MA2301

Algebra II

Jia Xiaodong

May 7, 2021

1 Rings

Rings are in fact quite common in our daily experience. Let us start with some definitions followed by examples.

Definition 1.1 (Rings)**.** A ring R is a set with two binary operations, addition $+ : R \times R \to R$, and multiplication $\times : R \times R \to R$, such that

- $(R, +)$ is an abelian group with identity 0_R,
- $\times$ is associative, i.e.
$$(a \times b) \times c = a \times (b \times c) = a \times b \times c,$$
- $\times$ is distributive over addition, i.e.
$$(a + b) \times c = a \times c + b \times c \qquad c \times (a + b) = c \times a + c \times b.$$

Often we write ab or $a \cdot b$ instead of $a \times b$. □

Definition 1.2 (Commutative rings)**.** A ring R is *commutative* if multiplication $\times$ is commutative, that is $ab = ba$ for all $a, b \in R$. □

Note that in rings addition is already abelian, however there is no such requirement for multiplication.

Definition 1.3 (Multiplicative identities)**.** A ring R has an (multiplicative) identity if there exists an element 1_R such that $1 \times a = a = a \times 1$ for all $a \in R$. □

Also our definition of rings do not require the existence of a multiplicative identity, only an additive one.

Definition 1.4 (Division rings)**.** A ring R with $1 \neq 0$ is called a *division ring* if $(R - \{0\}, \times)$ is a group, that is for any $x \in R - \{0\}$, we can find $y \in R - \{0\}$ such that $xy = yx = 1$. □

Definition 1.5 (Fields)**.** A commutative division ring is called a *field*. □

Definition 1.6 (Inverses)**.** Let $(R, +, \times)$ be a ring. For any $a \in R$, we denote the additive inverse of a as $-a$, that is, $a + (-a) = 0$. □

Example 1.1.

- A single element forms a trivial ring.
- $(\mathbb{Z}, +, \times)$ is a ring, same for $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$.
- $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ are fields, but not $\mathbb{Z}$.
- Let C be the set of continuous functions $f : \mathbb{R} \to \mathbb{R}$. C is a ring together with the usual operations of function addition and multiplication, and 1 and 0 correspond to the usual one and zero functions. However it is not a field. For example the multiplicative inverse of $f(x) = x$ is not in C.
- The real polynomials form a ring.
- Let V be a finite dimensional real vector space. Then the set of all linear transformations from V to V, with addition and composition, forms a ring. Here 1 and 0 correspond to the identity and zero transformation respectively.
- $(\mathbb{Z}/n\mathbb{Z}, +, \times)$ forms a ring.

◊

Theorem 1.1 (Basic properties of rings)**.** *Let $(R, +, \times)$ be a group. For all $a, b \in R$,*

i. $0 \times a = a \times 0 = a$.

ii. 1 *is unique.*

iii. $-a = (-1)a$.

iv. $(-a)b = a(-b) = -(ab)$.

v. $(-a)(-b) = ab$.

Proof.

i. $0 \times a = (0 + 0) \times a = 0 \times a + 0 \times a \implies 0 \times a = 0$.

ii. Suppose there exists $1' \in R$ such that $1'a = a = a1'$. Then $1' = 1' \cdot 1 = 1$.

iii. $(-1)a + a = (-1)a + 1a = (-1 + 1)a = 0a = 0$. The same goes for $a + (-1)a$.

iv.

■

Definition 1.7 (Zero divisors)**.** Let R be a ring. An element $a \in R$, $a \neq 0$ is called a *zero divisor* if there exists $b \in R$, $b \neq 0$, such that $ab = 0$ or $ba = 0$. □

Example 1.2. Let $R = \mathbb{Z}/4\mathbb{Z}$. Then $2 \in R$ is a zero divisor, because first of all $2 \neq 0$, and secondly $2 \cdot 2 = 4 = 0$ (in R). ◊

Definition 1.8 (Units)**.** Let R be a ring where $0 \neq 1$. An element $u \in R$ is called an *unit* if there exists $v \in R$ such that $uv = vu = 1$. The set of units in R is denoted as R^*. We also write $v = u^{-1}$. □

Example 1.3. Let $R = \mathbb{Z}/4\mathbb{Z}$. $1 \in R$ is an (uninteresting) unit. $3 \in R$ is also an unit, since $3 \cdot 3 = 9 = 1$. ◊

Theorem 1.2. *Let R be a ring where $0 \neq 1$. Then R^* is a group under multiplication.*

Proof. R^* is closed under multiplication. Take any $a, b \in R^*$, then by definition $a^{-1}, b^{-1} \in R^*$ as well. This means that $ab \in R^*$ since $abb^{-1}a^{-1} = 1$. The other three group axioms are easy to verify. ■

Theorem 1.3. *Let R be a ring where $0 \neq 1$. Then a element cannot be both a unit and a zero divisor.*

Proof. Suppose $u \in R$ is a unit and a zero divisor. So there exists $v, w \in R$, $w \neq 0$, such that $uv = vu = 1$ and $(ux = 0) \vee (xu = 0)$. However

$$0 = v0 = vux = 1x = x$$

which is a contradiction. Checking for the other case is similar. ■

Definition 1.9 (Polynomial rings)**.** Let R be a commutative ring with 1. Let x be a formal variable [1]. The *polynomial ring* $R[x]$ is defined as the set consisting of all $a_n x^n + \cdots + a_1 x + a_0$ for all $n \in \mathbb{N}$. Addition and multiplication is defined as per ordinary polynomials. □

Definition 1.10 (Group rings)**.** Let G be a finite group. Let R be a commutative ring with $1 \neq 0$. Then we define the group ring over R, as a set $R[G]$ containing elements of the form $\sum_{g \in G} a_g \cdot g$, for $a_g \in R$ and $g \in G$. Addition is done component wise, and multiplication is defined as

$$(a_1 g_1 + a_2 g_2)(a_1' g_1' + a_2' g_2') = a_1 a_1' g_1 g_1' + a_1 a_2' g_1 g_2' + a_2 a_1' g_2 g_1' + a_2 a_2' g_2 g_2'.$$ □

Example 1.4. Consider the group ring $\mathbb{C}[\mathbb{Z}/2\mathbb{Z}]$. Let $\mathbb{Z}/2\mathbb{Z} = \{e, x\}$ with $x^2 = e$, the identity element. This group ring consists of all elements of the form $ae + bx$, for all $a, b \in \mathbb{C}$. Furthermore, $\mathbb{C}[\mathbb{Z}/2\mathbb{Z}]$ is a 2 dimensional complex vector space with the basis $\{e, x\}$. ♢

Definition 1.11 (Subrings)**.** Let R be a ring. A subring S of R is an additive subgroup that is closed under multiplication. □

Example 1.5. Below are some examples of subrings.

- $\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$.
- $2\mathbb{Z} \subset \mathbb{Z}$. ♢

The second example above illustrates that a subring may not necessarily have a multiplicative identity.

Definition 1.12 (Integral domains)**.** Let R be a commutative ring where $1 \neq 0$. R is called an *integral domain* if R has no zero divisors. □

Example 1.6. Some examples and counterexamples.

- $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{C}$, are all integral domains. In fact, any field is an integral domain.
- $\mathbb{Z}$ is an integral domain.

[1] This means that x is some additional element added to R that does not have any other properties.

- $\mathbb{C}[\mathbb{Z}/2\mathbb{Z}]$ is not an integral domain. It is commutative, however $(1-2)(1+2) = 1-2^2 = 1-1 = 0$.
- If R is a ring and G is a finite non-trivial group, then $R[G]$ is not an integral domain.

◊

Theorem 1.4 (Cancellation for integral domains)**.** *Let R be an integral domain. Let $x, y, z \in R - \{0\}$. Then $xy = xz$ implies $y = z$.*

Proof. Assume $xy = xz$. Then $xy - xz = xy + x(-z) = 0$, and so $x(y-z) = 0$. Since there are no zero divisors and x is non-zero, necessarily $y - z = 0$. ■

Theorem 1.5. *Let R be a finite integral domain. Then R is a field.*

Proof. We need to show that for any $x \in R - \{0\}$, we can find $y \in R - \{0\}$ such that $xy = yx = 1$. Consider the map $f : R \to R$ given by $f(a) = xa$. If $f(a) = xa = xb = f(b)$ then by cancellation, $a = b$ and so f is injective. However since R is finite, this means that f is also bijective, and hence invertible, so there exists such a y for all x. ■

Definition 1.13 (Product ring)**.** Let S and R be rings. We define the *product ring $S \times R$* as the set $S \times R$ using the usual Cartesian product together with addition and multiplication done component-wise, i.e.

$$\begin{aligned}(a,b) + (c,d) &= (a+c, b+d)\\ (a,b)(c,d) &= (ac, bd)\end{aligned}$$

□

The identity element $1_{S\times R}$ of a product ring $S \times R$ is $(1_S, 1_R)$. Consider $S' = \{(s,0) \mid s \in S\}$, which is a subring of $S \times R$. The identity element of S is $(1_S, 0) \neq (1_S, 1_R)$. This also shows that the identity of a subring might not be equal to the identity of the original ring.

If S and R are non-trivial rings, then $S \times R$ always has zero divisors, i.e. it is not an integral domain. Take $(1,0)(0,x) = 0_{S\times R}$.

Theorem 1.6. *Let R be an integral domain. Then the polynomial ring $R[x]$ is also an integral domain.*

Proof. Let $f(x), g(x) \in R[x]$, with degrees n and m respectively. Then the degree of $f(x)g(x)$ is $n+m$. Let $f(x) = a_n x^n + \cdots + a_1 x + a_0 \in R[x]$ with $a_n \neq 0$ and $g(x) = b_, x^m + \cdots + b_1 x + b_0 \in R[x]$ with $b_n \neq 0$. Then $f(x)g(x) = a_n b_m x^{m+n} + \cdots$.

If n and m are non-zero, then we are done. Otherwise, since R is an integral domain, $f(x)g(x)$ will still be non-zero. ■

Corollary 1.6.1. $(R[x])^* = R^*$.

Proof. Following from the above, the degree of $f(x)g(x)$ cannot decrease. Therefore the only way to obtain a product equating to 1 is for both f and g to be degree 0. ■

1.1 Ring homomorphisms and ideals

Definition 1.14 (Homomorphisms)**.** Let R and S be rings. A *ring homomorphism* $\varphi : R \to S$ is a map such that it preserves addition and multiplication,

$$\varphi(a+b) = \varphi(a) + \varphi(b) \qquad \varphi(ab) = \varphi(a)\varphi(b)$$

□

Definition 1.15 (Isomorphisms)**.** If a ring homomorphism $\varphi : R \to S$ is a bijection, then we call it an *isomorphism*. We say that R is *isomorphic* to S, and write $R \cong S$. □

Definition 1.16 (Kernels and images)**.** Let $\varphi : R \to S$ be a ring homomorphism. Then the *kernel* of φ is defined as $\ker\varphi = \varphi^{-1}[\{0_s\}] = \{a \in R \mid \varphi(a) = 0_s\} \subseteq R$. The *image* of φ is defined as $\operatorname{img}\varphi = \varphi[R] = \{b \in R \mid \exists a \in R\ \varphi(a) = S\}$ □

Example 1.7. The identity map $\mathrm{id} : R \to R$ defined by $\mathrm{id}(a) = a$ is a trivial homomorphism. The zero map $z : R \to S$ defined by $z(a) = 0$ is also a trivial homomorphism. ◊

Example 1.8. Let $R \subseteq S$ be a subring, then the embedding map $i : R \to S$ is a ring homomorphism and $\ker i = \{0_R\}$. ◊

Example 1.9. $\varphi : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ defined by $\varphi(a) = \overline{a}$ is a homomorphism of abelian groups. Furthermore, by the definition of multiplication in $\mathbb{Z}/n\mathbb{Z}$, we know that $\varphi(ab) = \overline{ab} = \overline{a}\overline{b} = \varphi(a)\varphi(b)$ so in fact φ is a ring homomorphism. ◊

Example 1.10. Let R be a commutative ring with $1 \neq 0$. Consider the evaluation map $\varphi : R[x] \to R$ that maps $f(x) \mapsto f(a)$, for some $a \in R$. Then φ is a ring homomorphism. ◊

Lemma 1.7. *Let $\varphi : R \to S$ be a ring homomorphism. Then φ is injective iff $\ker\varphi = \{0\}$.*

Proof. This statement is true if we consider φ as a group homomorphism. This is still true for rings since injectivity is a set theoretic property. ■

This kind of reduction to group homomorphisms will be used again later on.

Lemma 1.8. *Let $\varphi : R \to S$ be a ring homomorphism.*

i. $\operatorname{img}\varphi = \varphi[R]$ *is a subring of S.*

ii. $\ker\varphi$ *is a subring of R.*

iii. φ is an isomorphism iff there exists a ring homomorphism $\psi : S \to R$ such that $\varphi \circ \psi = \mathrm{id}_R$ and $\psi \circ \varphi = \mathrm{id}_S$.

Proof. To check for subgroups, it suffices to check that $\varphi[R]$ $(\ker\varphi)$ is an abelian group under addition and $\varphi[R]$ $(\ker\varphi)$ is closed multiplication. Since φ is a also a group homomorphism, $\varphi[R]$ $(\ker\varphi)$ is an abelian group so we only need to check multiplication.

i. Let $a, b \in \varphi[R]$, so there must be $c, d \in R$ such that $\varphi(c) = a$ and $\varphi(d) = b$. Then $\varphi(cd) = \varphi(c)\varphi(d) = ab$ so it is closed under multiplication.

ii. Now let $a, b \in \ker\varphi$, then $\varphi(ab) = \varphi(a)\varphi(b) = 0 \cdot 0 = 0$, so it is closed under multiplication.

iii. Take the set theoretic inverse of φ, namely $\varphi^{-1} : S \to R$. We claim that φ^{-1} is a ring homomorphism. Take $s, t \in S$. There exists $j, k \in R$ such that $\varphi(j) = s$ and $\varphi(k) = t$. Then $\varphi(k + t) = \varphi(k) + \varphi(t) = s + t$ and $\varphi(kt) = \varphi(k)\varphi(t) = st$. Then taking the inverse on both sides we get the result we are after.

The other direction comes directly from definition. ■

If we ignore multiplication, then we can see that a ring homomorphism is an abelian group homomorphism. From what we know about groups, $\ker\varphi$ is a normal subgroup, so we should imagine that in the context of rings, the subring $\ker\varphi$ would have some special properties as well.

Definition 1.17. For two rings A, B and $a \in A$, define $aB = \{ab \mid b \in B\}$. We define Ba and other more complicated combinations in a similar way. □

Definition 1.18 (Ideals)**.** Let R be a ring and I be a subgroup of R. Then I is a *left (right) ideal* of R if I is a subring and for any $r \in R$, $rI \subseteq I$ $(Ir \subseteq I)$. If I is both a left and right ideal of R then we call it a *two-sided ideal* of R, or just an *ideal* in short. If an ideal $I \neq R$ we call it a *proper* ideal. □

Example 1.11. If R is a ring, then $\{0\} \subseteq R$ and $R \subseteq R$ are trivial ideals. ◊

Lemma 1.9. *Let $\varphi : R \to S$ be a ring homomorphism. Then $\ker\varphi$ is an ideal of R.*

Proof. We have already shown that $\ker\varphi$ is a subring. Take $r \in R$ and $k \in \ker\varphi$. Then $\varphi(rk) = \varphi(r)\varphi(k) = \varphi(r) \cdot 0 = 0$ so $rx \in \ker\varphi$. Similarly, we can show that $xr \in \ker\varphi$. Therefore we conclude that $r\ker(\varphi) \subseteq \ker(\varphi) \supseteq \ker(\varphi)r$. ■

Example 1.12. $n\mathbb{Z} \subseteq \mathbb{Z}$ is an ideal. ◊

Lemma 1.10. *Let $\varphi : R \to S$ be a ring homomorphism. If J is an ideal of R, then $\varphi[J]$ is an ideal of S.*

Proof. Take any $a, b \in \varphi[J]$. Then there exists $a', b' \in J$ such that $\varphi(a') = a$ and $\varphi(b') = b$. Now $a - b = \varphi(a') - \varphi(b') = \varphi(a' - b') \in \varphi[J]$.

Take any $s \in S$ and $a \in \varphi[J]$. There exists $a', \in J$ such that $\varphi(a') = a$ and $s' \in R$ such that $\varphi(s') = s$. Now since J is an ideal, $s'a' \in J$ and so $\varphi(s'a') = \varphi(s')\varphi(a') = sa \in \varphi[J]$. The same goes for the other side, as. ■

Lemma 1.11. *Let $\varphi : R \to S$ be a ring homomorphism. If J is an ideal of S, then $\varphi^{-1}[J]$ is an ideal of R.*

Proof. Similar. ■

Definition 1.19 (Quotient rings)**.** Let $I \subseteq R$ be an ideal. We define the quotient ring R/I as the quotient group R/I under addition, together with multiplication defined by $\overline{a}\overline{b} = \overline{ab}$.

Here we use $\overline{a} = a + I$ to denote an element (coset) in R/I. □

Theorem 1.12. *The definition for quotient rings is well defined and indeed forms a ring.*

Proof. Let $\overline{a}, \overline{b} \in R/I$. Suppose $\overline{a'}, \overline{b'} \in R/I$ such that $\overline{a} = \overline{a'}$ and $\overline{b} = \overline{b'}$. We need to check that $\overline{a'b'} = \overline{ab}$.

If $a' = a + i_1$ and $b' = b + i_2$ for some $i_1, i_2 \in I$, then $a'b' = ab + ai_2 + i_1 b + i_1 i_2 \in ab + I$, so in other words $\overline{a'b'} = \overline{ab}$.

Finally we need to prove associativity and distributivity. They are shown in a similar way, and we skip them. ■

Example 1.13. $R/\{0\} \cong R \cong R/R$ are trivial quotient rings. ♢

Example 1.14. $\mathbb{Z}/n\mathbb{Z}$ is a quotient ring. ♢

Theorem 1.13 (First isomorphism theorem)**.** *Let $I \subseteq R$ be an ideal.*

i. The natural quotient map $\pi : R \to R/I$, defined by the map $a \mapsto \overline{a}$ is a ring homomorphism.

ii. Let $\varphi : R \to S$ be a ring homomorphism, then $R/\ker\varphi \cong \varphi[R]$.

iii. Let $\varphi : R \to S$ be a ring homomorphism such that $I \subseteq \ker\varphi$. Then there exists an unique ring homomorphism $\overline{\varphi} : R/I \to S$ such that the following diagram commutes:

$$\begin{tikzcd}
R \arrow[r, "\varphi"] \arrow[d, "\pi"'] & S \\
R/I \arrow[ur, dashed, "\overline{\varphi}"'] &
\end{tikzcd}$$

Proof.

i. We already know that $\pi : R \to R/I$ is a homomorphism of abelian groups. $\pi(ab) = \overline{ab} = \overline{a}\overline{b} = \pi(a)\pi(b)$ so it is a ring homomorphism as well. It follows directly that $\ker\pi = I = 0 + I$.

ii. Consider the map $\gamma : R/\ker\varphi \to \varphi[R]$ given by $\overline{a} = a + \ker\varphi \mapsto \varphi(a)$. We already know it is an isomorphism of abelian groups. What we need to check is that it is compatible with multiplication. Indeed it is: $\gamma(\overline{a}\overline{b}) = \varphi(ab) = \varphi(a)\varphi(b) = \gamma(\overline{a})\gamma(\overline{b})$

iii. We claim that $\overline{\varphi} : R/I \to S$ defined by $\overline{a} \mapsto \varphi(a)$ satisfies the claim. Uniqueness follows from the fact that π is surjective, the definition is essentially forced by the requirements. Next we check that it is well-defined. Suppose there exists $a, a' \in R$ such that $\overline{a} = \overline{a}'$. However $\overline{\varphi}(\overline{a}) = a = a' = \overline{\varphi}(\overline{a}')$ so $\overline{\varphi}$ is well defined. Next we check that $\overline{\varphi}$ is a ring homomorphism. We have $\overline{\varphi}(\overline{a}\overline{b}) = \overline{\varphi}(\overline{ab}) = \varphi(ab) = \varphi(a)\varphi(b) = \overline{\varphi}(\overline{a})\overline{\varphi}(\overline{b})$. Similarly so for addition.

■

Theorem 1.14. *If $M_1 \cong M_2$ and $N_1 \cong N_2$ then $M_1/N_1 \cong M_2/N_2$.*

Proof. Let $\varphi : M_1 \to M_2$ be an isomorphism and $\pi : M_2 \to M_2/N_2$ be the quotient map. Then by the first isomorphism theorem

$$M_1/N_1 = M_1/\varphi^{-1}(N_2) = M_1/\ker(\pi \circ \varphi) \cong M_2/N_2.$$

■

Definition 1.20 (Universal property of quotient rings)**.** Let $I \subseteq R$ be an ideal. Then the quotient ring of R by I is defined to be a ring $\tilde{R}(\cong R/I)$ together with a ring homomorphism $\pi : R \to \tilde{R}$ such that for any ring homomorphism $\varphi : R \to S$ with $I \subseteq \ker \varphi$, there exists an unique ring homomorphism $\overline{\varphi} : \tilde{R} \to S$ making the following diagram commute

$$\begin{tikzcd}
R \arrow[r, "\varphi"] \arrow[d, "\pi"'] & S \\
\tilde{R} \arrow[ur, dashed, "\overline{\varphi}"'] &
\end{tikzcd}$$

□

Example 1.15. $\varphi : a + n\mathbb{Z} \mapsto a + nm\mathbb{Z}$.

$$\begin{tikzcd}
\mathbb{Z} \arrow[r, "\pi"] \arrow[d, "\pi'"'] & \mathbb{Z}/nm\mathbb{Z} \\
\mathbb{Z}/n\mathbb{Z} \arrow[ur, dashed, "\varphi"'] &
\end{tikzcd}$$

◇

Example 1.16. Let G be a finite group and let R be a commutative ring with $1 \neq 0$. Then consider the map $\pi : R[G] \to R$ given by $\sum a_g g \mapsto \sum a_g$. We can check that it is a homomorphism. We only show a proof for two terms, but it can be easily generalised.

$$\begin{aligned}
\pi((ag + bh)(a'g' + b'h')) &= \pi(aa'gg' + ab'gh' + ba'hg' + bb'hh') \\
&= aa' + ab' + ba' + bb' \\
&= (a + b)(a' + b') \\
&= \pi(ag + bh)\pi(a'g' + b'h').
\end{aligned}$$

Let $\varphi : G \to H$ be a group homomorphism. Then φ induces a ring homomorphism $\varphi' : R[G] \to R[H]$ defined by $\sum a_g g \mapsto \sum a_g \varphi(g)$. ◇

Example 1.17. Consider all integer solutions to $x^2 + y^2 = 3z^2$. Suppose $x, y, z \in \mathbb{Z}^3$ are coprime and forms a solution. If they are not coprime then we can factor out their divisors until they are coprime. Then this equation must hold in $\mathbb{Z}/3\mathbb{Z}$ as well. Consider the natural map $\pi : \mathbb{Z} \to \mathbb{Z}/3\mathbb{Z}$.

$$\pi(x)^2 + \pi(y)^2 = \pi(x^2 + y^2) = \pi(3z^2) = \pi(3)\pi(z^2) = 0$$

We can exhaustively check the nine combinations for $\pi(x)$ and $\pi(y)$ such that $\pi(x) = 0 = \pi(y)$. However this means that x and y have a common divisor 3. Thus there is no non-trivial integer solutions, i.e. the only solution is $0, 0, 0$. ◇

Theorem 1.15 (Other isomorphism theorems)**.** *Let R be a ring.*

i. Let $A \subseteq R$ be a subring and $B \subseteq R$ be an ideal. Then we have a ring isomorphism $(A + B)/B \cong A/(A \cap B)$.

ii. Let $I \subseteq J \subseteq R$ be ideals. Then $R/J \cong (R/I)/(J/I)$.

iii. Let I be an ideal. There exists an order ($\subseteq$) preserving bijection from subrings (ideals) of R containing I to the subrings (ideals) of R/I.

Proof.

i. For this to work of course $A+B$ and $A \cap B$ have to be ideals. The proofs for them are skipped. Next, we claim that the isomorphism is given by $\varphi : A \to (A+B)/B$ defined by $a \mapsto \overline{a} = a + B$. Furthermore, it is surjective. For any $\overline{x} \in (A+B)/B$, there exists $a \in A$ and $b \in B$ such that $a + b = x$. But we have $\varphi(a) = a + B = a + b + B = \overline{x}$. Finally, we show that $\ker \varphi = A \cap B$. Indeed, if $\overline{a} = \overline{0} = B$, then it must be that $a \in B$ as well. Hence theorem 1.13 provides the isomorphism between $\varphi[A] = (A+B)/B$ and $A/\ker\varphi = A/(A \cap B)$.

ii. Consider the following diagram. Since $I \subseteq \ker \pi_J = J$, by theorem 1.13 we have a unique map $\varphi : R/I \to R/J$ given by $a + I \mapsto a + J$.

$$\begin{tikzcd}
R \arrow[r, "\pi_J"] \arrow[d, "\pi_I"] & R/J \\
R/I \arrow[ur, dashed, "\varphi"'] &
\end{tikzcd}$$

Theorem 1.13 also tells us that $(R/I)/\ker\varphi \cong R/J$. But we also have

$$\begin{aligned}
\ker\varphi &= \{a + I \mid \varphi(a+I) = 0\} \\
&= \{a + I \mid \varphi(a+I) = \pi_J(a) = 0\} \\
&= \{a + I \mid a \in J\} \qquad\qquad = J/I.
\end{aligned}$$

iii. Consider the quotient map $\pi : R \to R/I$ defined by $a \mapsto \overline{a}$. Let S be a subring of R. Then the map φ defined by $S \mapsto \overline{S} = \pi[S]$ is the bijection in question. Checking that φ is an order preserving bijection is skipped.

To check that it preserves ideals, let I be an ideal and $\overline{r} \in R/I$. Then $\overline{r}\overline{I} = \pi(r)\pi[I] = \pi[rI] \subseteq \pi[I] = \overline{I}$. The other side is checked similarly. Next, if $\overline{I} = \varphi(R)$ is an ideal and $r \in R$, then $\varphi(rI) = \overline{r}\overline{I} \subseteq \overline{I}$, which means $rI \subseteq I$.

■

1.2 Properties of ideals

Definition 1.21 (Sum and product of ideals)**.** Let I and J be ideals of A. Then we define the sum $I + J = \{i + j \mid i \in I, j \in J\}$. We also define the product $IJ = \{\sum ij \mid i \in I, j \in J\}$ (the sum is finite). Using this we also have the shorthand $I^n = I \cdots I$ (repeated n times). □

Theorem 1.16. *Let R be a ring with $1 \neq 0$ and let I and J be ideals of R.*

i. $I + J$ is the smallest ideal of R containing both I and J.

ii. IJ is an ideal contained in $I \cap J$.

iii. If R is commutative and $I + J = R$, then $IJ = I \cap J$.

Proof.

i. Let $K = I + J$. First we show that K is an ideal. Take any $r \in R$. Then $Kr = \{(a+b)r \mid a \in I, b \in J\}$. However, since I and J are ideals, $ar \in I$ and $br \in J$ and so $(a+b)r = ar + br \in K$. The same goes for rK. Therefore K is an ideal.

   Next, suppose there is another ideal H that contains both, I and J. For any $a \in I$, $b \in J$, since $a, b \in H$ as we suppose, then it has to be closed under addition, i.e. $a + b \in H$. Thus $K \subseteq H$.

ii. Let $L = IJ$. First we show that L is an ideal. Take any $r \in R$, then $Lr = \{(\sum ab)r \mid a \in I, b \in J\}$. By distributivity, $(\sum ab)r = \sum abr$ and since $br \in J$ we find that $\sum abr \in IJ = L$. The same goes for rL. Therefore L is an ideal.

   Next we have to show that $L = IJ \subseteq I \cap J$. Take any element $l = \sum_i a_i b_i \in L$ with all $a_i \in I$ and $b_i \in J$. Consider a term of the finite sum, $a_j b_j$. Since I is an ideal, $a_j b_j \in I$. Since J is an ideal, $a_j b_j \in J$ as well. Therefore, the entire finite sum is in both I and J.

iii. We have already shown that $IJ \subseteq I \cap J$. We now show that in this case we also have $I \cap J \subseteq IJ$.

   Firstly, since $1 \in R = I + J$, this implies that there is some $a \in I$ and $b \in J$ such that $a + b = 1$. Now take any $i \in I \cap J$. Then $1i = i = (a+b)i = ai + ib \in IJ$. This shows that $I \cap J \subseteq IJ$ and so $I \cap J = IJ$.

■

Theorem 1.17. *Let R be an integral domain. Then $(a) = (b)$ iff $a = bu$ for some unit u.*

Proof. We have $a \in (b)$ so there exists $r \in R$ such that $a = br$. Similarly, we have $r' \in R$ such that $b = ar'$. Then $a = ar'r$. Suppose now that $a = 0$. Then $b = 0$ and we are done. If $a \neq 0$, then cancellation in integral domains theorem 1.4 imply that $r'r = 1$. ■

Definition 1.22. We define the left ideal generated by A to be the smallest left ideal containing A, and it is given by RA. The right ideal generated by A is the smallest right ideal containing A, and it is given by AR. □

To check that this expression is correct, we only have to note that $RA \subseteq I$ for all I such that $A \subseteq I$. Hence $RA \subseteq \bigcap I$, but since RA is also a left ideal (checking skipped) that contains A, we conclude that in fact $RA = \bigcap I$. The case for right ideals is similar. We can also generalize it to two-sided ideals.

Definition 1.23 (Generators of ideals)**.** Let R be a ring with $1 \neq 0$. Let $A \subseteq R$. We denote by (A) the ideal *generated* by A, which is the smallest ideal in R containing A:

$$(A) = \bigcap_{\substack{A \subseteq I \\ I \text{ is ideal}}} I = RAR = \Big\{ \sum r_1 a r_2 \mid r_1, r_2 \in R, a \in A \Big\}.$$

An ideal generated by a finite set is called *finitely generated.* □

Of course, if R is commutative, we have simply $RAR = RRA = RA = AR = (A)$.

Definition 1.24 (Principle ideals)**.** An ideal generated by a single element is called a *principle ideal.* □

By definition, principle ideals are also finitely generated.

Example 1.18. $(0) = 0$ and $(1) = R$ are trivial principle ideals. ♢

Example 1.19. In $\mathbb{Z}$, $(n) = n\mathbb{Z}$ is a principle ideal. ♢

Example 1.20. The ideal $(2, x) \subseteq \mathbb{Z}[x]$ is not principle. Assume instead that it is, i.e. $(2, x) = (r)$ for some $r \in \mathbb{Z}[x]$. In order to generate 2, r has to be degree 0, or in other words, $r \in \mathbb{Z}$. Since $r \in (2, x)$, we have $r = 2a + bx$ for some $a, b \in \mathbb{Z}$. However since r is degree 0, so must b, so $(r) = (2a)$, which clearly does not contain the element x. ♢

Example 1.21. Let $R = 2\mathbb{Z}[x]$ (without 1). R is not a finitely generated ideal. Assume the contrary. Then $R = (f_1, f_2, \dots, f_n)$. Then, suppose f_i has the highest degree and let that be m. Then, $2x^{m+1} \notin (f_1, f_2, \dots, f_n)$ since multiplication of f's will give us a leading coefficient of 4. ♢

Lemma 1.18. *Let R be a ring with $1 \neq 0$.*

i. Let $I \subseteq R$ be an ideal. Then $I = R$ iff I contains an unit.

ii. If R is commutative, then R is a field iff the only ideals of R are the trivial ideals R and (0).

Proof. Recall that a ring R is a field iff every non-zero element is an unit.

i. ($\implies$): $1 \in I = R$ is an unit.

($\impliedby$): Let $u \in I$ be an unit. Then $ru^{-1}u = r \in I$ for any $r \in R$. So $R \subseteq I$.

ii. ($\implies$): For any $I \neq (0)$ (i.e. I is a non-zero ideal), I contains an unit. Therefore $I = \mathbb{R}$.

($\impliedby$): Let $0 \neq u \in R$. We consider $I = (u) = Ru = R$. So there exists $v \in R$ such that $vu = 1$.

■

Definition 1.25 (Maximal ideals)**.** Let $I \subsetneq R$ by an (possibly one sided) ideal. Then I is called (non-trivial) *maximal* if any ideal $J \supseteq I$ is either I or R. □

Lemma 1.19. *Let R be a commutative ring with $1 \neq 0$. Then*

i. R is a field iff (0) is a maximal ideal.

ii. Let $I \subseteq R$ be an ideal. Then I is maximal iff R/I is a field.

Proof.

i. This follows directly from lemma 1.18.

ii. ($\implies$): Assume that I is maximal. Recall the bijection from the ideals of R to the ideals of R/I From the isomorphism theorems (theorem 1.15) we have an order preserving bijection from the ideals of R containing I to the ideals of R/I induced by the projection map $\pi : R \to R/I$. Since I is maximal, $\varphi(I) = I/I = (0)$ is also maximal and hence R/I is a field by part (i).

($\impliedby$): By convention R/I contains identities $1 \neq 0$. Similarly as above, using the same bijection we understand that I is a maximal ideal.

■

Example 1.22. The ideal generated by a prime number p is a maximal ideal. It suffices to show that $\mathbb{Z}/p\mathbb{Z}$ is a field. This is equivalent to showing that $\mathbb{Z}/p\mathbb{Z}$ is an integral domain by theorem 1.5. Consider non zero elements $\overline{a}, \overline{b} \in \mathbb{Z}/p\mathbb{Z}$, which means $p \nmid a$ and $p \nmid b$. Then $p \nmid ab$, and so $\overline{a}\overline{b} \neq \overline{0}$. ♢

Example 1.23. The ideal $(x-a) \subset \mathbb{C}[x]$ is maximal for any $a \in \mathbb{C}$. We consider the map $\varphi : \mathbb{C}[x] \to \mathbb{C}$ defined by $f \mapsto f(a)$. It is easy to check that $\ker \varphi = (x-a)$. Then by the isomorphism theorem, we have $\mathbb{C}[x]/(x-a) \cong \mathbb{C}$ is a field. Therefore $(x-a)$ is maximal. ♢

The following proof uses Zorn's Lemma so we will state it here[2].

Theorem 1.20 (Zorn's Lemma)**.** *Let Σ be a partially ordered set. If every chain in Σ has an upper bound then there is a maximal element in Σ.*

Theorem 1.21. *Let R be a ring with $1 \neq 0$. Then any proper ideal (left, right, or two-sided) I is contained in a maximal ideal of the same kind.*

Proof. Let $\Sigma = \{J \subseteq R \mid J \text{ is a proper ideal}, J \subseteq I\}$. Let $C \subseteq \Sigma$ be a chain with respect to the inclusion ($\subseteq$) partial order. Define $k = \bigcup_{J \in C} J$. First we claim that k is an ideal. To check that it is an abelian group under addition, let $x, y \in k$. Then $x \in J_1$ and $y \in J_2$. Without loss of generality we assume $J_1 \subseteq J_2$, then $x, y \in J_2$ and so $x - y \in J_2$. To check that $rK \subseteq K$ is fairly easy and we skip it.

Furthermore $k \neq R$ since all J's do not contain 1 (if they did, then we would have $R \subseteq J$). So any chain in Σ has an upper bound given by k.

By Zorn's Lemma, there is a maximal element $M \in \Sigma$. We claim that M is a maximal ideal containing I. Suppose we can find $M' \supsetneq M \supseteq I$, and $M' \neq R$, then M is not the maximal element in Σ. ■

Note that there can be multiple maximal ideals in the above theorem. We only show that there exists at least one that contains I. The following example highlights this.

Example 1.24. Let p and q be different primes in $\mathbb{Z}$, and consider $(pq) \subset \mathbb{Z}$. Then $(pq) \subset (p)$ and $(pq) \subset (q)$. So the maximal ideal containing (pq) is not unique. ♢

Definition 1.26 (Prime ideals)**.** Let R be commutative with $1 \neq 0$. An ideal $P \subsetneq R$ is called *prime* if for any $ab \in P$, we have $a \in P$ or $b \in P$ or both. □

Example 1.25. Let $p \in \mathbb{Z}$ be any prime number, then $(p) \subset \mathbb{Z}$ is a prime ideal. Let $ab \in (p)$, so $ab = cp$ for some $c \in \mathbb{Z}$. Then either $p \mid a$ or $p \mid b$, since p is a prime. So either $a \in (p)$ or $b \in (p)$. ♢

Example 1.26. (0) is a prime ideal in any integral domain R. If $ab = 0$, then $a = 0$ or $b = 0$ since R is an integral domain. So either $a \in (0)$ or $b \in (0)$. ♢

Lemma 1.22. *Let R be commutative with $1 \neq 0$. Let $P \subset R$ be a proper ideal. Then P is prime iff R/P is an integral domain.*

[2] For a reference on Zorn's Lemma see the MA3205 set theory notes.

Proof.

($\Longrightarrow$): Let $\overline{a}\overline{b} = \overline{0} \in R/P$. Now $\overline{a}\overline{b} = \overline{0}$ iff $ab \in P$ iff $a \in P$ or $b \in P$ iff either $\overline{a} = \overline{0}$ or $\overline{b} = \overline{0}$.

($\Longleftarrow$): Let $ab \in P$. Then $\overline{ab} = \overline{a}\overline{b} = \overline{0} \in R/P$. Since R/P is an integral domain, $\overline{a} = \overline{0}$ or $\overline{b} = \overline{0}$ and so either $a \in P$ or $b \in P$. ■

From lemma 1.19 we know that if I is maximal then R/I is a field. But a field is also an integral domain, so the above lemma also tells us that a maximal ideal is also a prime ideal in a commutative ring.

Example 1.27. Let R be a commutative ring with $1 \neq 0$ and suppose for every $a \in R$ there exists $n > 1$ such that $a^n = a$. Then every prime ideal in R is also maximal.

Let P be a prime ideal in R and consider R/P. For any $0 \neq \overline{a} \in R/P$ we have $\overline{a} - (\overline{a})^n = \overline{a}(1 - \overline{a}^{n-1}) = 0$. Since P is prime, R/P is an integral domain, so $1 - \overline{a}^{n-1} = 0$ and thus $\overline{a}(\overline{a})^{n-2} = \overline{1}$. This means R/P is a field and hence P is also maximal. ♢

Theorem 1.23. *Let R and S be two commutative rings with $1 \neq 0$. Let $\varphi : R \to S$ be a ring homomorphism.*

i. *If P is a prime ideal of S then $\varphi^{-1}[P]$ is either R or a prime ideal of R.*

ii. *Assume φ is surjective. Then if M is a maximal ideal of S, $\varphi^{-1}[M]$ is a maximal ideal of R.*

Proof.

i. We already know that $\varphi^{-1}[P]$ is an ideal of R. If it is exactly R then we are done. So suppose that it is not, and we now want to show that it is a prime ideal. It suffices to show that $R/\varphi[P]$ is an integral domain. By the first isomorphism theorem

$$\begin{tikzcd} R \arrow[r, "\pi\circ\varphi"] \arrow[d, "\pi'"'] & S/P \\ R/\varphi^{-1}[P] \arrow[ur, dashed, "\overline{\varphi}"'] & \end{tikzcd}$$

the map $\overline{\varphi}$ is injective. Since S/P is an integral domain, so is $R/\varphi^{-1}[P]$.

ii. Similar to the above with the observation that S/M is a field.

■

Theorem 1.24. *Let R be a commutative ring with $1 \neq 0$. Let P be a prime ideal of R, and let I and J be two ideals of R such that $I \cap J \subseteq P$. Then either $I \subseteq P$ or $J \subseteq P$.*

Proof. Suppose on the contrary $I \not\subseteq P$ and $J \not\subseteq P$. Then there is $i \in I$, $j \in J$, such that $i, j \notin P$. Then $ab \in IJ \subseteq I \cap J \subseteq P$ so either $a \in P$ or $b \in P$, a contradiction. ■

Definition 1.27 (Nilradicals)**.** Let R be commutative with $1 \neq 0$. We define the *nilradical* of R as the ideal of all nilpotent elements of the ring,

$$\mathfrak{N}(R) = \{x \in R \mid x^n = 0, n > 0\}.$$ □

Theorem 1.25. *$\mathfrak{N}(R)$ is an ideal.*

Proof. Firstly, let $x \in \mathfrak{N}(R)$ and $x^n = 0$. Then for any $r \in R$, we have $(rx)^n = r^n x^n = 0$ so $rx \in \mathfrak{N}R$.

Next, we check that it is an abelian group under addition. let $x, y \in \mathfrak{N}(R)$, with $x^n = 0$ and $y^m = 0$. Take $a = \max(n, m)$, then we have $(x-y)^{2a} = \sum_{i=0}^{2a} \binom{2a}{i} x^i y^{2a-i}(-1)^{2a-i}$. Since for each term either $i \geq n$ or $2n - 1 \geq n$, every term goes to zero. ■

Theorem 1.26. *Let R be a commutative ring with $1 \neq 0$. Let P be the set of all prime ideals of R. Then $\mathfrak{N}(R) = \bigcap P$.*

Proof. We first show that $\mathfrak{N}(R) \subseteq \bigcap P$. Let $x \in R$ such that $x^n = 0$ for some $n > 0$. Then $x^n = 0 \in Q$ for some prime Q. Since Q is prime, we have either $x \in Q$ or $x^{n-1} \in Q$. If the latter, we repeat the process inductively.

Next we show that $\mathfrak{N}(R) \supseteq \bigcap P$. We will show that if $x \notin \mathfrak{N}(R)$, then $x \notin Q$ for some prime ideal Q. Let $\Sigma = \{J \subsetneq R \mid x^m \notin J, m \in \mathbb{Z}^+\}$. Since we assume that x is not nilpotent, $\Sigma \supseteq (0)$ and so is not empty. Let $C \subseteq \Sigma$ be a chain with respect to the inclusion partial order. Similarly to theorem 1.21, $k = \bigcup_{J \in C} J$ is an ideal and $k \neq R$. Also, $x^m \notin k$ for any $m > 0$, since x^m is not in any of the J's. So $k \in \Sigma$ is an upper bound for C. So by Zorn's lemma, Σ has a maximal element M. We claim that M is a prime ideal. Let $a, b \in R$ such that $a \notin M$, $b \notin M$, but $ab \in M$. Then $(a) + M \supsetneq M \subsetneq (b) + M$, which by the maximality of M means that $x^c \in (a) + M$ and $x^d \in (b) + M$ for some $c, d > 0$. Then $x^{c+d} \in ((a) + M)((b) + M) = (ab) + M = M$, contradiction. Therefore M is a prime ideal, and by assumption $x^m \notin M$, which concludes the proof. ■

The nilradical is in fact just a special case of the radical. We see that $\sqrt{0} = \mathfrak{N}(R)$.

Definition 1.28 (Radical of ideals)**.** Let R be a commutative ring with $1 \neq 0$. Let I be an ideal in R and define the *radical* of I as

$$\sqrt{I} = \{r \in R \mid r^n \in I, n \in \mathbb{Z}^+\}.$$

An ideal I is called *radical* if $\sqrt{I} = I$. □

Theorem 1.27. *$\sqrt{I}$ is an ideal, and it contains I.*

Proof. $I \subseteq \sqrt{I}$ is free from the definition. The fact that it is an ideal is checked the exact same way as for $\mathfrak{N}(R)$. ■

Theorem 1.28. *Any prime ideal P is radical.*

Proof. Let P be a prime ideal. It suffices to show that $\sqrt{P} \subseteq P$. For any $r \in \sqrt{P}$ such that $r^n \in P$, since P is prime either $r \in P$ or $r^{n-1} \in P$. For the former case, we are done, and for the latter, we just repeat inductively to show that $r \in P$. ■

Theorem 1.29. *Let $n \in \mathbb{Z}^+$. (0) is radical in $\mathbb{Z}/n\mathbb{Z}$ iff n is square free, i.e. n is a product of distinct primes to the first power.*

Proof.

($\Longrightarrow$): We show the contrapositive. Suppose that n is not square free, for instance $n = p^2$ for some prime number p. Then $\overline{p}^2 = \overline{0} \in \mathbb{Z}/n\mathbb{Z}$ so $\sqrt{(0)} \ni p \notin (0)$.

($\Longleftarrow$): Suppose that n is square free, so $n = p_1 \cdots p_i$. If $\overline{r}^m = \overline{0}$, then $r^m = nk$ for some $k \in \mathbb{Z}$. We factor r into its prime factors to get $(q_1 \cdots q_j)^m = p_1 \cdots p_i k$. Taking an exponent does not create any new factors, which then means that $q_1, \ldots, q_j$ must already exist on the right hand side, $r \mid n$ in the first place. This means $\overline{r} = \overline{0}$. ■

Definition 1.29 (Local rings)**.** A ring is said to be *local* if it has an unique maximal ideal. □

Theorem 1.30. *Let R be a commutative ring with $1 \neq 0$. If R is a local ring with an unique maximal ideal M, then every element not in M is a unit.*

Proof. Any non unit u is contained in M, since (u) is a proper ideal and is contained in M. ■

Theorem 1.31. *If the set of non-units of R forms an ideal then R is a local ring.*

Proof. Let M be the idea of non-units of R. Consider any other maximal ideal M'. Since M' is maximal, it cannot contain any units. So $M = M'$. ■

2 Field of fractions and localisations

The quotient is a way of creating a new ring from a ring. Localisations are also a way of creating something new from a ring. In this section, we will give three increasingly general definitions. Also we only consider commutative rings with $1 \neq 0$.

First, we have some examples as motivations for what we are going to do.

Example 2.1. Recall the construction of $\mathbb{Q}$ from $\mathbb{Z} \times \mathbb{Z}$. We define a equivalence relation $\sim$ where $(a, b) \sim (c, d)$ iff $ad = cb$. Addition is defined as $(a, b) + (c, d) = (ad + cb, bd)$, and multiplication is defined as $(a, b)(c, d) = (ac, bd)$. Any non-zero element is invertible. ♢

Example 2.2. Let $0 \neq a \in \mathbb{Z}$. Define $\mathbb{Z}\left[\frac{1}{a}\right] = \left\{\frac{b}{a^n} \mid b \in \mathbb{Z}, n \in \mathbb{Z}^+\right\} \subset \mathbb{Q}$. In $\mathbb{Z}\left[\frac{1}{a}\right]$, a is invertible. ♢

Definition 2.1 (Field of fractions)**.** Let R be an integral domain with $1 \neq 0$. Let $D = R - \{0\}$. Define the *field of fractions* of R or the *quotient field* of R, denoted as $F(R)$.

Let $F = \{(a, b) \mid a \in R, b \in D\}$ and the equivalence relation $\sim$ defined as $(a, b) \sim (c, d)$ iff $ad = cb$. Then $F(R) = F/\sim$. We will write $\frac{a}{b}$ to denote the equivalence class containing (a, b).

We define addition $+$ on $F(R)$ as

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

and multiplication $\cdot$ on $F(R)$ as

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

□

Checking that $\sim$ is an equivalence relation has been left out.

Theorem 2.1. *$F(R)$ is a well-defined field, and R is embedded in Q via the ring homomorphism $\varphi : R \to F(R)$ given by $r \mapsto \frac{r}{1}$ (i.e. φ is injective).*

Proof. First we check that $F(R)$ is a well defined ring. We check that $+$ is well-defined. Let $\frac{a}{b} = \frac{a'}{b'}$, i.e. $ab' = a'b$. Then

$$\frac{a}{b} + \frac{c}{d} = \frac{ad+cb}{bd} \frac{a'}{b'} + \frac{c}{d} \qquad = \frac{a'd+cb'}{b'd}$$

Now we also have

$$\begin{aligned}(ad+cb)b'd &= ab'd^2 + cbb'd \\ &= a'bd^2 + cb'bd \\ &= (a'd+cb')bd.\end{aligned}$$

Thus $+$ is well-defined. We will skip the checking of the well-definiteness of $\times$, commutativity, associativity, distributivity, since they are either similar, or very routine. Let us assume that $F(R)$ is a ring.

Then, We may do some very quick checking to confirm that $\frac{b}{b} = 1$. This then means that $\frac{a}{b}\frac{b}{a} = \frac{ab}{ab} = 1$, and so any $\frac{a}{b}$ is invertible for $a, b \neq 0$. We can also check that $\frac{0}{b} = 0$. Thus $F(R)$ is a field.

Now consider the map φ. Through the definition of multiplication and addition, it is easy to show that φ is a ring homomorphism. Now from definition $\ker\varphi = \left\{r \in R \mid \frac{r}{1} = 0\right\}$. However, since for any $d \neq 0$ we have $\frac{r}{1} = 0 = \frac{0}{d}$, this would mean that $rd = 0$, but since R is an integral domain, $r = 0$. Thus $\ker\varphi = \{0\}$. ■

We relax the requirement that R is an integral domain, and allow D to be slightly smaller.

Definition 2.2 (Localisations, version 1)**.** Let R be a commutative ring with $1 \neq 0$. Let $D \subseteq R - \{0\}$ be such that D contains no zero divisors, D contains 1, and is closed under multiplication[3]. We denote the *localisation* of R with respect to D by $D^{-1}R$.

Let $A = \{(a,b) \mid a \in R, b \in D\}$ and the equivalence relation $\sim$ defined as $(a,b) \sim (c,d)$ iff $ad = bc$. Then $D^{-1}R = A/\sim$. We will write $\frac{a}{b}$ to denote the equivalence class containing (a,b).

We define addition $+$ on $F(R)$ as

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

and multiplication $\cdot$ on $F(R)$ as

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

□

Theorem 2.2. *Let R be a commutative ring with $1 \neq 0$. Let $D \subseteq R - \{0\}$ be such that D contains no zero divisors, D contains 1, and is closed under multiplication.*

i. $D^{-1}R$ is a well-defined ring with embedding $\pi : R \to D^{-1}R$ given by $r \mapsto \frac{r}{1}$.

ii. Any $d \in D$ is invertible in $D^{-1}R$ via the embedding π, i.e. $\frac{d}{1}$ is invertible.

iii. (Universal property) Let S be any commutative ring with $1 \neq 0$. Let $\varphi : R \to S$ be a ring homomorphism such that $\varphi[D] \subseteq S^*$[4]. Then there is an unique map $\varphi' : D^{-1}R \to S$ that

[3] Closure under multiplication is required to keep the “denominator” well defined.

[4] Recall S^* denotes the group of units of S.

makes the following diagram commute.

$$\begin{tikzcd}
R \arrow[r, "\varphi"] \arrow[d, "\pi"'] & S \\
D^{-1}R \arrow[ur, dashed, "\varphi'"'] &
\end{tikzcd} \tag{1}$$

Proof.

i. ii. Similar as the previous. For injectivity of π we use the fact that D has no zero divisors instead of it being an integral domain.

iii. Define φ' as $\frac{r}{d} \mapsto \varphi(r)\varphi(d)^{-1}$. We can see that the definition is forced, since $\frac{1}{d}$ is the inverse of d and therefore $\varphi\left(\frac{1}{d}\right)$ should be mapped to the inverse of $\varphi(d)$. Therefore it is unique. Let us check that it is well-defined. Let $\frac{r'}{d'} = \frac{r}{d} \in D^{-1}R$, i.e. $r'd = rd'$. This also means that $\varphi(r')\varphi(d) = \varphi(r)\varphi(d')$. Furthermore since d is invertible $\varphi(r')\varphi(d')^{-1} = \varphi(r)\varphi(d)^{-1}$. Thus

$$\begin{aligned}
\varphi'\left(\frac{r}{d}\right) &= \varphi(r)\varphi(d)^{-1} \\
&= \varphi(r')\varphi(d')^{-1} \\
&= \varphi'\left(\frac{r'}{d'}\right).
\end{aligned}$$

It is left to check that φ' is a homomorphism, this is skipped.

■

Example 2.3. Let $R = \mathbb{Z}$ and $D = \{p^n \mid n \geq 0\}$ where p is a prime. The localisation of R with respect to D is $D^{-1}R = \mathbb{Z}\left[\frac{1}{p}\right]$. ◊

Example 2.4. Let $R = \mathbb{R}[x]$ the polynomials over real numbers. We consider them as polynomial functions on $\mathbb{R}$. Let $D = \{x^n \mid n \in \mathbb{N}\}$. Then the localisation of R with respect to D is given by

$$D^{-1}R = \left\{\frac{f(x)}{x^n} \mid f(x) \in \mathbb{R}[x], n \in \mathbb{N}\right\}.$$

These are functions on the open subset $\mathbb{R} - \{0\} \to \mathbb{R}$. This is a reason behind the name, since it allows us to study functions "locally". ◊

Theorem 2.3. *Let R be an integral domain. Then the field of fractions $F(R)$ is the smallest field containing R.*

Proof. In other words, if S is a field containing R, then S contains a subring isomorphic to $F(R)$. Let $D = R - \{0\}$ and $F(R) = D^{-1}R$. Then since elements in S are all invertible, using theorem 2.2, we have

$$\begin{tikzcd}
R \arrow[r, "\varphi"] \arrow[d, "\pi"'] & S \\
F(R) = D^{-1}R \arrow[ur, dashed, "\varphi'"'] &
\end{tikzcd}$$

Then φ' is injective. In the field $F(R)$ the only ideals are (0) and $F(R)$ itself. Thus $\ker \varphi'$ being an ideal has to be either of them. It cannot be $F(R)$ since that will just map everything to $0 \in S$ and φ will not be an embedding (R is not trivial since we assume $1 \neq 0$). Thus $\ker \varphi' = (0)$ and it is injective, and so there is some subring $\varphi'[F(R)]$ isomorphic to $F(R)$. ■

We further relax the conditions on the absence of 0 and zero divisors. The new definition is directly expressed as a similar theorem.

Theorem 2.4 (Localisations, version 2)**.** *Let R be a commutative ring with $1 \neq 0$. Let D be a multiplicatively closed subset of R containing 1. Then there exists a ring $D^{-1}R$ together with a ring homomorphism $\pi : R \to D^{-1}R$ such that*

i. *$\pi[D]$ is invertible.*

ii. *Let S be any commutative ring with $1 \neq 0$. Let $\varphi : R \to S$ be such that $\varphi[D] \subseteq S^*$. Then there is an unique map $\varphi' : D^{-1}R \to S$ that makes the diagram commute*

$$\begin{tikzcd}
R \arrow[r, "\varphi"] \arrow[d, "\pi"'] & S \\
D^{-1}R \arrow[ur, dashed, "\varphi'"'] &
\end{tikzcd}$$

iii. *$D^{-1}R$ is unique up to isomorphism.*

Proof. We will construct $D^{-1}R$ and the ring homomorphism $\pi : R \to D^{-1}R$. Let $Q = \{(a, b) \mid a \in R, b \in D\}$ and the equivalence relation $\sim$ be defined by $(a, b) \sim (c, d)$ iff $x(ad - bc) = 0$ for some $x \in D$. Then $D^{-1}R = Q/\sim$. Addition and multiplication are defined as before. Now we check that this construction satisfies the points above.

Let us check that multiplication is well-defined. Let $\frac{a}{b} = \frac{a'}{b'}$, i.e. $x(ab' - a'b) = 0$ for some $s \in D$. This would also mean that $x(acb'd - a'cbd) = x(ab' - a'b)cd = 0$. Therefore

$$\frac{a}{b}\frac{c}{d} = \frac{ac}{bd} = \frac{a'c}{b'd} = \frac{a'}{b'}\frac{c}{d}.$$

Addition is similar. ■

Example 2.5. Let R be any commutative ring. Let $D = \{0, 1\}$. Then $D^{-1}R \cong \{a\}$, a ring with a single element. Since homomorphisms preserve the zero element, then $\pi(0) = 0_{D^{-1}R}$ is invertible, meaning $0_{D^{-1}R} = 1_{D^{-1}R}$. ◊

Example 2.6. Let $R = \mathbb{Z}/6\mathbb{Z}$. This is not an integral domain. Let $D = \{\overline{1}, \overline{2}, \overline{4}\}$. Then $D^{-1}R \cong \mathbb{Z}/3\mathbb{Z}$. We can confirm this computationally. Some strange examples include $\frac{\overline{5}}{\overline{1}} = \frac{\overline{2}}{\overline{1}}$, since $\frac{\overline{3}}{\overline{1}} = 0$. We also have $\frac{\overline{1}}{\overline{2}} = \frac{\overline{2}}{\overline{1}}$, since $\overline{2}(\overline{1} \cdot \overline{1} - \overline{2} \cdot \overline{2}) = \overline{2}(-\overline{3}) = 0$.

We may see that localisations are in a way the opposite of taking quotients. Taking localisations over the powers of 2, gives us $\mathbb{Z}/3\mathbb{Z}$. We can also verify that taking quotients over the power of 3, so $D' = \{\overline{3}, \overline{1}\}$, gives us $\mathbb{Z}/2\mathbb{Z}$. ◊

We see that if R is an integral domain, then $x(ad - bc) = 0$ means $ad = bc$ as long as $x \neq 0$. If D has no zero divisors, then $x(ad - bc) = 0$ also means that $ad = bc$. So the previous two definitions are satisfied by this new construction.

Theorem 2.5. *Let R, D, and π be defined as usual (theorem 2.4).*

i. $\ker \pi = \{r \in R \mid xr = 0, x \in D\}$.

ii. $D^{-1}R = 0$ *iff* $0 \in D$.

Proof.

i. π is defined by $r \mapsto \frac{r}{1}$. Then $\frac{r}{1} = \frac{0}{d} = \frac{0}{1}$ iff $x(r \cdot 1 - r \cdot 0) = xr = 0$.

ii. If $0 \in D$ then $\frac{a}{b} = \frac{0}{1}$ as $0(a \cdot 1 - b \cdot 0) = 0$. If $D^{-1}R = 0$, then $\frac{1}{1} = \frac{0}{1}$, so $x(1 \cdot 1 - 1 \cdot 0) = x = 0$ for some $x \in D$.

■

3 Integers and things

3.1 Chinese remainder theorem

We will show, for example, that there is a ring isomorphism

$$\mathbb{Z}/30\mathbb{Z} = \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}.$$

In other words, there is an unique integer x mod 30 such that

$$x \equiv 1 \bmod 2 \qquad x \equiv 2 \bmod 3 \qquad x \equiv 2 \bmod 5.$$

We can check that $x = 17$. So the isomorphism maps $\overline{17} \mapsto (\overline{1}, \overline{2}, \overline{2})$.

Definition 3.1 (Coprimality)**.** Let R be commutative with $1 \neq 0$. Two ideals A and B of R are *coprime* if $A + B = R$. □

Recall that in this case $A \cap B = AB$ (theorem 1.16).

Example 3.1. In $\mathbb{Z}$, consider the ideals (2) and (3). Since $3 - 2 = 1$ we can see that indeed $(2) + (3) = \mathbb{Z}$ and that they are coprime. ◊

Theorem 3.1. *Let R be commutative with $1 \neq 0$. Let $A_1, A_2, \ldots, A_k$ be pairwise coprime ideals of R. Then we have the isomorphism*

$$R/(A_1 \cap A_2 \cap \cdots \cap A_k) = R/(A_1 A_2 \ldots A_k) \cong R/A_1 \times R/A_2 \times \cdots \times R/A_k.$$

Proof. First we have a ring homomorphism $R \to R/A_1 \times R/A_2 \times \cdots \times R/A_k$ given by the map $r \mapsto (r + A_1, r + A_2, \ldots, r + A_k)$. The checking is skipped. We also see that $\ker \varphi = A_1 \cap A_2 \cap \cdots \cap A_k$.

First we consider the claim for $k = 2$. By assumption $A_1 + A_2 = R$. Let $a_1 \in A_1$ and $a_2 \in A_2$ such that $a_1 + a_2 = 1$. Then $\varphi(a_1) = (a_1 + A_1, a_1 + A_2) = (0 + A_1, 1 - a_2 + A_2) = (\overline{0}, \overline{1})$. Similarly $\varphi(a_1) = (\overline{1}, \overline{0})$. Then we see that φ is surjective, since for any $(\overline{x}, \overline{y})$ we have $\varphi(ya_1 + xa_2) = \overline{y}(\overline{0}, \overline{1}) + \overline{x}(\overline{1}, \overline{0}) = (\overline{x}, \overline{y})$. Thus the first isomorphism theorem (theorem 1.13) tells us that $R/(A_1 + A_2) = R/(A_1 \cap A_2) = R/\ker \varphi \cong \varphi[R] = R/A_1 \times R/A_2$.

Now let us consider the general case. We claim that

$$R/(A_1 \cdots A_k) \cong R/A_1 \times R/(A_2 \cdots A_k).$$

After this, we can inductively show that

$$R/(A_1 \cdots A_k) \cong R/A_1 \times R/(A_2 \cdots A_k) \cong R/A_1 \times R/A_2 \times \cdots \times R/A_k.$$

To show the claim, we just need to show that A_2 and $A_2 \dots A_k$ are coprime, and we can use the proof for $k = 2$ on them. Now recall that $A_1 + A_i = R$ for $i = 2, \dots, k$. So we can find $x_i \in A_1$ and $y_i \in A_i$ such that $x_i + y_i = 1$. Then

$$(x_2 + y_2)(x_3 + y_3) \cdots (x_k + y_k) = 1$$

We can expand the left hand side, and collect all the terms as such

$$(x_2 + y_2)(x_3 + y_3) \cdots (x_k + y_k) = \sum_{i=2}^{k} (m_i x_i) + y_2 y_3 \cdots y_k.$$

Now the summation is contained in A_1 since it is an ideal, and obviously $y_2 \cdots y_k \in A_2 \cdots A_k$. Therefore we have shown that $1 \in A_1 + A_2 \cdots A_k$, and this is equivalent to saying $A_1 + A_2 \cdots A_k = R$. ■

3.2 Principle integral domains

We have the following picture

Integral doms. ⊃ unique factor. doms. ⊃ principle ideal doms. ⊃ Euclidean doms. ⊃ fields

We shall focus on principle ideal domains (PIDs), and only within the context of integral domains.

Definition 3.2 (Principle ideal domains)**.** A principle ideal domain is an integral domain in which every ideal is principle. □

Recall that a principle ideal is an ideal generated by a single element. This also means that $(x) = Rx = \{ax \mid a \in R\}$. We will use this fact often.

Example 3.2. $\mathbb{Z}$ is a PID. Let $I \subseteq \mathbb{Z}$ be a non-zero ideal (there is nothing to show for the zero ideal). Let $r \neq 0 \in I$ be the element such that $|r|$ is minimal, since $\mathbb{Z}^+$ is well-ordered. Then we claim $I = (r)$. It suffices to show that $I \subseteq (r)$. Take $x \in I$, then we do Euclidean division, which gives $x = qr + y$ for some $q, y \in \mathbb{Z}$ and $|y| < |r|$. By the minimality of $|x|$, $y = 0$. So $x = qr \in (r) = Rr$. ♢

The generalisation of this example leads to Euclidean domains, as we will see later on.

Theorem 3.2. *Let R be a PID. Then any non-zero prime ideal P is also maximal.*

Proof. Let $P = (p)$. Let an ideal $I \supsetneq P$. We show that $I = R$. Let $I = (i)$. Then $i \notin P$, otherwise $I = P$. Since $P \subset I$, we have $p \in I = Ri$ so $p = iq$ for some q. Since P is prime, either $i \in P$ (impossible) or $q \in P = Rp$. Therefore $q = rp$ for some r. Thus $p = iq = irp$. We have cancellation in integral domains (theorem 1.4), so this means $1 = ir$. Therefore, i is a unit and so $I = R$ (lemma 1.18). ■

The above theorem together with lemma 1.22 tells us that in a PID, maximal and prime ideals are the same thing.

Definition 3.3 (Greatest common divisors)**.** Let R be a commutative ring with $1 \neq 0$. Let $a, b \in R$ with $b \neq 0$. We say a is a *multiple* of b, or b *divides* a, if $a = bc$ for some $c \in R$. We write $b \mid a$.

A *greatest common divisor* (gcd) of a and b is a non-zero element $d \in R$ such that $d|a \wedge d|b$ and for any d' such that $d'|a \wedge d'|b$, we have $d'|d$. Then we write $d = \gcd(a, b)$[5]. □

Theorem 3.3. *Let R be an integral domain. Let $a, b \in R$ with $b \neq 0$. Let d and e both be the gcd of a and b. Then $d = ue$ for some unit u.*

Proof. Both d and e are gcd we have $d \mid e$ and $e \mid d$. So there exists $x, y \in R$ such that $dx = e$ and $ey = d$. So $eyx = e$. Cancellation gives us $yx = 1$ since $e \neq 0$. ■

Theorem 3.4. *Let R be a PID. Let $a, b \in R$ with $b \neq 0$. Let the ideal $(a, b) = (d)$. Then $d = \gcd(a, b)$. Therefore, the gcd always exists in a PID.*

Proof. First we show that $d \mid a$ and $d \mid b$. Since $a \in (d)$, there exists a' such that $a = da'$. Similarly, there exists b' such that $b = db'$.

Next, let $d' \mid a$ and $d' \mid b$, we show that $d' \mid d$. We have $a = d'a''$ and $b = d'b''$, so $a \in (d')$ and $b \in (d')$. Then $(d) = (a, b) \subseteq (d')$. Then, $d \in (d')$, so $d' \mid d$. ■

The previous theorem also implies that the gcd exists as a linear combination $\gcd(a, b) = ax + by$, simply as a condition of existing in the ideal. We may also note that we don't really need R to be a PID, all we need is that $(a, b) = (d)$.

Definition 3.4 (Noetherian rings)**.** Let R be commutative (with $1 \neq 0$) is called *Noetherian* if for any chain of ideals $I_1 \subseteq I_2 \subseteq \cdots$ of R, we can find an $m > 0$ such that $I_m = I_k$ for all $k \geq m$. Equivalently, $\bigcup_i I_i = I_m$. □

The condition is known as the *ascending chain condition*. It is a finiteness assumption. If R is Noetherian, then any ideal is contained in the maximal ideal without Zorn's lemma.

Theorem 3.5. *Let R be a PID. Then R is Noetherian.*

Proof. Let $I_1 \subseteq I_2 \subseteq \cdots$ be a chain of ideals in R. Then $I = \bigcup_i I_i = (i)$ is an ideal. Then there is some m such that $i \in I_m$, which means $I \subseteq I_m$. So $I_m = \bigcup_i I_i$. ■

3.3 Euclidean domains

Definition 3.5 (Norms)**.** Let R be an integral domain. A *norm* on R is a function $N : R \to \mathbb{N}$ such that $N(0) = 0$. □

We often consider the case where $N(r) > 0$ for all $r \neq 0$. This is the case for $\mathbb{Z}$.

[5] We ignore the problem of uniqueness and existence for now.

Definition 3.6 (Euclidean domains)**.** The ring R is called an *Euclidean domain* if we can perform the following division operation with respect to some norm N: for any $a, b \in R$ with $b \neq 0$, we have $a = qb + r$ with either $r = 0$ or $N(r) < N(b)$. We call q the *quotient* and r the *remainder.* □

Example 3.3. Any field with any norm is an Euclidean domain since we can perform division. ♢

Example 3.4. $\mathbb{R}[x]$ with $N(f(x)) = \deg f(x)$ (the degree of the polynomial) is an Euclidean domain. We can just do long division of polynomials. ♢

Example 3.5. The Gaussian integers $\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\} \subset \mathbb{C}$ with $N(a + bi) = a^2 + b^2$ is an Euclidean domain. For $\alpha, \beta \in \mathbb{Z}[i]$, with $\beta \neq 0$, we can do division in $\mathbb{C}$ and write $\frac{\alpha}{\beta} = x + yi \in \mathbb{Q}[i]$. Let $m, n \in \mathbb{Z}$ such that $|x - m| \leq \frac{1}{2}$ and $|y - n| \leq \frac{1}{2}$. Then

$$\begin{aligned}
\alpha &= (x + yi)\beta \\
&= (m + ni + x - m + (y - n)i)\alpha \\
&= \underbrace{(m + ni)\alpha}_{\in \mathbb{Z}[i]} + \underbrace{(x - m + (y - n)i)\beta}_{\in \mathbb{Z}[i]}
\end{aligned}$$

We also have

$$\begin{aligned}
N((x - m + (y - n)i)\beta) &= N(x - m + (y - n)i)N(\beta) \\
&\leq \frac{1}{2}N(\beta)
\end{aligned}$$

as long as $N(\beta) \neq 0$. So we have managed to perform our Euclidean division. ♢

Theorem 3.6. *Let R be an Euclidean domain. Then R is a PID.*

Proof. Recall example 3.2. This proof is similar. Let $I \subseteq R$ be a non-zero ideal. Then $I = (r)$ where r is the element in I with minimal norm. ■

Since an Euclidean is a PID, the gcd of a and b ($\neq 0$) always exists. The standard algorithm to find this is the Euclidean alogirthm. We demonstrate this with an example. Essentially we are doing the same thing that the proof wants, to find the principle ideal (r) where $N(r)$ is minimal.

Example 3.6. Consider $\mathbb{Z}$ with norm $N(r) = |r|$. Let $a = 1071$ and $b = 462$. Let g be their gcd, so $(d) = (1071, 462)$. First, we have

$$1071 = 2 \times 462 + 147.$$

This means that $147 \in (1071, 462)$ and $1071 \in (462, 147)$. Therefore $(1071, 462) = (462, 147)$. We can do this again

$$147 = 7 \times 21 + 0$$

so $(462, 147) = (21, 0) = (21)$ and so 21 is the gcd. ♢

3.4 Unique factorisation domains

Unique factorisation domains model the unique prime factorisation that we can perform on the integers.

Definition 3.7 (Irreducibles and primes)**.** Let R be an integral domain. Let $r \in R$ be non-zero and non-unit. Then r is called *irreducible* in R if $r = ab$ implies that either a or b is an unit. Also, a non-zero element p is called a *prime* if (p) is a prime ideal. □

To expand on the definition of primes, if $p \mid ab$ ($ab \in (p)$), then either $p \mid a$ ($a \in (p)$) or $p \mid b$ ($b \in (p)$). The two definitions look similar, since irreducibility (inability to factor further) looks just like the ordinary definition of prime numbers. However they are not. Intuitively irreducibility is about the decomposition of a single element, but being a prime is about the decomposition of other elements (especially its powers). It just so happens that on $\mathbb{Z}$ they behave similarly.

Example 3.7. Consider $\mathbb{Z}[\sqrt{-5}]$. Then 3 is irreducible. Let $3 = \alpha\beta$, then applying the standard norm in $\mathbb{C}$, we have $9 = N(3) = N(\alpha)N(\beta)$. It is easy to see that $N(\alpha), N(\beta) \in \mathbb{Z}$. Therefore, $N(\alpha) = \pm 3$ or $N(\alpha) = \pm 1$ which are all the factors of 9. But a norm of ± 3 is impossible in $\mathbb{Z}[\sqrt{-5}]$ due to

However, 3 is not prime since $3^2 = (2 + \sqrt{-5})(2 - \sqrt{-5})$ yet neither of them belong in (3). ◊

Being a prime is stronger than being irreducible.

Theorem 3.7. *Let R be an integral domain. Then if p is a prime, then p is also irreducible.*

Proof. Let $p = ab$. Since p is a prime, either $p \mid a$ or $p \mid b$. If $p \mid a$, then $pr = a$ for some $r \in R$. So $p = ab = prb$, then since $p \neq 0$ cancellation gives us $rb = 1$. So b is a unit. The same goes for the case where $p \mid b$ instead. ■

The reason why the notions of irreducibility and primality behave in the same way on the integers is revealed in the next theorem.

Theorem 3.8. *Let R be a PID. Then p is a prime iff p is irreducible.*

Proof. We have already shown that a prime is irreducible in a more general setting. Now we show the other direction. Let p be irreducible. We want to show that (p) is a prime ideal. In PIDs prime ideals and maximal ideals are equivalent (see remark theorem 3.2), so it is enough to show that (p) is a maximal ideal. Let $M = (m) \supseteq (p)$ be a (non-trivial) maximal ideal containing (p) (theorem 1.21). Then $p \in (m)$, so $p = mr$ for some $r \in R$. However by assumption $M \neq R$, so m is not an unit. However, since p is irreducible, r must be a unit. So $m = pr^{-1} \in (p)$ which means $(p) = M$ is maximal. ■

This also shows that $\mathbb{Z}[\sqrt{-5}]$ is not a PID.

Definition 3.8 (Unique factorisation domains)**.** An unique factorisation domain (UFD) is an integral domain R such that every non-zero, non-unit element $r \in R$ can be written as a (finite) product of (not necessarily distinct) irreducible elements. Furthermore this factorisation is unique up rearrangement and multiplication by units. □

Uniqueness up to the multiplication of units mean that if we have two factorisations $p_1 \dots p_n = q_1 \dots q_n$ then $q_i = u_i p_i$ for some unit u_i.

Theorem 3.9. *Let R be an UFD. Then a non-zero, non-unit element p is a prime iff it is irreducible.*

Proof. UFDs are integral domains and we have shown that if p is prime then p is irreducible. Let p be irreducible, we want to show that p is prime. Let $ab \in (p)$. Then there is some $r \in R$ such that $pr = ab$. Now consider the unique factorisation of pr and ab into irreducible elements and compare the factors. Keeping in mind that p is irreducible, we have

$$pr_1 \dots r_i = a_1 \dots a_j b_1 \dots a_k.$$

By the uniqueness of the factorisation, then $p = uc$ for some unit u and $c \in \{a_1, \dots, a_k, b_1, \dots a_k\}$. Then $u^{-1}p = c$ and so $p \mid c$. Depending on where c comes from we either have $p \mid a$ or $p \mid b$ and so p is prime. ■

Theorem 3.10. *Let R be an UFD. Let $a, b \in R$ be non-zero elements. Suppose*

$$a = up_1^{a_1} \dots p_n^{a_n} \qquad\qquad b = vp_1^{b_1} \dots p_n^{b_n}$$

for irreducible $p_1, p_2, \dots, p_n$ and all $a_i, b_i \geq 0$. Then $\gcd(a, b) = p_1^{c_1} \dots p_n^{c_n}$ where $c_i = \min(a_i, b_i)$.

Proof. Let $d \mid a$. Then we claim that we can write $d = wp_1^{d_1} \dots p_n^{d_n}$ with $d_i \leq a_i$ and w is an unit. This is clear if we consider the unique factorisation of $dr = a$ (skipped). Then if d also divides b, we can also have $d_i \leq b_i$. So if both $d \mid a$ and $d \mid b$, we obtain $d_i \leq \min(a_i, b_i)$. ■

Theorem 3.11. *Let R be a PID. Then R is also an UFD.*

Proof. Recall that in any PID an element is prime iff it is irreducible iff its ideal is maximal.

Let $r \in R$ be non-zero and non-unit. First we show that r has an irreducible factor. If r is irreducible, then we are done. Otherwise, then we can write $r = ab$ where neither a nor b are units. If a or b are not reducible, then we repeat the process. Without any loss of generality let $a = a_1 b_1$. Then we obtain a chain of ideals $(r) \subseteq (a) \subseteq (a_1) \subseteq \dots$. But R is Noetherian by theorem 3.5, the chain has to stabilise, and this means r has an irreducible factor a_m where $(a_m) = \bigcup_i (a_i) \cup (a) \cup (r)$. Furthermore, this also shows that r has only finitely many irreducible factors.

The factorisation of r is also unique up to multiplication by units and rearrangement. Let us assume that there are two factorisations of r into irreducibles, say $r = p_1 \dots p_n = q_1 \dots q_m$. Then $q_1 \mid p_1 \dots p_n$, so $p_1 \dots p_n \in (q_1)$. Since q_1 is irreducible, (q_1) is a prime ideal, and so there exists some $p_i \in (q_1)$. Then there exists some $u_1 \in R$ such that $p_i = q_1 u_1$. Because p_i is irreducible and q_1 is not an unit, therefore u_1 must be an unit. So $q_1 = p_i u_1^{-1}$. After rearrangement, assume $i = 1$. Then $q_1 \dots q_m = u_1 q_1 p_2 \dots p_m$ and cancellation gives $q_2 \dots q_m = u_1 p_2 \dots p_m$. We may repeat this process to get $q_2 = u_2 q_2$, and so on. This also shows that $n = m$, since otherwise supposing $n \geq m$ we will eventually obtain $1 = u_1 \dots u_m p_{m+1} \dots p_n$ which cannot occur since non of the irreducibles are units. ■

3.5 Gaussian integers

Recall that the Gaussian integers are defined as $\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\} \in \mathbb{C}$, with norm $N(\alpha) = \alpha \cdot \overline{\alpha}$. It is an Euclidean domain, a PID, and an UFD. In this section we will go through some of its applications.

Lemma 3.12. *The units in $\mathbb{Z}[i]$ are ± 1 and $\pm i$.*

Proof. Let $u \in \mathbb{Z}[i]$ be a unit. Then $uu^{-1} = 1$. So (the multiplicative property of the norm is easily checked)

$$N(u)N(u^{-1}) = N(1) = 1,$$

so $N(u) = 1$. The only possible choices for u are hence ± 1 and $\pm i$. ■

From the above lemma it is apparent that $N(u) = 1$ iff u is an unit.

Lemma 3.13.

i. *Let $\alpha \in \mathbb{Z}[i]$ be irreducible. Then $N(\alpha) = p^2$ or $N(\alpha) = p$ for some prime $p \in \mathbb{Z}$.*

ii. *Let $\beta \in \mathbb{Z}[i]$ such that $N(\beta) = p$ for some prime $p \in \mathbb{Z}$, then β is irreducible.*

Proof.

i. Recall that in an Euclidean domain irreducible elements and prime elements are equivalent. Then (α) is a prime ideal. Consider the natural embedding $\varphi : \mathbb{Z} \to \mathbb{Z}[i]$. Then $\varphi^{-1}[(\alpha)] = (\alpha) \cap \mathbb{Z}$ is either a prime ideal or the whole ring (theorem 1.23) $\mathbb{Z}$. If it is the whole ring, then it contains units and so cannot be prime and hence cannot be irreducible. Therefore $(\alpha) \cap \mathbb{Z} = (p)$ for some prime $p \in \mathbb{Z}$ and so $p \in (a)$. Let $p = \alpha \cdot \alpha'$, then $N(p) = N(\alpha)N(\alpha') = p^2$. Since α is not an unit, $N(\alpha) \neq 1$, so we can exhaustively check that the only two factorisations are if $N(\alpha) = p^2$ or $N(\alpha) = p$.

ii. Let $\beta = \beta_1 \cdot \beta_2$. Then $p = N(\beta) = N(\beta_1)N(\beta_2)$. Since p is a prime, we must have either $N(\beta_1) = 1$ or $N(\beta_2) = 2$. So either β_1 or β_2 is a unit, and so β is irreducible. ■

The previous lemma also means that if we can decompose a prime $p \in \mathbb{Z}$ inside $\mathbb{Z}[i]$, then p can be decomposed into at most two irreducible factors. Furthermore, if α is irreducible with $N(\alpha) = p^2$, then $\alpha = up$ for some unit u.

Lemma 3.14. *Let $p \in \mathbb{Z}$ be a prime integer. Then*

i. *If $p \equiv 1 \pmod 4$ then $p \mid n^2 + 1$ for some $n \in \mathbb{Z}$.*

ii. *If $p \equiv 2 \pmod 4$ then $p = (1 + i)(1 - i)$.*

iii. *If $p \equiv 3 \pmod 4$ then p is irreducible in $\mathbb{Z}[i]$.*

Proof.

i. Consider the equation $x^2 + 1 = 0$ in $\mathbb{Z}/p\mathbb{Z}$. The statement is equivalent to saying that a solution exists for this equation. For now, we claim without proof that $(\mathbb{Z}/p\mathbb{Z}, *) \cong \mathbb{Z}/(p - 1)\mathbb{Z}$ is a cyclic group of order $p - 1$. Then, since $4 \mid p - 1$, we can find an element $r \in \mathbb{Z}/p\mathbb{Z}$ of order 4. Then $r^2 = -1$.

ii. This is an obvious fact.

iii. It suffices to show that there is no $\alpha = x + yi \in \mathbb{Z}[i]$ such that $N(\alpha) = p$, since this means $N(p) = p^2$ is irreducible, which in turn means that p is prime (using a similar argument to the proof for lemma above). We have $N(\alpha) = x^2 + y^2$. Consider $\overline{N(\alpha)} = \overline{x}^2 + \overline{y}^2 \in \mathbb{Z}/4\mathbb{Z}$. We can exhaustively check all cases that we cannot have $\overline{x}^2 + \overline{y}^2 = \overline{3}$.

■

Corollary 3.14.1. *Let $p \in \mathbb{Z}$ be a prime such that $p \equiv 1 \pmod 4$. Then p is reducible in $\mathbb{Z}[i]$.*

Proof. We know that $p = \alpha\alpha'$ for unique (because UFD) irreducible α and α'. Assume p is irreducible. We know from the previous lemma that $p \mid n^2 + 1$ for some $n \in \mathbb{Z}$. Prime elements are equivalent to irreducible elements, so $p \mid n + i$ or $p \mid n - i$. However if p divides a complex number then it must divide its complex conjugate, so in fact both $p \mid n + i$ and $p \mid n - i$. Then p divides the difference, $p \mid 2i$, which is a contradiction. ■

Furthermore, for the unique factorisation of $p = \alpha \cdot \alpha'$ into irreducibles, this means that $N(\alpha) = p$, that is there are a, b such that $a^2 + b^2 = p$.

Theorem 3.15. *The irreducible elements in $\mathbb{Z}[i]$ are of the form (up to units)*

- *$a \pm bi$ where $a^2 + b^2 = p$ for some prime $p \in \mathbb{Z}$,*
- *$1 \pm i$ (with a norm of 2),*
- *a prime $p \in \mathbb{Z}$ with $p \equiv 3 \pmod 4$ (with a norm of p^2).*

Proof. This is just a collection of previously established results, in the sequence of $1 \pmod 4$, $2 \pmod 4$, and $3 \pmod 4$, and these are the only possibilities. ■

Theorem 3.16 (Fermat's little theorem)**.** *Let $p \in \mathbb{Z}$ be a positive prime. Then $p = a^2 + b^2$ for some $a, b \in \mathbb{Z}$ iff $p \equiv 2, 1 \pmod 4$. Furthermore the expression is unique up to -1 and permutation.*

4 Polynomial rings

Let R be commutative with $1 \neq 0$. Recall that the polynomial ring $R[x] \ni f(x) = a_n x^n + \cdots + a_0$. We can also define $R[x, y] = (R[x])[y] = (R[y])[x]$ and we can define $R[x_1, \ldots, x_n]$ similarly.

Theorem 4.1. *Let R, S be commutative rings with $1 \neq 0$. Let $\varphi : R \to S$ be a ring homomorphism. Then for any $s_1, s_2, \ldots, s_n \in S$, there exists an unique ring homomorphism $\psi : R[x_1, \ldots, x_n] \to S$ such that $\psi(r) = \varphi(r)$ and $\psi(x_i) = s_i$.*

Proof. The uniqueness is clear from the definition. For existence, consider the following composition. First we have $\alpha : R[x_1, \ldots, x_n] \to S[x_1, \ldots, x_n]$ by mapping coefficients $a_i \mapsto \varphi(a_i)$. Next, $\beta : S[x_1, \ldots, x_n] \to S$ is the evaluation homomorphism performed n times, with $x_n \mapsto s_n$. ■

An example of what the above means is to consider $R[x] \to S$ given by $\sum_{i=0}^{m} a_i x^i \mapsto \sum_{i=0}^{m} \varphi(a_i) s_1^i$. A multi-variable example would be much messier as we can well imagine.

Example 4.1. We do not need S to be commutative to make it work, we only need the subring $\operatorname{img}\varphi$ to be commutative. Consider $\mathbb{R}[x]$ and $S = \mathrm{Mat}_{n,m}(\mathbb{R})$. We consider $\mathbb{R} \to S$ given by $r \mapsto r \cdot \mathbf{I}$. Then if we pick any $A \in S$, we have $\mathbb{R}[x] \to S$ given by the theorem above. This is reminiscent of our study of minimal and characteristic polynomials in linear algebra. ♢

Theorem 4.2. *Let R be commutative with $1 \neq 0$. Let $I \subseteq R$ be an ideal. Then $(R/I)[x] \cong R[x]/(I)$.*

Proof. We consider the quotient map $R[x] \to (R/I)[x]$ given by $\sum_{i=0}^{n} a_i x^i \mapsto \sum_{i=0}^{n} \overline{a}_i x^i$. It is clear that the kernel is $I[x] = (I)$. Then apply the isomorphism theorems. ■

Corollary 4.2.1. *If I is prime in R, then $(I) \subseteq R[x]$ is also prime.*

Proof. First lemma 1.22 tells us that R/I is an integral domain. Theorem 1.6 then states that $(R/I)[x] \cong R[x]/(I)$ is an integral domain. Again by lemma 1.22 this tells us that (I) is prime. ■

4.1 Polynomial rings over fields

Let F be a field. Then $F[x]$ is an Euclidean domain with the norm $N(f) = \deg f$. For example, think of $\mathbb{Q}[x]$ and polynomial division that we are all familiar with.

What follows is a summary of easy properties we can deduce from what we already know.

Theorem 4.3. *Let F be a field and $f(x) \in F[x]$.*

i. $F[x]$ is an Euclidean domain, a PID and a UFD.

ii. Let $f(x) \in F[x]$, then $f(a) = 0$ iff $(x - a) \mid f(x)$.

iii. Let $f(x) \in F[x]$ be of degree n. Then $f(x)$ has at most n (not necessarily distinct) roots in F.

iv. $F[x]/(f(x))$ is a field iff $f(x)$ is irreducible or prime or maximal.

v. Let $p(x)$ and $q(x)$ be irreducible in $F[x]$. If $(p(x)) \neq (q(x))$ or $p(x) \neq uq(x)$ for some unit $u \in F[x]$, then $(p^a(x))$ and $(q^b(x))$ are coprime, i.e. $(p^a(x)) + (q^b(x)) = F[x]$ (by definition).

vi. Let $f(x) = p_1^{a_1}(x) p_2^{a_2}(x) \ldots p_n^{a_n}(x)$ be its irreducible factorisation. Then $F[x]/(f(x)) \cong F[x]/(p_1^{a_1}(x)) \times \cdots \times F[x]/(p_n^{a_n}(x))$.

Proof.

i. We have briefly described how it is an Euclidean domain.

ii. By the Euclidean algorithm, we can consider $f(x) = q(x)(x - a) + r$ for $r \in F$ (r is a constant as it needs to have degree lower than $x - a$). Then $f(a) = q(a)(a - a) + r$ so $0 = f(a) = r$.

iii. If $f(x)$ has no roots, we are done. Otherwise, if a is a root, then $f(x) = q(x)(x-a)$ where $\deg q = a - 1$. Then we complete the proof using induction.

iv. We know in a PID a prime ideal is a maximal ideal. If the left hand side is a field it means $(f(x))$ is maximal. A maximal ideal is always prime in an integral domain (see remark after lemma 1.22).

v. We can write $p^a(x) = p^a(x)q^0(x)$ and $q^b(x) = p^0(x)q^b(x)$. In an UFD we can find their GCD by theorem 3.10, and find that their gcd is an unit. Then there is an linear combination of $p^a(x)$ and $q^b(x)$ that equals an unit (see remark theorem 3.4). The rest follow.s

vi. Follows from the Chinese remainder theorem.

■

Theorem 4.4. *Let F be a field and $G \subseteq F^*$ be a finite subgroup. Then G is cyclic. In particular, $(\mathbb{Z}/p\mathbb{Z})^*$ is cyclic for prime p.*

Proof. We know that G is abelian and finite. So, by the classification of finite abelian groups, we can write

$$G \cong Z/d_1\mathbb{Z} \times \cdots \times \mathbb{Z}/d_n\mathbb{Z}$$

with $d_n \mid d_{n-1} \mid \cdots \mid d_1$. Consider the polynomial $x^{d_1} - 1 \in F[x]$. It has at most d_1 roots in F (and G). We know by the property of cyclic groups that any element $a \in \mathbb{Z}/d_1\mathbb{Z}$ satisfies $a^{d_1} = 1$ and so is a root. However that also means that any elements from the other groups are also roots. Thus $n = 1$. ■

There are subtle differences between polynomials and functions. For example in the field $\mathbb{Z}/2\mathbb{Z}$, the function $f(x) = x + 1$ is equal to the function $g(x) = x^2 + 1$, yet they are totally different polynomials. The following theorem shows why this is the case: usually when we think of functions over $\mathbb{R}$ or $\mathbb{C}$ etc., the field is infinite, which allows the two notions to be interchanged.

Theorem 4.5. *Let F be a field. Let P be the ring of polynomial functions on F. We have a ring homomorphism $\varphi : F[x] \to P$ given by $f(x) \mapsto f(x)$. φ is an isomorphism iff F is infinite.*

Proof. φ is surjective by definition of P. So φ is an isomorphism iff $\ker \varphi = 0$. We see that $\ker \varphi = \{f \in F[x] \mid \forall a \in F\ [f(a) = 0]\}$. If $F = \{a_0, a_1, \ldots, a_n\}$ is finite, then $(x - a_1)\ldots(x - a_n) \in \ker \varphi$. If F is infinite, then any $f(x) \in F[x]$ has only finitely many roots, so $\ker \varphi = \{0\}$. ■

Theorem 4.6. *Let R be commutative with $1 \neq 0$. Then R is an UFD iff $R[x]$ is an UFD (iff $R[x_1, x_2]$ is an UFD etc.)*

Proof. Section 9.3 textbook. ■

Corollary 4.6.1. *$\mathbb{Z}[x]$ is an UFD but not a PID.*

4.2 Irreducible polynomials

Let F be a field. Recall $(F[x])^* = F^* = F - \{0\}$. This arises simply by an argument using the degrees of polynomials.

Theorem 4.7. *Let $f(x) \in F[x]$ be of degree 2 or 3. Then $f(x)$ is irreducible if $f(x)$ has no root in F.*

Proof. If $\deg f(x)$ is 2 or 3, and $f(x)$ is reducible, then $f(x) = a(x)b(x)$ for $\deg a(x), \deg b(x) < \deg f(x)$. The degree must decrease, because a factor with a degree of 0 means that it is in the field, and is thus a unit. So we conclude that if $f(x)$ is reducible then it has a root. ■

This also means that $F[x]/(f(x))$ is a field if $f(x)$ is of degree 2 or 3.

Lemma 4.8. *Let R be an integral domain with a prime ideal $I \subseteq R$. Let $p(x) \in R[x]$ be monic and non-constant. Then if $\overline{p(x)}$ is irreducible in $R/I[x] \cong R[x]/I$, then $p(x)$ is irreducible in $R[x]$.*

Proof. Assume the contrary that $p(x) = a(x)b(x)$ with $\deg a(x), \deg b(x) \geq 1$. The bounds on the degree appear from the fact that $p(x)$ is monic. Write $a(x) = a_k x^k + \cdots$ and $b(x) = b_l x^l + \cdots$. Then $a_k b_l = 1$. So $\overline{a_k}, \overline{b_l} \in R/I$ are not 0. Therefore $\deg \overline{a(x)}, \deg \overline{b(x)} \geq 1$. Recall that $(R/I[x])^* = (R/I)^*$. So $\overline{p(x)} = \overline{a(x)b(x)}$ is a factorisation which is a contradiction. ■

The requirement of being monic is to avoid any interference from R. For example the polynomial $2x + 2 = 2(x + 1)$ is reducible due to the leading coefficient.

Example 4.2. $x^2 + x + 1$ is irreducible in $\mathbb{Z}[x]$ since it is irreducible in $\mathbb{Z}/2\mathbb{Z}[x]$ since we can check that it has no roots in the field $\mathbb{Z}/2\mathbb{Z}$. ♢

Example 4.3. We want to show that $x^2 + xy + 1$ is irreducible in $\mathbb{Z}[x, y]$. If $x^2 + 1$ is irreducible in $\mathbb{Z}[x]$, then $x^2 = xy + 1$ is irreducible in $\mathbb{Z}[x, y]$. $x^2 + 1$ is indeed irreducible since it has no roots in $\mathbb{Z}/3\mathbb{Z}[x]$. ♢

Theorem 4.9 (Eisenstein's criterion)**.** *Let R be an integral domain with a prime ideal I. Let $f(x) = x^n + c_{n-1}x^{n-1} + \cdots + c_0 \in R[x]$ be monic and non-constant. Suppose $c_{n-1}, \ldots, c_0 \in I$ but $c_0 \notin I^2$, then $f(x)$ is irreducible in $R[x]$.*

Equivalently, if $p \in \mathbb{Z}$ be a prime number and $p \mid c_i$ for $i = 0, \ldots, n-1$, but $p^2 \nmid a_0$, then $f(x)$ is irreducible in $R[x]$.

Proof. Assume the contrary that $p(x) = a(x)b(x)$ with $\deg a(x), \deg b(x) \geq 1$. Let $a(x) = a_r x^r + a_{r-1}x^{r-1} + \cdots + a_0$ and $b(x) = b_s x^s + b_{s-1}x^{s-1} + \cdots + b_0$. We then consider the image of $f(x)$ in $R/I[x]$. We have $\overline{x}^n = \overline{f(x)} = \overline{a(x)b(x)} = \overline{x}^{r+s} + \cdots + (\overline{b}_0\overline{a}_2 + \overline{b}_1\overline{a}_1 + \overline{b}_2\overline{a}_0)\overline{x}^2 + (\overline{b}_0\overline{a}_1 + \overline{b}_1\overline{a}_0)x + \overline{a}_0\overline{b}_0$. This means that $\overline{a}_0\overline{b}_0 = \overline{0}$, and $\overline{b}_0\overline{a}_1 + \overline{b}_1\overline{a}_0 = \overline{0}$, and so on.

$\overline{a}_0\overline{b}_0 = \overline{0}$ means that either $\overline{a}_0 = \overline{0}$ (and $a_0 \in I$) or $\overline{b}_0 = \overline{0}$ (and $b_0 \in I$).

If $\overline{a}_0 = \overline{0}$, then $\overline{b}_0\overline{a}_1 + \overline{b}_1\overline{a}_0 = \overline{b}_0\overline{a}_1 = \overline{0}$. If $b_0 = \overline{0}$ then we arrive at a contradiction with $c_0 = a_0 b_0 \notin I^2$.

Therefore the next case is when $\overline{a}_1 = 0$. Then $\overline{b}_0\overline{a}_2 + \overline{b}_1\overline{a}_1 + \overline{b}_2\overline{a}_0 = \overline{b}_0\overline{a}_2 = \overline{0}$. Again we cannot have $\overline{b}_0 = 0$. We continue like this, and at the end (without generality we can assume $r \leq s$) we obtain

$$\overline{a}_r\overline{b}_0 + \sum_{\substack{i+j=r \\ j\neq 0}} \overline{a}_i\overline{b}_j = \overline{b}_0 = \overline{0}$$

So either way we arrive at the same contradiction. ■

Example 4.4. Let $f(x) = x^2 + 10x + 5 \in \mathbb{Z}[x]$, and $I = (5)$. Then $I^2 = (25)$, so $f(x)$ is irreducible. ♢

Example 4.5. Consider $R[x][y]$ for an integral domain R. Then $y^n - x$ is irreducible, since $x \in (x) \subseteq R[x]$ but $x \notin (x^2)$. ♢

5 Modules

5.1 Definition and examples

Modules are a tool to study rings. An analogy is using simple groups to study groups.

Definition 5.1 (Modules)**.** Let R be a ring. A left R-*module* is an abelian group M together with an *action map* $\cdot : R \times M \to M$ given by $(r, m) \mapsto rm$ such that for any $m, n \in M$, $r, s \in R$, we have

- $(r + s) \cdot m = rm + sm$
- $(rs) \cdot m = r(sm)$
- $r \cdot (m + n) = rm + rn$
- $1 \cdot m = m$ if $1 \in R$. □

Similarly, a right action map $\cdot : M \times R \to M$ given by $(m, r) \mapsto mr$ has to obey the same properties with appropriate adjustments. We will drop the dot most of the time.

We will be focusing more on left modules. However note that if R is commutative then left modules are the same as right modules.

To say that M admits an R-module structure is the same as finding the action map, or to say that there exists a ring homomorphism $R \to \mathrm{Hom}_{\mathrm{ab}}(M, M) = \mathrm{End}_{\mathrm{ab}}(M)$[6] such that $1_R \mapsto \mathrm{id}$.

Example 5.1. Let R be a ring.

- We have the trivial modules. We have the zero module for any ring R, with $M = \{0\}$. We can define $(r, 0) \mapsto 0$. Then R is a left R-module via left multiplication. We define the action $r \cdot m = rm$. R is also a right R-module via the right multiplication $m \cdot r = mr$.
- Let $I \subseteq R$ be a left ideal. Then I is a left R-module via the left multiplication $r \cdot m = rm \in I$. ♢

[6] The subscript *ab* denote "abelian".

Example 5.2. Let M be an abelian group.

- Then M is a $\mathbb{Z}$-module via the action $n \cdot m = m + m + \cdots + m$, addition done n times for $n \in \mathbb{Z}^+$. For negative n we take the additive inverse. For $n = 0$ it is just 0.
- $\text{End}_{\text{ab}}(M)$ is a ring with 1, and M is an $\text{End}_{\text{ab}}(M)$-module, where $f \cdot m = f(m)$ for $f \in \text{End}_{\text{ab}}(M)$. $\Diamond$

Example 5.3. Let $R = \mathbb{Z}$ and $M = \mathbb{Z}$. We know $r \cdot m = rm$ is an action of R on M. But if we define $r \cdot m = r + m$ then this is not an action of R on M. $\Diamond$

Example 5.4. Let F be a field. Then F-modules are vector spaces over F. $\Diamond$

Definition 5.2 (Submodules)**.** Let R be a ring and let M be a R-module. Let $N \subseteq M$ be an abelian subgroup. Then N is called a R-submodule of M if N is closed under the R-action. $\square$

Example 5.5. Let R be a ring.

- We have the trivial submodules. The zero module is a submodule of any R-module. If M is a R-module then M is a also submodule of M.
- If we consider the action of R as a R-module then a left ideal $I \subseteq R$ will be a R-submodule of R.
- Let $S \subset R$ be a subring containing 1. Then R is a S-module via the left multiplication. Then S is a S-submodule of R. $\Diamond$

Example 5.6. Let M be a $\mathbb{Z}$-module. This means that M is an abelian group. Then $N \subseteq M$ is a $\mathbb{Z}$-submodule if $N \subseteq M$ is a subgroup. $\Diamond$

Example 5.7. Let F be a field.

- Then a if M is a module, a F-submodule of M is a F-subspace of M.
- Consider $F[x] \supset F$. Let M be a $F[x]$-module. Then we can consider the action of only elements in $F \subset F[x]$. Then M is also a module over F (via restriction). So M is a vector space over F with an action of $x : M \to M$ given by $m \mapsto xm$. It is easy to check that x is a linear transformation. Similarly any $f(x) : M \to M$ given by $m \mapsto f(x)m$ is a linear transformation by the same argument. Thus, given a F-vector space M together with a linear transformation $T : M \to M$, we can define a $F[x]$-module structure on M via the map $F[x] \times M \to M$ given by $(f(x), m) \mapsto f(T)m = (a_nT^n + \ldots a_0)m$.

$\Diamond$

Example 5.8. Let $G = S_2$ the symmetric group over two letters. Let $R = \mathbb{C}[G] = \{ae + b(12) \mid a, b \in \mathbb{C}\}$ (here e is the identity element). Any R-module V is naturally a $\mathbb{C}$-vector space (which is also a $\mathbb{C}$-module) through the embedding $a \mapsto ae$.

Consider $V = \mathbb{C}e_1 \oplus \mathbb{C}e_2 \cong \mathbb{C}^2$. We define a R-action on V via $(ae + b(12))(se_1 + te_2) = a(se_1 + te_2) + b(se_2 + te_1)$. We can check that this defines a R-module on V, which we skip. This is also a linear transformation from V to V.

Let $W = \mathbb{C}(e_1 + e_2) \subseteq V$ be a one-dimensional subspace of V. Then W is a R-submodule of V. Similarly for $U = \mathbb{C}(e_1 - e_2)$. We actually have $V = W \oplus U$. We will soon visit this again. $\Diamond$

Theorem 5.1. *Let $\varphi : S \to R$ be a ring homomorphism (with $1_S \mapsto 1_R$ if identities exist). Let M be a R-module. Then M is a S-module via the action $s \cdot m = \varphi(s) \cdot m$ for any $s \in S$ and $m \in M$.*

Proof. Let $s, t \in S$. We check the requirements one by one. For the first trhee, we demonstrate it once and leave out the rest. $(s+t) \cdot m = \varphi(s+t) \cdot m = (\varphi(s)+\varphi(s)) \cdot m = \varphi(s) \cdot m + \varphi(t) \cdot m = s \cdot m + t \cdot m$. The last requirement is done as such: $1_S \cdot m = \varphi(1_S) \cdot m = 1_R \cdot m = m$. This also explains the need for $1_S \mapsto 1_R$ (which is not generally true) if the identity exist.

Another way to show this is to consider $\psi : R \to \mathrm{End}_{\mathrm{ab}}(M)$ defined by $r \mapsto \varphi_r$ such that $\varphi_r(m) = rm$. Then, composing with φ we have $S \xrightarrow{\varphi} R \xrightarrow{\psi} \mathrm{End}_{\mathrm{ab}}(M)$. ■

Example 5.9. Let $S \subseteq R$ be a subring such that $1 \in S$. Then any R-module M is also a S-module. ◊

Example 5.10. Let R be a ring with 1. We have a ring homomorphism $\mathbb{Z} \to R$ given by $n \mapsto 1 + \cdots + 1$ performed n times. Then any R-module M is also a $\mathbb{Z}$-module. ◊

Example 5.11. Let F be a field and M be a $F[x]$-module. Then M is also a F-module, or a vector space over F. ◊

5.2 Properties of modules

Theorem 5.2. *Let R with $1 \neq 0$. Let M be a R-module.*

i. *$0_R \cdot m = 0_M$, and $(-1) \cdot m = -m$ for any $m \in M$.*

ii. *Any intersection of submodules of M is still a submodule.*

iii. *Let $z \in R$ be central, i.e. $az = za$ for all $a \in R$. Then $zM = \{zm \mid m \in M\}$ is a submodule of M.*

Proof. We have $R \to \mathrm{End}_{\mathrm{ab}}(M)$.

i. We have $0_R \cdot m = (0_R + 0_R) \cdot m = 0_R \cdot m + 0_R \cdot m$ and by cancellation in abelian groups $0_R \cdot m = 0_M$. Next, we check $0 = (1+(-1)) \cdot m = 1 \cdot m + (-1) \cdot m$ similarly we see $(-1) \cdot m$ is an additive inverse of m.

ii. Let I be an indexing set, and let M_c with $c \in I$ be a R-submodule of M. Then $\bigcap_{c \in I} M_c = N$ is an abelian subgroup of M. It remains to be check that N is closed under the action, i.e. $\forall n \in N, \forall r \in R\ [rn \in N]$. Since $n \in M_c$ for any c and since M_c is a submodule, therefore $rn \in M_c$.

iii. We first check that zM is an abelian subgroup. Take any $z \in Z$ and $m, n \in M$, we have (skipping steps) $zm - zn = \cdots = z(m-n) \in zM$. Next we have to check that zM is closed under the action. Take any $r \in R$. Then $r(zm) = (rz)m = (zr)m = z(rm) \in zM$.

■

Theorem 5.3. *Let R with $1 \neq 0$ be a ring and M be a R-module. Let $I \subseteq \mathrm{Ann}_M(R)$ be an ideal. Then M is a R/I-module via the action $(r + I) \cdot m = r \cdot m$ for $r \in R$ and $m \in M$.*

Proof. We have $\varphi : R \to \mathrm{End}_{\mathrm{ab}}(M)$. Then $\mathrm{Ann}_M(R) = \ker\varphi$. Then if $I \subseteq \ker\varphi$, we have

$$\begin{tikzcd}
R \arrow[r, "\varphi"] \arrow[d, "\pi"'] & \mathrm{End}_{\mathrm{ab}}(M) \\
R/I \arrow[ur, dashed, "\overline{\varphi}"'] &
\end{tikzcd}$$

Another way to see this is to do the manual checking. First we check that the action is well-defined. Let $r + I = r' + I$, or in other words $r = r' + a$ for some $a \in I$. Then $(r + I) \cdot m = r \cdot m = (r' + a) \cdot m = r' \cdot m + a \cdot m = r' \cdot m = (r' + I) \cdot m$. We then check the requirements for modules. We will check only one of them, the others are similar or easy. We have $(r + I + s + I) \cdot m = (r + s) \cdot m = r \cdot m + s \cdot m = (r + I) \cdot m + (s + I) \cdot m$. ■

Definition 5.3 (Generated submodules)**.** Let R be a ring with $1 \neq 0$ and M be a R-module. Let $N \subseteq M$ be a subset. Define $RN = \{\sum_{\text{finite}} an \mid a \in R, n \in N\} \subseteq M$ is a submodule of M. Elements of RN are called *R-linear combinations*, and RN is called the R-submodule of M generated by N. M is called *finitely generated* if $M = RA$ for a finite set A, and M is called *cyclic* if $M = Rm$ for some element $m \in M$. □

It is quite easy to see that RN is a submodule of M. First of all it is an abelian group, and second, it is closed under the R-action, since we just multiply the "coefficients". To see why RN is generated by N, note that RN takes the form of a minimal abelian group containing N.

Theorem 5.4. *Let R be a ring, M be an R-module, and $M_1, M_2, \dots, M_n$ be submodules of M. Define the sum of modules*

$$\sum_{i=1}^{n} M_i = \left\{ \sum_{i=1}^{n} m_i \mid m_i \in M_i \right\}.$$

This is a R-submodule of M.

Proof. The sum remains as an abelian group. It is closed under the action by distributing across the summation. ■

In fact the sum can be completely arbitrary, as long as the elements are finite sums.

Theorem 5.5. *Let R be a ring, M be a R-module, I be an indexing set, and M_c be R-submodules of M for $c \in I$. Then*

$$\sum_{c \in I} M_c = \left\{ \sum_{c \in I} m_c \mid m_c \in M_c \text{ with finitely many non-zeros} \right\}$$

is a R-submodule of M.

Definition 5.4 (Module homomorphisms)**.** Let R be a ring with $1 \neq 0$. Let M and N be R-modules. A *R-module homomorphism* $\varphi : M \to N$ is a map such that

- $\forall m, m' \in M\ [\varphi(m + m') = \varphi(m) + \varphi(m')]$
- $\forall m \in M, r \in R\ [\varphi(r \cdot m) = r \cdot \varphi(m)]$. □

We define the set of R-module maps from M to N as $\mathrm{Hom}_R(M, N)$. We write $\mathrm{Hom}_R(M, M) = \mathrm{End}_R(M)$.

It should be noted that the R-actions in the definition above are performed on different modules.

Definition 5.5 (Module isomorphisms)**.** A module homomorphism is an *isomorphism* if it is a bijection. Equivalently, we can find its inverse (we have to show that the inverse is also an R-module map, which we skip). □

Definition 5.6 (Kernels and images)**.** Let $\varphi : M \to N$ be a R-module homomorphism. Define the kernel as $\ker\varphi = \{m \in M \mid \varphi(m) = 0\}$ and image as $\operatorname{img}\varphi = \varphi[M] = \{\varphi(m) \mid m \in M\}$. □

Example 5.12. Let F be a field. Let M and N be F-modules. Any F-module homomorphism is a linear transformation. ◊

Example 5.13. A $\mathbb{Z}$-module homomorphism is an abelian group homomorphism. ◊

Example 5.14. Let R act on R via left multiplication.

- Then $\varphi_a : R \to R$ defined by $r \mapsto ra$ for $a \in R$ is an R-module homomorphism.
- Let R act on R via left multiplication. Then $z \in R$ be central, then $\varphi_z : R \to R$ defined by $r \mapsto zr = rz$ is a R-module homomorphism. ◊

Theorem 5.6 (Properties of module homomorphisms)**.** *Let M, N, L be R-modules.*

i. *Let $\varphi : M \to N$ be a R-module homomorphism. Then* $\ker\varphi \subseteq M$ *and* $\operatorname{img}\varphi \subseteq N$ *are R-submodules.*

ii. *Let $\varphi, \psi \in \operatorname{Hom}_R(M, N)$. Then we define $\varphi + \psi : M \to N$ as $(\varphi + \psi)(m) = \varphi(m) + \psi(m)$. Then $\varphi + \psi \in \operatorname{Hom}_R(M, N)$.*

iii. *Let $\varphi \in \operatorname{Hom}_R(M, N)$ and $\psi \in \operatorname{Hom}_R(N, L)$, then $\varphi \circ \psi \in \operatorname{Hom}_R(M, L)$.*

iv. $\operatorname{Hom}_R(M, M) = \operatorname{End}_R(M)$ *is a ring with identity and M is naturally a* $\operatorname{End}_R(M)$*-module.*

Proof.

i. $\ker\varphi$ is an abelian subgroup of M, if we think of φ as just an abelian group homomorphism. We show that $\ker\varphi$ is closed under the R-action. Let $m \in \ker\varphi$ and $r \in R$. Then $\varphi(rm) = r\varphi(m) = 0$ so $rm \in \ker\varphi$. The same goes for $\operatorname{img}\varphi$. Let $n \in \operatorname{img}\varphi$ and $r \in R$. Then there exists some $m \in M$ such that $\varphi(m) = n$. Then $rn = r\varphi(m) = \varphi(rm)$ so $rn \in \operatorname{img}\varphi$.

ii. The sum of an abelian group homomorphism is still an abelian group homomorphism. We only need to check multiplication, $(\varphi + \psi)(rm) = \varphi(rm) + \psi(rm) = r\varphi(m) + r\psi(m) = r(\varphi(m) + \psi(m)) = r(\varphi + \psi)(m)$.

iii. Again addition comes free from abelian group homomorphisms. For multiplication, $(\varphi \circ \psi)(rm) = \varphi(r\psi(m)) = r\varphi(\psi(m))$.

iv. $(\operatorname{End}_R(M), +, \circ)$ is a ring. Most properties are free, for example function composition is associative so we will skip the proof. Also, the action of $\operatorname{End}_R(M)$ on M is given by $f \cdot m = f(m) \in M$. ■

The motivation behind the last point can be seen in linear algebra, where instead of studying vector spaces directly, another point of view is to study endomorphisms on vector spaces instead, which are matrices.

5.3 Quotient modules

Definition 5.7 (Quotient modules)**.** Let M be a R-module with a submodule $N \subseteq M$. We define the quotient R-module M/N as

- M/N as the quotient abelian group, and
- $r \cdot (m + N) = r \cdot m + N$ for any $r \in R$. □

Theorem 5.7. *The above definition is well-defined, and the quotient map $\pi : M \to M/N$ is a R-module homomorphism.*

Proof. Let $m + N = m' + N$. This is the same as saying $m - m' = n \in N$. Then $r \cdot (m + N) = rm + N = rm' + rn + N = rm' + N = r(m' + N)$. To show that $\cdot$ defines a R-module structure on M/N, we can simply perform the computation on representatives, as we have already shown that on the representatives it is well-defined.

π is already an abelian group homomorphism. From the definition it is clear that $r\pi(m) = \overline{rm} = \pi(rm)$. ■

Example 5.15. Let F be a field. Then M/N is the quotient vector space. ◊

Example 5.16. Let $I \subseteq R$ be a left ideal. Then I is also a R-submodule of R. Then $\pi : R \to R/I$ is a R-module homomorphism. ◊

Theorem 5.8 (Universal property of quotient modules)**.** *Let R be a ring with $1 \neq 0$. Let M be a R-module with a submodule $N \subseteq M$. Then for any R-module L and a R-module map $\varphi : M \to L$ such that $N \subseteq \ker \varphi$, we have an unique R-module map $\overline{\varphi} : M/N \to L$ such that the following diagram commutes:*

$$\begin{tikzcd}
M \arrow[r, "\varphi"] \arrow[d, "\pi"'] & L \\
M/N \arrow[ur, dashed, "\overline{\varphi}"'] &
\end{tikzcd}$$

Proof. The uniqueness is forced from the definition. We know this to be true for abelian groups. We only need to check the R-action. This is essentially free since $\varphi(m) = \overline{\varphi}(m + N)$, and we skip it. ■

Theorem 5.9 (Isomorphism theorems for modules)**.** *Let R be a ring with $1 \neq 0$ and L a R-module.*

i. Let M, N be R-modules and let $\varphi : M \to N$ be a module homomorphism. Then $M/\ker \varphi \cong \varphi[M]$.

ii. Let M, N be submodules of L. Then $(M + N)/M \cong N/(N \cap M)$.

iii. Let M, N be submodules of L. Then $L/M \cong (L/N)/(M/N)$ as R-modules.

iv. Let M, N be R-modules such that $N \subseteq M$. Then we have a bijection via the quotient map from the set of submodules of M containing N to the set of submodules of M/N.

Proof. We already know these statements for abelian groups, all that we need to do is to check that the module structure is compatible. ■

Theorem 5.10. *Let M be a R-module. Let $M = Rm$ for some $m \in M$. Then $R/I \cong M$ for some left ideal $I \subseteq R$.*

Proof. Define $\varphi : R \to M$ given by $r \mapsto rm$. We claim that φ is a R-module homomorphism. Firstly, $\varphi(s + r) = (s + r)m = sm + rm = \varphi(s) + \varphi(r)$ by the properties of the R-action. Similarly, $\varphi(rs) = rsm = r\varphi(s)$. So by the isomorphism theorem, $R/\ker\varphi \cong M$ where $\ker\varphi$ is a R-submodule or left ideal of R. ■

In spirit, this means that for cyclic modules we can study them as quotients of the ring with some left ideal.

5.4 Simple modules

Definition 5.8 (Simple modules)**.** Let R be a ring with $1 \neq 0$. A non-zero R-module M is called *simple* or *irreducible* if the zero module and M are the only submodules of M. □

Example 5.17. Let F be a field. Then one-dimensional vector spaces over F are simple, since the subspace's dimension has to be either zero or one. ♢

Example 5.18. Simple $\mathbb{Z}$-modules are of the form $\mathbb{Z}/p\mathbb{Z}$ for a prime number p. ♢

Example 5.19. Consider $R = \mathrm{Mat}_n(F)$ for a field F. Let $V = F^n$ be a R-module with action given by matrix multiplication. Then V is simple. Let non-zero $W \subseteq N$ be a R-submodule. Then we can find a vector $0 \neq w \in W$, and for any $v \in V$ we can find a matrix $A \in R$ such that $Aw = v$. Thus $v \in W$. ♢

Theorem 5.11 (Schur's lemma)**.** *Let R be a ring with $1 \neq 0$. Let M, N be simple R-modules. Then $\varphi : M \to N$ is either an isomorphism or the zero map. In particular, $\mathrm{End}_R(M)$ is a division ring.*

Proof. $\ker\varphi$ is either zero or M. Similarly with $\mathrm{img}\,\varphi$. Then we just have to consider all possibilities. $\ker\varphi = \{0\}$ and $\mathrm{img}\,\varphi = N$ gives us an isomorphism. $\ker\varphi = M$ and $\mathrm{img}\,\varphi = \{0\}$ gives us the zero map. $\ker\varphi = \{0\}$ and $\mathrm{img}\,\varphi = \{0\}$ cannot occur since we exclude the zero module from being simple modules. ■

Example 5.20.

- $\mathrm{End}_F(F) \cong F$.
- $\mathrm{End}_{\mathbb{Z}}(\mathbb{Z}/p\mathbb{Z}) \cong \mathbb{Z}/p\mathbb{Z}$.
- $\mathrm{End}_{\mathrm{Mat}_n(F)}(F^n) \cong F$. ♢

5.5 Ring Algebras

Definition 5.9 (R-algebras)**.** Let R be commutative ring with $1 \neq 0$. A *R-algebra* A is a ring with identity together with a ring homomorphism $f : R \to A$, such that $1_R \mapsto 1_A$ and $f[R]$ is in the centre of A. □

Example 5.21. Let R be a commutative ring with $1 \neq 0$.

- The polynomial ring $R[x]$ is a R-algebra.
- The group ring $R[G]$ is a R-algebra for a finite group G.
- Let M be a R-module. Then $\mathrm{End}_R(M)$ is a R-algebra with $f : R \to \mathrm{End}_R(M)$ defined as $r \mapsto (f_r : M \to M, m \mapsto rm)$. f_r is a R-module homomorphism.
- A R-algebra A is a ring with identity that is also a R-module, where the R-action is compatible with the ring operations, for example $ra = f(r)a$

$\Diamond$

Example 5.22. Let F be a field.

- Then $M \cong F^n$ is a F-vector space. Then $\mathrm{End}_F(F^n) \cong \mathrm{Mat}_n(F)$ which is a F-algebra.
- Let A be a F-algebra. Then any A-module is automatically a F-vector space.

$\Diamond$

Example 5.23. A ring is just a $\mathbb{Z}$-algebra. $\Diamond$

5.6 Direct sums of modules

Definition 5.10 (Direct sum of modules)**.** Let R be a ring with $1 \neq 0$. Let $M_1, M_2, \ldots, M_n$ be R-modules. We define their direct sum $M = M_1 \oplus M_2 \oplus \cdots \oplus M_n$ as

- $M = M_1 \times M_2 \times \cdots \times M_n$ as abelian groups (i.e. the product group), and
- the R-action on M is defined as $r(m_1, m_2, \ldots, m_n) = (rm_1, rm_2, \ldots, rm_n)$.

□

Theorem 5.12. *Let $M_1, M_2, \ldots, M_n$ be R-modules.*

i. We have R-module homomorphisms $f_i : M_i \to M$ given by $m_i \mapsto (0, \ldots, m_i, \ldots, 0)$, and $g_i : M \to M_i$ given by $(m_1, \ldots, m_i, \ldots m_n) \mapsto m_i$.

ii. We have $M_1 \oplus M_2 \cong M_2 \oplus M_1$. So we can write $M_1 \oplus \cdots \oplus M_n = \bigoplus_{i=1}^n M_i$ without ambiguity.

Proof.

i. It is straightforward to see that they are module homomorphisms since we define the R-action component-wise.

ii. The isomorphism is because all we are doing here is swapping coordinates. The operations are performed component-wise, so the order does not really matter. ■

Example 5.24. $R^{\oplus n} = \bigoplus_{i=1}^n R$ is called the *free R-module of rank n*. We often write $e_i = (0, \ldots, 1, \ldots 0)$ be the element with the i-th component to be 1. They form a R-basis of $R^{\oplus n}$, that is, any $m \in R^{\oplus n}$ can be written uniquely as a linear combination of $\{e_1, e_2, \ldots, e_n\}$. $\Diamond$

Theorem 5.13 (Universal property of direct sums)**.** *Let R be a ring with 1. Let $M_1, M_2, \dots, M_n$ be R-modules. Then for any collection of R-module homomorphisms $\{\varphi_i : M_i \to N\}$, there is an unique R-module homomorphism $\varphi : \bigoplus_{i=1}^n M_i \to N$ such that the following diagram commutes*

$$\begin{tikzcd} M_i \arrow[r, "f_i"] \arrow[d, "\varphi_i"'] & \bigoplus_{i=1}^n M_i \arrow[dl, dashed, "\varphi"] \\ N & \end{tikzcd}$$

Proof. We define $\varphi : \bigoplus_{i=1}^n M_i \to N$ by $(m_1, \dots m_n) = m_1 e_1 + \dots + m_n e_n \mapsto \varphi_1(m_1) + \dots + \varphi_n(m_n)$. This definition is forced by the diagram. This also shows uniqueness. φ is a module homomorphism again due to the operations being defined component-wise. ■

This means that any morphism from the direct sum to any module N can be constructed just by looking at those from each M_i to N.

Theorem 5.14. *Let R be a ring with 1. Let $N_1, N_2, \dots, N_k$ be submodules of M. Then the following are equivalent:*

i. The natural map[7] *$\varphi : \bigoplus_{i=1}^k N_i \to \sum_{i=1}^k N_i = \{\sum_{i=1}^k n_i \mid n_i \in N_i\} \subseteq M$ is an isomorphism of R-modules.*

ii. Any $x \in \sum_{i=1}^k N_i$ can be written uniquely as $x = a_1 + \dots + a_k$ with $a_i \in N_i$.

iii. The intersection $N_j \cap \sum_{i \neq j} N_i = \{0\}$ for any j.

Proof.

$(i) \implies (ii)$: Surjectivity comes directly by definition. Thus φ is an isomorphism iff its kernel is trivial. Now $\ker \varphi = \{(n_1, n_2, \dots, n_k) \in \bigoplus_{i=1}^k N_i \mid \sum_{i=1}^k n_i = 0\} = \{(0, \dots, 0)\}$. This means that $0 \in M$ has an unique expression $0_M = 0_{N_1} + \dots + 0_{N_k}$. Let us assume that we have $x = a_1, a_2, \dots, a_k = b_1, b_2, \dots, b_k$, with $a_i, b_i \in N_i$. Then $0 = (a_1 - b_1) + \dots + (a_k - b_k)$, but by the unique expression of 0_M, all $a_i = b_i$.

$(ii) \implies (iii)$: Assume the contrary, that $0 \neq x \in N_j \cap \sum_{i \neq j} N_i$ for some j. But this means $x = 0_{N_1} + \dots + 0_{N_{j-1}} + x + 0_{N_{j+1}} + \dots + 0_{N_k} = a_i + \dots + a_{j-1} + 0 + a_{j+1} + \dots + a_k$. But this means that x has two distinct expressions.

$(iii) \implies (i)$: Let $(n_1, n_2, \dots, n_k) \in \ker \varphi$. Then $\sum_{i=1}^k n_i = 0$, or after rearranging, $N_j \ni -n_j = \sum_{i \neq j} n_i \in \sum_{i \neq j} N_i$. But by (iii) this means that $n_j = 0$ for any j, and therefore $\ker \varphi = \{0\}$. ■

Example 5.25. We have $\mathbb{Z}/6\mathbb{Z} \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ by the Chinese remainder theorem. We also have $\mathbb{Z}/6\mathbb{Z} \cong \mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/3\mathbb{Z}$ as modules. This is because inside $\mathbb{Z}/6\mathbb{Z}$, we have $\mathbb{Z}/2\mathbb{Z} = \{\overline{0}, \overline{3}\}$ and $\mathbb{Z}/3\mathbb{Z} = \{\overline{0}, \overline{2}, \overline{4}\}$, and their intersection is trivial. This kind of decomposition makes life easier because the end result is two simple $\mathbb{Z}$-modules. ♢

[7] It is induced by the embedding of N_i to $\sum N_i$ and to $\bigoplus N_i$. More directly, it is component-wise map from $\bigoplus N_i$ to the N_i's and then adding them up.

5.7 Free modules

Definition 5.11 (Free modules)**.** Let R be a ring with 1. Let $A = \{a_1, a_2, \ldots, a_n\}$ be a finite set. Then the *free R-module over A* (or *generated by A*) is a R-module $F(A)$ together with a function $A \to F(A)$ such that for any function $\varphi_A : A \to M$ for a R-module M, there exists a R-module homomorphism $\varphi : F(A) \to M$ such that the diagram commutes:

$$\begin{tikzcd} A \arrow[r] \arrow[d, "\varphi_A"'] & F(A) \arrow[dl, dashed, "\varphi"] \\ M & \end{tikzcd}$$

□

Theorem 5.15. *Let M be a R-module. Let $m_1, m_2, \ldots, m_n \in M$. Then there exists an unique R-module homomorphism $\varphi : R^{\oplus n} \to M$ such that $e_i \mapsto m_i$.*

Proof. The uniqueness comes by the definition of φ as $\varphi(a_1e_1 + \cdots + a_ne_n) = a_1m_1 + \cdots + a_nm_n$. Because the R-action is done component wise, it is also clear that φ is a R-module homomorphism. ■

In particular, if $M = R\{m_1, m_2, \ldots, m_n\}$, then φ is surjective, or M is a quotient module of $R^{\oplus n}$ (see isomorphism theorems). Also, sometimes we just write R^n instead of $R^{\oplus n}$ (not to be confused with $R \cdots R$).

Theorem 5.16. *$F(A)$ exists and furthermore $F(A) \cong R^{\oplus n}$.*

Proof. Take $F(A) = R^{\oplus n}$ and the function $A \to R^{\oplus n}$ be defined as $a_i \mapsto e_i$. Then for any $\varphi_A : A \to M$ defined by $a_i \mapsto m_i$, the induced map $\varphi : R^{\oplus n} \to M$ is precisely the unique R-module map $e_i \mapsto m_i$. The uniqueness follows by the universal property. ■

Recall the following theorem from linear algebra. If F is a field, then $F^n \cong F^m$ iff $n = m$. We want to show this for modules as well.

Theorem 5.17. *Let R be a commutative ring with 1. Let $M_1, M_2, \ldots, M_n$ be R-modules with submodules $N_1 \subseteq M_1$, $N_2 \subseteq M_2$, etc. Then taking the direct sum commutes with taking the quotient: $\bigoplus_i M_i / \bigoplus_i N_i \cong \bigoplus_i (M_i/N_i)$.*

Proof. As sets $\bigoplus M_i = M_1 \times \cdots \times M_n$. Then consider the following diagram.

$$\begin{tikzcd} M_i \arrow[r, "f_i"] \arrow[d, "\pi_i"] & \bigoplus_i M_i \arrow[d, "\pi"] \\ M_i/N_i \arrow[r] & \bigoplus_i (M_i/N_i) \end{tikzcd}$$

We see that π is defined by $(m, 1, \ldots, 2n) \mapsto (m_1 + N_1, \ldots, m_n + N_n)$. This is a surjective map. Furthermore $\ker \pi = \bigoplus_i N_i$. By the first isomorphism theorem we obtain our result. ■

Corollary 5.17.1. *Let $I \subseteq R$ be a left ideal. Then $R^{\oplus n}/IR^{\oplus n} \cong (R/IR)^{\oplus n}$*

Proof. It suffices to show that $IR^{\oplus n} = (IR)^{\oplus n}$. This can be seen directly from the definition $IR^{\oplus n} = \{\sum_{\text{finite}} ir \mid i \in I, r \in R^{\oplus n}\}$. ■

Theorem 5.18. *Let R be a commutative ring with 1. Then $R^{\oplus n} \cong R^{\oplus m}$ iff $n = m$.*

Proof. Let I be a maximal ideal of R. Then if $\varphi : R^{\oplus n} \to R^{\oplus m}$ is an isomorphism, since $\varphi(iR^{\oplus n}) = i\varphi(R^{\oplus n})$ for any $i \in R$, we must have $IR^{\oplus n} \cong IR^{\oplus m}$. Then taking quotients,

$$(R/I)^{\oplus n} = (R/IR)^{\oplus n} \cong R^{\oplus n}/IR^{\oplus n} \cong R^{\oplus m}/IR^{\oplus m} \cong (R/I)^{\oplus n}$$

Both sides are actually R/I-modules. But R/I is a field (lemma 1.19), and from linear algebra we get that $n = m$. ■

5.8 Direct products

Definition 5.12 (Direct product of modules)**.** Let R be a ring with $1 \neq 0$. Let $M_1, M_2, \ldots, M_n$ be R-modules. We define their direct product $M = M_1 \otimes M_2 \otimes \cdots \otimes M_n$ as

- $M = M_1 \times M_2 \times \cdots \times M_n$ as abelian groups (i.e. the product group), and
- the R-action on M is defined as $r(m_1, m_2, \ldots, m_n) = (rm_1, rm_2, \ldots, rm_n)$.

□

Theorem 5.19 (Universal property of direct products)**.** *Let R be a ring with 1. Let $M_1, M_2, \ldots,$ M_n be R-modules. Then for any R-module homomorphism $\varphi_i : N \to M_i$, there is an unique R-module homomorphism $\varphi : N \to \prod_{i=1}^n M_i$ such that the following diagram commutes:*

$$\begin{tikzcd}
M_i & \prod_{i=1}^n M_i \arrow[l, "g_i"'] \\
N \arrow[u, "\varphi_i"] \arrow[ur, dashed, "\varphi"']
\end{tikzcd}$$

The direct product exists and is unique up to isomorphism.

Proof. Define $\varphi : N \to \prod_{i=1}^n M_i$ as $s \mapsto (\varphi_1(s), \ldots \varphi_n(s))$. It remains to be checked that φ is a R-module homomorphism. $\varphi(rs) = (\varphi_1(rs), \ldots, \varphi_n(rs)) = (r\varphi_1(s), \ldots, r\varphi_n(s)) = r\varphi(s)$. $\varphi(s + s') = \varphi(s) + \varphi(s')$ since addition is done component-wise. ■

The direct product and the direct sum may seem identical but they are not when dealing with infinite number of terms. For example $\bigoplus_{i\in\mathbb{Z}^+} \mathbb{Z} \ncong \prod_{i\in\mathbb{Z}^+} \mathbb{Z}$ (Section 10.3 ex 24).

Let $M_1, \ldots, M_s, N_1, \ldots, N_t$ be R-modules. We want to study $\operatorname{Hom}_R(\bigoplus_{i=1}^s M_i, \bigoplus_{i=1}^t N_t)$, $\operatorname{End}_R(\bigoplus_{i=1}^s M_i)$, and $\operatorname{End}_R(R^{\oplus n})$.

Example 5.26. Let F be a field. Then $\operatorname{End}_F(F^n) \cong \operatorname{Mat}_n(F)$. ◊

Definition 5.13. Let $M_1, \ldots M_s$ be R-modules. Define the following matrix ring $\widetilde{\operatorname{End}}_R(\bigoplus_{i=1}^s M_i)$

- As a set of $s \times s$ matrices $\{(a_{ij}) \mid a_{ij} \in \operatorname{Hom}_R(M_j, M_i)$.

- Addition is defined as $(a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij})$
- Multiplication is defined as $(a_{ij})(b_{ij}) = (\sum_{k=1}^{s} a_{ik} \circ b_{kj})$.

□

Theorem 5.20. *We have* $\mathrm{End}_R(\bigoplus_{i=1}^{s} M_i) \cong \widetilde{\mathrm{End}}_R(\bigoplus_{i=1}^{s} M_i)$.

Proof. Define the map $\mathrm{End}_R(\bigoplus_{i=1}^{s} M_i) \to \widetilde{\mathrm{End}}_R(\bigoplus_{i=1}^{s} M_i)$ as $\varphi \mapsto (\varphi_{ij})$. See the following diagram.

$$\begin{tikzcd}
M_j \arrow[r, shift left, "f_j"] \arrow[d, "\varphi_{ij}"'] & \bigoplus_{i=1}^{s} M_i \arrow[l, shift left, "g_j"] \arrow[d, "\varphi"] \\
M_i \arrow[r, shift left, "f_j"] & \bigoplus_{i=1}^{s} M_i \arrow[l, shift left, "g_i"]
\end{tikzcd}$$

This map is a bijection. The fact that it is a ring homomorphism follows from linear algebra (similar to linear transformations). ■

In particular, we have $\mathrm{End}_R(R^{\oplus n}) \cong \mathrm{Mat}_n(R)$ for a commutative ring R. Furthermore, even though we used End_R, but we see that we only need properties of the direct sum in the theorem. Thus we can also have something like

$$\begin{gathered}
\mathrm{Hom}_R(R^{\oplus s}, R^{\oplus t}) \times \mathrm{Hom}_R(R^{\oplus k}, R^{\oplus s}) \to \mathrm{Hom}_R(R^{\oplus k}, R^{\oplus t}) \\
\cong \\
\mathrm{Mat}_{t,s}(R) \times \mathrm{Mat}_{s,k}(R) \to \mathrm{Mat}_{t,k}(R)
\end{gathered}$$

5.9 Noetherian modules

Definition 5.14 (Noetherian modules)**.** Let R be a ring with 1. A R-module M is called *Noetherian* if any ascending chain of submodules of M stabilises, that is

$$N_1 \subseteq N_2 \subseteq \cdots$$

then $N_k = N_{k+1} = \cdots$ for some k. □

Theorem 5.21. *Let M be a R-module. Then M is Noetherian iff any submodule of M is finitely generated.*

Proof. Recall that being finitely generated means that N is made up of R-linear combinations of finitely many elements, i.e. $N = R\{n_1, \ldots, n_k\}$.

($\Longrightarrow$): We assume M is Noetherian. Let $N \subseteq M$ be a submodule. Assume that N is not finitely generated. Then we can find a sequence of elements in N, denoted $\{n_i\}_i^{\infty}$, such that $n_i \notin R\{n_1, \ldots, n_{i-1}\} \neq N$. Then we have an ascending chain of modules that never stabilises.

($\Longleftarrow$): Suppose any submodule of M is finitely generated. Then for any ascending chain of submodules of M, say $N_1 \subseteq N_2 \subseteq \cdots$, let $N = \bigcup_i^{\infty} N_i$. This is a submodule of M, hence is finitely generated. Let $N = R\{n_1, \ldots, n_k\}$. Then we can find N_k such that $n_1, \ldots n_k \in N_k$, so the chain stabilises at N_k. ■

Just to be sure that we did not prove something silly, there does exist finitely generated modules M with non-finitely generated submodules N, but they will generally be quite strange looking:

Example 5.27. Take a polynomial ring with infinitely many variables $R = F[x_1, x_2, \dots]$. Then let $M = R = R\{1\}$. Let N be the submodule consisting of polynomials without the constant terms. This is not finitely generated since there are infinitely many degree one polynomials. ♢

Theorem 5.22. *Let R be a ring with 1. Let M be a R-module with a submodule N. Then M is Noetherian iff both N and M/N are Noetherian.*

Proof.

$(\implies)$: Assume M be Noetherian. Then all submodules of M are finitely generated. Submodules of N are also submodules of M, thus all submodules of N are finitely generated and so N is Noetherian. Any submodule $L/N \subseteq M/N$ for some $L \subseteq M$ is going to be finitely generated since if $L = R\{l_1, \dots, l_n\}$ then $L/N = R\{l_1 + N, \dots, l_n + N\}$. So M/N is also Noetherian.

$(\impliedby)$: Assume both N and M/N are Noetherian. Then any submodule of N and any submodule of M/N are finitely generated. We want to show that any $L \subseteq M$ is finitely generated. We know $\overline{L}$ (the image of L in M/N[8]) is finitely generated, so let $\overline{L} = R\{a_1 + N, \dots a_s + N\}$. We also know that $L \cap N$ is finitely generated, so let $L \cap N = R\{b_1, \dots, b_t\}$.

We claim that L is generated by $\{a_1, \dots, a_s, b_1, \dots, b_t\}$. Let $x \in L$. Then $x+N = r_1a_1+\cdots+r_sa_s+N$ with $r_i \in R$. Then $x-(r_1a_1+\cdots+r_sa_s) \subseteq N\cap L$. Therefore $x = r_1a_1+\cdots+r_sa_s+r'_1b_1+\cdots+r'_tb_t$ for $r'_i \in R$. ■

Definition 5.15 (Noetherian rings)**.** Let R be a ring with 1. We say R is left (right) Noetherian if R is Noetherian as a left (right) R-module via left (right) multiplication. we say R is Noetherian if it is both left and right Noetherian. □

Recall that if R is a PID, then R is Noetherian.

Theorem 5.23. *Let R be left Noetherian. Then any finitely generated R-module is Noetherian.*

Proof. We first show $R^{\oplus n}$ is Noetherian. First consider $R^{\oplus 2}$. It contains $R \times \{0\} = N$ as a submodule and $M = R^{\oplus 2}/N \cong R$ as a quotient. Thus $R^{\oplus 2}$ is Noetherian by theorem 5.22. We then proceed by induction for all n.

Let $M = R\{m_1, \dots, m_n\}$ be a finitely generated R-module. We have a surjective R-module homomorphism $R^{\oplus n} \to M$ given by $e_i \mapsto m_i$ (see remark after theorem 5.15). So M is a quotient of a Noetherian module (again, see the remark), and is hence Noetherian. ■

In particular, we have the special case where a finitely generate module over a PID, it is Noetherian, which means that all of its sub-modules are finitely generated as well.

Corollary 5.23.1. *Any subgroup of a finitely generated abelian group (equivalently a $\mathbb{Z}$-module) is finitely generated.*

Definition 5.16 (Cokernels)**.** Let $\varphi : M \to N$ be a module homomorphism. Define the *cokernel* of φ as $\operatorname{coker}\varphi = N/\operatorname{img}\varphi$. □

[8] This is something like L/N but in general N might not be in L so it is not exactly L/N.

Example 5.28. Let $\varphi : R^{\oplus 3} \to R^{\oplus 3}$ for a PID R be the matrix

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \alpha \end{pmatrix}$$

Then $\operatorname{coker} \varphi = R^{\oplus 2} \oplus R/(\alpha)$. ◊

Theorem 5.24. *Let R be left Noetherian. Then any finitely generated R-module M is also finitely presented, that is there is some $\varphi : R^{\oplus m} \to R^{\oplus n}$ such that $M \cong \operatorname{coker} \varphi$.*

Proof. For any surjective module homomorphism $f : R^{\oplus n} \to M$ (the fact that at least one exists is comes from the proof of the previous theorem), since $R^{\oplus n}$ is Noetherian, $\ker f$ is a submodule and so it is finitely generated. So now we can consider a surjective homomorphism $g : R^{\oplus m} \to \ker f$. Denote the embedding map $\pi : \ker f \to R^{\oplus n}$. Then let $\varphi : R^{\oplus m} \to R^{\oplus n} = \pi \circ g$. Thus $\operatorname{img} \varphi = \ker f$, and $\operatorname{coker} \varphi = R^{\oplus n} / \ker f \cong M$ by the isomorphism theorem for modules. ■

Theorem 5.25. *Let R be a ring with 1. Let M be a Noetherian module. Then M has a simple quotient, that is, there is $N \subseteq M$ such that M/N is a simple R-module.*

Proof. Recall that M/N is simple if it is non-zero and if $\{0\}$ and M/N are the only submodules.

If M is simple, we are done. Otherwise, we can find $0 \neq N_1 \subsetneq M$. If M/N_1 is simple, then we are done. Otherwise, we can find a non-zero submodule $N_2' \subsetneq M/N_1$. Let $\pi : M \to M/N_1$ be the quotient map. Then let $N_2 = \pi^{-1}[N_2'] \subsetneq M$. By the isomorphism theorem, $M/N_2 \cong (M/N_1)/(N_2/N_1) = (M/N_1)/N_2'$. We continue the process for M/N_2. Then we can obtain an ascending chain of submodules $N_1 \subset N_2 \subset \dots$ but this process must terminate after finitely many steps by the Noetherian assumption. ■

Example 5.29. Let $R = F[x]$ for a field F. This is Noetherian since it is a Euclidean domain.

- But R as an R-module has no simple submodules since by a degree argument xM is always a submodule of M.
- Let $R = F[x]$. Then the module $M = \operatorname{Hom}_F(F[x], F)$ with the action defined by $(g(x)f)(h(x)) = f(g(x)h(x))$ has no simple quotients.

◊

Theorem 5.26 (Hilbert basis theorem)**.** *Let R be a commutative Noetherian ring. Then $R[x]$ is Noetherian.*

Proof. Assume the contrary, i.e. $R[x]$ is not Noetherian. We know that this means there is an ideal I that is not finitely generated as a $R[x]$-module. We then construct a sequence $\{f_i\}_{i=1}^{\infty}$ in I such that $f_i \notin R\{f_1, \dots, f_{i-1}\}$ is of minimal degree. Let a_i be the leading coefficient of f_i. Consider the chain of ideals in R generated by the coefficients $(a_1) \subseteq (a_1, a_2) \subseteq \dots$. This stabilizes, say at $(a_1, \dots, a_m)$. So $a_{m+1} = \sum_{i=1}^{m} r_i a_i$. Then we consider $g = \sum_{i=1}^{m} r_i f_i (x^{\deg f_{m+1} - \deg f_i}) \in R[x]\{f_1, \dots, f_m\}$. Then since $f_{m+1} \notin R[x]\{f_1, \dots, f_m\}$, so $f_{m+1} - g \notin R[x]\{f_1, \dots, f_m\}$ either. However, it has degree lower than $\deg f_{m+1}$. This contradicts with the choice of f_{m+1}. ■

In particular, $R[x_1, \dots, x_n]$ is also Noetherian.

6 Modules over PIDs

6.1 Smith normal form

We can have seen multiple references to matrices already. We can study modules by matrix computation. The following theorem gives a motivation as to why matrix computation is useful.

Theorem 6.1. *Let R be a ring with 1. Let $\varphi : R^{\oplus m} \to R^{\oplus n}$ be a R-module homomorphism. Let $S : R^{\oplus m} \to R^{\oplus m}$ and $T : R^{\oplus n} \to R^{\oplus n}$ be isomorphisms of R-modules. Then $\operatorname{coker}\varphi \cong \operatorname{coker}(T \circ \varphi \circ S)$.*

Proof. Consider the picture below. We want to show $\operatorname{coker}\varphi = M \cong \operatorname{coker} T \circ \varphi \circ S$.

$$\begin{tikzcd}
 & R^{\oplus m} \arrow[r, "\varphi"] & R^{\oplus n} \arrow[r, two heads, "\pi"] & (M = \operatorname{coker}\varphi) \\
R^{\oplus m} \arrow[r, "S"] & R^{\oplus m} \arrow[r, "\varphi"] & R^{\oplus n} \arrow[u, "T"] &
\end{tikzcd}$$

Consider the map $f : R^{\oplus n} \xrightarrow{T^{-1}} R^{\oplus n} \xrightarrow{\pi} M$. Since π is surjective and T is an isomorphism, f is surjective. Furthermore $R^{\oplus n}/\ker f \cong M$ by the isomorphism theorem. In fact $\ker f = T[\ker \pi]$. We now claim that $\operatorname{img}(T \circ \varphi \circ S) = \ker f$, so $\operatorname{coker}(T \circ \varphi \circ S) \cong R^{\oplus n}/(\operatorname{img} T \circ \varphi \circ S) \cong R^{\oplus n}/\ker f \cong M$.

First, $\operatorname{img}(T\circ\varphi\circ S) \subseteq \ker f$ iff $f[\operatorname{img} T\circ\varphi\circ S] = \{0\}$. By definition we have $f[T\circ\varphi\circ S[R^{\oplus m}]] = \pi \circ T^{-1} \circ T \circ \varphi \circ S[R^{\oplus m}] = \pi \circ \varphi \circ S[R^{\oplus m}]$. Since $\pi \circ \varphi = 0$, we have our result.

Next, $\ker f = \ker(T^{-1}\circ\pi) = T[\ker \pi] = T[\operatorname{img}\varphi]$ since T^{-1} is an isomorphism. On the other hand, $\operatorname{img}(T\circ\varphi\circ S) = T[\operatorname{img}\varphi\circ S] = T[\operatorname{img}\varphi]$ since S is an isomorphism. So $\ker f \subseteq \operatorname{img}(T\circ\varphi\circ S)$. ■

Example 6.1. Consider $\mathbb{Z}^2$. Let $M = \operatorname{coker}\varphi$ with $\varphi = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$. Then let

$$\varphi = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \qquad T = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \qquad S = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \qquad T \circ \varphi \circ S = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

We see that $\operatorname{coker}\varphi \cong \operatorname{coker}(T \circ \varphi \circ S) \cong \mathbb{Z}$. ♢

There is one problem now. How are we to find invertible T and S? Recall that in a PID R, if non-zero $a, b \in R$ and $(\alpha) = (a, b)$, then $\alpha = \gcd(a, b)$ and $\alpha = ac + bd$ for some $c, d \in R$. This gives us the following easy fact:

Theorem 6.2. *Let R be a PID. Let $a, b \in R$ be coprime. So we can find $c, d \in R$ such that $ac + bd = 1$. Then*

$$\begin{pmatrix} a & b \\ -d & c \end{pmatrix}, \begin{pmatrix} c & -b \\ d & a \end{pmatrix}, \begin{pmatrix} a & -d \\ b & c \end{pmatrix}, \begin{pmatrix} c & d \\ -b & a \end{pmatrix}$$

are all invertible in $\operatorname{Mat}_2(R)$ (the first two are mutual inverses and so are the other two).

We also see that it is quite easy to generalise this to something like

$$\begin{pmatrix} 1 & 0 & \cdots & & & \\ 0 & \ddots & & & & \\ \vdots & & a & b & & \\ & & -d & c & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}$$

Corollary 6.2.1. *Let R be a PID. Let $a, b \in R$ and $\alpha = \gcd(a, b)$. Then there exists invertible matrices $S, T \in \mathrm{Mat}_2(R)$ such that*

$$\begin{pmatrix} a & b \\ ? & ? \end{pmatrix} S = \begin{pmatrix} \alpha & 0 \\ ? & ? \end{pmatrix} \qquad T \begin{pmatrix} a & ? \\ b & ? \end{pmatrix} = \begin{pmatrix} \alpha & ? \\ 0 & ? \end{pmatrix}$$

Proof. Let $a = \alpha a'$ and $b = \alpha b'$ and $\alpha = ac + bd$ for some $c, d \in R$. Then substituting and cancelling gives $a'c + b'd = 1$. We have

$$\begin{pmatrix} a & b \\ ? & ? \end{pmatrix} \begin{pmatrix} c & -b' \\ d & a' \end{pmatrix} = \begin{pmatrix} ac + bd & -ab' + ba' \\ ? & ? \end{pmatrix} = \begin{pmatrix} \alpha & 0 \\ ? & ? \end{pmatrix}$$

$$\begin{pmatrix} c & d \\ -b' & a' \end{pmatrix} \begin{pmatrix} a & ? \\ b & ? \end{pmatrix} = \begin{pmatrix} ac + bd & ? \\ -b'a + a'b & ? \end{pmatrix} = \begin{pmatrix} \alpha & ? \\ 0 & ? \end{pmatrix}.$$

■

We can also generalise this to something like

$$\begin{pmatrix} a & b & c \\ ? & ? & ? \end{pmatrix} S = \begin{pmatrix} \alpha & 0 & c \\ ? & ? & ? \end{pmatrix}.$$

Theorem 6.3. *Let R be a PID. Let $a, b \in R$. Then there exist invertible matrices $S, T \in \mathrm{Mat}_2(R)$ such that*

$$T \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} S = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}$$

with $\alpha \mid \beta$.

Proof. Let $a = \alpha a'$ and $b = \alpha b'$ and $\alpha = \gcd(a, b) = ac + bd$ for some $c, d \in R$. Then substituting and cancelling gives $a'c + b'd = 1$. We consider the following operations

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} a & 0 \\ b & b \end{pmatrix}$$
$$\begin{pmatrix} c & d \\ -b' & a' \end{pmatrix} \begin{pmatrix} a & 0 \\ b & b \end{pmatrix} = \begin{pmatrix} \alpha & db \\ 0 & a'b \end{pmatrix}$$
$$\begin{pmatrix} \alpha & db \\ 0 & a'b \end{pmatrix} \begin{pmatrix} 1 & -db' \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \alpha & 0 \\ 0 & a'b \end{pmatrix}$$

Then $\alpha \mid a'b$ since $\alpha \mid b$. ■

In fact, $\alpha = \gcd(a, b)$, and $\alpha\beta = ab$ is the determinant of the middle matrix.

Theorem 6.4 (Smith normal form)**.** *Let R be a PID. Let $\varphi : R^{\oplus m} \to R^{\oplus n}$ be a R-module homomorphism. We consider $\varphi \in \mathrm{Mat}_{n,m}(R)$. Then there exists invertible $S \in \mathrm{Mat}_m(R)$ and $T \in \mathrm{Mat}_n(R)$ such that*

$$S\varphi T = \begin{pmatrix} \alpha_1 & 0 & \cdots & & & & \\ 0 & \alpha_2 & & & & & \\ \vdots & & \ddots & & & & \\ & & & \alpha_k & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix}$$

where $\alpha_1 \mid \alpha_2 \mid \alpha_3 \dots$ and all $\alpha_i \neq 0$. The entries α_i are called the invariant factors *of φ. The matrix is called the* Smith normal form*.*

Right now the name is not well defined since the normal form depends on S and T. However we can actually prove the independence of S and T by considering the principle minors (determinants of a smaller square sub-matrix) of φ.

Proof. We will show the construction for a 2×3 matrix and it can be easily generalised further. Using the operations available to us from the above, along with elementary operations, we first sort out the first row:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \to \begin{pmatrix} a'_{11} = \gcd(a_{11}, a_{12}) & 0 & a_{13} \\ ? & ? & ? \end{pmatrix} \to \begin{pmatrix} a'_{11} & a_{13} & 0 \\ ? & ? & ? \end{pmatrix}$$
$$\to \begin{pmatrix} a''_{11} = \gcd(a'_{11}, a_{13}) & 0 & 0 \\ ? = a'_{21} & ? & ? \end{pmatrix}$$

Next, if $a'_{11} \mid a_{21}$, we can add a multiple of the first row to the second row to cancel it out. The precise matrix used is $\begin{pmatrix} 1 & 0 \\ -a'_{21}/a''_{11} & 1 \end{pmatrix}$. Note that doing this does not change the first row. Otherwise we do the same gcd taking operation as before:

$$\begin{pmatrix} a''_{11} = \gcd(a'_{11}, a_{13}) & 0 & 0 \\ ? = a'_{21} & ? & ? \end{pmatrix} \to \begin{cases} \begin{pmatrix} a''_{11} & 0 & 0 \\ 0 & ? & ? \end{pmatrix}, & \text{if } a''_{11} \mid a'_{21} \\ \begin{pmatrix} a'''_{11} = \gcd(a''_{11}, a'_{21}) & ? & ? \\ 0 & ? & ? \end{pmatrix}, & \text{otherwise} \end{cases}.$$

It would seem that if we do not have divisibility then we are in a tough spot since the first row has been messed up again. However, if $a''_{11} \nmid a'_{21}$, then the number of irreducible factors in a'''_{11} is less than the number of irreducible factors in a''_{11} (R is a UFD). Therefore we start from the beginning with this new matrix, and we know that this will terminate after finitely many steps. Finally, when we get to the first case where $a''_{11} \mid a'_{21}$ we just repeat the entire process but for the second row only. ■

6.2 Structure theorem

Theorem 6.5 (Fundamental theorem for finitely generated modules over a PID)**.** *Let R be a PID and let M be a finitely generated R-module. Then there is some $k \geq 0$ such that*

$$M \cong R^{\oplus k} \oplus R/(\alpha_1) \oplus \cdots \oplus R/(\alpha_r)$$

as R-modules, with $\alpha_1 \mid \alpha_2 \mid \dots$ and all $\alpha_i \neq 0$ and $\alpha_i \notin R^*$ (not units). The α_i are called the invariant factors *of M and k is called the* free rank *of M.*

The names are not well-defined yet, and we shall leave a proof of uniqueness till later.

Proof. Let $\varphi : R^{\oplus m} \to \mathbb{R}^{\oplus n}$ and let $M \cong \operatorname{coker}(\varphi) \cong \operatorname{coker}(S \circ \varphi \circ T)$ where S and T are invertible and $S \circ \varphi \circ T = \operatorname{diag}(\alpha_1, \dots, \alpha_r, u_1, \dots, u_l, 0 \dots)$ gives us the Smith normal form, only that we rearrange the terms such that all u_i are units, but all a_i are not units.

Then, $\operatorname{img}(S \circ \varphi \circ T) = (\alpha_1) \oplus \cdots \oplus (\alpha_r) \oplus R \oplus \cdots \oplus R \oplus 0 \oplus \cdots \oplus 0$. Therefore $M \cong \operatorname{coker}(S \circ \varphi \circ T) \cong R^{\oplus k} \oplus R/(\alpha_1) \oplus \cdots \oplus R/(\alpha_r)$. Note that k can be possibly 0. ■

Example 6.2. Let F be a field (which is also a PID). Any finitely generated F-module is isomorphic to $F^{\oplus k}$ for some $k \geq 0$. This is because in a field every non-zero element is an unit. Therefore this also shows that every F-vector space has a basis. ♢

Example 6.3. Let F be a field. Let $G \subseteq F^*$ be a finite subgroup. Then G is cyclic.

Since G is a finitely generated abelian group, $G \cong \mathbb{Z}/n_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/n_r\mathbb{Z}$ for $n_1 \mid n_2 \mid \dots$ where all n_i are not units or zero.

Recall that $F[x]$ is an Euclidean domain. Any $f(x)$ of degree n has at most n roots. Consider the polynomial $x^{n_r} - 1 \in F[x]$. It has at most n_r roots in F and hence also in $G \subseteq F^*$. Therefore $G \cong \mathbb{Z}/n_1\mathbb{Z}$ only. ♢

Example 6.4. Let A be the abelian group generated by x, y, z, subject to the relations

$$2x + 4y + 4z = 0 \qquad 6x + 3y - 6z = 0 \qquad 4x + 14y + 20z = 0$$

Consider $\pi : \mathbb{Z}^3 \to A$ given by $\mathbf{e}_1 \mapsto x$, $\mathbf{e}_2 \mapsto y$, $\mathbf{e}_3 \mapsto z$. Then by definition $\ker \pi$ is generated by

$$\begin{pmatrix} 2 \\ 4 \\ 4 \end{pmatrix}, \begin{pmatrix} 6 \\ 3 \\ -6 \end{pmatrix}, \begin{pmatrix} 4 \\ 14 \\ 20 \end{pmatrix}.$$

Then we consider $\varphi : \mathbb{Z}^3 \to \mathbb{Z}^3$ where $\mathbf{e}_1 \mapsto (2,4,4)^T$, $\mathbf{e}_2 \mapsto (6,3,-6)^T$, $\mathbf{e}_3 \mapsto (4,14,20)^T$. Then $\operatorname{coker} \varphi \cong \mathbb{Z}^3 / \operatorname{img} \varphi \cong \mathbb{Z}^3 / \ker \pi \cong A$.

$$\begin{pmatrix} 2 & 6 & 4 \\ 4 & 3 & 14 \\ 4 & -6 & 20 \end{pmatrix} \to \begin{pmatrix} 2 & 0 & 0 \\ 4 & 3 - 3 \cdot 4 & 14 - 4 \cdot 2 \\ 4 & -6 - 3 \cdot 3 & 20 - 4 \cdot 2 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 4 & -9 & 6 \\ 4 & -18 & 12 \end{pmatrix} \to \begin{pmatrix} 2 & 0 & 0 \\ 0 & -9 & 6 \\ 0 & -18 & 12 \end{pmatrix} \to \cdots$$

$$\to \begin{pmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 2 & 0 & 0 \\ -3 & -3 & 0 \\ 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} -1 & -3 & 0 \\ -3 & -3 & 0 \\ 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} -1 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

So $A \cong \operatorname{coker} \varphi \cong \mathbb{Z}/\mathbb{Z} \oplus \mathbb{Z}/6\mathbb{Z} \oplus \mathbb{Z}/0\mathbb{Z} = \mathbb{Z}/6\mathbb{Z} \oplus \mathbb{Z}$. ♢

Theorem 6.6 (Elementary factors)**.** *Let R be a PID and let M be a finitely generated R-module. Then $M \cong R^{\oplus k} \oplus R/(p^{a_1}) \oplus \dots \oplus R/(p_n^{a_n})$ for some $k \geq 0$ and not necessarily distinct primes p_i and $a_i > 0$. Then $p_i^{a_i}$ are called the* elementary factors *of M.*

Proof. Recall the Chinese remainder theorem for PID.

$$R/(\alpha_i) = R/(p_1^{a_1} \dots p_n^{a_n}) \cong R/(p^{a_1}) \oplus \dots \oplus R/(p_n^{a_n}).$$

So if $M \cong R^{\oplus k} \oplus R/(\alpha_1) \oplus \dots$, we just decompose each α_i into primes. ■

Definition 6.1 (Torsion)**.** Let M be a R-module for any ring R. We define the *torsion* of M as

$$\mathrm{Tor}(M) = \{m \in M \mid rm = 0, 0 \neq r \in R\}.$$

M is called *torsion free* if $\mathrm{Tor}(m) = \{0\}$. □

Theorem 6.7. *If R is an integral domain then* $\mathrm{Tor}(M)$ *is a submodule.*

Proof. First we check that $\mathrm{Tor}(M)$ is an abelian subgroup. Let $m, n \in \mathrm{Tor}(M)$ such that $rm = 0$ and $sn = 0$. Then $sr(m - n) = s(rm) - r(sn) = 0$, so $m - n \in \mathrm{Tor}(M)$.

Next we check that $\mathrm{Tor}(M)$ is closed under the R-action. For any $s \in R$, $r(sm) = s(rm) = 0$ so $sm \in \mathrm{Tor}(M)$. ■

Lemma 6.8. *Let R be a ring, let $\varphi : M \to N$ be a R-module homomorphism. Then $\varphi[\mathrm{Tor}(m)] \subseteq \mathrm{Tor}(N)$.*

Proof. Let $m \in \mathrm{Tor}(M)$, so $rm = 0$ with $0 \neq r \in R$. Then $\varphi(rm) = r\varphi(m) = 0$. So $\varphi(m) \in \mathrm{Tor}(N)$. ■

Theorem 6.9. *Let R be a PID. Let M be a finitely generated module over R. Then M is torsion free iff M is free.*

Proof. Let $M \cong R^{\oplus k} \oplus R/(\alpha_1) \oplus \dots \oplus R/(\alpha_r)$ with $\alpha \mid \alpha \mid \dots$ and all α_i are not zero or units. From the previous lemma, isomorphisms preserve torsions. Therefore the torsion of M is easy to compute from the right hand side, in fact we have $\mathrm{Tor}(M) = R/(\alpha_1) \oplus \dots \oplus R/(\alpha_r)$. Thus M is torsion free iff $M \cong R^{\oplus k}$ with $k > 0$. ■

Example 6.5. The abelian group $\mathbb{Q}$ is torsion free but not free (and hence is not finitely generated).

Let $\frac{p}{q} \in \mathbb{Q}$. Then $n \cdot \frac{p}{q} \neq 0$ for any $0 \neq n \in \mathbb{Z}$ so $\mathbb{Q}$ is torsion free. Now suppose it is free. Let A be a set such that $F(A) \cong \mathbb{Q}$ ($\mathbb{Q}$ is generated by A).

If $A = \{x\}$, then suppose the isomorphism $\varphi : \mathbb{Z} \cong F(A) \to \mathbb{Q}$ maps $x \mapsto \frac{p}{q}$ with $\gcd(p, q) = 1$. Then $\varphi[F(A)] = \{\mathbb{Z} \cdot \frac{p}{q}\} \neq \mathbb{Q}$, the map cannot be surjective.

Otherwise, consider the embedding $\mathbb{Z}^2 \to F(A)$, given by $(1, 0) \mapsto x$ and $(0, 1) \mapsto y$ for $x, y \in A$. Then consider the composition with an isomorphism from $F(A)$ to $\mathbb{Q}$. The overall map is defined by $(1, 0) \mapsto x \mapsto \frac{p_1}{q_1}$, $(0, 1) \mapsto y \mapsto \frac{p_2}{q_2}$. This map is injective, since the embedding and isomorphism are both injective. However, $(q_1 p_2, p_1, q_2) \mapsto q_1 p_2 \cdot \frac{p_1}{q_1} - p_1 q_2 \frac{p_2}{q_2} = 0$, which is a contradiction since the kernel is non-trivial. ♢

Lemma 6.10. *Let R be a PID. Let $p \in R$ be a prime element. M/pM is a finite dimensional F-vector space. Similarly, pM/p^2M is a finite dimensional F-vector space.*

Proof. First of all recall that since R is a PID, (p) is maximal, and so $F = R/(p)$ is a field. Also, for a R-module M, pM is a R-submodule of M.

We define the F-action on M/pM as

$$(a + (p))(m + pM) = am + pM.$$

We can also think of this as a R action on M/pM where (p) acts trivially and the action factors through. Checking that this is a well-defined action is skipped.

To show that M/pM is a finite dimensional F-vector space, it suffices to show that F-linear combination of $\overline{m}_1, \ldots, \overline{m}_n$, span the space, where $m_1, \ldots, m_n \in M$ are the generators of M. Since any $m = \sum_i r_i m_i$, therefore for any $\overline{m} \in M/pM$, $\overline{m} = \sum_i r_i \overline{m}_i = \sum_i \overline{r}_i \overline{m}_i$. ■

Lemma 6.11. *Let R be a PID and let $p, q \in R$ be distinct primes. Let $F = R/(p)$ and $a, b \geq 1$.*

i. Let $M = R/(q^a)$. Then as F-modules, we have $p^tM/p^{t+1}M \cong 0$.

ii. Let $M = R/(p^a)$. Then as F-modules, we have $p^tM/p^{t+1}M = \begin{cases} F, & \text{if } t < a \\ 0, & \text{if } t \geq a \end{cases}$.

iii. Let $M = R/(p^a) \oplus R/(q^b)$. Then as F-modules, we have $p^tM/p^{t+1}M \cong \begin{cases} F, & \text{if } t < a \\ 0, & \text{if } t \geq a \end{cases}$.

Proof.

i. We have for any $t \geq 0$

$$p^tM = p^t(R/(q^a)) = ((p^t) + (q^a))/(q^a) = (p^t, q^a)/(q^a) = (\gcd(p^t, q^a))/(q^a) = R/(q^a).$$

Then $p^tM/p^{t+1}M$ is trivial since the two sides of the quotient are the same thing.

ii. Similar to the above,

$$p^tM = (\gcd(p^t, p^a))/(p^a) = (p^{\min(t,a)})/(p^a).$$

If $t \geq a$, then $p^tM/p^{t+1}M = (p^{\min(t,a)})/(p^a) = 0$. If $t < a$, then $p^tM/p^{t+1}M = ((p^t)/(p^a))/((p^{t+1})/(p^a)) \cong (p^t)/(p^{t+1})$. We can define an homomorphism $R \to (p^t)/(p^{t+1})$ given by $1_R \mapsto p^t + (p^{t+1})$. The kernel of this homomorphism is (p), therefore we have $p^tM/p^{t+1}M \cong R/(p) = F$.

iii. Taking quotient commutes with taking the direct sum, so we only need to combine the previous two points.

$$\frac{p^t(R/(p^a)) \oplus p^t(R/(q^b))}{p^{t+1}(R/(q^a)) \oplus p^{t+1}(R/(q^b))} \cong \frac{p^t(R/(p^a))}{p^{t+1}(R/(a^a))} \oplus \frac{p^t(R/(q^b))}{p^{t+1}(R/(q^b))}$$

■

Theorem 6.12 (Structure theorem)**.** *Let R be a PID. Let M be a finitely generated R-module. Suppose*

$$M \cong R^{\oplus k} \oplus \bigoplus_{i=1}^{s} R/(p_i^{a_i}) \cong R^{\oplus k'} \oplus \bigoplus_{i=1}^{r} R/(q_i^{b_i}).$$

Then we have $k = k'$ and $p_i^{a_i} = u_i q_i^{b_i}$ for some units u_i up to rearrangement.

Proof. Firstly,

$$\operatorname{Tor}(M) \cong \bigoplus_{i=1}^{s} R/(p_i^{a_i}) \cong \bigoplus_{i=1}^{r} R/(q_i^{b_i}).$$

Then $M/\operatorname{Tor}(M) \cong R^k \cong R^{k'}$. Since R is commutative, by theorem 5.18, $k = k'$.

For simplicity of notation let us consider the following example for the next part of the proof. Let

$$M \cong R/(p_1) \oplus R/(p_1^2) \oplus R/(p_2^2) \cong R/(q_1^{b_1}) \oplus \cdots \oplus R/(q_2^{b_r})$$

where $p_1 \neq p_2$ but q_i are not necessarily distinct. There is no loss of generality here.

Let $F_1 = R/(p_1)$, let π_{p_1} be a homomorphism such that $\pi_{p_1}[M] = M/p_1M$. We have a F-vector space isomorphism by the previous theorem

$$\begin{aligned}\pi_{p_1}[R/(q_1^{b_1})] \oplus \cdots \oplus \pi_{p_1}[R/(q_r^{b_r})] \cong M/p_1M &\cong \pi_{p_1}[R/(p_1)] \oplus \pi_{p_1}[R/(p_1^2)] \oplus \pi_{p_1}[R/(p_2^2)] \\ &\cong F \oplus F \oplus 0.\end{aligned}$$

So there are exactly two elementary factors that are some power of p_1. Assume $q_1^{b_1} = p_1^{b_1}$ and $q_2^{b_2} = q_2^{b_2}$. By a similar consideration of $F_2 = R/(p_2)$, we can also conclude that there is exactly one elementary factor that is some power of p_2. Assume $q_3^{b_3} = p_3^{b_3}$. Then considering again $F_q = R/(q)$ for other primes, we can conclude that there are no other factors.

Now let $\pi_{p_1^2}$ be a homomorphism such that $\pi_{p_1^2}[M] = p_1M/p_1^2M$. Then

$$\begin{aligned}\pi_{p_1^2}[R/(p_1^{b_1})] \oplus \pi_{p_1^2}[R/(p_1^{b_2})] \oplus \pi_{p_1^2}[R/(p_2^{b_3})] &\cong p_1M/p_1^2M \\ &\cong \pi_{p_1^2}[R/(p_1)] \oplus \pi_{p_1^2}[R/(p_1^2)] \oplus \pi_{p_1^2}[R/(p_2^2)] \\ &\cong 0 \oplus F \oplus 0.\end{aligned}$$

This allows us to conclude that $b_1 = 1$ and $b_2 > 1$ up to permutation. We continue with $\pi_{p_1^3}$ and so on. ■

Corollary 6.12.1. *Let R be a PID and let M be a finitely generated R-module. Then the invariant factors of M is unique up to multiplication by units.*

Proof. The above theorem with the Chinese remainder theorem in reverse. ■

Example 6.6. We can use this to classify all abelian groups of order $180 = 2^2 \times 3^2 \times 5$. Since an abelian group $M \cong \mathbb{Z}/(p_1^{a_1}) \oplus \cdots \oplus \mathbb{Z}/(p_r^{a_r})$ for not necessarily distinct p_i, the order of $M = p_1^{a_1} \dots p_r^{a_r}$. So we have the following possibilities for elementary factors

- $2, 2, 3, 3, 5$

- $2^2, 3, 3, 5$
- $2^2, 3^2, 5$
- $2, 2, 3^2, 5$

So up to isomorphism there are only four abelian groups of order 180. ♢

6.3 Rational canonical form

Let F be a field. We consider the PID $F[x]$ (which is also an Euclidean domain). Any $F[x]$-module is automatically a F-vector space. We consider a $F[x]$-module V which is finite dimensional over F. Note that V is both a F-module and a $F[x]$-module, but considering it over F gives us the notion of dimension.

Fix a F-basis of V. We consider $V \cong F^n$ as F-modules where $\dim_F V = n$. Then as an $F[x]$-module, we consider x acting on V as the matrix $T \in \mathrm{Mat}_n(F)$. If we choose a different basis, then the matrix representation for x by the usual change of basis gTg^{-1} where g is the transition matrix.

Then since V is finite dimensional over F, it is finitely generated as a both a F and a $F[x]$-module. By the structure theorem, we have

$$V \cong F[x]/(f_1(x)) \oplus \cdots \oplus F[x]/(f_r(x))$$

(We do not include $(F[x])^k$ since that is infinite dimensional) with $f_1(x) \mid f_2(x) \mid \ldots$ where the f_i are not zero or units. In this section we will explore how to compute the invariant factors.

Definition 6.2 (Eigenvalues and eigenvectors)**.** A scalar $\lambda \in F$ is called an *eigenvalue* of T if there exists a non-zero $v \in V$ such that $Tv = \lambda v$. The non-zero vector v is called the *eigenvector* of T with eigenvalue λ. We call $\ker(\lambda I - T)$ the *eigenspace* of T, where I is the identity. □

Note that we use no notions of matrices in the above definition. We can just as easily replace T with x and think about it in terms of actions.

We recall some very quick facts from linear algebra.

Lemma 6.13. *Let $T \in \mathrm{Mat}_{n,m}(F)$. Then λ is an eigenvalue of T iff $\ker(\lambda I - T) \neq \{0\}$*

Lemma 6.14. *Let $T \in \mathrm{Mat}_{n,m}(F)$. Then $\ker(\lambda I - T) \neq \{0\}$ iff $\det(\lambda I - T) = 0$.*

Proof. Recall for a F-linear transformation $\lambda I - T : V \to V$ is an isomorphism iff it is injective iff it is surjective. ■

Definition 6.3 (Characteristic polynomial)**.** Let x be a formal variable. Then $c_T(x) = \det(xI - F) \in F[x]$ is called the *characteristic polynomial* of T. □

Lemma 6.15. *The characteristic polynomial is independent from the choice of the F-basis.*

Proof. Let g be a transition matrix. $c_{gTg^{-1}}(x) = \det(xI - gTg^{-1}) = \det(g(xI - T)g^{-1}) = (\det g)c_T(x)(\det g^{-1})$. ■

Theorem 6.16. *Let $xI - T \in \mathrm{Mat}_n(F[x])$ where $T \in \mathrm{Mat}_n(F) \subset \mathrm{Mat}_n(F[x])$. Let V be the $F[x]$-module where x acts as the matrix T with respect to a chosen F-basis $\{e_i\}_{i=1}^n$. Then $V \cong \mathrm{coker}(xI - T)$ as $F[x]$-modules.*

Proof. Let $\{\tilde{e}_i\}_{i=1}^n$ be the $F[x]$-basis of $F[x]^n$. Define $\pi : F[x]^n \to V$ by $\tilde{e}_i \mapsto e_i$. Recall that since $F[x]^n$ is a free $F[x]$-module, this defines an unique $F[x]$-module homomorphism $x\tilde{e}_i \mapsto xe_i$.

First, $\mathrm{img}(xI - T) \subseteq \ker \pi$, or equivalently, $\pi \circ (xI - T) = 0$. It suffices to show that $(\pi \circ (xI - T))(\tilde{e}_i) = 0$ for all $\tilde{e}_i$. Assume $T = (a_{ij})$. Then

$$\begin{aligned} \pi((xI - T)(\tilde{e}_j)) &= \pi(x\tilde{e}_j - \sum_{i=1}^n a_{ij}\tilde{e}_i) \\ &= Te_j - \sum_{i=1}^n a_{ij}e_i \\ &= 0. \end{aligned}$$

This means that from the isomorphism $F[x]^n / \ker \pi \to V$, we can have an surjective induced map $\mathrm{coker}(xI - T) = F[x]^n / \mathrm{img}(xI - T) \to V$.

Next, we claim that the F-dimension of $\mathrm{coker}(xI - T) \leq n = \dim V$. The result follows directly from this claim. We do this by showing $\sum_{j=1}^n b_j\tilde{e}_j + \mathrm{img}(xI - T) = \mathrm{coker}(xI - T)$ for $b_j \in F$. Consider the following computation:

$$\begin{aligned} x\tilde{e}_j + \mathrm{img}(xI - T) &= T\tilde{e}_j + (xI - T)\tilde{e}_j + \mathrm{img}(xI - T) \\ &= T\tilde{e}_j + \mathrm{img}(xI - T) \\ x^2\tilde{e}_j + \mathrm{img}(xI - T) &= x(T\tilde{e}_j + \mathrm{img}(xI - T)) \\ &= T^2\tilde{e}_j + \mathrm{img}(xI - T) \\ &\vdots \end{aligned}$$

So $\sum_{j=1}^n m_j(x)\tilde{e}_j + \mathrm{img}(xI - T) = \sum_{j=1}^n m_j(T)\tilde{e}_j + \mathrm{img}(xI - T)$. ■

Very often we want to choose a basis of V such that T looks nicer. Let $V \cong F[x]/(f(x))$ where $f(x) = \sum_{i=0}^n a_i x^i$ (we often assume that $f(x)$ is monic). Choose the F-basis for V as $\{1, \overline{x}, \overline{x}^2, \ldots, \overline{x}^{n-1}\}$. Then, x acts as the matrix

$$T = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & -a_{k-1} \end{pmatrix}$$

with respect to this basis. Further computation reveals that $\det(xI - T) = f(x)$.

Definition 6.4 (Rational canonical form)**.** Let V be a finite dimensional $F[x]$-module where x acts as $T \in \mathrm{Mat}_n(F)$ with respect to a chosen F-basis. Then we have $V \cong \bigoplus_{i=1}^r F[x]/(f_i(x))$.

The *rational canonical form* of T is of the form

$$\begin{pmatrix} T_1 & & & \\ & T_2 & & \\ & & \ddots & \\ & & & T_r \end{pmatrix}, \text{ where } T_i \text{ takes the form } \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & -a_{k-1} \end{pmatrix}$$

□

Example 6.7. A more detailed computation. Consider

$$T = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{C}).$$

To find the rational canonical form, we find the Smith normal form of $xI - T$.

$$\begin{aligned} \begin{pmatrix} x-2 & -1 & -1 \\ 0 & x-2 & -1 \\ 0 & 0 & x-3 \end{pmatrix} &\xrightarrow{C_1 \leftrightarrow C_2} \begin{pmatrix} -1 & x-2 & -1 \\ x-2 & 0 & -1 \\ 0 & 0 & x-3 \end{pmatrix} \\ &\xrightarrow{C_2 = C_2 + (x-2)C_1} \begin{pmatrix} -1 & 0 & 0 \\ x-2 & (x-2)^2 & 1-x \\ 0 & 0 & x-3 \end{pmatrix} \\ &\xrightarrow{R_2 = R_2 + (x-2)R_1} \begin{pmatrix} -1 & 0 & 0 \\ 0 & (x-2)^2 & 1-x \\ 0 & 0 & x-3 \end{pmatrix} \\ &\xrightarrow{R_2 = R_2 + R_3} \begin{pmatrix} -1 & 0 & 0 \\ 0 & (x-2)^2 & -2 \\ 0 & 0 & x-3 \end{pmatrix} \\ &\to \cdots \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (x-3)(x-2)^2 \end{pmatrix} \end{aligned}$$

So $V \cong \mathrm{coker}(xI - T) \cong \mathbb{C}[x]/(x^3 - 7x^2 + 16x - 12)$, and the rational canonical form of T is

$$\begin{pmatrix} 0 & 0 & 12 \\ 1 & 0 & -16 \\ 0 & 1 & 7 \end{pmatrix}.$$

♢

Definition 6.5. We say A is similar to B if $A = gBg^{-1}$ for $g \in \mathbf{GL}_n(F) = (\mathrm{End}_F(F^n))^*$ (invertible matrices). □

Lemma 6.17. *Let $T, S \in \mathrm{Mat}_n(F)$.*

i. The rational canonical form of T is unique.

ii. T is similar to its rational canonical form.

iii. T is similar to S iff they have the same rational canonical form.

Proof.

i. Follows from the uniqueness of invariant factors.

ii. The rational canonical form is obtained by a change of basis.

iii. Again, just a change of basis.

■

Definition 6.6 (Minimal polynomials)**.** Let $T \in \mathrm{Mat}_n(F)$. We consider the $F[x]$-module $V = F^n$ where x acts as T. Then let $(p_T(x)) = \mathrm{Ann}_{F[x]}(V) = \{r \in F[x] \mid rm = 0, m \in V\}$. We define the *minimal polynomial* of T as $m_T(x) = up_T(x)$ for some $u \in F^*$ such that $up_T(x)$ is monic. □

Theorem 6.18. *Let $f(x) \in F[x]$ such that $f(T) = 0 \in \mathrm{Mat}_n(F)$. Then $p_T(x) \mid f(x)$.*

Proof. Let $V = F^n$ be the $F[x]$-module where x acts as T. Then since $f(T) = 0$, we have $f(x) \in \mathrm{Ann}_{F[x]}(V) = (p_T(x))$, so $p_T(x) \mid f(x)$. ■

Theorem 6.19. *Recall we have $V \cong \bigoplus_{i=1}^r F[x]/(f_i(x))$ with $f_1 \mid f_2 \mid \dots$ and $f_i \notin F$. We have*

i. $p_T(x) = f_r(x)$, assuming both are monic.

ii. (Hamilton-Cayley theorem) $p_T(x) \mid c_T(x)$. In other words, $c_T(T) = 0$.

Proof.

i. We have $\mathrm{Ann}_{F[x]}\left(\bigoplus F[x]/(f_i(x))\right) = (f_r(x))$ due to the division criteria $f_1 \mid f_2 \mid \dots$. Therefore $f_r(x) = p_T(x)$ if they are both monic.

ii. Recall $\det(xI - T) = f_1(x) \cdots f_r(x)$. Then $p_T(x) = f_r(x) \mid c_T(x)$, so by definition $c_T(T) = 0$.

■

Theorem 6.20. *Let F be any field. Let $A \in \mathrm{Mat}_N(F)$ be such that $A^2 = A$. Then A is similar to a diagonal matrix which has only 0's and 1'a along the diagonal.*

Proof. Let $m_A(x)$ be the minimal polynomial of A. We know that $A^2 - A = 0$. This implies that $m_A(x) \mid x(x-1)$.

Let V be the $F[x]$-module where x acts as the matrix A. Then $V \cong \bigoplus_{i=1}^r F[x]/(f_i(x))$, where $f_1 \mid \cdots \mid f_r = m_A(x) \mid x(x-1)$ so we can write $V \cong F[x]/(x) \oplus \cdots F[x]/(x) \oplus F[x]/(x-1) \oplus \cdots F[x]/(x-1)$. The rational canonical form of A is thus just a diagonal matrix with only 0's and 1'. ■

Theorem 6.21. *Let A be a strictly triangular matrix (entries along the diagonal are 0). Then A is nilpotent.*

Proof. Consider the characteristic polynomial $c_A(x) = \det(xI - A) = x^n$. Then since $m_A(x) \mid c_A(x)$, we have $c_A(A) = A^n = 0$. ■

6.4 Jordan canonical form

The invariant factors form of the structure theorem gives rise to the rational canonical form. We shall now see that the elementary factors form will give rise to the Jordan canonical form.

Definition 6.7 (Algebraically closed fields)**.** Let F be a field. We say that F is *algebraically closed* if any non-constant $f(x) \in F[x]$ has a root in F. □

Example 6.8. $\mathbb{R}$ is not algebraically closed, whereas $\mathbb{C}$ is algebraically closed. ◊

Theorem 6.22. *Let F be an algebraically closed field.*

i. F is infinite.

ii. Let $f(x) \in F[x]$ be irreducible. Then $f(x) = (x - \lambda)u$ for some $\lambda \in F$ and an unit $u \in F^*$. Hence every irreducible polynomial is of degree 1.

Proof.

i. Suppose that F is finite and assume $F = \{a_1, a_2, \dots, a_n\}$. Then $f(x) = (x - a_1)\cdots(x - a_n) + 1$ has no roots in F.

ii. Let $f(x)$ be non-constant. Let λ be a root. Then $(x - \lambda) \mid f(x)$, or in other words $f(x) = p(x)(x - \lambda)$ for some $p(x) \in F[x]$. Then $p(x)$ has to be an unit.

■

Corollary 6.22.1. *Let F be algebraically closed. Then any non-constant $f(x)$ is a product of linear factors and units.*

Theorem 6.23. *Let F be an algebraically closed field. Then any $T \in \mathrm{Mat}_n(F)$ has an eigenvalue.*

Proof. Consider the characteristic polynomial $c_T(x) = \det(xI - T)$. Let λ be a root. Since F is algebraically closed we can always find such a λ. Then $\det(\lambda I - T) = 0$ which means λ is an eigenvalue of T. ■

Let F be an algebraically closed field. Let $V \cong F[x]/(p(x))^n$ for some irreducible $p(x) = x - a$. Then we can choose a F-basis $\{1, x - a, (x - a)^2, \dots, (x - a)^{n-1}\}$. Then x acts as the matrix

$$T = \begin{pmatrix} a & 1 & 0 & \cdots & 0 \\ 0 & a & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \\ \vdots & & \ddots & \ddots & 1 \\ 0 & & & 0 & a \end{pmatrix}$$

in the given basis. We can also see that $\det(xI - T) = (x - a)^n$.

Definition 6.8 (Jordan canonical form)**.** Let F be an algebraically closed field. Let $T \in \mathrm{Mat}_n(F)$. Let $V = F^n$ be a $F[x]$-module where x acts as T. Then $V \cong \mathrm{coker}(xI - T) \cong \bigoplus_{i=1}^r F[x]/(p_i^{n_i}) \cong \bigoplus_{i=1}^r F[x]/((x-a_i)^{n_i})$ with not necessarily distinct $a_i \in F$. Define the *Jordan canonical form* of T as

$$\begin{pmatrix} T_1 & & & \\ & T_2 & & \\ & & \ddots & \\ & & & T_r \end{pmatrix}, \text{ where } T_i \text{ takes the form } T_i = \begin{pmatrix} a_i & 1 & 0 & \cdots & 0 \\ 0 & a_i & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \\ \vdots & & \ddots & \ddots & 1 \\ 0 & & & 0 & a_i \end{pmatrix}$$

□

This is only well-defined up to a permutation of blocks, which is the same as a change of basis. Similar to last time, we have a few easy facts:

Theorem 6.24. *Let F be an algebraically closed field. Let $T, S \in \mathrm{Mat}_n(F)$.*

i. T is similar to its Jordan canonical form.

ii. T is similar to S iff T and S have the same Jordan canonical form up to permutation.

Example 6.9. Refer to example 6.7. We determined the Smith normal form of

$$T = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{C}).$$

to be

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (x-3)(x-2)^2 \end{pmatrix}.$$

Therefore $V \cong \mathrm{coker}\, A \cong \mathbb{C}[x]/(x-3)(x-2)^2$. The Jordan canonical form of T is therefore

$$\begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

◊

Example 6.10. Consider the $\mathbb{Q}[x]$-module $V \cong \mathbb{Q}^3$ where x acts as

$$T = \begin{pmatrix} 2 & -2 & 14 \\ 0 & 3 & -7 \\ 0 & 0 & 2 \end{pmatrix}.$$

Then $V \cong \mathrm{coker}(xI - T)$. We can convert $xI - T$ to the Smith normal form

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & x-2 & 0 \\ 0 & 0 & (x-3)(x-2) \end{pmatrix}$$

So the $\mathbb{C}[x]$ module $V \cong \mathbb{C}^3$ where x acts as T has the following isomorphism $V \cong \mathbb{C}[x]/(1) \oplus \mathbb{C}[x]/(x-2) \oplus \mathbb{C}[x]/(x-3)(x-2)$. By the Chinese remainder theorem, we then have $V \cong \mathbb{C}[x]/(x-2) \oplus \mathbb{C}[x]/(x-2) \oplus \mathbb{C}[x]/(x-3)$. The Jordan canonical form of T is therefore

$$J = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 2 & 3 \end{pmatrix}.$$

A matrix T like this is also called *diagonalisable* (in $\mathbb{C}$), since the Jordan canonical form is a diagonal matrix. $\Diamond$

Example 6.11. Let

$$T = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{C}).$$

The Smith normal form of $xI - T$ is

$$\begin{pmatrix} 1 & 0 \\ 0 & x^2+1 \end{pmatrix}.$$

Then we have $V \cong \mathbb{C}[x]/(x-i) \oplus \mathbb{C}[x]/(x+i)$. We see that T is diagonalisable in $\mathbb{C}$ but is not diagonalisable in $\mathbb{R}$ (T has no eigenvalues in $\mathbb{R}$). $\Diamond$

The characteristic polynomial is actually insufficient in determining the Jordan canonical form. For example, consider the following two matrices

$$T = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad\qquad S = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

They are obviously not similar. However, $c_T(x) = c_S(x) = x^2$. We still need to go through the computation of the Smith normal form.

Theorem 6.25. *Let F be an algebraically closed field. Let $T \in \mathrm{Mat}_n(F)$ with a Jordan canonical form J. Then T is diagonalisable iff J is diagonal iff the minimal polynomial has no multiple roots.*

Proof. The first equivalence is by definition. The second equivalence arises because the invariant factors are all degree 1. ■

A priori, we do not know that a matrix that already exists in the correct shape is its own canonical form, since the canonical forms were defined in terms of invariant factors and elementary factors. The following easy theorem assures us that they are correct. We can also generalise the calculation block-wise to larger canonical forms.

Theorem 6.26. *Let F be any field. Let*

$$T = \begin{pmatrix} 0 & 0 & -a_0 \\ 1 & 0 & -a_1 \\ 0 & 1 & -a_2 \end{pmatrix} \in \mathrm{Mat}_3(F) \qquad\qquad S = \begin{pmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{pmatrix} \in \mathrm{Mat}_3(F).$$

The rational canonical form of T is just T. The Jordan canonical form of S is just S.

Proof. A simple calculation shows that they are correct. ■

The following theorem summarises everything so far:

Theorem 6.27. *Let $A, B \in \mathrm{Mat}_n(F)$ for any field F. Let V, W, be $F[x]$-modules where x acts as A and B respectively. The following statements are equivalent:*

i. A and B are similar in $\mathrm{Mat}_n(F)$.

ii. A and B have the same rational canonical form.

iii. $V \cong W$ as $F[x]$-modules.

iv. V and W as $F[x]$-modules have the same rational factors.

v. V and W as $F[x]$-modules have the same elementary factors.

vi. If F is algebraically closed then A and B have the same Jordan canonical form.

6.5 Applications

Definition 6.9 (Nilpotent matrices)**.** Let F be a field. Let $A \in \mathrm{Mat}_n(F)$ be such that $A^k = 0$ for some $k > 0$. Such a matrix A is called *nilpotent.* □

In general there are infinitely many nilpotent matrices. What we want to do is to classify them in some clever way.

We consider the group of $n \times n$ invertible matrices over F, $\mathbf{GL}_n(F)$, acting on the set $\mathrm{Mat}_n(F)$ with $gA = gAg^{-1}$. So if A is nilpotent, so is the conjugate. Hence instead of classifying the set of nilpotent matrices, we can classify the orbits of nilpotent matrices. There might be infinitely many nilpotent matrices but there are finitely many orbits.

Theorem 6.28. *Let $A \in \mathrm{Mat}_n(F)$ be nilpotent. Then $c_A(x) = x^n$ and $m_T(x) = x^r$ for some $r > 0$.*

Proof. Since A is nilpotent, $A^k = 0$ for some $k > 0$. Let $f(x) = x^k$. We have $f(A) = 0$. So $m_A(x) \mid f(x)$, and $m_A = x^r$ for some $r > 0$. Now let V be the $F[x]$-module where x acts as T. So $V \cong \bigoplus_{i=1}^r F[x]/(f_i(x))$, with $f_1 \mid f_2 \mid \cdots$. Recall that $f_r(x) = m_A(x)$. But every $f_i \mid m_A(x)$. Since $c_A(x)$ is degree n, therefore $c_A(x) = a^n$. ■

Definition 6.10 (Partitions)**.** A partition of a positive integer n is a list of positive integer(s) such that their sum is n. The set of all partitions of n is therefore $\{(a_1, \dots, a_r) \mid \sum_i a_i = n, a_1 \geq \cdots \geq a_r\}$. □

Theorem 6.29. *The $\mathbf{GL}_n(F)$-orbits of nilpotent matrices in $\mathrm{Mat}_n(F)$ is in bijection with the set of partitions of n.*

Proof. Two matrices are similar iff they have the same rational canonical form. So for each $\mathbf{GL}_n(F)$-orbits of $\mathrm{Mat}_n(F)$, we can pick out a representative in rational canonical form. For a matrix in rational canonical form to be nilpotent, we have $c_A(x) = f_1(x) \dots f_r(x) = x^n$ where all $f_i = x^{a_i}$ and $a_1 \geq a_2 \cdots$, and this gives us a partition $n = \sum_i a_i$. Similarly, given a partition we can construct the rational canonical form. ■

Example 6.12. There are three orbits of nilpotent matrices in $\mathrm{Mat}_3(F)$. They are the orbits of

$$\begin{pmatrix}0&0&0\\0&0&0\\0&0&0\end{pmatrix} \qquad \begin{pmatrix}0&0&0\\1&0&0\\0&0&0\end{pmatrix} \qquad \begin{pmatrix}0&0&0\\1&0&0\\0&1&0\end{pmatrix}.$$

$\Diamond$

Next, we consider the conjugacy classes in $\mathbf{GL}_2(F_2)$ where $F_2 \cong \mathbb{Z}/2\mathbb{Z}$ is the field with three elements.

Theorem 6.30. *We consider the conjugation action of* $\mathbf{GL}_2(F_2)$ *on* $\mathrm{Mat}_2(F_2)$. *There are six orbits of this conjugation action.*

Proof. For any $A \in \mathrm{Mat}_2(F_2)$, we have the decomposition of the $F[x]$-module where x acts as A, $V \cong F[x]/(f_1(x)) \oplus F[x]/(f_2(x))$ where $f_1 \mid f_2$. It suffices to consider the invariant factors, since two matrices are similar iff they have the same rational canonical form.

We consider irreducible polynomials of degree at most 2:

$$x \qquad x-1=x+1 \qquad x^2+x+1.$$

So we can have either

- $f_1(x) = f_2(x) = x$
- $f_1(x) = f_2(x) = x-1$

(due to the need for $f_1 \mid f_2$), or we can have $V \cong F[x]/m_A(x)$ for

- $m_A(x) = (x-1)^2$
- $m_A(x) = (x-1)x$
- $m_A(x) = x^2+x+1$
- $m_A(x) = x^2$

■

Corollary 6.30.1. *There are three conjugacy classes of* $\mathbf{GL}_2(F_2)$.

Proof. We only need to determine which of the rational canonical forms are invertible.

- $f_1(x) = f_2(x) = x$ gives a rational canonical form of

  $$\begin{pmatrix}0&0\\0&0\end{pmatrix}$$

  which is not invertible.

- $f_1(x) = f_2(x) = x-1$ gives a rational canonical form of

  $$\begin{pmatrix}1&0\\0&1\end{pmatrix}$$

  which is invertible.

- $m_A(x) = (x-1)^2$ gives a rational canonical form of

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

which is invertible.

- $m_A(x) = (x-1)x$ gives a rational canonical form of

$$\begin{pmatrix} 0 & 0 \\ 1 & -1 \end{pmatrix}$$

which is not invertible.

- $m_A(x) = x^2 + x + 1$ gives a rational canonical form of

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

which is invertible.

- $m_A(x) = x^2$ gives a rational canonical form of

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

which is not invertible.

■

Next, we consider the orbits of $\mathbf{GL}_3(\mathbb{Q})$ on the set $A = \{g \in \mathbf{GL}_3(\mathbb{Q}) \mid g^6 = \mathrm{id}\}$. Note that the conjugation action is well defined. If $g^6 = \mathrm{id}$, then $(hgh^{-1})^6 = \mathrm{id}$ as well.

Theorem 6.31. *There are eight orbits of the $\mathbf{GL}_3(\mathbb{Q})$ conjugation action on the set $A = \{g \in \mathbf{GL}_3(\mathbb{Q}) \mid g^6 = \mathrm{id}\}$.*

Proof. Let $f(x) = x^6 - 1 = (x-1)(x^2+x+1)(x+1)(x^2-x+1)$. Then for any $B \in A$, we have $f(B) = 0$. So $m_B(x) \mid f(x)$. On the other hand, we know $\deg m_B(x) \leq 3$. We will go through all the possibilities for $m_B(x)$. This gives us f_r in the usual decomposition $\bigoplus_{i=1}^r F[x]/(f_i(x))$, where $f_1 \mid f_2 \mid \cdots \mid f_r$. This will allow us to rule out certain cases.

- $m_B(x) = x - 1$
- $m_B(x) = x + 1$
- $m_B(x) = x^2 - x + 1$. This is impossible since it is irreducible and we want a degree 1 polynomial to be a factor.
- $m_B(x) = x^2 + x + 1$. This is also impossible.
- $m_B(x) = (x-1)(x+1)$. This gives rise to two possibilities, with $f_1 = (x \pm 1)$.
- $m_B(x) = (x-1)(x^2 - x + 1)$
- $m_B(x) = (x-1)(x^2 + x + 1)$
- $m_B(x) = (x+1)(x^2 - x + 1)$
- $m_B(x) = (x+1)(x^2 + x + 1)$

■

7 Tensor product of modules

We start with a motivating example.

Example 7.1. Let $V, W \cong \mathbb{R}^3$ be two vector spaces. We know

$$V \oplus W = \{(a,b,c,d,e,f) \mid (a,b,c) \in V, (d,e,f) \in W\}.$$

This gives us a new vector space of dimension $6 = 3+3$. Naively we can ask for a construction that would give us a vector space of dimension $9 = 3 \times 3$. Consider the following operation

$$V \otimes_{\mathbb{R}} W = \left\{ \begin{pmatrix} ad & ae & af \\ bd & be & bf \\ cd & ce & cf \end{pmatrix} \mid (a,b,c) \in V, (d,e,f) \in W \right\}.$$

Let $k \in \mathbb{R}$, $v, v_1, v_2 \in V$, $w, w_1, w_2 \in W$. We have the following identities:

i. $(v_1 + v_2) \otimes_{\mathbb{R}} w = v_1 \otimes_{\mathbb{R}} w + v_2 \otimes_{\mathbb{R}} w$.

ii. $v \otimes_{\mathbb{R}} (w_1 + w_2) = v \otimes_{\mathbb{R}} w_1 + v \otimes_{\mathbb{R}} w_2$.

iii. $(kv) \otimes_{\mathbb{R}} w = v \otimes_{\mathbb{R}} (kw) = k(v \otimes_{\mathbb{R}} w)$.

◊

We will be using these three identities as axioms to define the tensor product $\otimes$ more generally.

Theorem 7.1. *Let $F(V \times W)$ be the free abelian group of infinite rank over the set $V \times W$. Let $k \in \mathbb{R}$, $v, v_1, v_2 \in V$, $w, w_1, w_2 \in W$. Let A be the subgroup of $F(V \times W)$ generated by*

$$\{(v_1 + v_2, w) - (v_1, w) - (v_2, w), (v, w_1 + w_2) - (v, w_1) - (v, w_2), (kv, w) - (v, kw)\}$$

Then the natural map $F(V \times W)/A \to V \otimes_{\mathbb{R}} W$ given by $(v,w) \mapsto v \otimes_{\mathbb{R}} w$ is an isomorphism of abelian groups.

Proof. We consider the map of abelian groups $F(V \times W) \to V \otimes_{\mathbb{R}} W$ given by $(v,w) \mapsto v \otimes_{\mathbb{R}} w$. This is well-defined, since $F(V \times W)$ is free with basis $(v,w) \in V \times W$. This is clearly surjective. We want to show that the kernel is A.

We define a R-module structure on $F(V \times W)/A$ as $k(v,w) = (kv,w) = (v,kw)$ for $(v,w) \in F(V \times W)/A$. ■

We look at another example. Recall that if $S \subseteq R$ is a subring with 1, then any R-module M is automatically a S-module via restriction. However, if N is a S-module, it is not so obvious how to make N into a R-module such that the restriction from the R-module N is again the same module N.

In general this is not possible. Let $\mathbb{Z} \subset \mathbb{Q}$. Let $M = \mathbb{Z}/2\mathbb{Z}$ be a $\mathbb{Z}$-module. Then M cannot be a $\mathbb{Q}$-module, since any $\mathbb{Q}$-module is a $\mathbb{Q}$-vector space, and there is no $\mathbb{Q}$-vector space with cardinality 2.

Hence, we find the next best thing.

Theorem 7.2 (Universal property of tensor products)**.** *Let $S \subseteq R$ be a subring of R containing 1. Let N be a S-module. We define the R-module $R \otimes_S N$ together with a S-module homomorphism $\iota : N \to R \otimes_S N$ via the following universal property: for any R-module M considered as a S-module, any S-module homomorphism $\varphi : N \to M$ there is an unique R-module homomorphism $\overline{\varphi} : R \otimes_S N \to M$ such that the following diagram commutes:*

$$\begin{tikzcd}
N \arrow[r, "\iota"] \arrow[dr, "\varphi"'] & R \otimes_S N \arrow[d, dashed, "\overline{\varphi}"] \\
 & M
\end{tikzcd}$$

We claim that $R \otimes_S N$ exists and is unique up to isomorphism of R-modules.

Proof. The uniqueness is standard. To prove the existence, start with the free abelian group $F(R \times N)$. Define the subgroup A to be generated by

- $(r_1 + r_2, n) - (r_1, n) - (r_2, n)$,
- $(r, n_1 + n_2) - (r, n_1) - (r, n_2)$,
- $(rs, n) - (r, sn)$

We claim that $F(R \times N)/A$ with the embedding $N \to F(R \times N)/A$ given by $n \mapsto (1, n)$ is the R-module we want.

We check the diagram. For any $\varphi : N \to M$, we have a map $F(R \times N) \to M$ given by $(r, n) \mapsto r\varphi(n)$.

$$\begin{tikzcd}
n \in N \arrow[r, mapsto] \arrow[d, "\varphi"'] & (1,n) \in F(R \times N) \arrow[r, mapsto] \arrow[dl, mapsto] & (1,n) \in F(R \times N)/A \arrow[dll, dashed, mapsto] \\
1 \cdot \varphi(n) \in M & &
\end{tikzcd}$$

To check that the map factors through $F(R \times N)/A$, we need to check that A maps to 0 in M. We just check each of the generators, take for example the third one

$$(rs, n) - (r, sn) \mapsto rs\varphi(n) - r\varphi(sn) = rs\varphi(n) - rs\varphi(n) = 0.$$

It is straightforward to check that $\overline{\varphi}$ is the unique R-module map to make everything commute. For the R-module structure on $F(R \times N)/A$, we define $r(r', n) = (rr', n)$. Checking that this is indeed a R-module structure is skipped. ■

We write $r \otimes_S n$ as the element $r(\iota(n)) = (r, n) \in F(R \times N)/A$. If the ring S is understood, we write $\otimes_S$ as $\otimes$.

Example 7.2. Let $Z \subset \mathbb{Q}$ and $M = \mathbb{Z}/2\mathbb{Z}$ be a Z-module. Then for any $\mathbb{Q}$-module V and $\mathbb{Z}$-module homomorphism $N \to V$, there is a corresponding map $\mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Z}/2\mathbb{Z} \to V$. But any homomorphism $N \to V$ is the zero map since V is a vector space and is torsion free as a Z-module. This means that $Q \otimes_{\mathbb{Z}} \mathbb{Z}/2\mathbb{Z}$ is just 0.

We can also consider $\mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Z}/2\mathbb{Z}$ concretely using the construction. We have $r \otimes_{\mathbb{Z}} n = \frac{1}{2}r2 \otimes_{\mathbb{Z}} n = \frac{1}{2}r \otimes_{\mathbb{Z}} 2n = \frac{1}{2}r \otimes_{\mathbb{Z}} 0 = 0$. ♢

Example 7.3. For any R-module N, we have $R \otimes_R N \cong N$. This arises from the universal property, since $\mathrm{id} : N \to N$ satisfies it as well. ◊

Example 7.4. For any $\mathbb{R}$-vector space $V \cong \mathbb{R}^n$, then $\mathbb{C} \otimes_{\mathbb{R}} V \cong \mathbb{C}^n$. ◊

What follows is a more general definition without the vector space structure.

Definition 7.1 (Balanced maps)**.** Let M be a right R-module and N be a left R-module. For any abelian group L, a map $M \times N \to R$ is said to be *R-balanced* if

- $\varphi(m, n_1 + n_2) = \varphi(m, n_1) + \varphi(m, n_2)$
- $\varphi(m_1 + m_2, n) = \varphi(m_1, n) + \varphi(m_2, n)$
- $\varphi(mr, n) = \varphi(m, rn)$. □

Theorem 7.3. *Let M be a right R-module and N be a left R-module. The tensor product $M \otimes_R N$ is defined as an abelian group together with a R-balanced map $\iota : M \times N \to M \otimes_R N$ such that for any R-balanced map φ, there exists an unique map $\overline{\varphi} : M \otimes_R N \to L$ such that the following diagram commutes:*

$$\begin{tikzcd}
M \times N \arrow[r, "\iota"] \arrow[d, "\varphi"'] & M \otimes_R N \arrow[dl, dashed, "\overline{\varphi}"] \\
L &
\end{tikzcd}$$

Then $M \otimes_R N$ exists and is unique up to isomorphism.

Proof. As usual, we define $M \otimes_R N$ as the free abelian group $F(M \times N)$ modulo the relations

- $(m_1 + m_2, n) - (m_1, n) - (m_2, n)$
- $(m, n_1 + n_2) - (m, n_1) - (m, n_2)$
- $(mr, n) - (m, rn)$ for $r \in R$.

We define $\iota : M \times N \to M \otimes_R N$ as $(m, n) \mapsto m \otimes_R n$. We can check that this is R-balanced, and for any R-balanced map $\varphi : M \times N \to L$, we can define $M \otimes_R N \to L$ by $m \otimes_R n \mapsto \varphi(m, n)$. It remains to be shown that everything is well-defined and satisfies the diagram (skipped). ■

Theorem 7.4. *Let M be a right R-module and N be a left R-module. Let L be an R-module. Then we have a bijection*

$$\mathrm{Hom}_{\mathbb{Z}}(M, \mathrm{Hom}_R(N, L)) \to \{R\text{-balanced maps } M \times N \to L\} \to \mathrm{Hom}_{\mathbb{Z}}(M \otimes_R N, L).$$

Proof. Let $\varphi \in \mathrm{Hom}_{\mathbb{Z}}(M, \mathrm{Hom}_R(N, L))$. Then we define $\varphi' : M \times N \to L$ by $(m, n) \mapsto \varphi(m)(n)$. Let $\psi : M \times N \to L$ be R-balanced. Then for any $n \in N$, we define $\psi'(\cdot, n) : M \to L$. ■

Definition 7.2 (Bilinear maps)**.** Let R be a commutative ring with 1. Let M, N, L be R-modules. A map of sets $\varphi : M \times N \to L$ is called *R-bilinear* if

- $\varphi(r_1 m_1 + r_2 m_2, n) = r_1 \varphi(m_1, n) + r_2 \varphi(m_2, n)$
- $\varphi(m, r_1 n_2 + r_2 n_2) = r_1 \varphi(m, n_1) + r_2 \varphi(m, n_2)$.

- $\varphi(m, rn) = r\varphi(m, n) = \varphi(rm, n)$.

It is automatically balanced. □

Theorem 7.5. *Let R be a commutative ring with 1. Let M, N, L be R-modules. Their tensor product $M \otimes_R N$ is a R-module together with a R-bilinear map $\iota : M \times N \to M \otimes_R N$ such that for any R-bilinear map $\varphi : M \times N \to L$, there exists an unique R-module homomorphism $\overline{\varphi} : M \otimes_R N \to L$ such that the following diagram commutes:*

$$\begin{tikzcd}
M \times N \arrow[r, "\iota"] \arrow[d, "\varphi"'] & M \otimes_R N \arrow[dl, dashed, "\overline{\varphi}"] \\
L &
\end{tikzcd}$$

Proof. Same as before (with a small modification to the relation), let $M \otimes_R N = F(M \times N)/\sim$ where $\sim$ is the relation

- $(m_1 + m_2, n) - (m_1, n) - (m_2, n)$
- $(m, n_1 + n_2) - (m, n_1) - (m, n_2)$
- $(rm, n) - (m, rn)$

Then we equip $M \otimes_R N$ with a R-module structure as $r(m \otimes_R n) = (rm) \otimes_R n = m \otimes_R (rn)$. Checking that this is a well-defined R-module is skipped. Next, for any $\varphi : M \times N \to L$, we can define $\overline{\varphi} : M \otimes_R N \to L$ by $m \otimes_R n \mapsto \varphi(m, n)$. Checking that this is a R-module homomorphism is also left out. ■

Example 7.5. The standard bilinear form $\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ given by the dot product $v, w \mapsto vw^T$ is $\mathbb{R}$-bilinear. We have the following diagram

$$\begin{tikzcd}
\mathbb{R}^3 \times \mathbb{R}^3 \arrow[r, "\iota"] \arrow[d, "\langle \cdot \cdot \rangle"'] & \mathbb{R}^3 \otimes_{\mathbb{R}} \mathbb{R}^3 \cong \mathrm{Mat}_3(\mathbb{R}) \arrow[dl, dashed, "\mathrm{tr}"] \\
\mathbb{R} &
\end{tikzcd}$$

where tr is the trace operation. ◊

Definition 7.3 (Bimodules)**.** Let S and R be commutative rings with 1. An *S-R-bimodule* M is an abelian group that is a left S-module and a right R-module simultaneously, such that

$$(sm)r = s(mr) = smr$$

for $r \in R$, $s \in S$, and $m \in M$. □

Example 7.6. Some simple examples:

- R is a R-R-bimodule via left and right multiplication.
- If M is a right R-module, then M is a $\mathbb{Z}$-R-bimodule.
- Let R be commutative. Then any R-module M is a R-R-bimodule.

◊

Theorem 7.6. *Let R and S be rings with 1. Let M be a S-R-bimodule and L be a S-module and N be a left R-module. Then the tensor product $M \otimes_R N$ is a S-module together with a R-balanced map $\iota : M \times N \to M \otimes_R N$ where $\iota(sm, n) = s(\iota(m, n))$, such that for any R-balanced map $\varphi : M \times N \to L$ where $\varphi(sm, n) = s(\varphi(m, n))$, we have an unique S-module map $\overline{\varphi} : M \otimes_R N \to L$ such that the following diagram commutes:*

$$\begin{tikzcd}
M \times N \arrow[r, "\iota"] \arrow[d, "\varphi"'] & M \otimes_R N \arrow[dl, dashed, "\overline{\varphi}"] \\
L &
\end{tikzcd}$$

Then $M \otimes_R N$ exists and is unique up to isomorphism.

Proof. As usual, we define the abelian group $M \otimes_R N = F(M \times N)/\sim$. Then we define a S-module structure on $M \otimes_R N$ using the S-module structure on M as $s(m \otimes_R n) = (sm) \otimes_R n$. Then we check everything is well-defined (skipped). Next, we define $\overline{\varphi} : M \otimes_R N \to L$ as $m \otimes_R n \mapsto \varphi(m, n)$. We check that this is a S-module homomorphism (skipped). ■

Example 7.7. Some more examples.

- Consider $\mathbb{Z}$-modules $\mathbb{Z}/2\mathbb{Z}$ and $\mathbb{Z}/3\mathbb{Z}$. We claim $\mathbb{Z}/2\mathbb{Z} \otimes \mathbb{Z}/3\mathbb{Z} \cong 0$. Note that 4 acts as the identity on $\mathbb{Z}/3\mathbb{Z}$. Then for any $m \in \mathbb{Z}/2\mathbb{Z}$ and $n \in \mathbb{Z}/3\mathbb{Z}$, we have $m \otimes n = m \otimes 4n = m4 \otimes n = 0$.
- Consider $\mathbb{Z}$-modules $\mathbb{Z}/2\mathbb{Z}$ and $\mathbb{Z}/2\mathbb{Z}$. Then $\mathbb{Z}/2\mathbb{Z}\otimes\mathbb{Z}/2\mathbb{Z} \cong \mathbb{Z}/2\mathbb{Z}$. Consider the $\mathbb{Z}$-bilinear map $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \to \mathbb{Z}/2\mathbb{Z}$ given by $(a, b) \mapsto ab$. By definition, we have a $\mathbb{Z}$-module homomorphism $\mathbb{Z}/2\mathbb{Z} \otimes \mathbb{Z}/2\mathbb{Z} \to \mathbb{Z}/2\mathbb{Z}$ such that $a \otimes b \mapsto ab$. This is surjective. On the left hand side, we have the elements

  $$1 \otimes 1 \qquad\qquad 1 \otimes 0 = 0 \otimes 1 = 0 \otimes 0.$$

  Then, an argument about cardinalities mean that there must be exactly two elements.

◊

7.1 Properties of tensor products

Theorem 7.7. *Let M be a right R-module and N be a left-R module. Let S be a right N module and L be a leftS module. Then we have an isomorphism of abelian groups $(M \otimes_R N) \otimes_S L \cong M \otimes_R (N \otimes_S L)$.*

Theorem 7.8. *Let M_1 and M_2 be right R-modules. Let N_1 and N_2 be left R-modules. Then we have the following isomorphism of abelian groups $(M_1 \oplus M_2) \otimes_R N_1 \cong (M_1 \otimes_R N_1) \oplus (M_2 \oplus N_1)$ and $M_1 \otimes (N_1 \oplus N_2) \cong (M_1 \otimes_R N_1) \oplus (M_1 \otimes_R N_2)$.*

Example 7.8. Recall that $\mathbb{C}$ is a $\mathbb{R}$-vector space. We have

$$\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C} = \mathbb{R}^2 \otimes_R \mathbb{R}^2 \cong \mathbb{R}^4.$$

But, $\mathbb{C} \otimes_{\mathbb{C}} \mathbb{C} \cong \mathbb{C}$. To see this we consider the $\mathbb{C}$-bilinear map $\mathbb{C} \to \mathbb{C} \to \mathbb{C}$ given by $(a, b) \mapsto ab$. Then this induces a map $\mathbb{C} \otimes_{\mathbb{C}} \mathbb{C} \to \mathbb{C}$ given by $a \otimes_{\mathbb{C}} b \mapsto ab$. This is surjective. To show isomorphism, it suffices to show that the left hand side is at most dimension 1. Indeed, $\mathbb{C} \otimes_{\mathbb{C}} \mathbb{C}$ is $\mathbb{C}$-spanned by $1 \otimes_{\mathbb{C}} 1$, since we have $a \otimes_{\mathbb{C}} b = 1 \otimes_{\mathbb{C}} ab1 = ab(1 \otimes_{\mathbb{C}} 1)$. ◊