism is a normal subgroup of G, so $\ker \varphi$ is trivial or G itself. If $\ker \varphi = G$, then $\mathrm{Im}\, \varphi$ is trivial, contradicting Sylow's second theorem. If $\ker \varphi = 1$, then $\mathrm{Im}\, \varphi$ has order 132, which is impossible.

Thus $n_3 = 22$. This means that G has $22 \cdot (3-1) = 44$ elements of order 3, and further G has $12 \cdot (11-1) = 120$ elements of order 11. However, the sum of these two totals is more than the total of 132 elements, so this is a contradiction. Hence G is not simple.

6 Matrix groups

6.1 Definitions

Definition. Let F be a field, such as $\mathbb{C}$ or $\mathbb{Z}/p\mathbb{Z}$. Let $GL_n(F)$ be set of $n \times n$ invertible matrices over F, which is called the *general linear group*. Let $SL_n(F)$ be set of $n \times n$ matrices with determinant one over F, which is called the *special linear group*. $SL_n(F)$ is the kernel of the determinant homomorphism on $GL_n(F)$, so $SL_n(F) \triangleleft GL_n(F)$.
Let $Z \triangleleft GL_n(F)$ denote the subgroup of *scalar matrices*, the group of nonzero multiples of the identity. The group $PGL_n(F) = GL_n(F)/Z$ is called the *projective general linear group*. Let $PSL_n(F) = SL_n(F)/Z \cap SL_n(F)$. By the second isomorphism theorem, $PSL_n(F)$ is isomorphic to $Z \cdot SL_n(F)/Z$, which is a subgroup of $PGL_n(F)$.

Example. Consider the finite group $G = GL_n\left(\mathbb{Z}/p\mathbb{Z}\right)$. A list of n vectors in $\mathbb{Z}/p\mathbb{Z}$ are the columns of a matrix $A \in G$ if and only if the vectors are linearly independent. Hence, by considering dimensionality of subspaces generated by each column,

$$\begin{aligned} |G| &= (p^n - 1)(p^n - p)(p^n - p^2) \cdots (p^n - p^{n-1}) \\ &= p^{1+2+\cdots+(n-1)}(p^n - 1)(p^{n-1} - 1) \cdots (p - 1) \\ &= p^{\binom{n}{2}} \prod_{i=1}^{n} (p^i - 1) \end{aligned}$$

Hence the Sylow p-subgroups have size $p^{\binom{n}{2}}$. Let U be the set of upper triangular matrices with ones on the diagonal. This forms a Sylow p-subgroup of G, since there are $\binom{n}{2}$ entries in a given upper triangular matrix, and there are p choices for such an entry.

6.2 Möbius maps in modular arithmetic

Recall that $PGL_2(\mathbb{C})$ acts on $\mathbb{C} \cup \{\infty\}$ by Möbius transformations. Likewise, $PGL_2\left(\mathbb{Z}/p\mathbb{Z}\right)$ acts on $\mathbb{Z}/p\mathbb{Z} \cup \{\infty\}$ by Möbius transformations. For a matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2\left(\mathbb{Z}/p\mathbb{Z}\right); \quad A : z \mapsto \frac{az+b}{cz+d}$$

Since the scalar matrices act trivially, we obtain an action on the projective general linear group instead of the general linear group. We can represent ∞ as an integer, say, p, for the purposes of constructing a permutation representation.

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

Proof. Suppose that $\frac{az+b}{cz+d} = z$ for all $z \in \mathbb{Z}/p\mathbb{Z} \cup \{\infty\}$. Since $z = 0$, we have $b = 0$. Since $z = \infty$, we find $c = 0$. Thus the matrix is diagonal. Finally, since $z = 1$, $\frac{a}{d} = 1$ hence $a = d$. Thus the matrix is scalar. So the permutation representation from $PGL_2\left(\mathbb{Z}/p\mathbb{Z}\right)$ has trivial kernel, giving injectivity as required.

If $p = 2$ or $p = 3$ we can compute the orders of relevant groups manually and show that the permutation representation is an isomorphism. □

Lemma. Let p be an odd prime. Then

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

Proof. By the example above,

$$\left|GL_2\left(\mathbb{Z}/p\mathbb{Z}\right)\right| = p(p^2-1)(p-1)$$

The homomorphism $GL_2\left(\mathbb{Z}/p\mathbb{Z}\right) \to \left(\mathbb{Z}/p\mathbb{Z}\right)^\times$ given by the determinant is surjective. Since $SL_2\left(\mathbb{Z}/p\mathbb{Z}\right)$ is the kernel of this homomorphism, we have

$$\left|SL_2\left(\mathbb{Z}/p\mathbb{Z}\right)\right| = p(p-1)(p+1)$$

Now, if $\begin{pmatrix}\lambda & 0\\ 0 & \lambda\end{pmatrix}$ is an element of the special linear group, then $\lambda^2 \equiv 1 \bmod p$. Then, $p \mid (\lambda-1)(\lambda+1)$ hence $\lambda \equiv \pm 1 \bmod p$. Thus,

$$Z \cap SL_2\left(\mathbb{Z}/p\mathbb{Z}\right) = \{\pm 1\}$$

and the elements are distinct since $p > 2$. Hence the order of the projective special linear group is half the order of the special linear group as required. □

Example. Let $G = PSL_2\left(\mathbb{Z}/5\mathbb{Z}\right)$. Then by the previous lemma, $|G| = 60$. Let $G \curvearrowright \mathbb{Z}/5\mathbb{Z} \cup \{\infty\}$ by Möbius transformations. The permutation representation $\varphi : G \to \mathrm{Sym}(\{0,1,2,3,4,\infty\})$ is injective, since the permutation representation of $PGL_2\left(\mathbb{Z}/p\mathbb{Z}\right)$ is known to be injective by a previous lemma.

We claim that $\mathrm{Im}\,\varphi \subseteq A_6$. Let $\psi = \mathrm{sgn} \circ \varphi$. If we can show ψ is trivial, $\mathrm{Im}\,\varphi \subseteq A_6$. Let $h \in G$, and suppose h has order $2^n m$ for odd m. If $\psi(h^m) = 1$, then since ψ is a group homomorphism we have $\psi(h)^m = 1$ giving $\psi(h) \neq -1 \implies \psi(h) = 1$. So to show ψ is trivial, it suffices to show $\psi(g) = 1$ for all $g \in G$ with order a power of 2. By the second Sylow theorem, if g has order a power of 2, it is contained in a Sylow 2-subgroup. Then it suffices to show that $\psi(H) = 1$ for all Sylow 2-subgroups H. But since $\ker\psi$ is normal and all Sylow 2-subgroups are conjugate, it suffices to show $\psi(H) = 1$ for a single Sylow 2-subgroup H. The Sylow 2-subgroup must have order 4. Hence consider

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

Both of these elements square to the identity element inside the projective special linear group. This generates a group of order 4 which is necessarily a Sylow 2-subgroup. We can explicitly compute the action of H on $\{0, 1, 2, 3, 4, \infty\}$.

$$\varphi\left(\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}\right) = (1\ 4)(2\ 3); \quad \varphi\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right) = (0\ \infty)(1\ 4)$$

These are products of two transpositions, hence even permutations. Thus $\psi(H) = 1$, proving the claim that $G \leq A_6$. We can prove that for any $G \leq A_6$ of order 60, we have $G \cong A_5$; this is a question from the example sheets.

6.3 Properties

The following properties will not be proven in this course.

- $PSL_n\left(\mathbb{Z}/p\mathbb{Z}\right)$ is simple for all $n \geq 2$ and p prime, except where $n = 2$ and $p = 2, 3$. Such groups are called finite groups of *Lie type*.
- The smallest non-abelian simple groups are $A_5 \cong PSL_2\left(\mathbb{Z}/5\mathbb{Z}\right)$, then $PSL_2\left(\mathbb{Z}/7\mathbb{Z}\right) \cong GL_3\left(\mathbb{Z}/2\mathbb{Z}\right)$ which has order 168.

7 Finite abelian groups

7.1 Products of cyclic groups

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

The proof for this theorem will be provided later in the course. Note that the isomorphism provided for by the theorem is not unique. An example of such behaviour is the following lemma.

Lemma. Let m, n be coprime integers. Then $C_m \times C_n \cong C_{mn}$.

Proof. Let g, h be generators of C_m and C_n. Then consider the element $(g, h)^k = (g^k, h^k)$, which has order mn. Thus $\langle (g, h) \rangle$ has order mn. So every element in $C_m \times C_n$ is expressible in this way, giving $\langle (g, h) \rangle = C_m \times C_n$. □

Corollary. Let G be a finite abelian group. Then $G \cong C_{n_1} \times \cdots \times C_{n_k}$ where each n_i is a power of a prime.

Proof. If $n = p_1 a^1 \cdots p^r a^r$ where the p_i are distinct primes, then applying the above lemma inductively gives C_n as a product of cyclic groups which have orders that are powers of primes. We can apply this to the theorem that every finite abelian group is isomorphic to a product of cyclic groups to find the result. □

Later, we will prove the following refinement of this theorem.

Theorem. Let G be a finite abelian group. Then $G \cong C_{d_1} \times \cdots \times C_{d_t}$ where $d_i \mid d_{i+1}$ for all i.

Remark. The integers $n_1, \ldots, n_k$ in the corollary above are unique up to ordering. The integers $d_1, \ldots, d_t$ are also unique, assuming that $d_1 > 1$. The proofs will be omitted.

Example. The abelian groups of order 8 are exactly C_8, $C_2 \times C_4$, and $C_2 \times C_2 \times C_2$. The abelian groups of order 12 are, using the corollary above, $C_2 \times C_2 \times C_3$, $C_4 \times C_3$, and using the above theorem, $C_2 \times C_6$ and C_{12}. However, $C_2 \times C_3 \cong C_6$ and $C_3 \times C_4 \cong C_{12}$, so the groups derived are isomorphic.

Definition. The *exponent* of a group G is the least integer $n \geq 1$ such that $g^n = 1$ for all $g \in G$. Equivalently, the exponent is the lowest common multiple of the orders of elements in G.

Example. The exponent of A_4 is $\operatorname{lcm}\{2, 3\} = 6$.

Corollary. Let G be a finite abelian group. Then G contains an element which has order equal to the exponent of G.

Proof. If $G \cong C_{d_1} \times \cdots \times C_{d_t}$ for $d_i \mid d_{i+1}$, every $g \in G$ has order dividing d_t. Hence the exponent is d_t, and we can choose a generator of C_{d_t} to obtain an element in G of the same order. □

8 Rings

8.1 Definitions

Definition. A *ring* is a triple $(R, +, \cdot)$ where R is a set and $+, \cdot$ are binary operations $R \times R \to R$, satisfying the following axioms.

(i) $(R, +)$ is an abelian group, and we will denote the identity element 0 and the inverse of x as $-x$;
(ii) $(R, \cdot)$ satisfies the group axioms except for the invertibility axiom, and we will denote the identity element 1 and the inverse of x as x^{-1} if it exists;
(iii) for all $x, y, z \in R$ we have $x \cdot (y + z) = x \cdot y + x \cdot z$ and $(y + z) \cdot x = y \cdot x + z \cdot x$.

If multiplication is commutative, we say that R is a *commutative* ring. In this course, we will study only commutative rings.

Remark. For all $x \in R$,

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

Further,

$$0 = 0 \cdot x = (1 + -1) \cdot x = x + (-1 \cdot x) \implies -1 \cdot x = -x$$

Definition. A subset $S \subseteq R$ is a *subring*, denoted $S \leq R$, if $(S, +, \cdot)$ is a ring with the same identity elements.

Remark. It suffices to check the closure axioms for addition and multiplication; the other properties are inherited.

Example. $\mathbb{Z} \leq \mathbb{Q} \leq \mathbb{R} \leq \mathbb{C}$ are rings. The set $\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\}$ is a subring of $\mathbb{C}$. This is known as the ring of Gaussian integers. The set $\mathbb{Q}[\sqrt{2}] = \left\{a + b\sqrt{2} : a, b \in \mathbb{Q}\right\}$ is a subring of $\mathbb{R}$.

Example. The set $\mathbb{Z}/n\mathbb{Z}$ is a ring.

Example. Let R, S be rings. Then the *product* $R \times S$ is a ring under the binary operations

$$(a, b) + (c, d) = (a + c, b + d); \quad (a, b) \cdot (c, d) = (a \cdot c, b \cdot d)$$

The additive identity is $(0_R, 0_S)$ and the multiplicative identity is $(1_R, 1_S)$. Note that the subset $R \times \{0\}$ is preserved under addition and multiplication, so it is a ring, but it is not a subring because the multiplicative identity is different.

8.2 Polynomials

Definition. Let R be a ring. A *polynomial* f over R is an expression

$$f = a_0 + a_1 X + a_2 X^2 + \cdots + a_n X^n$$

for $a_i \in R$. The term X is a formal symbol, no substitution of X for a value will be made. We could alternatively define polynomials as finite sequences of terms in R. The *degree* of a polynomial f is the largest n such that $a_n \neq 0$. A degree-n polynomial is *monic* if $a_n = 1$. We write $R[X]$ for the set of all such polynomials over R. Let $g = b_0 + b_1 X + \cdots + b_n X^n$. Then we define

$$f + g = (a_0 + b_0) + (a_1 + b_1)X + \cdots + (a_n + b_n)X^n; \quad f \cdot g = \sum_i \left(\sum_{j=0}^{i} a_j b_{i-j} \right) X^i$$

Then $(R[X], +, \cdot)$ is a ring. The identity elements are the constant polynomials 0 and 1. We can identify the ring R with the subring of $R[X]$ of constant polynomials.

Definition. An element $r \in R$ is a *unit* if r has a multiplicative inverse. The units in a ring, denoted $R^\times$, form an abelian group under multiplication. For instance, $\mathbb{Z}^\times = \{\pm 1\}$ and $\mathbb{Q}^\times = \mathbb{Q} \setminus \{0\}$.

Definition. A *field* is a ring where all nonzero elements are units and $0 \neq 1$.

Example. $\mathbb{Z}/n\mathbb{Z}$ is a field only if n is a prime.

Remark. If R is a ring such that $0 = 1$, then every element in the ring is equal to zero. Indeed, $x = 1 \cdot x = 0 \cdot x = 0$. Thus, the exclusion of rings with $0 = 1$ in the definition of a field simply excludes the trivial ring.

Proposition. Let $f, g \in R[X]$ such that the leading coefficient of g is a unit. Then there exist polynomials $q, r \in R[X]$ such that $f = qg + r$, where the degree of r is less than the degree of g.

Remark. This is the Euclidean algorithm for division, adapted to polynomial rings.

Proof. Let n be the degree of f and m be the degree of g, so

$$f = a_n X^n + \cdots + a_0; \quad g = b_m X^m + \cdots + b_0$$

By assumption, $b_m \in R^\times$. If $n < m$ then let $q = 0$ and $r = f$. Conversely, we have $n \geq m$. Consider the polynomial $f_1 = f - a_n b_m^{-1} g X^{n-m}$. This has degree at most $n - 1$. Hence, we can use induction on n to decompose f_1 as $f_1 = q_1 g + r$. Thus $f = (q_1 + a_n b_m^{-1} X^{n-m})g + r$ as required. □

Remark. If R is a field, then every nonzero element of R is a unit. Therefore, the above algorithm can be applied for all polynomials g unless g is the constant polynomial zero.

Example. Let R be a ring and X be a set. Then the set of functions $X \to R$ is a ring under

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

The set of continuous functions $\mathbb{R} \to \mathbb{R}$ is a subring of the ring of all functions $\mathbb{R} \to \mathbb{R}$, since they are closed under addition and multiplication. The set of polynomial functions $\mathbb{R} \to \mathbb{R}$ is also a subring, and we can identify this with the ring $\mathbb{R}[X]$.

Example. Let R be a ring. Then the *power series ring* $R[\![X]\!]$ is the set of power series on X. This is defined similarly to the polynomial ring, but we permit infinitely many nonzero elements in the expansion. The power series is defined formally; we cannot actually carry out infinitely many additions in an arbitrary ring. We instead consider the power series as a sequence of numbers.

Example. Let R be a ring. Then the ring of *Laurent polynomials* is $R[X, X^{-1}]$ with the restriction that $a_i \neq 0$ for finitely many i.

8.3 Homomorphisms

Definition. Let R and S be rings. A function $\varphi : R \to S$ is a *ring homomorphism* if
(i) $\varphi(r_1 + r_2) = \varphi(r_1) + \varphi(r_2)$;
(ii) $\varphi(r_1 \cdot r_2) = \varphi(r_1) \cdot \varphi(r_2)$;
(iii) $\varphi(1_R) = 1_S$.
We can derive that $\varphi(0_R) = 0_S$ from (i).
A ring homomorphism is an *isomorphism* if it is bijective. The *kernel* of a ring homomorphism is $\ker \varphi = \{r \in R : \varphi(r) = 0\}$.

Lemma. Let R, S be rings. Then a ring homomorphism $\varphi : R \to S$ is injective if and only if $\ker \varphi = \{0\}$.

Proof. Let $\varphi : (R, +) \to (S, +)$ be the induced group homomorphism on addition. The result then follows from the corresponding fact about group homomorphisms. □

8.4 Ideals

Definition. A subset $I \subseteq R$ 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 that an ideal is *proper* if $I \neq R$.

Lemma. Let $\varphi : R \to S$ be a ring homomorphism. Then $\ker \varphi$ is an ideal of R.

Proof. We know that $\ker \varphi$ is a subgroup by the equivalent fact from groups. If $r \in R$ and $x \in \ker \varphi$, then

$$\varphi(rx) = \varphi(r)\varphi(x) = \varphi(r) \cdot 0 = 0$$

Hence $rx \in \ker \varphi$. □

Remark. If I contains a unit, then the multiplicative identity lies in I. Then all elements lie in I. In particular, if I is a proper ideal, $1 \notin I$. Hence a proper ideal I is not a subring of R.

Lemma. The ideals in $\mathbb{Z}$ are precisely the subsets of the form $n\mathbb{Z}$ for any $n = 0, 1, 2, \ldots$.

Proof. First, we can check directly that any subset of the form $n\mathbb{Z}$ is an ideal. Now, let I be any nonzero ideal of $\mathbb{Z}$ and let n be the smallest positive element. Then $n\mathbb{Z} \subseteq I$. Let $m \in I$. Then by the Euclidean algorithm, $m = qn + r$ for $q, r \in \mathbb{Z}$ and $r \in \{0, 1, \ldots, n-1\}$. Then $r = m - qn$. We know $qn \in I$ since $n \in I$, so $r \in I$. If $r \neq 0$, this contradicts the minimality of n as chosen above. So $I = n\mathbb{Z}$ exactly. □

Definition. For an element $a \in R$, we write (a) to denote the subset of R given by multiples of a; that is, $(a) = \{ra : r \in R\}$. This is an ideal, known as the ideal *generated by* a. More generally, if $a_1, \ldots, a_n \in R$, then $(a_1, \ldots, a_n)$ is the set of elements in R given by linear combinations of the a_i. This is also an ideal.

Definition. Let $I \trianglelefteq R$. Then I is *principal* if there exists some $a \in R$ such that $I = (a)$.

8.5 Quotients

Theorem. Let $I \trianglelefteq R$. Then the set R/I of cosets of I in $(R, +)$ forms the *quotient ring* under the operations

$$(r_1 + I) + (r_2 + I) = (r_1 + r_2) + I; \quad (r_1 + I) \cdot (r_2 + I) = (r_1 \cdot r_2) + I$$

This ring has the identity elements

$$0_{R/I} = 0_R + I; \quad 1_{R/I} = 1_R + I$$

Further, the map $R \to {}^{R}\!/_{I}$ defined by $r \mapsto r + I$ is a ring homomorphism called the *quotient map*. The kernel of the quotient map is I. Hence any ideal is the kernel of some homomorphism.

Proof. From the analogous result from groups, the addition defined on the set of cosets yields the group $\left({}^{R}\!/_{I}, +\right)$. If $r_1 + I = r_1' + I$ and $r_2 + I = r_2' + I$, then $r_1' = r_1 + a_1$ and $r_2' = r_2 + a_2$ for some $a_1, a_2 \in I$. Then

$$r_1' r_2' = (r_1 + a_1)(r_2 + a_2) = r_1 r_2 + a_1 r_2 + r_1 a_2 + a_1 a_2$$

Hence $(r_1' r_2') + I = (r_1 r_2) + I$. The remainder of the proof is trivial. □

Example. In the integers $\mathbb{Z}$, the ideals are $n\mathbb{Z}$. Hence we can form the quotient ring ${}^{\mathbb{Z}}\!/_{n\mathbb{Z}}$. The ring ${}^{\mathbb{Z}}\!/_{n\mathbb{Z}}$ has elements $n\mathbb{Z}, 1 + n\mathbb{Z}, \ldots, (n-1) + n\mathbb{Z}$. Addition and multiplication behave like in modular arithmetic modulo n.

Example. Consider the ideal (X) inside the polynomial ring $\mathbb{C}[X]$. This ideal is the set of polynomials with zero constant term. Let $f(X) = a_n X^n + \cdots + a_0$ be an arbitrary element of $\mathbb{C}[X]$. Then $f(X) + X = a_0 + X$. Thus, there exists a bijection between ${}^{\mathbb{C}[X]}\!/_{(X)}$ and $\mathbb{C}$, defined by $f(x) + (X) \mapsto f(0)$, with inverse $a \mapsto a + (X)$. This bijection is a ring homomorphism, hence ${}^{\mathbb{C}[X]}\!/_{(X)} \cong \mathbb{C}$.

Example. Consider $(X^2 + 1) \lhd \mathbb{R}[X]$. For $f(X) = a_n X^n + \cdots + a_0 \in \mathbb{R}[X]$, we can apply the Euclidean algorithm to write $f(X)$ as $q(X)(X^2 + 1) + r(X)$ where the degree of r is less than two. Hence $r(X) = a + bX$ for some real numbers a and b. Thus, any element of ${}^{\mathbb{R}[X]}\!/_{(X^2+1)}$ can be written $a+bX+(X^2+1)$. Suppose a coset can be represented by two representatives: $a+bX+(X^2+1) = a' + b'X + (X^2 + 1)$. Then,

$$a + bX - a' - b'X = (a - a') - (b - b')X = g(X)(X^2 + 1)$$

Hence $g(X) = 0$, giving $a - a' = 0$ and $b - b' = 0$. Hence the coset representative is unique. Consider the bijection φ between this quotient ring and the complex numbers given by $a+bX+(X^2+1) \mapsto a+bi$. We can show that φ is a ring homomorphism. Indeed, it preserves addition, and $1 + (X^2 + 1) \mapsto 1$, so it suffices to check that multiplication is preserved.

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

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

8.6 Isomorphism theorems

Theorem (first isomorphism theorem)**.** Let $\varphi : R \to S$ be a ring homomorphism. Then,

$$\ker\varphi \lhd R; \quad \operatorname{Im}\varphi \leq S; \quad {}^{R}\!/_{\ker\varphi} \cong \operatorname{Im}\varphi$$

Proof. We have $\ker\varphi \lhd R$ from above. We know that $\operatorname{Im}\varphi \leq (S,+)$. Now we show that $\operatorname{Im}\varphi$ is closed under multiplication.

$$\varphi(r_1)\varphi(r_2) = \varphi(r_1r_2) \in \operatorname{Im}\varphi$$

Finally,

$$1_S = \varphi(1_R) \in \operatorname{Im}\varphi$$

Hence $\operatorname{Im}\varphi$ is a subring of S. Let $K = \ker\varphi$. Then, we define $\Phi : {}^{R}\!/_{K} \to \operatorname{Im}\varphi$ by $r + K \mapsto \varphi(r)$. By appealing to the first isomorphism theorem from groups, this is well-defined, a bijection, and a group homomorphism under addition. It therefore suffices to show that Φ preserves multiplication and maps the multiplicative identities to each other.

$$\Phi(1_R + K) = \varphi(1_R) = 1_S; \quad \Phi((r_1+K)(r_2+K)) = \Phi(r_1r_2+K) = \varphi(r_1r_2) = \varphi(r_1)\varphi(r_2)$$

The result follows as required. □

Theorem (second isomorphism theorem)**.** Let $R \leq S$ and $J \lhd S$. Then,

$$R \cap J \lhd R; \quad R + J = \{r + a : r \in R, a \in J\} \leq S; \quad {}^{R}\!/_{R\cap J} \cong {}^{R+J}\!/_{J} \leq {}^{S}\!/_{J}$$

Proof. By the second isomorphism theorem for groups, $R + J \leq (S,+)$. Further, $1_S = 1_S + 0_S$, and since R is a subring, $1_S + 0_S \in R + J$ hence $1_S \in R \cap J$. If $r_1, r_2 \in R$ and $a_1, a_2 \in J$, we have

$$(r_1 + a_1)(r_2 + a_2) = \underbrace{r_1r_2}_{\in R} + \underbrace{r_1a_2}_{\in J} + \underbrace{r_2a_1}_{\in J} + \underbrace{r_2a_2}_{\in J} \in R + J$$

Hence $R + J$ is closed under multiplication, giving $R + J \leq S$.

Let $\varphi : R \to {}^{S}\!/_{J}$ be defined by $r \mapsto r + J$. This is a ring homomorphism, since it is the composite of the inclusion homomorphism $R \subseteq S$ and the quotient map $S \to {}^{S}\!/_{J}$. The kernel of φ is the set $\{r \in R : r + J = J\} = R \cap J$. Since this is the kernel of a ring homomorphism, $R \cap J$ is an ideal in R. The image of φ is $\{r + J \mid r \in R\} = {}^{R+J}\!/_{J} \leq {}^{S}\!/_{J}$. By the first isomorphism theorem, ${}^{R}\!/_{R\cap J} \cong {}^{R+J}\!/_{J}$ as required. □

Remark. If $I \lhd R$, there exists a bijection between ideals in ${}^{R}\!/_{I}$ and the ideals of R containing I. Explicitly,

$$K \mapsto \{r \in R \mid r + I \in K\}; \quad J \mapsto {}^{J}\!/_{I}$$

Theorem (third isomorphism theorem)**.** Let $I \lhd R$ and $J \lhd R$ with $I \subseteq J$. Then,

$${}^{J}\!/_{I} \lhd {}^{R}\!/_{I}; \quad {}^{R/I}\!/_{J/I} \cong {}^{R}\!/_{J}$$

Proof. Let $\varphi : {}^{R}\!/_{I} \to {}^{R}\!/_{J}$ defined by $r + I \mapsto r + J$. We can check that this is a surjective ring homomorphism by considering the third isomorphism theorem for groups. Its kernel is $\{r + I : r \in J\} = {}^{J}\!/_{I}$, which is an ideal in ${}^{R}\!/_{I}$, and we conclude by use of the first isomorphism theorem. □

Remark. ${}^{J}\!/_{I}$ is not a quotient ring, since J is not in general a ring; this notation should be interpreted as a set of cosets.

Example. Consider the surjective ring homomorphism $\varphi : \mathbb{R}[X] \to \mathbb{C}$ which is defined by

$$f = \sum_n a_n X^n \mapsto f(i) = \sum_n a_n i^n$$

Its kernel can be found by the Euclidean algorithm, yielding $\ker \varphi = (X^2 + 1)$. Applying the first isomorphism theorem, we immediately find ${}^{\mathbb{R}[X]}\!/_{(X^2+1)} \cong \mathbb{C}$.

Example. Let R be a ring. Then there exists a unique ring homomorphism $i : \mathbb{Z} \to R$. Indeed, we must have

$$0_{\mathbb{Z}} \mapsto 0_R; \quad 1_{\mathbb{Z}} \mapsto 1_R$$

This inductively defines

$$n \mapsto \underbrace{1_R + \cdots + 1_R}_{n \text{ times}}$$

The negative integers are also uniquely defined, since any ring homomorphism is a group homomorphism.

$$-n \mapsto -(\underbrace{1_R + \cdots + 1_R}_{n \text{ times}})$$

We can show that any such construction is a ring homomorphism as required. Then, the kernel of the ring homomorphism is an ideal of $\mathbb{Z}$, hence it is $n\mathbb{Z}$ for some n. Hence, by the first isomorphism theorem, any ring contains a copy of ${}^{\mathbb{Z}}\!/_{n\mathbb{Z}}$, since it is isomorphic to the image of i. If $n = 0$, then the ring contains a copy of $\mathbb{Z}$ itself, and if $n = 1$, then the ring is trivial since $0 = 1$. The number n is known as the *characteristic* of R.

For example, $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ have characteristic zero. The rings ${}^{\mathbb{Z}}\!/_{p\mathbb{Z}}, {}^{\mathbb{Z}}\!/_{p\mathbb{Z}}[X]$ have characteristic p.

8.7 Integral domains

Definition. An *integral domain* is a ring R with $0 \neq 1$ such that for all $a, b \in R$, $ab = 0$ implies $a = 0$ or $b = 0$. A *zero divisor* in a ring R is a nonzero element $a \in R$ such that $ab = 0$ for some nonzero $b \in R$. A ring is an integral domain if and only if it has no zero divisors.

Example. All fields are integral domains. Any subring of an integral domain is an integral domain. For instance, $\mathbb{Z}[i] \leq \mathbb{C}$ is an integral domain.

Example. The ring $\mathbb{Z} \times \mathbb{Z}$ is not an integral domain. Indeed, $(1, 0) \cdot (0, 1) = (0, 0)$.

Lemma. Let R be an integral domain. Then $R[X]$ is an integral domain.

Proof. We will show that any two nonzero elements produce a nonzero element. In particular, let

$$f = \sum_n a_n X^n; \quad g = \sum_n b_n X^n$$

Since these are nonzero, the leading coefficients a_n and b_m are nonzero. Here, the leading term of the product fg has form $a_n b_m X^{n+m}$. Since R is an integral domain, $a_n b_m \neq 0$, so fg is nonzero. Further, the degree of fg is $n+m$, the sum of the degrees of f and g. □

Lemma. Let R be an integral domain, and $f \neq 0$ be a nonzero polynomial in $R[X]$. We define $\text{roots}(f) = \{a \in R : f(a) = 0\}$. Then $|\text{roots}(f)| \leq \deg(f)$.

Proof. Exercise on the example sheets. □

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

Proof. G is a finite abelian group. If G is not cyclic, we can apply a previous structure theorem for finite abelian groups to show that there exists $H \leq G$ such that $H \cong C_{d_1} \times C_{d_1}$ for some integer $d_1 \geq 2$. The polynomial $f(X) = X^{d_1} - 1 \in F[X]$ has degree d_1, but has at least d_1^2 roots, since any element of H is a root. This contradicts the previous lemma. □

Example. $\left(\mathbb{Z}/p\mathbb{Z}\right)^\times$ is cyclic.

Proposition. Any finite integral domain is a field.

Proof. Let $0 \neq a \in R$, where R is an integral domain. Consider the map $\varphi : R \to R$ given by $x \mapsto ax$. If $\varphi(x) = \varphi(y)$, then $a(x-y) = 0$. But $a \neq 0$, hence $x - y = 0$. Hence φ is injective. Since R is finite, φ is a bijection, hence it has an inverse φ^{-1}, which yields the multiplicative inverse of a by considering $\varphi^{-1}(a)$. This may be repeated for all a. □

Theorem. Any integral domain R is a subring of a field F, and every element of F can be written in the form ab^{-1} where $a, b \in R$ and $b \neq 0$. Such a field F is called the *field of fractions* of R.

Proof. Consider the set $S = \{(a, b) \in R : b \neq 0\}$. We can define an equivalence relation

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

This is reflexive and commutative. We can show directly that it is transitive.

$$\begin{aligned}(a, b) \sim (c, d) \sim (e, f) &\implies ad = bc;\ cf = de \\ &\implies adf = bcf = bde \\ &\implies af = be \\ &\implies (a, b) \sim (e, f)\end{aligned}$$

Hence $\sim$ is indeed an equivalence relation. Now, let $F = S/\sim$, and we write $\frac{a}{b}$ for the class $[(a,b)]$. We define the ring operations

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

These can be shown to be well-defined. Thus, F is a ring with identities $0_F = \frac{0_R}{1_R}$ and $1_F = \frac{1_R}{1_R}$. If $\frac{a}{b} \neq 0_F$, then $a \neq 0$. Thus, $\frac{b}{a}$ exists, and $\frac{a}{b} \cdot \frac{b}{a} = 1$. Hence F is a field.

We can identify R with the subring of F given by $\frac{r}{1}$ for all $r \in R$. This is clearly isomorphic to R. Further, any element of F can be written as $\frac{a}{b} = ab^{-1}$ as required. □

This is analogous to the construction of the rationals using the integers.

Example. Consider $\mathbb{C}[X]$. This has field of fractions $\mathbb{C}(X)$, called the field of *rational functions* in X.

8.8 Maximal ideals

Definition. An ideal $I \lhd R$ is *maximal* if $I \neq R$ and, if $I \subseteq J \lhd R$, we have $J = I$ or $J = R$.

Lemma. A nonzero ring R is a field if and only if its only ideals are zero or R.

Proof. Suppose R is a field. If $0 \neq I \lhd R$, then I contains a nonzero element, which is a unit since R is a field. Hence $I = R$.

Now, suppose a ring R has ideals that are only zero or R. If $0 \neq x \in R$, consider (x). This is nonzero since it contains x. By assumption, $(x) = R$. Thus, the element 1 lies in (x). Hence, there exists $y \in R$ such that $xy = 1$, and hence this y is the multiplicative inverse as required. □

Proposition. Let $I \lhd R$. Then I is maximal if and only if R/I is a field.

Proof. R/I is a field if and only if its ideals are either zero, denoted I/I, or R/I itself. By the correspondence theorem, I and R are the only ideals in R which contain I. Equivalently, $I \lhd R$ is maximal. □

8.9 Prime ideals

Definition. An ideal $I \lhd R$ is *prime* if $I \neq R$ and, for all $a, b \in R$ such that $ab \in I$, we have $a \in I$ or $b \in I$.

Example. The ideals in the integers are (n) for some $n \geq 0$. $n\mathbb{Z}$ is a prime ideal if and only if n is prime or zero. The case for $n = 0$ is trivial. If $n \neq 0$ we can use the property that $p \mid ab$ implies either $p \mid a$ or $p \mid b$. Conversely, if n is composite, we can write $n = uv$ for $u, v > 1$. Then $uv \in n\mathbb{Z}$ but $u, v \notin n\mathbb{Z}$.

Proposition. Let $I \lhd R$. Then I is prime if and only if R/I is an integral domain.

Proof. If I is prime, then for all $ab \in I$ we have $a \in I$ or $b \in I$. Equivalently, for all $a+I, b+I \in R/I$, we have $(a+I)(b+I) = 0+I$ if $a+I = 0+I$ or $b+I = 0+I$. This is the definition of an integral domain. □

Remark. If I is a maximal ideal, then R/I is a field. A field is an integral domain. Hence any maximal ideal is prime.

Remark. If the characteristic of a ring is n, then $\mathbb{Z}/n\mathbb{Z} \leq R$. In particular, if R is an integral domain, then $\mathbb{Z}/n\mathbb{Z}$ must be an integral domain. Equivalently, $n\mathbb{Z} \lhd \mathbb{Z}$ is a prime ideal. Hence n is zero or prime. Thus, in an integral domain, the characteristic must either be zero or prime. A field always has a characteristic, which is either zero (in which case it contains $\mathbb{Z}$ and hence $\mathbb{Q}$) or prime (in which case it contains $\mathbb{Z}/p\mathbb{Z} = \mathbb{F}_p$ which is already a field).

9 Factorisation in integral domains

In this section, let R be an integral domain.

9.1 Prime and irreducible elements

Recall that an element $a \in R$ is a unit if it has a multiplicative inverse in R. Equivalently, an element a is a unit if and only if $(a) = R$. Indeed, if $(a) = R$, then $1 \in (a)$ hence there exists a multiple of a equal to 1. We denote the set of units in R by $R^\times$.

Definition. An element $a \in R$ *divides* $b \in R$, written $a \mid b$, if there exists $c \in R$ such that $b = ac$. Equivalently, $(b) \subseteq (a)$.
Two elements $a, b \in R$ are *associates* if $a = bc$ where c is a unit. Informally, the two elements differ by multiplication by a unit. Equivalently, $(a) = (b)$.

Definition. An element $r \in R$ is *irreducible* if r is not zero or a unit, and $r = ab$ implies a is a unit or b is a unit. An element $r \in R$ is *prime* if r is not zero or a unit, and $r \mid ab$ implies $r \mid a$ or $r \mid b$.

Remark. These properties depend on the ambient ring R; for instance, 2 is prime and irreducible in $\mathbb{Z}$, but neither prime nor irreducible in $\mathbb{Q}$. The polynomial $2X$ is irreducible in $\mathbb{Q}[X]$, but not in $\mathbb{Z}[X]$.

Lemma. $(r) \lhd R$ is a prime ideal if and only if $r = 0$ or r is prime.

Proof. Suppose (r) is a prime ideal with $r \neq 0$. Since prime ideals are proper, r cannot be a unit. Suppose $r \mid ab$, or equivalently, $ab \in (r)$. By the definition of a prime ideal, $a \in (r)$ or $b \in (r)$. Hence, $r \mid a$ or $r \mid b$. By definition of a prime element, r is prime.

Conversely, first note that the zero ideal $(0) = \{0\}$ is a prime ideal, since R is an integral domain. Suppose r is prime. We know $(r) \neq R$ since r is not a unit. If $ab \in (r)$, then $r \mid ab$, so $r \mid a$ or $r \mid b$, giving $a \in (r)$ or $b \in (r)$ as required for (r) to be a prime ideal. □

Lemma. Prime elements are irreducible.

Proof. Let r be prime. Then r is nonzero and not a unit. Suppose $r = ab$. Then, in particular, $r \mid ab$, so $r \mid a$ or $r \mid b$ by primality. Let $r \mid a$ without loss of generality. Hence $a = rc$ for some element $c \in R$. Then, $r = ab = rcb$, so $r(1-cb) = 0$. Since R is an integral domain, and $r \neq 0$, we have $cb = 1$, so b is a unit. □

Example. The converse does not hold in general. Let

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

Since R is a subring of the field $\mathbb{C}$, it is an integral domain. We can define the *norm* $N : R \to \mathbb{Z}$ by $N(a + b\sqrt{-5}) = a^2 + 5b^2 \geq 0$. Note that this norm is multiplicative: $N(z_1 z_2) = N(z_1)N(z_2)$.

We claim that the units are exactly ± 1. Indeed, if $r \in R^\times$, then $rs = 1$ for some element $s \in R$. Then, $N(r)N(s) = N(1) = 1$, so $N(r) = N(s) = 1$. But the only elements $r \in R$ with $N(r) = 1$ are $r = \pm 1$.

We will now show that the element $2 \in R$ is irreducible. Suppose $2 = rs$ for $r, s \in R$. By the multiplicative property of N, $N(2) = 4 = N(r)N(s)$ can only be satisfied by $N(r), N(s) \in \{1, 2, 4\}$. Since $a^2 + 5b^2 = 2$ has no integer solutions, R has no elements of norm 2. Hence, either r or s has unit norm and is thus a unit by the above discussion. We can show similarly that $3, 1+\sqrt{-5}, 1-\sqrt{-5}$ are irreducible, as there exist no elements of norm 3.

We can now compute directly that $(1 + \sqrt{-5})(1 - \sqrt{-5}) = 6 = 2 \cdot 3$, hence $2 \mid (1 + \sqrt{-5})(1 - \sqrt{-5})$. But $2 \nmid (1 + \sqrt{-5})$ and $2 \nmid (1 - \sqrt{-5})$, which can be checked by taking norms. Hence, 2 is irreducible but not a prime.

In order to construct this example, we have exhibited two factorisations of 6 into irreducibles: $(1 + \sqrt{-5})(1 - \sqrt{-5}) = 6 = 2 \cdot 3$. Since $R^\times = \{\pm 1\}$, these irreducibles in the factorisations are not associates.

9.2 Principal ideal domains

Definition. An integral domain R is a *principal ideal domain* if all ideals are principal ideals. In other words, for all ideals I, there exists an element r such that $I = (r)$.

Example. $\mathbb{Z}$ is a principal ideal domain.

Proposition. In a principal ideal domain, all irreducible elements are prime.

Proof. Let $r \in R$ be irreducible, and suppose $r \mid ab$. If $r \mid a$, the proof is complete, so suppose $r \nmid a$. Since R is a principal ideal domain, the ideal (a, r) is generated by a single element $d \in R$. In particular, since $r \in (d)$, we have $d \mid r$ so $r = cd$ for some $c \in R$.

Since r is irreducible, either c or d is a unit. If c is a unit, $(a,r) = (r)$, so in particular $r \mid a$, which contradicts the assumption that $r \nmid a$, so c cannot be a unit. Thus, d is a unit. In this case, $(a,r) = R$. By definition of (a,r), there exist $s,t \in R$ such that $1 = sa + tr$. Then, $b = sab + trb$. We have $r \mid sab$ since $r \mid ab$, and we know $r \mid trb$. Hence $r \mid b$ as required. □

Lemma. Let R be a principal ideal domain. Then an element r is irreducible if and only if (r) is maximal.

Proof. Suppose r is irreducible. Since r is not a unit, $(r) \neq R$. Suppose $(r) \subseteq J \subseteq R$ where J is an ideal in R. Since R is a principal ideal domain, $J = (a)$ for some $a \in R$. In particular, $r = ab$ for some $b \in R$, since $(r) \subseteq J$. Since r is irreducible, either a or b is a unit. But if a is a unit, we have $J = R$. If b is a unit, then a and r are associates so they generate the same ideal. Hence, (r) is maximal.

Conversely, suppose (r) is maximal. Note that r is not a unit, since $(r) \neq R$. Suppose $r = ab$. Then $(r) \subseteq (a) \subseteq R$. But since (r) is maximal, either $(a) = (r)$ or $(a) = R$. If $(a) = (r)$, then b is a unit. If $(a) = R$, then a is a unit. Hence r is irreducible. Note that this direction of the proof did not require that R was a principal ideal domain, however R must still be an integral domain. □

Remark. Let R be a principal ideal domain, and $0 \neq r \in R$. Then, (r) is maximal if and only if r is irreducible, which is true if and only if r is prime, which is equivalent to the fact that (r) is prime. Hence, the maximal ideals are the nonzero prime ideals.

Definition. An integral domain is a *Euclidean domain* if there exists a function $\varphi : R \setminus \{0\} \to \mathbb{Z}_{\geq 0}$ such that, for all $a, b \in R$.
(i) if $a \mid b$ then $\varphi(a) \leq \varphi(b)$;
(ii) if $b \neq 0$ then $\exists q, r \in R$ such that $a = bq + r$ and either $r = 0$ or $\varphi(r) < \varphi(b)$.
Such a φ is called a *Euclidean function*.

Example. $\mathbb{Z}$ is a Euclidean domain, where the Euclidean function φ is the absolute value function.

Proposition. Euclidean domains are principal ideal domains.

Proof. Let R have Euclidean function φ. Let $I \lhd R$ be a nonzero ideal. Let $b \in I \setminus \{0\}$ that minimises $\varphi(b)$. Then $(b) \subseteq I$. For any element $a \in I$, we can use the Euclidean algorithm to show $a = bq + r$ where $r = 0$ or $\varphi(r) < \varphi(b)$. But since $r = a - bq \in I$, $\varphi(r)$ cannot be lower than the minimal element $\varphi(b)$. Thus $r = 0$, so $a = bq$. Hence, $I = (b)$, so all ideals are principal. □

Remark. In the above proof, only the second property of the Euclidean function was used. The first property is included in the definition since it will allow us to easily describe the units in the ring.

$$R^\times = \{u \in R : u \neq 0, \varphi(u) = \varphi(1)\}$$

It can be shown that, if there exists a function φ satisfying (ii), there exists a (possibly not unique) function φ' satisfying (i) and (ii).

Example. Let F be a field. Then $F[X]$ is a Euclidean domain with Euclidean function $\varphi(f) = \deg(f)$. We have already proven the requisite properties of Euclidean functions.

The ring $R = \mathbb{Z}[i]$ is a Euclidean domain with $\varphi(u + iv) = N(u + iv) = u^2 + v^2$. Since the norm is multiplicative, $N(zw) = N(z)N(w)$ which immediately gives property (i) in the definition. Consider $z, w \in \mathbb{Z}[i]$ where $w \neq 0$. Consider $\frac{z}{w} \in \mathbb{C}$. This has distance less than 1 from the nearest element q of R. Let $r = z - wq \in R$. Then $z = wq + r$ where

$$\varphi(r) = |r|^2 = |z - wq|^2 < |w|^2 = \varphi(w)$$

So property (ii) is satisfied.

Hence $F[X]$ and $\mathbb{Z}[i]$ are principal ideal domains.

Example. Let A be a nonzero $n \times n$ matrix over a field F. Let $I = \{f \in F[X] : f(A) = 0\}$. I is an ideal. Indeed, if $f, g \in I$, then $(f - g)(A) = f(A) - g(A) = 0$, and for $f \in I$ and $g \in F[X]$, we have $(f \cdot g)(A) = f(A) \cdot g(A) = 0$ as required. Since $F[X]$ is a principal ideal domain, $I = (f)$ for some polynomial $f \in F[X]$. All units in $F[X]$ are the nonzero constant polynomials. Hence, the polynomial of smallest degree in I is unique up to multiplication by a unit, so without loss of generality we may assume f is monic. This yields the minimal polynomial of A.

Example. Let $\mathbb{F}_2$ be the finite field of order 2, which is isomorphic to $\mathbb{Z}/2\mathbb{Z}$. Let $f(X)$ be the polynomial $X^3 + X + 1 \in \mathbb{F}_2[X]$.

We claim that f is irreducible. Suppose $f = gh$ where the degrees of g, h are positive. Since the degree of f is 3, one of g, h must have degree 1. Hence f has a root. But we can check that $f(0) = f(1) = 1$ so f has no root in $\mathbb{F}_2$. Hence f is irreducible as required.

Since $\mathbb{F}_2[X]$ is a principal ideal domain, we have that $(f) \lhd \mathbb{F}_2[X]$ is a maximal ideal. Hence, $\mathbb{F}_2[X]/(f)$ is a field. We can verify that this field has order 8, using the Euclidean algorithm. Any element in this quotient has coset representative $aX^2 + bX + c$ for $a, b, c \in \mathbb{F}_2$. We can show that all 8 of these possibilities yields different polynomials. So we have constructed a field of order 8. This technique will be explored further in Part II Galois Theory.

Example. The ring $\mathbb{Z}[X]$ is not a principal ideal domain. Consider the ideal $I = (2, X) \lhd \mathbb{Z}[X]$. We can write

$$I = \{2f_1(X) + Xf_2(X) : f_1, f_2 \in \mathbb{Z}[X]\} = \{f \in \mathbb{Z}[X] : 2 \mid f(0)\}$$

Suppose $I = (f)$ for some element f. Since $2 \in I$, we must have $2 = fg$ for some polynomial g. By comparing degrees, the degrees of f and g must be zero, since $\mathbb{Z}$ is an integral domain. Hence f is an integer, so $f = \pm 1$ or $f = \pm 2$. If $f = \pm 1$ then $I = \mathbb{Z}[X]$, and if $f = \pm 2$ then $I = 2\mathbb{Z}[X]$. These both lead to contradictions, since $1 \notin I$ and $X \in I$.

9.3 Unique factorisation domains

Definition. An integral domain is a *unique factorisation domain* if
(i) every nonzero, non-unit element is a product of irreducibles;
(ii) if $p_1 \cdots p_m = q_1 \cdots q_n$ where p_i, q_i are irreducible, then $m = n$, and p_i, q_i are associates, up to reordering.

Proposition. Let R be an integral domain satisfying property (i) above (every nonzero, non-unit element is a product of irreducibles). Then R is a unique factorisation domain if and only if every irreducible is prime.

Proof. Suppose R is a unique factorisation domain. Let $p \in R$ be irreducible, and $p \mid ab$. Then $ab = pc$ for some $c \in R$. Writing a, b, c as products of irreducibles, it follows from uniqueness of factorisation that $p \mid a$ or $p \mid b$. Hence p is prime.

Conversely, suppose every irreducible is prime. Suppose $p_1 \cdots p_m = q_1 \cdots q_n$ where p_i, q_i are irreducible and hence prime. Since $p_1 \mid q_1 \cdots q_n$, we have $p_1 \mid q_i$ for some i. After reordering, we may assume that $p_1 \mid q_1$, so $p_1 u = q_1$ for $u \in R$. Since q_1 is irreducible, u is a unit since p_1 cannot be a unit. Hence p_1, q_1 are associates. Cancelling p_1 from both sides, we find $p_2 \cdots p_m = u q_2 \cdots q_n$. We may absorb this unit into q_2 without loss of generality. Inductively, all p_i and q_i are associates, for each i. Hence R is a unique factorisation domain. □

Definition. Let R be a ring. Suppose, for all nested sequences of ideals in R written $I_1 \subseteq I_2 \subseteq \cdots$, there exists N such that $I_n = I_{n+1}$ for all $n \geq N$. Then, we say that R is a *Noetherian* ring.

This condition is known as the 'ascending chain condition'. In other words, we cannot infinitely nest distinct ideals in a Noetherian ring.

Lemma. Principal ideal domains are Noetherian rings.

Proof. Let $I = \bigcup_{i=1}^{\infty} I_i$. Then, I is an ideal in R. Since R is a principal ideal domain, $I = (a)$ for some $a \in R$. Then $a \in \bigcup_{i=1}^{\infty} I_i$, so in particular $a \in I_N$ for some N. But then for all $n \geq N$, $(a) \subseteq I_N \subseteq I_n \subseteq I_{n+1} \subseteq I = (a)$. So all inclusions are equalities, so in particular $I_n = I_{n+1}$. □

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

Proof. First, we verify property (i), that every nonzero, non-unit element is a product of irreducibles. Let $x \neq 0$ be an element of R which is not a unit. Suppose x does not factor as a product of irreducibles. This implies in particular that x is not irreducible. By definition, we can write x as the product of two elements 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, so without loss of generality we can suppose x_1 is not a product of irreducibles. We have $(x) \subset (x_1)$. This inclusion is strict, since y_1 is not a unit. Now, we can write $x_1 = x_2 y_2$ where x_2 is not a unit, and inductively we can create $(x) \subset (x_1) \subset (x_2) \subset \cdots$. But R is Noetherian, so this is a contradiction. So every nonzero, non-unit element is indeed a product of irreducibles.

By the proposition above, it suffices to show that every irreducible is prime. This has already been shown previously. Hence R is a unique factorisation domain. □

Example. We have shown that all Euclidean domains are principal ideal domains, and all principal ideal domains are unique factorisation domains, and all unique factorisation domains are integral domains. We now provide examples for counterexamples to the converses.

The ring $\mathbb{Z}/4\mathbb{Z}$ is not an integral domain since 2 is a zero divisor.

The ring $\mathbb{Z}[\sqrt{-5}] \leq \mathbb{C}$ is integral, but not a unique factorisation domain.

The ring $\mathbb{Z}[X]$ has been shown to be not a principal ideal domain. We can show using later results that this is a unique factorisation domain.

We can construct the ring $\mathbb{Z}\left[\frac{1+\sqrt{-19}}{2}\right]$, which can be shown to be not a Euclidean domain, but is a principal ideal domain. This proof is beyond the scope of Part IB Groups, Rings and Modules, but will be proved in Part II Number Fields.

Finally, $\mathbb{Z}[i]$ is a Euclidean domain, and is hence a principal ideal domain, a unique factorisation domain, and an integral domain.

Definition. Let R be an integral domain.

(i) $d \in R$ is a *common divisor* of $a_1, \dots, a_n \in R$ if $d \mid a_i$ for all i;
(ii) $d \in R$ is a *greatest common divisor* of $a_1, \dots, a_n$ if for all common divisors d', we have $d' \mid d$;
(iii) $m \in R$ is a *common multiple* of $a_1, \dots, a_n$ if $a_i \mid m$ for all i;
(iv) $m \in R$ is a *least common multiple* of $a_1, \dots, a_n$ if for all common multiples m', we have $m \mid m'$.

Remark. Greatest common divisors and lowest common multiples are unique up to associates, if they exist.

Proposition. In unique factorisation domains, greatest common divisors and least common multiples always exist.

Proof. Let $a_i = u_i \prod_j p_j^{n_{ij}}$ where the p_j are irreducible and pairwise non-associate, u_i is a unit, and $n_{ij} \in \mathbb{Z}_{\geq 0}$. We claim that $d = \prod_j p_j^{m_j}$, where $m_j = \min_{1 \leq i \leq n} n_{ij}$, is the greatest common divisor. Certainly d is a common divisor. If d' is a common divisor, then d' can be written as a product of irreducibles, which will be denoted $d' = w \prod_j p_i^{t_j}$. We can see that $t_j \leq n_{ij}$ for all i, so in particular, $t_j \leq m_j$. This implies $d' \mid d$. Hence d is a greatest common divisor. The argument for the least common multiple is similar, replacing minima with maxima. □

9.4 Factorisation in polynomial rings

Theorem. Let R be a unique factorisation domain. Then $R[X]$ is also a unique factorisation domain.

The proof for this theorem will require a number of key lemmas. In this subsection, R will denote a unique factorisation domain, with field of fractions F. We have $R[X] \leq F[X]$. Since polynomial rings over fields are Euclidean domains, $F[X]$ is a principal ideal domain, and hence a unique factorisation domain. This does not immediately imply that $R[X]$ is a unique factorisation domain, however.

Definition. The *content* of a polynomial $f = \sum_{i=0}^{n} a_i X^i \in R[X]$ is $c(f) = \gcd\{a_0, \dots, a_n\}$. This is well-defined up to multiplication by a unit.
We say that f is *primitive* if $c(f)$ is a unit.

Lemma. The product of primitive polynomials is primitive. Further, for $f, g \in R[X]$, $c(fg)$ and $c(f)c(g)$ are associates.

Proof. Let $f = \sum_{i=0}^{n} a_i X^i$ and $g = \sum_{i=0}^{m} b_i X^i$. Suppose fg is not primitive, so $c(fg)$ is not a unit. This implies that there exists a prime p such that $p \mid c(fg)$. Since f, g are primitive, $p \nmid c(f)$ and $p \nmid c(g)$.

Suppose p does not divide all of the a_k or the b_ℓ. Let k, ℓ be the smallest values such that $p \nmid a_k$ and $p \nmid b_\ell$. Then, the coefficient of $X^{k+\ell}$ in fg is given by

$$\sum_{i+j=k+\ell} a_i b_j = \underbrace{\cdots + a_{k-1} b_{\ell+1}}_{\text{divisible by } p} + a_k b_\ell + \underbrace{a_{k+1} b_{\ell-1} + \cdots}_{\text{divisible by } p}$$

Thus $p \mid a_k b_\ell$. This is a contradiction as we have $p \mid a_k$ or $p \mid b_\ell$.

To prove the second part, let $f = c(f) f_0$ for some $f_0 \in R[X]$. Here, f_0 is primitive. Similarly, $g = c(g) g_0$ for a primitive g_0. Thus $fg = c(f)c(g) f_0 g_0$. The expression $f_0 g_0$ is a primitive polynomial by the first part, so $c(fg)$ is equal to $c(f)c(g)$ up to associates. □

Corollary. If $p \in R$ is prime in R, then p is prime in $R[X]$.

Proof. Since R is an integral domain, we have $R[X]^\times = R^\times$, so p is not a unit. Let $f \in R[X]$. Then $p \mid f$ in $R[X]$ if and only if $p \mid c(f)$ in R. Thus, if $p \mid gh$ in $R[X]$, we have $p \mid c(gh) = c(g)c(h)$. In particular, since p is prime in R, we have $p \mid c(g)$ or $p \mid c(h)$, so $p \mid g$ or $p \mid h$. So p is prime in $R[X]$. □

Lemma. Let $f, g \in R[X]$, where g is primitive. Then if $g \mid f$ in $F[X]$, then $g \mid f$ in $R[X]$.

Proof. Let $f = gh$, where $h \in F[X]$. We can find a nonzero $a \in R$, such that $ah \in R[X]$. In particular, we can multiply the denominators of the coefficients of h to form a. Now, $ah = c(ah)h_0$ where h_0 is primitive. Then $af = c(ah)h_0 g$. Since h_0 and g are primitive, so is $h_0 g$. Thus, taking contents, $a \mid c(ah)$. This implies $h \in R[X]$. Hence $g \mid f$ in $R[X]$. □

Lemma (Gauss' lemma)**.** Let $f \in R[X]$ be primitive. Then if f is irreducible in $R[X]$, we have that f is irreducible in $F[X]$.

Proof. Since $f \in R[X]$ is irreducible and primitive, its degree must be larger than zero. Hence f is not a unit in $F[X]$. Suppose f is not irreducible in $F[X]$, so $f = gh$ for $g, h \in F[X]$ with degree larger than zero. Let $\lambda \in F^\times$ such that $\lambda^{-1} g \in R[X]$ is primitive. For example, let $b \in R$ such that $bg \in R[X]$ to clear denominators, then $bg = c(bg)g_0$, giving $\lambda = c(bg)b^{-1}$. Replacing g by $\lambda^{-1} g$ and h by λh, we still have a factorisation of f. Hence, we may assume without loss of generality that $g \in R[X]$ and

is primitive. By the previous lemma, we have that $h \in R[X]$, with degrees larger than zero. This contradicts irreducibility. □

Remark. We will see that the reverse implication in Gauss' lemma also holds.

Lemma. Let $g \in R[X]$ be primitive. If g is prime in $F[X]$, then g is prime in $R[X]$.

Proof. It suffices to show that if $f_1, f_2 \in R[X]$, then $g \mid f_1 f_2$ implies $g \mid f_1$ or $g \mid f_2$. Since g is prime in $F[X]$, $g \mid f_1$ or $g \mid f_2$ in $F[X]$. By the previous lemma, $g \mid f_1$ or $g \mid f_2$ in $R[X]$ as required. □

We can now prove the first theorem of this subsection, that polynomial rings over unique factorisation domains are unique factorisation domains.

Proof. Let $f \in R[X]$. Then, $f = c(f)f_0$ for f_0 primitive in $R[X]$. Since R is a unique factorisation domain, $c(f)$ is a product of irreducibles in R. If an element of R is irreducible, it is irreducible as an element of $R[X]$. Hence, it suffices to find a factorisation of f_0.

Suppose f_0 is not irreducible, so $f_0 = gh$ for $g, h \in R[X]$. Since f_0 is primitive, g and h are primitive, and the degrees of g, h are larger than zero. By induction on the degree, we can factor f_0 as a product of primitive irreducibles in $R[X]$.

It now suffices to show uniqueness of the factorisation. By a previous proposition, it in fact suffices to show that every irreducible element of $R[X]$ is prime. Let f be irreducible. Write $f = c(f)f_0$, where f_0 is primitive. Since f is irreducible, f must be constant or primitive.

Suppose f is constant. Since f is irreducible in $R[X]$, it must be irreducible in R. As R is a unique factorisation domain, f is prime in R. By a previous corollary, f is prime in $R[X]$.

Now, suppose f is primitive. Since f is irreducible in $R[X]$, we can use Gauss' lemma to show that f is irreducible in $F[X]$. Thus, f is prime in $F[X]$, as $F[X]$ is a unique factorisation domain. Finally, we can see that f is prime in $R[X]$ by the previous lemma. □

Remark. We know that the prime elements in an integral domain are irreducible. This implies that the implications in the last paragraph above are in fact equivalences. In particular, in Gauss' lemma, the implication is an equivalence.

Example. The above theorem implies that $\mathbb{Z}[X]$ is a unique factorisation domain.

Let $R[X_1, \dots, X_n]$ be the ring of polynomials in n variables. We can rewrite this as $R[X_1] \dots [X_n]$, so by induction this is a unique factorisation domain if R is.

9.5 Eisenstein's criterion

Proposition. Let R be a unique factorisation domain, and $f(X) = \sum_{i=0}^{n} a_i X^i \in R[X]$ be a primitive polynomial. Let $p \in R$ be irreducible (or, equivalently, prime) such that

(i) $p \nmid a_n$;
(ii) $p \mid a_i$ for all $i < n$; and
(iii) $p^2 \nmid a_0$.

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

Proof. Suppose $f = gh$ for $g, h \in R[X]$ not units. Since f is primitive, g, h must have positive degree. Let $g(X) = \sum_{i=0}^{k} r_i X^i$ and $h(X) = \sum_{i=0}^{\ell} s_i X^i$, so $k + \ell = n$. Then $p \nmid a_n = r_k s_\ell$, so $p \nmid r_k$ and $p \nmid s_\ell$. Further, $p \mid a_0 = r_0 s_0$ so $p \mid r_0$ or $p \mid s_0$. Without loss of generality, we may assume $p \mid r_0$. There exists a minimal $j \leq k$ such that $p \mid r_i$ for all $i < j$ but $p \nmid r_j$.

$$a_j = r_0 s_j + r_1 s_{j-1} + \cdots + r_{j-1} s_1 + r_j s_0$$

By assumption, a_j is divisible by p since $j < n$. Further, the first j terms in the expansion are divisible by p. Thus, $p \mid r_j s_0$. By assumption, $p \nmid r_j$, so $p \mid s_0$. In particular, $p^2 \mid r_0 s_0 = a_0$, contradicting the third criterion. □

Example. Let $f(X) = X^3 + 2X + 5 \in \mathbb{Z}[X]$. We will show this is irreducible as a polynomial over $\mathbb{Q}$. If f is not irreducible in $\mathbb{Z}[X]$, then it factorises as $f(X) = (X + a)(X^2 + bX + c)$ up to multiplication by units. Here, $ac = 5$. But $\pm 1, \pm 5$ are not roots of f, so this is irreducible in $\mathbb{Z}[X]$. By Gauss' lemma, f is irreducible in $\mathbb{Q}[X]$, since $\mathbb{Q}$ is the field of fractions of $\mathbb{Z}$. In particular, $\mathbb{Q}[X]/(f)$ is a field, since the ideal (f) is maximal.

Example. Let $p \in \mathbb{Z}$ be a prime, and let $f(X) = X^n - p$. By Eisenstein's criterion, f is irreducible in $\mathbb{Z}[X]$. It is then irreducible in $\mathbb{Q}[X]$ by Gauss' lemma.

Example. Consider $f(X) = X^{p-1} + X^{p-2} + \cdots + X + 1 \in \mathbb{Z}[X]$, where p is prime. Eisenstein's criterion does not apply directly. Consider

$$f(X) = \frac{X^p - 1}{X - 1}; \quad Y = X - 1$$

By using this substitution of Y,

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

We can apply Eisenstein's criterion to this new polynomial, since $p \mid \binom{p}{i}$ 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 $\mathbb{Z}[X]$. Of course, $f(X)$ is therefore irreducible in $\mathbb{Q}[X]$ as before.

10 Algebraic integers

10.1 Gaussian integers

Recall the ring of Gaussian integers $\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\} \leq \mathbb{C}$. There is a norm function $N : \mathbb{Z}[i] \to \mathbb{Z}_{\geq 0}$ given by $a + bi \mapsto a^2 + b^2$, and $N(xy) = N(x)N(y)$. This norm is a Euclidean function, giving the Gaussian integers the structure of a Euclidean domain and hence a principal ideal domain and a unique factorisation domain. In particular, the primes are the irreducibles. The units in $\mathbb{Z}[i]$ are $\pm 1, \pm i$, since they are the only elements of unit norm.

Example. 2 is not irreducible in $\mathbb{Z}[i]$, since it factors as $(1 + i)(1 - i)$. 5 is not irreducible, since it factors as $(2 + i)(2 - i)$. These are nontrivial factorisations since the norms of the factors are not unit length.

3 is a prime, since it is irreducible. Indeed, $N(3) = 9$, so if 3 were reducible it would factor as ab where $N(a) = N(b) = 3$. But $\mathbb{Z}[i]$ has no elements of norm 3. Similarly, 7 is a prime.

Proposition. Let $p \in \mathbb{Z}$ be a prime. Then, the following are equivalent.
(i) p is not prime in $\mathbb{Z}[i]$;
(ii) $p = a^2 + b^2$ for $a, b \in \mathbb{Z}$;
(iii) $p = 2$ or $p \equiv 1 \bmod 4$.

Proof. Suppose p is not prime in $\mathbb{Z}[i]$. So let $p = xy$ for $x, y \in \mathbb{Z}[i]$ not units. Then, $p^2 = N(p) = N(x)N(y)$. Since x, y are not units, $N(x), N(y) > 1$ and in particular $N(x) = N(y) = p$. Writing $x = a + bi$ for $a, b \in \mathbb{Z}$, we have $p = N(x) = a^2 + b^2$, which is the condition in (ii).

Now, suppose $p = a^2 + b^2$. The only squares modulo 4 are 0 and 1. Since $p \equiv a^2 + b^2 \bmod 4$, we have that p cannot be congruent to 3, modulo 4.

Finally, suppose $p = 2$ or $p \equiv 1 \bmod 4$. We have already observed above that 2 is not prime. It hence suffices to consider the case where $p \equiv 1 \bmod 4$. We have that $\left(\mathbb{Z}/p\mathbb{Z}\right)^\times$ is cyclic of order $p - 1$ by a previous theorem. Hence, if $p \equiv 1 \bmod 4$, we have that $4 \mid p - 1$, and hence $\left(\mathbb{Z}/p\mathbb{Z}\right)^\times$ contains an element of order 4. In particular, there exists $x \in \mathbb{Z}$ with $x^4 \equiv 1 \bmod p$, but $x^2 \not\equiv 1 \bmod p$. Then $x^2 \equiv -1 \bmod p$, or in other words, $p \mid (x^2 + 1)$. But this factorises as $p \mid (x + i)(x - i)$. We can see that $p \nmid x + i$, $p \nmid x - i$, so p cannot be prime. □

Remark. The proof that (iii) implies (ii) is entirely nontrivial. It required lots of theory in order to reach the result, even though its statement did not require even the notion of a complex number.

Theorem. The primes in $\mathbb{Z}[i]$ are, up to associates,
(i) $a + bi$, where $a, b \in \mathbb{Z}$ and $a^2 + b^2 = p$ is a prime in $\mathbb{Z}$ with $p = 2$ or $p \equiv 1 \bmod 4$; and
(ii) the primes p in $\mathbb{Z}$ satisfying $p \equiv 3 \bmod 4$.

Proof. First, we must check that all such elements are prime. For (i), note that $N(a+bi) = p$ is prime, so $a + bi$ is irreducible. We can use the above proof to deduce that primes in $\mathbb{Z}$ of form (ii) are primes in $\mathbb{Z}[i]$.

It now suffices to show that any prime in the Gaussian integers satisfies one of the two above conditions. Let z be prime in $\mathbb{Z}[i]$. We note that $\overline{z}$ is also irreducible. Now, $N(z) = z\overline{z}$, which is a factorisation of the norm into irreducibles.

Let p be a prime in $\mathbb{Z}$ dividing $N(z)$. If $p \equiv 3 \bmod 4$, p is prime in $\mathbb{Z}[i]$. So $p \mid z$ or $p \mid \overline{z}$ so p is associate to z or $\overline{z}$.

Otherwise, $p = a^2 + b^2 = (a + bi)(a - bi)$ where $a \pm bi$ are prime in $\mathbb{Z}[i]$ as they have norm p. So we have $p = (a+bi)(a-bi) \mid z\overline{z}$, so z is an associate of $a+bi$ or $a-bi$ by uniqueness of factorisation. □

Remark. In the above theorem, if $p = a^2 + b^2$, $a + bi$ and $a - bi$ are not associate unless $p = 2$.

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

Proof. Suppose $n = a^2 + b^2$. So $n = N(a + bi)$. Hence n is a product of norms of primes in the Gaussian integers. By the classification above, those norms are

(i) the primes $p \in \mathbb{Z}$ with $p \not\equiv 3 \bmod 4$; and

(ii) squares of primes $p \in \mathbb{Z}$ with $p \equiv 3 \bmod 4$.

The result follows. □

Example. We can write $65 = 5 \cdot 13$ as the sum of two primes since $5, 13 \equiv 1 \bmod 4$. We first factorise 5 and 13 into primes in the Gaussian integers.

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

Thus, the factorisation of 65 into irreducibles in $\mathbb{Z}[i]$ is

$$\begin{aligned} 65 &= (2+3i)(2+i)(2-3i)(2-i) \\ &= [(2+3i)(2+i)]\overline{[(2+3i)(2+i)]} \\ &= N((2+3i)(2-i)) \\ &= N(1+8i) = 1^2 + 8^2 \end{aligned}$$

This was dependent on the choice of grouping of terms. Alternatively,

$$65 = N((2+i)(2-3i)) = N(7+4i) = 7^2 + 4^2$$

10.2 Algebraic integers

Definition. A number $\alpha \in \mathbb{C}$ is *algebraic* if α is a root of some nonzero polynomial $f \in \mathbb{Q}[X]$. α is an *algebraic integer* if it is a root of some monic polynomial $f \in \mathbb{Z}[X]$.

Let $R \leq S$, and $\alpha \in S$. We write $R[\alpha]$ to denote the smallest subring of S containing R and α. Alternatively, $R[\alpha]$ is the intersection of all subrings of S containing R and α. Further, $R[\alpha] = \operatorname{Im} \varphi$ where $\varphi : R[X] \to S$ is the homomorphism $g(X) \mapsto g(\alpha)$.

Definition. Let α be an algebraic number. Consider the homomorphism $\varphi : \mathbb{Q}[X] \to \mathbb{C}$ where $g(X) \mapsto g(\alpha)$. Since $\mathbb{Q}[X]$ is a a principal ideal domain, $\ker \varphi = (f)$ for some $f \in \mathbb{Q}[X]$. This ideal contains a nonzero element since α is an algebraic number, hence f is nonzero. Multiplying f by a unit, we may assume f is monic without loss of generality. This unique f is known as the *minimal polynomial* of α.

Corollary. All minimal polynomials are irreducible. By the first isomorphism theorem, $\mathbb{Q}[X]/(f) \cong \mathbb{Q}[\alpha] \leq \mathbb{C}$. Any subring of a field is an integral domain. Hence (f) is a prime ideal in $\mathbb{Q}[X]$, and hence f is irreducible. In particular, this implies that $\mathbb{Q}[\alpha]$ is a field.

Proposition. Let α be an algebraic integer, and $f \in \mathbb{Q}[X]$ be its minimal polynomial. Then $f \in \mathbb{Z}[X]$, and $(f) = \ker \theta \lhd \mathbb{Z}[X]$ where $\theta : \mathbb{Z}[X] \to \mathbb{C}$ is given by $g(X) \mapsto g(\alpha)$.

Remark. If α is an algebraic integer, then the polynomial in the definition can be taken to be minimal without loss of generality. $\mathbb{Z}[X]$ is not a principal ideal domain, so the above argument cannot work verbatim.

Proof. Let f be the minimal polynomial of α. Let $\lambda \in \mathbb{Q}^\times$ such that λf has coefficients in $\mathbb{Z}$ and is primitive. Then $\lambda f(\alpha) = 0$, so $\lambda f \in \ker \theta$.

Let $g \in \ker \theta$, so in particular $g \in \mathbb{Z}[X]$. Then $g \in \ker \varphi$, and hence $\lambda f \mid g$ in $\mathbb{Q}[X]$. By a previous lemma, $\lambda f \mid g$ in $\mathbb{Z}[X]$. Thus, $\ker \theta = (\lambda f)$.

Now, since α is an algebraic integer, we know that there exists a monic polynomial $g \in \ker \theta$ such that $g(\alpha) = 0$. Then $\lambda f \mid g$ in $\mathbb{Z}[X]$, so $\lambda = \pm 1$ as both f, g are monic. Hence, $f \in \mathbb{Z}[X]$, and $(\lambda f) = (f) = \ker \theta$. □

Let $\alpha \in \mathbb{C}$ be an algebraic integer. Then, applying the isomorphism theorem to θ, $\mathbb{Z}[X]/(f) \cong \mathbb{Z}[\alpha]$. For example:

$$\begin{aligned}
\mathbb{Z}[X]/(X^2+1) &\cong \mathbb{Z}[i] \\
\mathbb{Z}[X]/(X^2-2) &\cong \mathbb{Z}\left[\sqrt{2}\right] \\
\mathbb{Z}[X]/(X^2+X+1) &\cong \mathbb{Z}\left[\frac{-1+\sqrt{-3}}{2}\right] \\
\mathbb{Z}[X]/(X^n-p) &\cong \mathbb{Z}[\sqrt[n]{p}]
\end{aligned}$$

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

Proof. Let $\alpha \neq 0$, since the case where $\alpha = 0$ is trivial. Then the minimal polynomial of α has coefficients in $\mathbb{Z}$. Since α is rational, the minimal polynomial is $X - \alpha$. Hence $\alpha \in \mathbb{Z}$ as it is a coefficient of the minimal polynomial. □

11 Noetherian rings

11.1 Definition

Recall the definition of a Noetherian ring.

Definition. A ring R is *Noetherian* if, for all sequences of nested ideals $I_1 \subseteq I_2 \subseteq \cdots$, there exists $N \in \mathbb{N}$ such that for all $n > N$, $I_n = I_{n+1}$.

Lemma. Let R be a ring. Then R satisfies the ascending chain condition (so R is Noetherian) if and only if all ideals in R are finitely generated.

We have already shown that principal ideal domains are Noetherian, since they satisfy this 'ascending chain' condition. This now will immediately follow from the lemma.

Proof. First, suppose that all ideals in R are finitely generated. Let $I_1 \subseteq I_2 \subseteq \cdots$ be an ascending chain of ideals. Consider $I = \bigcup_{i=1}^{\infty} I_i$, which is an ideal. I is finitely generated, so $I = (a_1, \ldots, a_n)$.

These elements belong to a nested union of ideals. In particular, we can choose $N \in \mathbb{N}$ such that all a_i are contained within I_N. Then, for $n \geq N$, we find

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

So the inclusions are all equalities, so $I_N = I_n$.

Conversely, suppose that R is Noetherian. Suppose that there exists an ideal $J \lhd R$ which is not finitely generated. Let $a_1 \in J$. Then since J is not finitely generated, $(a_1) \subset J$. We can therefore choose $a_2 \in J \setminus (a_1)$, and then $(a_1) \subset (a_1, a_2) \subset J$. Continuing inductively, we contradict the ascending chain condition. □

11.2 Hilbert's basis theorem

Theorem. Let R be a Noetherian ring. Then $R[X]$ is Noetherian.

Proof. Suppose there exists an ideal J that is not finitely generated. Let $f_1 \in J$ be an element of minimal degree. Then $(f_1) \subset J$. So we can choose $f_2 \in J \setminus (f_1)$, which is also of minimal degree. Inductively we can construct a sequence $f_1, f_2, \dots$, where the degrees are non-decreasing. Let a_i be the leading coefficient of f_i, for all i. We then obtain a sequence of ideals $(a_1) \subseteq (a_1, a_2) \subseteq (a_1, a_2, a_3) \subseteq \cdots$ in R. Since R is Noetherian, there exists $m \in \mathbb{N}$ such that for all $n \geq m$, we have $a_n \in (a_1, \dots, a_m)$. Let $a_{m+1} = \sum_{i=1}^{m} \lambda_i a_i$, since a_{m+1} lies in the ideal $(a_1, \dots, a_m)$. Now we define

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

The degree of g is equal to the degree of f_{m+1}, and they have the same leading coefficient a_{m+1}. Then, consider $f_{m+1} - g \in J$ and $\deg(f_{m+1} - g) < \deg f_{m+1}$. By minimality of the degree of f_{m+1}, $f_{m+1} - g \in (f_1, \dots, f_m)$, hence $f_{m+1} \in (f_1, \dots, f_m)$. This contradicts the choice of f_{m+1}, so J is in fact finitely generated. □

Corollary. $\mathbb{Z}[X_1, \dots, X_n]$ is Noetherian. Similarly, $F[X_1, \dots, X_n]$ is Noetherian for any field F, since fields satisfy the ascending chain condition.

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

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

where $\mathcal{F} \subseteq R$ is a (possibly infinite) set of polynomials. Such a set is referred to as an *algebraic variety*. Let

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

We can check that $I \lhd R$. Since R is Noetherian, $I = (g_1, \dots, g_r)$. Hence

$$V = \{(a_1, \dots, a_n) \in \mathbb{C}^n : g(a_1, \dots, a_n) = 0, \forall g \in I\}$$

Lemma. Let R be a Noetherian ring, and $I \lhd R$. Then R/I is Noetherian.

Proof. Let $J_1' \subseteq J_2' \subseteq \cdots$ be a chain of ideals in R/I. By the ideal correspondence, J_i' corresponds to an ideal J_i that contains I, so $J_i' = J_i/I$. So $J_1 \subseteq J_2 \subseteq \cdots$ is a chain of ideals in R. Since R is Noetherian, there exists $N \in \mathbb{N}$ such that for all $n \geq N$, we have $J_N = J_n$, and so $J_N' = J_n'$. Hence R/I satisfies the ascending chain condition. □

Example. The ring of Gaussian integers $\mathbb{Z}/(X^2+1)$ is Noetherian. If $R[X]$ is Noetherian, then $R[X]/(X) \cong R$ is Noetherian. This is a converse to the Hilbert basis theorem.

The ring of polynomials in countably many variables is not Noetherian.

$$\mathbb{Z}[X_1, X_2, \dots] = \bigcup_{n \in \mathbb{N}} \mathbb{Z}[X_1, \dots, X_n]$$

In particular, consider the ascending chain $(X_1) \subset (X_1, X_2) \subset (X_1, X_2, X_3) \subset \cdots$.

Let $R = \{f \in \mathbb{Q}[X] : f(0) \in \mathbb{Z}\} \leq \mathbb{Q}[X]$. Even though $\mathbb{Q}[X]$ is Noetherian, R is not. Indeed, consider $(X) \subset \left(\frac{1}{2}X\right) \subset \left(\frac{1}{4}X\right) \subset \left(\frac{1}{8}X\right) \subset \cdots$. These inclusions are strict, since $2 \in R$ is not a unit.

12 Modules

12.1 Definitions

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

(i) $(M, +)$ is an abelian group with identity $0 = 0_M$;
(ii) $(r_1 + r_2) \cdot m = r_1 \cdot m + r_2 \cdot m$;
(iii) $r \cdot (m_1 + m_2) = r \cdot m_1 + r \cdot m_2$;
(iv) $r_1 \cdot (r_2 \cdot m) = (r_1 \cdot r_2) \cdot m$;
(v) $1_R \cdot m = m$;

Remark. Closure is implicitly required by the types of the $+$ and $\cdot$ operations.

Example. A module over a field is precisely a vector space.

A $\mathbb{Z}$-module is precisely the same as an abelian group, since

$$\cdot : \mathbb{Z} \times A \to A; \quad n \cdot a = \begin{cases} \underbrace{a + \cdots + a}_{n \text{ times}} & \text{if } n > 0 \\ 0 & \text{if } n = 0 \\ -\left(\underbrace{a + \cdots + a}_{-n \text{ times}}\right) & \text{if } n < 0 \end{cases}$$

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

$$\cdot : F[X] \times V \to V; \quad f \cdot v = (f(\alpha))(v)$$

Note that the structure of the $F[X]$-module depends on the choice of α. We can write $V = V_\alpha$ to disambiguate.

For any ring R, we can consider R^n as an R-module via

$$r \cdot (r_1, \dots, r_n) = (r \cdot r_1, \dots, r \cdot r_n)$$

In particular, the case $n = 1$ shows that any ring R can be considered an R-module where the scalar multiplication in the ring and the module agree.

For an ideal $I \lhd R$, we can regard I as an R-module, since I is preserved under multiplication by elements in R. The quotient ring R/I is also an R-module, defining multiplication as $r \cdot (s+I) = rs+I$.

Let $\varphi : R \to S$ be a ring homomorphism. Then any S-module can be regarded as an R-module. We define $r \cdot m = \varphi(r) \cdot m$. In particular, this applies when R is a subring of S, and φ is the inclusion map. So any module over a ring can be viewed as a module over any subring.

Definition. Let M be an R-module. Then $N \subseteq M$ is an *R-submodule of M*, written $N \leq M$, if $(N, +) \leq (M, +)$, and for all $rn \in N$ for all $r \in R$ and $n \in N$.

Example. By considering R as an R-module, a subset of R is an R-submodule if and only if it is an ideal. If $R = F$ is a field, this definition corresponds to the definition of a vector subspace.

Definition. Let $N \leq M$ be R-modules. Then, the *quotient* M/N is defined as the quotient of groups under addition, and with scalar multiplication defined as $r \cdot (m + N) = rm + N$. This is well-defined, since N is preserved under scalar multiplication. This makes M/N an R-module.

Remark. Submodules are analogous both to subrings and to ideals.

Definition. Let M, N be R-modules. Then $f : M \to N$ is a *R-module homomorphism* if it is a homomorphism of $(M, +)$ and $(N, +)$, and scalar multiplication is preserved: $f(r \cdot m) = r \cdot f(m)$. An *$R$-module isomorphism* is an R-module homomorphism that is a bijection.

Example. If $R = F$ is a field, F-module homomorphisms are exactly linear maps.

Theorem. Let $f : M \to N$ be an R-module homomorphism. Then
(i) $\ker f = \{m \in M : f(m) = 0\} \leq M$;
(ii) $\operatorname{Im} f = \{f(m) \in N : m \in M\} \leq N$;
(iii) $M/\ker f \cong \operatorname{Im} f$.

Theorem. Let $A, B \leq M$ be R-submodules. Then
(i) $A + B = \{a + b : a \in A, b \in B\} \leq M$;
(ii) $A \cap B \leq M$;
(iii) $A/A \cap B \cong A + B/B$.

Theorem. For $N \leq L \leq M$ are R-submodules, then

$$M/N_{/L/N} \cong M_{/L}$$

For $N \leq M$, there is a correspondence between submodules of $M_{/N}$ and submodules of M containing N. These isomorphism theorems can be proved exactly as before. Note that these results apply to vector spaces; for example, the first isomorphism theorem immediately gives the rank-nullity theorem.

12.2 Finitely generated modules

Definition. Let M be an R-module. If $m \in M$, then we write $Rm = \{rm : r \in R\}$. This is an R-submodule of M, known as the submodule *generated by* m.
If $A, B \leq M$, we can define $A + B = \{a + b : a \in A, b \in B\}$, known as the *sum of submodules*. In particular, this sum is commutative.

Definition. A module M is *finitely generated* if it is the sum of finitely many submodules generated by a single element. In other words, $M = Rm_1 + \cdots + Rm_n$.

This is the analogue of finite dimensionality in linear algebra.

Lemma. An R-module M is finitely generated if and only if there exists a surjective R-module homomorphism $f : R^n \to M$ for some n.

Proof. If M is finitely generated, we have $M = Rm_1 + \cdots + Rm_n$. We define $f : R^n \to M$ by $(r_1, \ldots, r_n) \mapsto r_1m_1 + \cdots + r_nm_n$. This is surjective.

Conversely, suppose such a surjective homomorphism f exists. Let $e_i = (0, \ldots, 1, \ldots, 0)$ be the element of R^n with all entries zero except for 1 in the ith place. Let $m_i = f(e_i)$. Then, since f is surjective, any element $m \in M$ is contained in the image of f, so is of the form $f(r_1, \ldots, r_n) = r_1m_1 + \cdots + r_nm_n$. □

Corollary. Any quotient by a submodule of a finitely generated module is finitely generated.

Proof. Let $N \leq M$, where M is finitely generated. Then there exists a surjective R-module homomorphism $f : R^n \to M$. Then $q \circ f$, where q is the quotient map, is also a surjective homomorphism. So $M_{/N}$ is finitely generated. □

Example. It is not always the case that a submodule of a finitely generated module is finitely generated. Let R be a non-Noetherian ring, and I an ideal in R that is not finitely generated (in the ring sense). R is a finitely generated R-module, since $R1 = R$. I is a submodule of R, which is not finitely generated (in the module sense).

Remark. If R is Noetherian, it is always the case that submodules of finitely generated R-modules are finitely generated. This will be shown on the example sheets.

12.3 Torsion

Definition. Let M be an R-module.
(i) $m \in M$ is *torsion* if there exists $0 \neq r \in R$ such that $rm = 0$;
(ii) M is a *torsion module* if every element is torsion;
(iii) M is a *torsion-free module* if 0 is the only torsion element.

Example. The torsion elements in a $\mathbb{Z}$-module (which is an abelian group) are precisely the elements of finite order. If F is a field, any F-module is torsion-free.

12.4 Direct sums

Definition. Let $M_1, \dots, M_n$ be R-modules. Then the *direct sum* of $M_1, \dots, M_n$, written $M_1 \oplus \cdots \oplus M_n$, is the set $M_1 \times \cdots \times M_n$, with the operations of addition and scalar multiplication defined componentwise. We can show that the direct sum of (finitely many) R-modules is an R-module.

Example. $R^n = R \oplus \cdots \oplus R$, where we take the direct sum of n copies of R.

Lemma. Let $M = \bigoplus_{i=1}^n M_i$, and for each M_i, let $N_i \leq M_i$. Then $N = \bigoplus_{i=1}^n N_i$ is a submodule of M. Further,

$$M/N = \bigoplus_{i=1}^n M_i \Big/ \bigoplus_{i=1}^n N_i \cong \bigoplus_{i=1}^n M_i/N_i$$

Proof. First, we can see that this N is a submodule. Applying the first isomorphism theorem to the surjective R-module homomorphism $M \to \bigoplus_{i=1}^n M_i/N_i$ given by $(m_1, \dots, m_n) \mapsto (m_1 + N_1, \dots, m_n + N_n)$, the result follows as required, since the kernel is N. □

12.5 Free modules

Definition. Let $m_1, \dots, m_n \in M$. The set $\{m_1, \dots, m_n\}$ is *independent* if $\sum_{i=1}^n r_i m_i = 0$ implies that the r_i are all zero.

Definition. A subset $S \subseteq M$ *generates M freely* if:
(i) S generates M, so for all $m \in M$, we can find finitely many entries s_i and coefficients r_i such that $m = \sum_{i=1}^k r_i s_i$;
(ii) any function $\psi : S \to N$, where N is an R-module, extends to an R-module homomorphism $\theta : M \to N$.

Remark. In (ii), such an extension θ is always unique if it exists, by (i).

Definition. An R-module M freely generated by some subset $S \subseteq M$ is called *free*. We say that S is a *free basis* for M.

Remark. Free bases in the study of modules are analogous to bases in linear algebra. All vector spaces are free modules, but not all modules are free.

Proposition. For a finite subset $S = \{m_1, \dots, m_n\} \subseteq M$, 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 $r_1 m_1 + \cdots + r_n m_n$ for some $r_i \in R$;
(iv) the R-module homomorphism $R^n \to M$ given by $(r_1, \dots, r_n) \mapsto r_1 m_1 + \cdots + r_n m_n$ is bijective, so is an isomorphism.

Proof. Not all implications are shown, but they are similar to arguments found in Part IB Linear Algebra. We show (i) implies (ii). Let S generate M freely. Suppose S is not independent. Then there exist r_i such that $\sum_{i=1}^n r_i m_i = 0$ but not all r_i are zero. Let $r_j \neq 0$. Since S generates M freely, consider the module homomorphism $\psi : S \to R$ given by

$$\psi(m_i) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

Then

$$0 = \theta(0) = \theta\left(\sum_{i=1}^n r_i m_i\right) = \sum_{i=1}^n r_i \theta(m_i) = r_j \neq 0$$

This is a contradiction, so S is independent.

To show (ii) implies (iii), it suffices to show uniqueness. If there exist two ways to write an element as a linear combination, consider their difference to find a contradiction from (ii).

We can show (iii) implies (i). Then it remains to show (iii) and (iv) are equivalent. □

Example. A non-trivial finite abelian group is not a free $\mathbb{Z}$-module.

The set $\{2, 3\}$ generates $\mathbb{Z}$ as a $\mathbb{Z}$-module. This is not a free basis, since they are not independent: $2 \cdot 3 - 3 \cdot 2 = 0$. However, it contains no subset that is a free basis. This is different to vector spaces, where we can always construct a basis from a subset of a spanning set.

Proposition (invariance of dimension)**.** Let R be a nonzero ring. If $R^m \cong R^n$ as R-modules, then $m = n$.

Proof. Let $I \lhd R$, and M an R-module. We define $IM = \{\sum a_i m_i : a_i \in I, m_i \in M\}$. Since I is an ideal, we can show that IM is a submodule of M. The quotient module M/IM is an R-module, but we can also show that it is an R/I-module, by defining scalar multiplication as

$$(r + I) \cdot (m + IM) = (r \cdot m + IM)$$

We can check that this is well-defined; this follows from the fact that for $b \in I$, $b \cdot (m + IM) = bm + IM$, but $b \in I$ so $bm \in IM$.

Now, suppose that $R^m \cong R^n$. Then let $I \lhd R$ be a maximal ideal in R. We can prove the existence of such an ideal under the assumption of the axiom of choice, and in particular using Zorn's lemma. By the above discussion, we find an isomorphism of R/I-modules

$$\left(R/I\right)^m \cong R^m/IR^m \cong R^n/IR^n \cong \left(R/I\right)^n$$

This is an isomorphism of vector spaces over R/I which is a field, since I is maximal. Hence, using the corresponding result from linear algebra, $n = m$. □

12.6 Row and column operations

We will assume that R is a Euclidean domain in this subsection, and let φ be a Euclidean function for R. We will consider an $m \times n$ matrix with entries in R.

Definition. The *elementary row operations* on a matrix are
(i) add $\lambda \in R$ multiplied by the jth row to the ith row, where $i \neq j$;
(ii) swap the ith row and the jth row;
(iii) multiply the ith row by $u \in R^\times$.
Each of these operations can be realised by left-multiplication by some $m \times m$ matrix. These operations are all invertible, so their matrices are all invertible.

We can define elementary column operations in an analogous way, using right-multiplication by an $n \times n$ matrix instead.

Definition. Two $m \times n$ matrices A, B are *equivalent* if there exists a sequence of elementary row and column operations that transforms one matrix into the other. If they are equivalent, then there exist invertible matrices P, Q such that $B = QAP$.

Definition. A $k \times k$ *minor* of an $m \times n$ matrix A is the determinant of a $k \times k$ submatrix of A, which is a matrix of A produced by removing $m - k$ rows and $n - k$ columns.
The kth Fitting ideal $\mathrm{Fit}_k(A) \lhd R$ is the ideal generated by the $k \times k$ minors of A.

Lemma. The kth Fitting ideal of a matrix is invariant under elementary row and column operations.

Proof. It suffices by symmetry to show that the elementary row operations do not change the Fitting ideal. For the first elementary row operation on a matrix A, suppose we add $\lambda \in R$ multiplied by the jth row to the ith row, yielding a matrix A'. In particular, $a_{ik} \mapsto a_{ik} + \lambda a_{jk}$ for all k. Let C be a $k \times k$ submatrix of A and C' the corresponding submatrix of A'.

If row i was not chosen in C, then C and C' are the same matrix. Hence the corresponding minors are equal. If row i and row j were both chosen in C, we have that C, C' differ by a row operation. Since the determinant is invariant under this elementary row operations, the corresponding minors are equal.

If row i was chosen but row j was not chosen, by expanding the determinant along the ith row, we find

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

where we can show that D is a $k \times k$ submatrix of A that includes row j but not row i. By definition, $\det D \in \mathrm{Fit}_k(A)$ and $\det C \in \mathrm{Fit}_k(A)$, so certainly $\det C' \in \mathrm{Fit}_k(A)$. Hence $\mathrm{Fit}_k(A') \subseteq \mathrm{Fit}_k(A)$. By the invertibility of the elementary row operations, $\mathrm{Fit}_k(A') \supseteq \mathrm{Fit}_k(A)$.

The proofs for the other elementary row operations are left as an exercise. □

12.7 Smith normal form

Theorem. An $m \times n$ matrix $A = (a_{ij})$ over a Euclidean domain R is equivalent to a matrix of the form

$$\begin{pmatrix} d_1 & & & & & \\ & \ddots & & & & \\ & & d_t & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & \end{pmatrix}; \quad d_1 \mid d_2 \mid \cdots \mid d_t$$

The d_i are known as *invariant factors*, and they are unique up to associates.

Proof. If $A = 0$, the matrix is already in Smith normal form. Otherwise, we can swap columns and rows such that $a_{11} \neq 0$. We will reduce $\varphi(a_{11})$ as much as possible until it divides every other element in the matrix, using the following algorithm.

If $a_{11} \nmid a_{1j}$ for some $j \geq 2$, then $a_{1j} = qa_{11} + r$ where $q, r \in R$ and $\varphi(r) < \varphi(a_{11})$. We can subtract q multiplied by column 1 from column j. Swapping such columns leaves $a_{11} = r$. If $a_{11} \nmid a_{i1}$ for some $i \geq 2$, then repeat the above process using row operations. Now, $a_{11} \mid a_{ij}$ for all i, j. These steps are repeated until a_{11} divides all entries of the first row and first column. This algorithm will always terminate, for example because the Euclidean function takes values in $\mathbb{Z}_{\geq 0}$ and $\varphi(a_{11})$ strictly decreases in each iteration.

Now, we can subtract multiples of the first row and column from the others to give

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix}$$

If $a_{11} \nmid a_{ij}$ for $i, j \geq 2$, then add the ith row to the first row. There is now an element in the first row that does a_{11} not divide. We can then perform column operations as above to decrease $\varphi(a_{11})$. We will then restart the algorithm. After finitely many steps, this algorithm will terminate and a_{11} will divide all elements a_{ij} of the matrix.

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix}; \quad a_{11} \equiv d_1 \mid a_{ij}$$

We can now apply the algorithm to A', since column and row operations not including the first row or column do not change whether $a_{11} \mid a_{ij}$.

We now demonstrate uniqueness of the invariant factors. Suppose A has Smith normal form with invariant factors d_i where $d_1 \mid \cdots \mid d_t$. Then, for all k, $\mathrm{Fit}_k(A)$ can be evaluated in Smith normal form by invariance of the Fitting ideal under row and column operations. Hence $\mathrm{Fit}_k(A) = (d_1 d_2 \cdots d_k) \lhd R$. Thus, the product $d_1 \cdots d_k$ depends only on A, and is unique up to associates. Cancelling, we can see that each d_i depends only on A, up to associates. □

Example. Consider the matrix over $\mathbb{Z}$ given by

$$A = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$$

Using elementary row and column operations,

$$\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \xrightarrow{c_1 \mapsto c_1 + c_2} \begin{pmatrix} 1 & -1 \\ 3 & 2 \end{pmatrix} \xrightarrow{c_2 \mapsto c_1 + c_2} \begin{pmatrix} 1 & 0 \\ 3 & 5 \end{pmatrix} \xrightarrow{r_2 \mapsto -3r_1 + r_2} \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}$$

This is in Smith normal form as $1 \mid 5$.

Alternatively, $(d_1) = (2, -1, 1, 2) = (1)$. So $d_1 = \pm 1$. Further, $(d_1 d_2) = (\det A) = (5)$. So $d_1 d_2 = \pm 5$ and hence $d_2 = \pm 5$.

12.8 The structure theorem

Lemma. Let R be a Euclidean domain with Euclidean function φ (or, indeed, a principal ideal domain). Any submodule of the free module R^m is generated by at most m elements.

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

$$I = \{r \in R : \exists r_2, \ldots, r_m \in R, (r, r_2, \ldots, r_m) \in N\}$$

Since N is a submodule, this is an ideal. Since R is a principal ideal domain, $I = (a)$ for some $a \in R$. Let $n = (a, a_2, \ldots, a_m) \in N$. For $(r_1, \ldots, r_m) \in N$, we have $r_1 = ra$ for some r. Hence $(r_1, \ldots, r_m) - rn = (0, r_2 - ra_2, \ldots, r_m - ra_m)$, which lies in $N' = N \cap \{0\} \times R^{m-1} \leq R^{m-1}$, hence $N = Rn + N'$. By induction, N' is generated by $n_2, \ldots, n_m$, hence $(n, n_2, \ldots, n_m)$ generate N. □

Theorem. Let R be a Euclidean domain, and $N \leq R^m$. Then there is a free basis $x_1, \ldots, x_m$ for R^m such that N is generated by $d_1 x_1, \ldots, d_t x_t$ for some $d_i \in R$ and $t \leq m$, and such that $d_1 \mid \cdots \mid d_t$.

Proof. By the above lemma, we have $N = Ry_1 + \cdots + Ry_n$ for some $y_i \in R^m$ for some $n \leq m$. Each y_i belongs to R^m so we can form the $m \times n$ matrix A which has columns y_i. A is equivalent to a matrix A' in Smith normal form with invariant factors $d_1 \mid \cdots \mid d_t$.

A' is obtained from A by elementary row and column operations. Switching row i and row j in A corresponds to reassigning the standard basis elements e_i and e_j to each other. Adding a multiple of row i to row j corresponds to replacing $e_1, \ldots, e_m$ with a linear combination of these basis elements

which is a free basis. In general, each row operation simply changes the choice of free basis used for R^m. Analogously, each column operation changes the set of generators y_i for N.

Hence, after applying these row and column operations, the free basis e_i of R^m is converted into $x_1, \dots, x_m$, and N is generated by $d_1 x_1, \dots, d_t x_t$. □

Theorem (structure theorem for finitely generated modules over Euclidean domains)**.** Let R be a Euclidean domain, and M a finitely generated module over R. Then

$$M \cong {}^{R}\!/_{(d_1)} \oplus \cdots \oplus {}^{R}\!/_{(d_t)} \oplus \underbrace{R \oplus \cdots \oplus R}_{k \text{ copies}} \cong {}^{R}\!/_{(d_1)} \oplus \cdots \oplus {}^{R}\!/_{(d_t)} \oplus R^k$$

for some $0 \neq d_i \in R$ and $d_1 \mid \cdots \mid d_t$, and where $k \geq 0$. The d_i are called invariant factors.

Proof. Since M is a finitely generated module, there exists a surjective R-module homomorphism $\varphi : R^m \to M$ for some m. By the first isomorphism theorem, $M \cong {}^{R^m}\!/_{\ker \varphi}$. By the previous theorem, there exists a free basis $x_1, \dots, x_m$ for R^m such that $\ker \varphi \leq R^m$ is generated by $d_1 x_1, \dots, d_t x_t$ and where $d_1 \mid \cdots \mid d_t$. Then,

$$\begin{aligned} M &\cong \frac{\underbrace{R \oplus \dots R}_{k \text{ copies}}}{d_1 R \oplus \cdots \oplus d_t R \oplus \underbrace{0 \oplus \cdots \oplus 0}_{m-t \text{ copies}}} \\ &\cong {}^{R}\!/_{(d_1)} \oplus \cdots \oplus {}^{R}\!/_{(d_t)} \oplus \underbrace{R \oplus \cdots \oplus R}_{m-t \text{ copies}} \end{aligned}$$

□

Remark. After deleting those d_i which are units, the invariant factors of M are unique up to associates. The proof is omitted.

Corollary. Let R be a Euclidean domain. Then any finitely generated torsion-free module is free.

Proof. Since M is torsion-free, there are no submodules of the form ${}^{R}\!/_{(d)}$ with d nonzero, since then multiplying an element of M by d would give zero. Hence, by the structure theorem, $M \cong R^m$ for some m. □

Example. Consider $R = \mathbb{Z}$, and the abelian group $G = \langle a, b \rangle$ subject to the relations $2a + b = 0$ and $-a + 2b = 0$, so $G \cong {}^{\mathbb{Z}^2}\!/_{N}$ where N is the $\mathbb{Z}$-submodule of $\mathbb{Z}^2$ generated by $(2, 1)$ and $(-1, 2)$. Consider

$$A = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$$

which has Smith normal form $d_1 = 1$ and $d_2 = 5$. Hence, by changing basis for $\mathbb{Z}^2$, we can let N be generated by $(1, 0)$ and $(0, 5)$. Hence,

$$G \cong {}^{\mathbb{Z} \oplus \mathbb{Z}}\!/_{\mathbb{Z} \oplus 5\mathbb{Z}} \cong {}^{\mathbb{Z}}\!/_{5\mathbb{Z}}$$

12.9 Primary decomposition theorem

More generally, applying the structure theorem to $\mathbb{Z}$-modules, we obtain the structure theorem for finitely generated abelian groups:

Theorem. Let G be a finitely generated abelian group. Then

$$G \cong C_{d_1} \times \cdots \times C_{d_t} \times \mathbb{Z}^r$$

where $d_1 \mid \cdots \mid d_t$ in $\mathbb{Z}$, and $r \geq 0$.

We have replaced the submodule notation $\mathbb{Z}/n\mathbb{Z}$ and $\oplus$ with the group notation C_n and $\times$. The previous theorem for the structure of finite abelian groups is a special case of this theorem, where $r = 0$. We have also seen that any finite abelian group can be written as a product of cyclic groups of prime power order. This also has a generalisation for modules. The previous result relied on the lemma $C_{mn} \cong C_m \times C_n$ where m and n are coprime. There is an analogous result for principal ideal domains.

Lemma. Let R be a principal ideal domain, and $a, b \in R$ with unit greatest common divisor. Then, treating these quotients as R-modules,

$$R/(ab) \cong R/(a) \oplus R/(b)$$

Proof. Since R is a principal ideal domain, $(a, b) = (d)$ for some $d \in R$. The greatest common divisor of a, b is a unit, so d is a unit, giving $(a, b) = R$. Hence, there exist $r, s \in R$ such that $ra + sb = 1$. This is a generalisation of Bézout's theorem.

Now, we define an R-module homomorphism $\psi : R \to R/(a) + R/(b)$ by $\psi(x) = (x+(a), x+(b))$. Then $\psi(sb) = (sb+(a), sb+(b)) = (1-ra+(a), sb+(b)) = (1+(a), (b))$, and similarly $\psi(ra) = ((a), 1+(b))$. Hence, $\psi(sbx + rby) = (x + (a), y + (b))$ so ψ is surjective.

Clearly we have $(ab) \subset \ker \psi$, so it suffices to show the converse. If $x \in \ker \psi$, then $x \in (a)$ and $x \in (b)$, so $x \in (a) \cap (b)$. Since $x = x(ra + sb) = r(ax) + s(bx)$, we must have that $s(bx) \in (a)$ and $r(ax) \in (b)$, so $x \in (ab)$. Hence $\ker \psi = (ab)$, and the result follows from the first isomorphism theorem for modules. □

Lemma (primary decomposition theorem)**.** Let R be a Euclidean domain and M a finitely generated R-module. Then

$$M \cong R/(p_1^{n_1}) \oplus \cdots \oplus R/(p_k^{n_k}) \oplus R^m$$

where the quotients are considered as R-modules, where p_i are primes in R, which are not necessarily distinct, and where $m \geq 0$.

Proof. By the structure theorem,

$$M \cong R/(d_1) \oplus \cdots \oplus R/(d_t) \oplus \underbrace{R \oplus \cdots \oplus R}_{k \text{ copies}} \cong R/(d_1) \oplus \cdots \oplus R/(d_t) \oplus R^m$$

where $d_1 \mid \cdots \mid d_t$. So it suffices to show that each $R/(d_i)$ can be written as a product of factors of the form $R/(p_j^{n_j})$. Since R is a unique factorisation domain and a principal ideal domain, d_i can be written as a product $up_1^{\alpha_1} \cdots p_r^{\alpha_r}$ where u is a unit and the p_j are pairwise non-associate primes. By the previous lemma,

$$R/(d_i) \cong R/(p_1^{\alpha_1}) \oplus \ldots R/(p_r^{\alpha_r})$$

□

12.10 Rational canonical form

Let V be a vector space over a field F, and $\alpha : V \to V$ be a linear map. Let V_α denote the $F[X]$-module V where scalar multiplication is defined by $f(X) \cdot v = f(\alpha)(v)$.

Lemma. If V is finite-dimensional as a vector space, then V_α is finitely generated as an $F[X]$-module.

Proof. Consider a basis $v_1, \ldots, v_n$ of V, so $v_1, \ldots, v_n$ generate V as an F-vector space. Then, these vectors generate V_α as an $F[X]$-module, since $F \leq F[X]$. □

Example. 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. With respect to this basis, α has the matrix form

$$\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \tag{$*$}$$

Example. Suppose $V_\alpha \cong F[X]/(X-\lambda)^n$ as an $F[X]$-module. Consider the basis $1, X-\lambda, (X-\lambda)^2, \ldots, (X-\lambda)^{n-1}$ for $F[X]/(X-\lambda)^n$ as an F-vector space. Here, $\alpha - \lambda\,\mathrm{id}$ has matrix $(*)$ from the previous example. Hence, α has matrix $(*) + \lambda I$.

Example. Suppose $V_\alpha \cong F[X]/(f)$ where $f \in F[X]$ as an $F[X]$-module, such that f is monic. Let

$$f(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_0$$

With respect to basis $1, X, \ldots, X^{n-1}$, α has matrix

$$C(f) = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ 0 & 0 & 1 & \cdots & 0 & -a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}$$

since f is monic and the last column represents X^n. The above matrix is known as the *companion matrix* of the monic polynomial.

Theorem (Rational canonical form)**.** Let F be a field, V be a finite-dimensional F-vector space, and $\alpha : V \to V$ be a linear map. Then the $F[X]$-module V_α decomposes as

$$V_\alpha \cong {}^{F[X]}\!/_{(f_1)} \oplus \cdots \oplus {}^{F[X]}\!/_{(f_t)}$$

for some monic polynomials $f_i \in F[X]$, and $f_1 \mid \cdots \mid f_t$. Moreover, with respect to a suitable basis, α has matrix

$$\begin{pmatrix} C(f_1) & & & \\ & C(f_2) & & \\ & & \ddots & \\ & & & C(f_t) \end{pmatrix} \tag{$**$}$$

Proof. We know that V_α is finitely generated as an $F[X]$-module, since V is finite-dimensional. Since $F[X]$ is a Euclidean domain, the structure theorem applies, and

$$V_\alpha \cong {}^{F[X]}\!/_{(f_1)} \oplus \cdots \oplus {}^{F[X]}\!/_{(f_t)} \oplus F[X]^m$$

for some m, where $f_1 \mid \cdots \mid f_t$. Since V is finite-dimensional, $m = 0$. As F is a field, without loss of generality we may multiply each f_i by a unit to ensure that they are monic. Then, using the previous example, we can construct the companion matrices for each polynomial and obtain the matrix as required. □

Remark. If α is represented by an $n \times n$ matrix A, there exists a change of basis matrix P such that PAP^{-1} has form $(*)$ as stated in the theorem, so A is similar to such a block diagonal matrix of companion matrices. Note further that $(**)$ can be used to find the minimal and characteristic polynomials of α; the minimal polynomial is f_t, and the characteristic polynomial is $f_1 \cdots f_t$. In particular, the minimal polynomial divides the characteristic polynomial, and this implies the Cayley–Hamilton theorem.

Example. Consider $\dim V = 2$. Here, $\sum \deg f_i = 2$, so there are two cases: one polynomial of degree two, or two polynomials of degree one. Consider $V_\alpha \cong {}^{F[X]}\!/_{(X-\lambda)} \oplus {}^{F[X]}\!/_{(X-\mu)}$. Since one of the f_i must divide the other, we have $\lambda = \mu$. If we have one polynomial of degree two, we have $V_\alpha \cong {}^{F[X]}\!/_{(f)}$, where f is the characteristic polynomial of α.

Corollary. Let A, B be invertible 2×2 non-scalar matrices over a field F. Then A, B are similar if and only if their characteristic polynomials are equal.

Proof. Certainly if A, B are similar they have the same characteristic polynomial, which is proven in Part IB Linear Algebra. Conversely, if the matrices are non-scalar, the modules V_α, V_β are of the form ${}^{F[X]}\!/_{(f)}$ by the previous example, so they are both similar to the companion matrix of f, where f is the characteristic polynomial of A or B. □

Definition. The *annihilator* of an R-module M is

$$\mathrm{Ann}_R(M) = \{r \in R : \forall m \in M,\ rm = 0\} \lhd R$$

Example. Let $I \lhd R$. Then the annihilator of ${}^{R}\!/_{I}$ is $\mathrm{Ann}_R\left({}^{R}\!/_{I}\right) = I$.

Let A be a finite abelian group. Then, considering A as a $\mathbb{Z}$-module, $\mathrm{Ann}_{\mathbb{Z}}(A) = (e)$ where e is the *exponent* of the group, which is the lowest common multiple of the orders of elements in the group.

Let V_α be as above. Then $\mathrm{Ann}_{F[X]}(V_\alpha) = (f)$ where f is the minimal polynomial of α.

12.11 Jordan normal form

Jordan normal form concerns matrix similarity in $\mathbb{C}$. The following results are therefore restricted to this particular field.

Lemma. The primes (or equivalently, irreducibles) in $\mathbb{C}[X]$ are the polynomials $X - \lambda$ for $\lambda \in \mathbb{C}$, up to associates.

Proof. By the fundamental theorem of algebra, any non-constant polynomial with complex coefficients has a complex root. By the Euclidean algorithm, we can show that having a root λ is equivalent to having a linear factor $X - \lambda$. Hence the irreducibles have degree one, and thus are $X - \lambda$ exactly, up to associates. □

Theorem. Let $\alpha : V \to V$ be an endomorphism of a finite-dimensional $\mathbb{C}$-vector space V. Let V_α be the set V as a $\mathbb{C}[X]$-module, where scalar multiplication is defined by $f \cdot v = f(\alpha)(v)$. Then, there exists an isomorphism of $\mathbb{C}[X]$-modules

$$V_\alpha \cong {}^{\mathbb{C}[X]}\!/_{((X-\lambda_1)^{n_1})} \oplus \cdots \oplus {}^{\mathbb{C}[X]}\!/_{((X-\lambda_t)^{n_t})}$$

where $\lambda_i \in \mathbb{C}$ are not necessarily distinct. In particular, there exists a basis for this vector space such that α has matrix in block diagonal form

$$\begin{pmatrix} J_{n_1}(\lambda_1) & & & \\ & J_{n_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{n_t}(\lambda_t) \end{pmatrix}$$

where each *Jordan block* $J_{n_i}(\lambda_i)$ is an $n_i \times n_i$ matrix of the form

$$J_{n_i}(\lambda_i) = \begin{pmatrix} \lambda_i & 0 & 0 & \cdots & 0 \\ 1 & \lambda_i & 0 & \cdots & 0 \\ 0 & 1 & \lambda_i & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i \end{pmatrix}$$

Proof. Note $\mathbb{C}[X]$ is a Euclidean domain using the degree function, and V_α is finitely generated as a $\mathbb{C}[X]$-module. These are the assumptions of the primary decomposition theorem. Applying this, we find the module decomposition as required, noting that the primes in $\mathbb{C}[X]$ are the linear polynomials. Note that the free factor $\mathbb{C}[X]$ cannot appear in the decomposition since V is finite-dimensional.

We have already seen that for a module $W_\alpha \cong {}^{F[X]}\!/_{((X-\lambda)^n)}$, multiplication by X is represented by the matrix $J_n(\lambda)$ with respect to the basis $1, (X-\lambda), \dots, (X-\lambda)^{n-1}$. Hence the result follows by considering the union of these bases. □

Remark. If α is represented by a matrix A, then A is similar to a matrix in Jordan normal form. This is the form of the result often used in linear algebra.

The Jordan blocks are uniquely determined up to reordering. This can be proven by considering the dimensions of the *generalised eigenspaces*, which are $\ker((\alpha - \lambda \operatorname{id})^m)$ for some $m \in \mathbb{N}$.

The minimal polynomial of α is $\prod_\lambda (X-\lambda)^{c_\lambda}$ where c_λ is the size of the largest λ-block. The characteristic polynomial of α is $\prod_\lambda (X-\lambda)^{a_\lambda}$ where a_λ is the sum of the sizes of the λ-blocks.

The number of λ-blocks is the dimension of the eigenspace of λ.

12.12 Modules over principal ideal domains (non-examinable)

The structure theorem above was proven for Euclidean domains. This also holds for principal ideal domains. Some of the ideas relevant to this proof are illustrated in this subsection.

Theorem. Let R be a principal ideal domain. Then any finitely generated torsion-free R-module is free.

If R is a Euclidean domain, this was proven as a corollary to the structure theorem.

Lemma. Let R be a principal ideal domain and M be an R-module. Let $r_1, r_2 \in R$ be not both zero, and let d be their greatest common divisor. Then,

(i) there exists $A \in SL_2(R)$ such that

$$A \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} = \begin{pmatrix} d \\ 0 \end{pmatrix}$$

(ii) if $x_1, x_2 \in M$, then there exist $x_1', x_2' \in M$ such that $Rx_1 + Rx_2 = Rx_1' + Rx_2'$, and $r_1 x_1 + r_2 x_2 = d x_1' + 0 \cdot x_2'$.

Proof. Since R is a principal ideal domain, $(r_1, r_2) = (d)$. Hence, by definition, $d = \alpha r_1 + \beta r_2$ for some $\alpha, \beta \in R$. Let $r_1 = s_1 d$ and $r_2 = s_2 d$. Then $\alpha s_1 + \beta s_2 = 1$. Now, let

$$A = \begin{pmatrix} \alpha & \beta \\ -s_2 & s_1 \end{pmatrix} \implies \det A = 1; \quad A \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} = \begin{pmatrix} d \\ 0 \end{pmatrix}$$

as required.

For the second part, let $x_1' = s_1 x_1 + s_2 x_2$ and $x_2' = -\beta x_1 + \alpha x_2$. Then $Rx_1' + Rx_2' \subseteq Rx_1 + Rx_2$. The matrix defining x_1', x_2' in terms of x_1, x_2 is invertible since its determinant is a unit; we can solve for x_1, x_2 in terms of x_1', x_2'. So $Rx_1' + Rx_2' = Rx_1 + Rx_2$. Then by direct computation we can see that $r_1 x_2 + r_2 x_2 = d x_1' + 0 \cdot x_2'$. □

The structure theorem for principal ideal domains follows the same method; it is deduced for Smith normal form. That theorem also holds for principal ideal domains. The above lemma allows one to

prove Smith normal form for principal ideal domains. In a Euclidean domain, we used the Euclidean function for a notion of size in order to perform induction; in a principal ideal domain we can count the irreducibles in a factorisation.

Proof of theorem. Let $M = Rx_1 + \cdots + Rx_n$ where n is minimal. If $x_1, \ldots, x_n$ are independent, then M is free as required. Suppose that the x_i are not independent, so there exists r_i such that $\sum r_i x_i = 0$ but not all of the r_i are zero. By reordering, we can suppose that $r_1 \neq 0$. By using part (ii) of the previous lemma, after replacing x_1 and x_2 by suitable x_1', x_2', we may assume that $r_1 \neq 0$ and $r_2 = 0$. By repeating this process with x_1 and x_i for all $i \geq 2$, we obtain $r_1 \neq 0$ and $r_2 = \cdots = r_n = 0$, so $r_1 x_1'' = 0$ for some nonzero $x_1'' \in M$. But M is torsion-free, so r_1 must be zero, and this is a contradiction. □