MA3201 Algebra II

Thang Pang Ern

These notes are based off **Prof. Emile Okada's** MA3201 Algebra II materials. Additional references are cited in the bibliography.

This set of notes was last updated on February 12, 2026. Please send me an email at `thangpangern@u.nus.edu` if you would like to contribute a nice discussion to the notes or point out a typo.

Contents

Chapter 1

Introduction to Rings

After surviving MA2202 Algebra I, we now move from groups to a richer algebraic structure where two operations coexist and interact: *rings*. While group theory focuses on a single operation and its symmetries, ring theory incorporates both addition and multiplication, allowing us to study familiar objects such as integers, polynomials, matrices, and functions within a unified framework.

Just like groups, rings lie at the heart of modern algebra. They provide the natural language for number theory (MA3265 Elementary Number Theory, MA5202 Algebraic Number Theory, algebraic geometry (MA4273 Algebraic Geometry), representation theory (MA5218 Representation Theory), and many areas of analysis.

1.1 Basic Definitions and Examples

Definition 1.1 (ring). A ring R is a set equipped with two binary operations $+$ and $\cdot$ (called addition and multiplication) satisfying the following axioms:

(i) $(R,+)$ is an Abelian group

(ii) $\cdot$ is associative, i.e. $(a\cdot b)\cdot c = a\cdot(b\cdot c)$ for all $a,b,c\in R$

(iii) the distributive laws hold in R, i.e. for all $a,b,c\in R$, we have

$$(a+b)\cdot c = (a\cdot c)+(b\cdot c) \quad \text{and} \quad a\cdot(b+c) = (a\cdot b)+(a\cdot c)$$

In Definition 1.1, we usually write ab in place of $a\cdot b$ or $a\times b$ for $a,b\in R$. The additive identity of R will always be denoted by 0.

Definition 1.2 (commutative ring). R is commutative if $\cdot$ is commutative.

Definition 1.3 (multiplicative identity). R is said to have a multiplicative identity (or contain a 1) if there exists an element $1 \in R$ such that

$$1 \cdot a = a \cdot 1 = a \quad \text{for all } a \in R.$$

Definition 1.4 (division ring). A ring R with identity 1, where $1 \neq 0$, is a division ring if every non-zero element $a \in R$ has a multiplicative inverse, i.e.

$$\text{for any } 0 \neq a \in R \quad \text{there exists } b \in R \text{ such that } ab = ba = 1.$$

Definition 1.5 (field). A commutative division ring is a field.

Example 1.1 (basic examples and non-examples). Some of the mentioned definitions may look abstract so this is a good juncture where we take a look at some examples and non-examples of rings and fields, as well as the structures in between.

- **(i)** $\mathbb{Z}$ is a ring. In fact it is a commutative ring but not a division ring. Hence, $\mathbb{Z}$ is not a field.
- **(ii)** $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ are fields
- **(iii)** Let n be a positive integer. $\mathbb{Z}/n\mathbb{Z}$ is a commutative ring, but in general, it may not have multiplicative inverses. If $n = p$ for some prime p, then $\mathbb{Z}/p\mathbb{Z}$ is a field.
- **(iv)** $2\mathbb{Z}$ is a commutative ring without identity
- **(v)** The trivial ring $R = \{0\}$ is a commutative ring with identity $1 = 0$

Example 1.2. The following sets can be regarded as rings:

$$\mathcal{P}_n(\mathbb{R}) = \text{the set of polynomials of degree at most } n \text{ with coefficients in } \mathbb{R}$$
$$\mathcal{C}^0(\mathbb{R}) = \text{the set of continuous functions on } \mathbb{R}$$
$$\mathcal{C}^1(\mathbb{R}) = \text{the set of differentiable functions on } \mathbb{R}$$

Example 1.3 (Dummit and Foote p. 267 Question 4). Prove that if R and S are non-zero rings with identity $1 \neq 0$, then $R \times S$ is never a field.

Solution. Let R and S be non-zero rings with identity. Consider the element $(1,0) \in R \times S$. Since $R \neq 0$, we have $1 \neq 0$ in R, so $(1,0) \neq (0,0)$. Similarly, $(0,1) \neq (0,0)$ since $S \neq 0$.

However, in the direct product ring, $(1,0)(0,1) = (0,0)$. Hence, $(1,0)$ and $(0,1)$ are non-zero zero divisors in $R \times S$. In particular, $R \times S$ is not an integral domain. Since every field is an integral domain, it follows that $R \times S$ cannot be a field. □

Definition 1.6 (endomorphism ring). Let V be a vector space over $\mathbb{R}$. Then, the endomorphism ring $\text{End}_{\mathbb{R}}(V)$ is defined as follows:

$$\text{End}_{\mathbb{R}}(V) = \{\varphi : \varphi \in \text{Hom}(\mathbb{R}, \mathbb{R})\}$$

Here, $\mathrm{Hom}(\mathbb{R},\mathbb{R})$ is the set of homomorphisms from $\mathbb{R}$ to itself.

We will formally relook at Definition 1.6 in Definition 4.9.

Example 1.4 (endomorphism ring). The endomorphism ring $\mathrm{End}_{\mathbb{R}}(V)$ of a vector space over $\mathbb{R}$ is an example of a non-commutative ring. To see why, consider

$$\mathbf{A}=\begin{bmatrix}1&1\\0&0\end{bmatrix} \text{ and } \mathbf{B}=\begin{bmatrix}0&0\\0&1\end{bmatrix} \quad \text{so} \quad \mathbf{AB}=\begin{bmatrix}0&1\\0&0\end{bmatrix} \text{ and } \mathbf{BA}=\begin{bmatrix}0&0\\0&0\end{bmatrix}.$$

Example 1.5 (quaternions). Recall from MA2202 that the quaternions Q_8 is an example of a non-commutative ring. In fact, this is historically the first example of such a ring. To see why, note that $i,j,k\in Q_8$ where $ij=k$ but $ji=-k\neq k$.

Proposition 1.1. Let R be a ring. Then, the following hold:

(i) $0a=a0=0$ for all $a\in R$

(ii) $(-a)b=a(-b)=-ab$ for all $a,b\in R$

(iii) $(-a)(-b)=ab$ for all $a,b\in R$

(iv) If R has a multiplicative identity, it is unique and $-a=(-1)a$

Example 1.6 (Dummit and Foote p. 230 Question 1). Let R be a ring with multiplicative identity 1. Show that $(-1)^2=1$ in R.

Solution. Let $x\in R$. Then,

$$x(-1)^2=(-x)(-1)=x \quad \text{and} \quad (-1)^2x=(-1)(-x)=x.$$

This means that whenever $x\in R$ is multiplied by $(-1)^2$, we obtain the same result. By the uniqueness of the multiplicative identity (Definition 1.3), which is 1, we conclude that $x=1$. □

Example 1.7 (quadratic field). Let

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

It is a well-known fact that $\mathbb{Q}(\sqrt{2})$ is a ring. We claim that $\mathbb{Q}(\sqrt{2})$ is a field, i.e. we need to deduce that for any $a+b\sqrt{2}\in\mathbb{Q}(\sqrt{2})$, there exists $c+d\sqrt{2}\in\mathbb{Q}(\sqrt{2})$ such that their product equals to 1. In other words,

$$\frac{1}{a+b\sqrt{2}} \text{ is of the form } c+d\sqrt{2} \quad \text{where } c,d\in\mathbb{Q}.$$

This is simply the process of *rationalising the denominator* that one would recall from secondary school. That is,

$$\frac{1}{a+b\sqrt{2}}=\frac{a-b\sqrt{2}}{a^2-2b^2} \quad \text{so we can choose} \quad c=\frac{a}{a^2-2b^2} \text{ and } d=-\frac{b}{a^2-2b^2}.$$

This is permitted because $a^2 - 2b^2 \neq 0$. To see why, it suffices to justify that $a \neq b\sqrt{2}$. This is clear because a, b are rational numbers and $\sqrt{2}$ is irrational, but the product of a rational and an irrational cannot be rational.

As we have seen from Example 1.1, $2\mathbb{Z}$ is an example of a ring without identity. In Example 1.29, we will encounter another ring without identity. In fact, rings without identity are known as rngs.

Example 1.8 (Dummit and Foote p. 232 Question 20). Let R be the collection of sequences $(a_1, a_2, \ldots)$ of integers $a_1, a_2, \ldots$ where all but finitely many of the a_i are 0. This known as the direct sum of infinitely many copies of $\mathbb{Z}$. Prove that R is a ring under componentwise addition and multiplication which does not have an identity.

Solution. Define, for $a = (a_1, a_2, \ldots), b = (b_1, b_2, \ldots) \in R$,

$$a + b = (a_1 + b_1, a_2 + b_2, \ldots) \quad \text{and} \quad ab = (a_1 b_1, a_2 b_2, \ldots).$$

Then, $a + b \in R$ and $ab \in R$. Next, as addition is component-wise, then associativity and commutativity follow from those in $\mathbb{Z}$, so $(R, +)$ is an Abelian group. Next, multiplication is associative and the distributive laws hold. As such, R is a ring.

We then prove that R has no multiplicative identity. Suppose on the contrary that $e = (e_1, e_2 \ldots) \in R$ is an identity. Let $r^{(k)} \in R$ be the sequence with 1 in the k^{th} coordinate and 0 elsewhere. Then, $er^{(k)} = r^{(k)}$. Comparing the k^{th} coordinate yields $e_k \cdot 1 = 1$, so $e_k = 1$ for all $k \geq 1$. This implies $e = (1, 1, 1, \ldots)$ which is not in R because it has infinitely many non-zero entries. This contradiction shows that no identity exists in R. □

Definition 1.7 (zero divisor). Let R be a ring. A non-zero element $a \in R$ is a zero divisor if there exists a non-zero $b \in R$ such that either $ab = 0$ or $ba = 0$.

Definition 1.8 (unit). Let R be a ring with a multiplicative identity. An element $u \in R$ is a unit if

$$\text{there exists some } v \in R \quad \text{such that} \quad uv = vu = 1.$$

The set of units in R is denoted by R^* or $R^\times$ (we would usually use the latter).

Note that a zero divisor may not be a unit. For example, consider the ring $\mathbb{Z}/6\mathbb{Z}$, where the representative $\overline{2}$ is a zero divisor because $\overline{2} \cdot \overline{3} = \overline{0}$ in $\mathbb{Z}/6\mathbb{Z}$ but it is not a unit because $\overline{2}$ has no multiplicative inverse modulo 6.

Example 1.9 (Dummit and Foote p. 230 Question 2). Let R be a ring with multiplicative identity 1. Prove that if u is a unit in R then so is $-u$.

Solution. Since u is a unit, then there exists $v \in R$ such that $uv = vu = 1$. Then,

$$(-u)(-v) = uv = 1 \quad \text{and} \quad (-v)(-u) = vu = 1$$

so $-u$ is also a unit. □

Example 1.10. The ring $\mathbb{Z}$ has no zero divisors and its only units are ± 1. To see why the second property holds, suppose u is a unit of $\mathbb{Z}$. Then, for any $a \in \mathbb{Z}$, we have $au = ua = 1$. We shall find solutions to $au = 1$ in $\mathbb{Z}$. We have either $(a,u) = (1,1)$ or $(a,u) = (-1,-1)$ so the units of $\mathbb{Z}$ are indeed ± 1.

Proposition 1.2. $R^\times$ is a group under multiplication, referred to as the group of units of R (Definition 1.8).

Example 1.11. The group of units of $\mathbb{Z}/n\mathbb{Z}$ is $(\mathbb{Z}/n\mathbb{Z})^\times$. Recall from MA1100 Basic Discrete Mathematics or MA2202 Algebra I that

$$(\mathbb{Z}/n\mathbb{Z})^\times = \{a \in \mathbb{Z}/n\mathbb{Z} : \gcd(a,n) = 1\}.$$

All elements in $(\mathbb{Z}/n\mathbb{Z})^\times$ are zero divisors. In sum, every non-zero element of $\mathbb{Z}/n\mathbb{Z}$ is either a unit or a zero divisor.

Definition 1.9 (integral domain). A commutative ring with multiplicative identity 1 is an integral domain if it has no zero divisors.

Example 1.12 (examples of integral domains). Again, we take a look at some examples.

(i) $\mathbb{Z}$ is an integral domain

(ii) The ring $\mathbb{Z}[\sqrt{2}] = \left\{a + b\sqrt{2} : a,b \in \mathbb{Z}\right\}$ is an integral domain[1]

(iii) Every field is an integral domain. To see why, suppose $a,b \in F$ for some field F such that $a \neq 0$ and $ab = 0$. Considering $ab = 0$, multiplying both sides by a^{-1} (which is permitted because the multiplicative inverse of a exists by Definition 1.4), we obtain $b = 0$.

Definition 1.10 (Gaussian integer). Let

$\mathbb{Z}[i] = \{a + bi : a,b \in \mathbb{Z}\} \subseteq \mathbb{C}$ denote the set of Gaussian integers (Figure 1.10).

[1] Recall from Example 1.7 that we used round brackets $(\cdot)$, but in this example, we used parentheses $[\cdot]$. In general, square brackets $[\cdot]$ are used to indicate ring adjunction, whereas parentheses $(\cdot)$ are used to indicate field adjunction. Thus, $\mathbb{Z}[\sqrt{2}]$ is a ring and not a field, whereas over a field F and an algebraic element α, one has $F[\alpha] = F(\alpha)$, so the distinction appears.

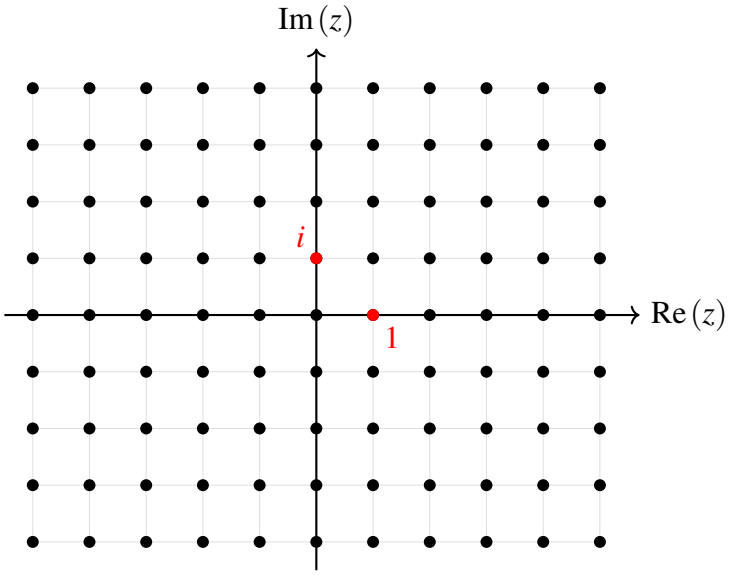


Figure 1.1: The Gaussian integers form the integer lattice in the complex plane

Example 1.13. The Gaussian integers $\mathbb{Z}[i]$ is a commutative ring with identity and its unit elements are $\pm 1, \pm i$. At this stage, we are unable to use any tool to find the unit elements in $\mathbb{Z}[i]$. Having said that, we will learn in Chapter 2.1 that $\mathbb{Z}[i]$ forms what is called a Euclidean domain and we can explot some properties in order to find the units of $\mathbb{Z}[i]$.

Proposition 1.3 (cancellation property). Suppose x, y, z are elements in an integral domain and $xy = xz$. Then, either $x = 0$ or $y = z$.

Proof. We have $x(y-z) = 0$ so $x = 0$ or $y - z = 0$. As we are working in an integral domain (Definition 1.9), it follows that $x = 0$ or $y = z$. □

Proposition 1.4. Any finite integral domain is a field.

Proof. Let R be a finite integral domain and $a \in R$ be non-zero. By the cancellation law, the map $x \mapsto ax$ is injective. To see why, for any $a \neq 0$ and any $x, y \in R$, suppose $ax = ay$, then $a(x-y) = 0$ so $x = y$. Next, since R is finite, this map is surjective, i.e.

$$\text{there exists } b \in R \quad \text{such that} \quad ab = 1, \text{ i.e. } a \text{ is a unit in } R.$$

Since a was an arbitrary non-zero element, then R is a field. □

Example 1.14 (Dummit and Foote p. 231 Question 11). Let R be a ring with multiplicative identity 1. Prove that if R is an integral domain and $x^2 = 1$ for some $x \in R$, then $x = \pm 1$.

Solution. We have $x^2 - 1 = 0$ so $(x+1)(x-1) = 0$. Since R is an integral domain, either $x + 1 = 0$ or $x - 1 = 0$, which follows that $x = \pm 1$. □

Corollary 1.1 (Sadhukhan). If p is a prime, then $\mathbb{Z}/p\mathbb{Z}$ is a field.

We have seen this example in **(iii)** of Example 1.1. One can prove Corollary 1.1 using Euclid's lemma. Moreover, one can prove a stronger result (generally covered in MA3265 Elementary Number Theory too) that $\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime.

Definition 1.11 (subring). A subring of a ring R is a subgroup of R that is closed under multiplication.

Proposition 1.5 (subring criterion). A non-empty subset S of a ring R is a subring if S is closed under subtraction and multiplication, i.e.

$$a, b \in S \quad \text{implies} \quad a - b \in S \text{ and } ab \in S.$$

Example 1.15. We give some examples.

(i) $\mathbb{Z}$ is a subring of $\mathbb{Q}$ and $\mathbb{Q}$ is a subring of $\mathbb{R}$

(ii) $n\mathbb{Z} = \{nk \in \mathbb{Z} : k \in \mathbb{Z}\}$ is a subring of $\mathbb{Z}$

(iii) $\mathbb{Z}[i]$ is a subring of $\mathbb{C}$

Example 1.16 (Dummit and Foote p. 230 Question 3). Let R be a ring with identity and let S be a subring of R containing the identity. Prove that if u is a unit in S then u is a unit in R. Show by example that the converse is false.

Solution. Since $1_R \in S$, then S has a multiplicative identity which is 1 so $1_S = 1_R$. If u is a unit in S, then there exists an element $v \in S$ such that $uv = vu = 1_S = 1_R$. Since $v \in R$, then u is a unit in R as well.

For the second part, note that $\mathbb{Z}$ is a subring of $\mathbb{Q}$. $2 \in \mathbb{Q}$ is a unit since $2 \cdot \frac{1}{2} = 1$. However, 2 is not a unit in $\mathbb{Z}$ (recall from Example 1.10 that the units of $\mathbb{Z}$ are ± 1). □

Example 1.17 (Dummit and Foote p. 230 Question 5). Decide which of the following **(a)**-**(f)** are subrings of $\mathbb{Q}$:

(a) the set of all rational numbers with odd denominators (when written in lowest terms)

(b) the set of all rational numbers with even denominators (when written in lowest terms)

(c) the set of non-negative rational numbers

(d) the set of squares of rational numbers

(e) the set of all rational numbers with odd numerators (when written in lowest terms)

(f) the set of all rational numbers with even numerators (when written in lowest terms)

Solution.

(a) Let

$$\frac{a}{2m+1} \text{ and } \frac{b}{2n+1} \quad \text{be contained in the set,}$$

where each fraction is in its lowest term. Then,

$$\frac{a}{2m+1} - \frac{b}{2n+1} = \frac{2an+a-2bm-b}{(2m+1)(2n+1)}.$$

Since the denominator is a product of odd numbers, the set is closed under subtraction. Similarly, one can show that the set is closed under multiplication so by the subring criterion (Proposition 1.5), the set is a subring of $\mathbb{Q}$.

(b) No as it is not closed under subtraction/addition. Note that $\frac{1}{6}$ is contained in the set but $\frac{1}{6}+\frac{1}{6}=\frac{1}{3}$, for which the denominator is odd!

(c) No, since only 0 has an additive inverse. In particular, let x be contained in the set. Then, $x \in \mathbb{Q}_{\geq 0}$. If $x > 0$, then by Proposition 1.1, x does not have an additive inverse so the set is not a subring of $\mathbb{Q}$.

(d) Note that $1 = 1^2$ is in the set but $1+1=2$ is not the square of a rational. That is, there does not exist $x \in \mathbb{Q}$ such that $x^2 = 2$. So, the set is not a subring of $\mathbb{Q}$.

(e) No, as $\frac{1}{3}$ is in its lowest terms but when we sum $\frac{1}{3}$ with itself, we obtain $\frac{2}{3}$ which does not have an odd numerator. Hence, the set is not a subring of $\mathbb{Q}$.

(f) It is a simple exercise to show that this set is a subring of $\mathbb{Q}$. □

Proposition 1.6 (Dummit and Foote p. 230 Question 4). The intersection of any non-empty collection of subrings of a ring is also a subring.

Example 1.18 (Dummit and Foote p. 231 Question 12). Prove that any subring of a field which contains the identity is an integral domain.

Solution. Let F be a field and S be a subring of F. Suppose $x, y \in S$ such that $xy = 0$. Since $x, y \in F$ and the zero element in S is the same as that in F, then either $x = 0$ or $y = 0$. As such, S is an integral domain. □

Example 1.19 (Dummit and Foote p. 231 Question 18). Prove that $\{(r,r) \mid r \in R\}$ is a subring of $R \times R$.

Solution. Note that $0 \in R$ so R is non-empty. Hence, $(0,0) \in R \times R$ so $R \times R$ is non-empty.

Next, let $(a_1, b_1), (a_2, b_2) \in R \times R$. Then,

$$(a_1,b_1) - (a_2,b_2) = (a_1 - a_2, b_1 - b_2) \quad \text{and} \quad (a_1,b_1)(a_2,b_2) = (a_1a_2, b_1b_2).$$

These are both contained in $R \times R$. By the subring criterion (Proposition 1.5), it follows that the mentioned set is a subring of $R \times R$. □

On p. 232, Question 19 of *Abstract Algebra* by Dummit and Foote [1], it generalises Example 1.19. It mentions that if I is any non-empty index set and R_i is a ring for each $i \in I$, then

the direct product $\prod_{i \in I} R_i$ is a ring under componentwise addition and multiplication.

Definition 1.12 (center). The center of a ring R is defined to be the set of all $z \in R$ which commute with r for all $r \in R$. In other words,

$$Z(R) = \{z \in R : zr = rz \text{ for all } r \in R\}.$$

We note that Definition 1.12 is very much analogous to the center of a group G from MA2202 Algebra I.

Example 1.20 (Dummit and Foote p. 231 Question 7). See Definition 1.12 for the definition of the center of a ring. Here, we work with rings with a multiplicative identity.

(i) Prove that the center of a ring is a subring that contains the identity.

(ii) Prove that the center of a division ring is a field.

Solution.

(i) Note that $0 \in Z(R)$ since $0 \cdot r = r \cdot 0 = 0$ for all $r \in R$. As such, $Z(R)$ is non-empty. Next, take $z_1, z_2 \in Z(R)$, so $z_1 r = r z_1$ and $z_2 r = r z_2$. Hence,

$$(z_1 - z_2) r = z_1 r - z_2 r = r z_1 - r z_2 = r(z_1 - z_2)$$

and

$$(z_1 z_2) r = z_1 (z_2 r) = z_1 (r z_2) = r(z_1 z_2)$$

so by the subring criterion (Proposition 1.5), it follows that $Z(R)$ is a subring of R. If R contains the multiplicative identity 1_R, then by definition of $Z(R)$, we have $1_R \cdot r = r \cdot 1_R = r$. Hence, $1_R \in Z(R)$ so the center contains the identity.

(ii) Say we are given some division ring S. Recall Definition 1.4 which mentions that in every division ring, every non-zero element has a multiplicative inverse. For the center to be a field, it suffices to show that elements in the center commute. This follows from the fact that every field is a commutative ring. □

Definition 1.13 (nilpotent element). Let R be a ring with multiplicative identity. Then, $x \in R$ is nilpotent if there exists $m \in \mathbb{Z}^+$ such that $x^m = 0$.

Example 1.21 (Dummit and Foote p. 258 Question 27). Let R be a commutative ring with $1 \neq 0$. Prove that if a is a nilpotent element of R (Definition 1.13), then $1 - ab$ is a unit for all $b \in R$.

Solution. Let R be a commutative ring with $1 \neq 0$, and let $a \in R$ be nilpotent. Then there exists an integer $n \geq 1$ such that $a^n = 0$. Fix any $b \in R$ and set $x = ab$. Since R is commutative, we have

$$x^n = (ab)^n = a^n b^n = 0,$$

so x is nilpotent. Consider the elements

$$u = 1 - x = 1 - ab \quad \text{and} \quad v = 1 + x + x^2 + \cdots + x^{n-1}.$$

Then, using the telescoping (geometric series) identity in a ring,

$$uv = (1-x)\left(1 + x + x^2 + \cdots + x^{n-1}\right) = 1$$

Similarly,

$$vu = (1 + x + x^2 + \cdots + x^{n-1})(1-x) = 1 - x^n = 1.$$

Hence, v is a two-sided inverse of $u = 1 - ab$, so $1 - ab$ is a unit in R. Therefore, if a is nilpotent, then $1 - ab \in R^\times$ for all $b \in R$. □

Example 1.22 (Dummit and Foote p. 231 Question 13). See Definition 1.13 for the definition of a nilpotent element in a ring.

(a) Show that if $n = a^k b$ for some $a, b \in \mathbb{Z}$, then $\overline{ab}$ is a nilpotent element of $\mathbb{Z}/n\mathbb{Z}$.

(b) **(i)** If $a \in \mathbb{Z}$, show that $\overline{a} \in \mathbb{Z}/n\mathbb{Z}$ is nilpotent if and only if every prime divisor of n is also a divisor of a.

(ii) In particular, determine the nilpotent elements of $\mathbb{Z}/72\mathbb{Z}$ explicitly.

(c) Let R be the ring of functions from a non-empty set X to a field F. Prove that R contains no non-zero nilpotent elements.

Solution.

(a) Suppose $n = a^k b$. Then, because multiplication is commutative in $\mathbb{Z}$, then $(ab)^k = a^k b^k$, which is equal to $a^k b \cdot b^{k-1} = nb^{k-1}$. This is equal to 0 modulo n. As such, $\overline{ab}$ is a nilpotent element in $\mathbb{Z}/n\mathbb{Z}$, or equivalently, we have $(ab)^m \equiv 0 \pmod{m}$ for some $m \in \mathbb{Z}^+$.

(b) **(i)** We first deal with the forward direction. Suppose $a \in \mathbb{Z}$ and $\overline{a} \in \mathbb{Z}/n\mathbb{Z}$ is nilpotent. Then, $a^m \equiv 0 \pmod{n}$ for some $m \in \mathbb{Z}^+$, i.e. a^m is divisible by n. Let p be a prime divisor of n. Then, $p \mid n$, so $p \mid a^m$. As such, $p \mid a$.

We then prove the reverse direction. Suppose

$$n = p_1^{e_1} \ldots p_k^{e_k} \quad \text{and} \quad a = p_1^{d_1} \ldots p_k^{d_k} m \quad \text{where } 1 \leq e_i, d_i \text{ and } m \in \mathbb{Z}.$$

Let $t = \max\{e_i\}$. Then,

$$a^t = p_1^{d_1 t} \ldots p_k^{d_k t} m^t = nb \quad \text{for some } b \in \mathbb{Z}.$$

It follows that $a^t \equiv 0 \pmod{n}$.

(ii) Since the prime factorisation of 72 is $2^3 \cdot 3^2$, then the nilpotent elements are

$$0, 6, 12, 18, 24, 30, 36, 42, 48, 54, 60, 66,$$

which are namely the multiples of 6.

(c) Suppose on the contrary that R contains a non-zero nilpotent element. So, there exists $x \in X$ such that $\alpha(x) \neq 0$. Let m be the smallest $\mathbb{Z}^+$ such that $[\alpha(x)]^m = 0$. Then, we can write $[\alpha(x)]^{m-1}\alpha(x) = 0$, which is the product of two non-zero elements, contradicting the fact that F is a field. □

Example 1.23 (Dummit and Foote p. 231 Question 14). Let x be a nilpotent element of the commutative ring R.

(a) Prove that x is either zero or a zero divisor.

(b) Prove that rx is nilpotent for all $r \in R$.

(c) Prove that $1+x$ is a unit in R.

(d) Deduce that the sum of a nilpotent element and a unit is a unit.

Solution.

(a) We have $x^m = 0$ for some $m \in \mathbb{Z}^+$. If $m = 1$, then $x = 0$. On the other hand, if $m > 1$, then given that $x \neq 0$, we have $x^{m-1} \neq 0$ but $x^m = 0$, so x is a zero divisor.

(b) Again, suppose $x^m = 0$ for some $m \in \mathbb{Z}^+$. So, $(rx)^m = r^m x^m = 0$ where we sued the fact that R is commutative in the first equality.

(c) We recall the identity

$$(1+x)\left(1-x+x^2-x^3+x^4+\ldots\right) = 1.$$

Since $x^m = 0$, then $x^k = 0$ for all $k \geq m$. As such, we can reduce $1-x+x^4-x^3+x^4+\ldots$ to a finite sum, i.e.

$$(1+x)\sum_{i=0}^{m-1}(-1)^i x^i = 1.$$

Hence, $1+x$ is a unit in R.

(d) Let u be a unit and x be nilpotent in R. Then, write $u+x = u\left(1+u^{-1}x\right)$. We know that u is a unit. Knowing that the product of units yields another unit, it suffices to show that $1+u^{-1}x$ is also a unit. Actually, this follows from **(c)** by replacing x with $u^{-1}x$. The result follows. □

Definition 1.14 (Boolean ring). Let R be a ring with multiplicative identity. Then, R is Boolean if $a^2 = a$ for all $a \in R$.

Example 1.24 (Dummit and Foote p. 231 Question 15). See Definition 1.14 for the definition of a Boolean ring. Prove that every Boolean ring is commutative.

Solution. Suppose R is a Boolean ring and $a, b \in R$. The trick is to consider $(a+b)^2$ in R, so $(a+b)^2 = a+b$. Hence,

$$a^2+ab+ba+b^2 = a+b \quad \text{which implies} \quad a+ab+ba+b = a+b$$

since $a^2 = a$ and $b^2 = b$. As such, $ab+ba = 0$, so $ab = -ba$. Now, it suffices to show that $-ba = ba$, which would in turn prove that $ab = ba$. This follows from the fact that R is a Boolean ring. □

Example 1.25 (Dummit and Foote p. 231 Question 16). Prove that the only Boolean ring that is an integral domain is $\mathbb{Z}/2\mathbb{Z}$.

Solution. Let R be a Boolean ring that is an integral domain. Consider $a \in R$ which is non-zero. Then, $a^2 = a$ so $a(a-1) = 0$. Since R is an integral domain, then $a = 0$ or $a = 1$. We conclude that $R = \{0,1\}$, which precisely behaves like $\mathbb{Z}/2\mathbb{Z}$.[2] □

Example 1.26 (Dummit and Foote p. 232 Question 21). Let X be any nonempty set and let $\mathcal{P}(X)$ be the set of all subsets of X (the *power set of* X). Define addition and multiplication on $\mathcal{P}(X)$ by

$$A+B = (A-B)\cup(B-A) \quad \text{and} \quad A \times B = A \cap B$$

i.e. addition is symmetric difference and multiplication is intersection.

(a) Prove that $\mathcal{P}(X)$ is a ring under these operations ($\mathcal{P}(X)$ and its subrings are often referred to as *rings of sets*).

(b) Prove that this ring is commutative, has an identity and is a Boolean ring.

Solution.

(a) Tedious but very straightforward.

(b) Note that

$$A \cdot B = A \cap B = B \cap A = B \cdot A,$$

so $\mathcal{P}(X)$ is commutative. Since $A \cdot X = A \cap X = A$ and $X \cdot A = X \cap A = A$, then $\mathcal{P}(X)$ has an identity. Lastly, $\mathcal{P}(X)$ is Boolean since $A \cdot A = A \cap A = A$. □

Example 1.27 (Dummit and Foote p. 232 Question 22). Give an example of an infinite Boolean ring.

Solution. Recall from MA1100 Basic Discrete Mathematics that if X is an infinite set, then $\mathcal{P}(X)$ is also an infinite set. Recall from **(b)** of Example 1.26 where we showed that $\mathcal{P}(X)$ is a Boolean ring under symmetric difference and intersection. □

[2] This is a manifestation of ring isomorphism.

Example 1.28 (Dummit and Foote p. 231 Question 17). Let R and S be rings. Prove that the direct product $R \times S$ is a ring under componentwise addition and multiplication. Prove that $R \times S$ is commutative if and only if both R and S are commutative. Prove that $R \times S$ has an identity if and only if both R and S have identities.

Solution. The first part of proving $R \times S$ is a ring is trivial. For the second part, suppose $R \times S$ is commutative, That is, for any $r_1, r_2 \in R$ and $s_1, s_2 \in S$, we have

$$(r_1, s_1)(r_2, s_2) = (r_2, s_2)(r_1, s_1),$$

which implies $r_1 r_2 = r_1 r_1$ and $s_1 s_2 = s_2 s_1$. This shows that R and S are commutative. In a similar fashion, one can deduce the reverse direction.

Again, the last part is trivial. If R and S have identities, then

$$(1,1)(r,s) = (r,s) \quad \text{and} \quad (r,s)(1,1) = (r,s)$$

so $(1,1) \in R \times S$ is an identity. Next, suppose $(u,v) \in R \times S$ is an identity. Then, for all $r \in R$ and $s \in S$, we have

$$(ur, vs) = (u,v)(r,s) = (r,s) = (r,s)(u,v) = (ru, sv).$$

As such, $ur = ru = r$ and $vs = sv = s$ so by the uniqueness of identities, both R and S have an identity. □

Definition 1.15 (discrete valuation). Let K be a field. A discrete valuation on K is a function $\nu : K^\times \to \mathbb{Z}$ satisfying the following properties:

(i) $\nu(ab) = \nu(a) + \nu(b)$, i.e. ν is a homomorphism from the multiplicative group of non-zero elements of K to $\mathbb{Z}$

(ii) ν is surjective

(iii) $\nu(x+y) \geq \min\{\nu(x), \nu(y)\}$ for all $x, y \in K^\times$ with $x + y \neq 0$

The set $R = \{x \in K^\times : \nu(x) \geq 0\} \cup \{0\}$ is called the valuation ring of ν.

Example 1.29 (Dummit and Foote p. 232 Question 26). Let K be a field. Recall the definition of a discrete valuation on K (Definition 1.15). Let

$$R = \{x \in K^\times : \nu(x) \geq 0\} \cup \{0\}$$

which is called the valuation ring of ν.

(a) Prove that R is a subring of K which contains the identity[3].

(b) Prove that for each non-zero element $x \in K$, either x or x^{-1} is in R.

[3] In general, a ring R is called a discrete valuation ring if there exists a field K and some discrete valuation ν on K such that R is the valuation ring of ν.

(c) Prove that an element x is a unit of R if and only if $\nu(x) = 0$.

Solution.

(a) Let $x_1, x_2 \in R$. We will use Definition 1.1. First, we show that $0, 1 \in R$. By Definition 1.15, $0 \in R$. Next, $\nu(1 \cdot 1) = \nu(1) + \nu(1)$ by **(i)** of Definition 1.15, so $\nu(1) = 0$, which implies $1 \in R$.

We then prove that R is closed under addition. Suppose $x_1, x_2 \in R$. If either $x_1 = 0$ or $x_2 = 0$, then we are done. Suppose $x_1, x_2 \in K^\times$ such that $\nu(x_1), \nu(x_2) \geq 0$. If $x_1 + x_2 = 0$, then the result holds, otherwise, by **(iii)**[4] of Definition 1.15, $\nu(x_1 + x_2) \geq \min\{\nu(x_1), \nu(x_2)\} \geq 0$, so R is closed under addition.

We also need to justify that the additive inverse of R exists. If $x = 0$, then we are done, otherwise note that $\nu(-x) = \nu(-1) + \nu(x)$. By **(i)** of Definition 1.15, $\nu(-1) = 0$ so $\nu(-x) \geq 0$. Hence, the additive inverse of R, which is $-x$, exists.

We then prove that R is closed under multiplication. Suppose $x_1, x_2 \in R$. Again, if either $x_1 = 0$ or $x_2 = 0$, then we are done. Suppose $x_1, x_2 \in K^\times$ such that $\nu(x_1), \nu(x_2) \geq 0$. By **(i)** of Definition 1.15, it follows that R is closed under multiplication.

(b) Let K be a field. Recall from **(a)** that R is a subring of K. Say $x \in R$. Since x is non-zero, then $\nu(x) \geq 0$. By **(i)** of Definition 1.15, $\nu(1) = \nu(x) + \nu\left(x^{-1}\right)$, and from **(a)** we know that $\nu(1) = 0$, so $\nu\left(x^{-1}\right) = -\nu(x) \leq 0$, so x^{-1} is not contained in R. By a symmetric argument, it follows that if $x^{-1} \in R$, then x is not contained in R.

(c) We first prove the forward direction. Suppose x is a unit of R. Then, there exists $y \in R$ such that $xy = 1$. By **(i)** of Definition 1.15, $\nu(x) + \nu(y) = \nu(1) = 0$, where $\nu(1) = 0$ follows from **(a)**. Since $\nu \geq 0$, then we must have $\nu(x) = \nu(y) = 0$.

For the reverse direction, suppose $\nu(x) = 0$. Then, consider $\nu\left(x \cdot x^{-1}\right) = \nu(x) + \nu\left(x^{-1}\right)$, which implies $\nu(1) = \nu\left(x^{-1}\right)$. Since $\nu(1) = 0$ by **(a)**, then $\nu\left(x^{-1}\right) = 0 \geq 0$ so $x^{-1} \in R$. In particular, x is a unit of R since its inverse, x^{-1}, satisfies $x \cdot x^{-1} = 1$. □

Example 1.30. With reference to Example 1.29 and Definition 1.15 on discrete valuation, a specific example of a discrete valuation ring is obtained when p is prime, $K = \mathbb{Q}$ (the field of rational numbers) and

$$\nu_p : \mathbb{Q}^\times \to \mathbb{Z} \quad \text{by} \quad \nu_p\left(\frac{a}{b}\right) = \alpha \text{ where } \frac{a}{b} = p^\alpha \cdot \frac{c}{d}.$$

Here, p does not divide c and d. Prove that the corresponding valuation ring R is the ring of all rational numbers whose denominators are relatively prime to p. Describe the units of this valuation ring.

[4] This is known as the ultrametric inequality.

Solution. The valuation ring R is defined to be

$$R = \left\{x \in \mathbb{Q}^\times : v_p(x) \geq 0\right\} \cup \{0\}.$$

Take a non-zero rational number $x = \frac{a}{b}$ written in lowest terms, i.e. $\gcd(a,b) = 1$ and $b > 0$. Write the factorisations $a = p^\alpha a_0$ and $b = p^\beta b_0$, where $\alpha, \beta \geq 0$ with p not dividing a_0 and b_0. So,

$$\frac{a}{b} = p^{\alpha-\beta} \cdot \frac{a_0}{b_0}.$$

Hence, the p-adic valuation of $\frac{a}{b}$ is equal to $\alpha - \beta$. By Definition 1.15, $v_p\left(\frac{a}{b}\right) \geq 0$ if and only if $\alpha \geq \beta$. Since $\gcd(a,b) = 1$, then we cannot have both $\alpha \geq 1$ and $\beta \geq 1$, which forces $\alpha = 0$ or $\beta = 0$.

If p does not divide b, then $\beta = 0$. On the other hand, if p divides b, then $\beta \geq 1$ so $\alpha = 0$. This implies $v_p\left(\frac{a}{b}\right) < 0$, which leads to a contradiction. As such, we must have p not dividing b. Hence, we can write R as

$$R = \left\{\frac{a}{b} \in \mathbb{Q} : \gcd(a,b) = 1 \text{ and } p \text{ does not divide } b\right\} \cup \{0\}.$$

This is precisely the ring rational numbers whose denominators are relatively prime to p. As for the units, they are all $\frac{a}{b} \in \mathbb{Q}$ such that $\gcd(a,b) = 1$, p not dividing both a and b. □

1.2 Polynomial Rings, Matrix Rings, and Group Rings

Definition 1.16 (polynomial ring). Let R be a commutative ring with multiplicative identity 1. Let x be a formal variable. We define the polynomial ring $R[x]$ as follows:

$$R[x] = \{a_0 + a_1 x + \ldots + a_n x^n \in R : n \in \mathbb{Z}_{\geq 0}\}$$

Addition and multiplication on $R[x]$ are defined in the obvious/naive way.

Each polynomial should be regarded as a formal expression instead of a function.

Example 1.31 (Dummit and Foote p. 237 Question 1). Let R be a commutative ring with 1. Let $p(x) = 2x^3 - 3x^2 + 4x - 5$ and $q(x) = 7x^3 + 33x - 4$. In each of **(a)**, **(b)**, and **(c)**, compute $p(x) + q(x)$ and $p(x)q(x)$ under the assumption that the coefficients of the two given polynomials are taken from the specified ring (where the integer coefficients are taken modulo n in **(b)** and **(c)**.

(a) $R = \mathbb{Z}$ **(b)** $R = \mathbb{Z}/2\mathbb{Z}$ **(c)** $R = \mathbb{Z}/3\mathbb{Z}$

Solution. It is easy to deduce that

$$p(x) + q(x) = 9x^3 - 3x^2 + 37x - 9$$
$$p(x)q(x) = 14x^6 - 21x^5 + 94x^4 + 212x^3 - 189x^2 + 181x + 20$$

which answers **(a)**. To answer **(b)** and **(c)**, we simply reduce the coefficients modulo 2 and 3 respectively. □

Lemma 1.1. If R is an integral domain, then $R[x]$ is an integral domain.

Proof. In $R[x]$, take two non-zero polynomials

$$f(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_0 \quad \text{and} \quad g(x) = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_0$$

where a_m and b_n are both non-zero. So, f and g are polynomials of degree m and n respectively. Consider the product $h(x) = f(x)g(x)$. The coefficient of x^{m+n} in h comes only from the product of the leading terms $a_m x^m$ and $b_n x^n$, so the leading coefficient of h is $a_m b_n$. Since R is an integral domain and a_m, b_n are both non-zero, then the leading coefficient of h is non-zero, so $h(x) \neq 0$. By Definition 1.9, $R[x]$ is an integral domain. □

Definition 1.17 (matrix ring). Let R be a ring with multiplicative identity 1 (or $\mathbf{I}$). Define the matrix ring $\mathcal{M}_{n\times n}(R)$ to be the set consisting of $\left(a_{ij}\right)_{n\times n}$ where $a_{ij} \in R$. Addition and multiplication on $\mathcal{M}_{n\times n}(R)$ are defined as per matrix multiplication in MA2001 Linear Algebra I.

Example 1.32. If $R = \mathbb{R}$ (recall that $\mathbb{R}$ is a field from Example 1.1) so $\mathbb{R}$ is a multiplicative identity with 1. Then, $\mathcal{M}_{n\times n}(\mathbb{R})$ is the usual matrix algebra. We have the usual subring of diagonal matrices, as well as the subring of upper triangular matrices.

Definition 1.18 (group ring). Let R be a commutative ring with multiplicative identity 1. Let G be a finite group. Define the group ring $R[G]$ to be the set consisting of

$$\sum_{g\in G} a_g g \quad \text{where } a_g \in R.$$

The addition on $R[G]$ is defined in the obvious way.

Example 1.33. We discuss multiplication in $R[G]$. We have

$$(a_g g + a_h h)\left(a_{g'} g' + a_{h'} h'\right) = a_g a_{g'} gg' + a_h a_{g'} hg' + a_g a_{h'} gh' + a_h a_{h'} hh'.$$

Here, gg', hg', gh', hh' denote group multiplication in G.

Example 1.34. We shall discuss the structure of $\mathbb{R}[\mathbb{Z}/2\mathbb{Z}]$. Note that this group ring consists of formal linear combinations of the group elements with coefficients in $\mathbb{R}$. Hence,

$$\mathbb{R}[\mathbb{Z}/2\mathbb{Z}] = \{a_0 e_0 + a_1 e_1 : a_0, a_1 \in \mathbb{R}\}.$$

Here, $e_0, e_1 \in \mathbb{Z}/2\mathbb{Z}$, with e_0 being the identity element of $\mathbb{Z}/2\mathbb{Z}$.

Lemma 1.2. Let R be a commutative ring with multiplicative identity 1 and G be a finite group. Then, the following hold:

(i) Let $e \in G$ be its identity element. Then, $1e$ is the identity of the ring $R[G]$.

(ii) Let $e \neq g \in G$. Then, $1-g$ is a zero divisor.

(iii) Let $H \leq G$. Then, $R[H]$ is a subring of $R[G]$.

(iv) $R[G]$ is commutative if and only if G is commutative

Definition 1.19 (product of rings). Let S and R be two rings. Define their product $S \times R$ to be the same as the product of sets. Addition and multiplication are defined component-wise, i.e.

$$(a,b)+(c,d) = (a+c, b+d) \quad \text{and} \quad (a,b)\cdot(c,d) = (ac, bd).$$

Definition 1.20 (formal power series). Let R be a commutative ring with 1. Define the set $R[[x]]$ of formal power series in the indeterminate x with coefficients from R to be all formal infinite sums

$$\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots.$$

Define addition and multiplication of power series in the same way as for power series with real or complex coefficients i.e. extend polynomial addition and multiplication to power series as though they were polynomials of infinite degree:

$$\sum_{n=0}^{\infty} a_n x^n + \sum_{n=0}^{\infty} b_n x^n = \sum_{n=0}^{\infty} (a_n + b_n) x^n$$
$$\sum_{n=0}^{\infty} a_n x^n \times \sum_{n=0}^{\infty} b_n x^n = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_k b_{n-k} \right) x^n$$

In Definition 1.20, the term formal is used here to indicate that convergence is not considered, so that formal power series need not represent functions on R.

Example 1.35 (Dummit and Foote p. 238 Question 4). Let R be a commutative ring with 1. Consider Definition 1.20 on formal power series.

(a) Prove that $R[[x]]$ is a commutative ring with 1.

(b) Show that $1-x$ is a unit in $R[[x]]$ with inverse $1+x+x^2+\cdots$.

(c) Prove that

$$\sum_{n=0}^{\infty} a_n x^n \text{ is a unit in } R[[x]] \quad \text{if and only if} \quad a_0 \text{ is a unit in } R.$$

Solution.

(a) We first show that $R[[x]]$ is an additive Abelian group. Note that coefficient-wise addition is associative and commutative because addition in R is. That is,

$$(f+g)+h = \sum_{n\geq 0} \big((a_n+b_n)+c_n\big)x^n = \sum_{n\geq 0} \big(a_n+(b_n+c_n)\big)x^n = f+(g+h),$$

and similarly $f+g=g+f$. The additive identity is $0=\sum_{n\geq 0} 0\cdot x^n$, and the additive inverse of $f=\sum a_n x^n$ is $-f=\sum(-a_n)x^n$.

Next, let

$$f=\sum a_n x^n \quad \text{and} \quad g=\sum b_n x^n \quad \text{and} \quad h=\sum c_n x^n.$$

Write

$$fg=\sum_{n\geq 0} d_n x^n \quad \text{where } d_n=\sum_{i=0}^{n} a_i b_{n-i}.$$

Then the coefficient of x^n in $(fg)h$ equals

$$\sum_{j=0}^{n} d_j c_{n-j}=\sum_{j=0}^{n}\left(\sum_{i=0}^{j} a_i b_{j-i}\right)c_{n-j}=\sum_{\substack{i,r,s\geq 0\\ i+r+s=n}} a_i b_r c_s,$$

where the last sum is finite (only triples with $i+r+s=n$). Similarly, the coefficient of x^n in $f(gh)$ is

$$\sum_{\substack{i,r,s\geq 0\\ i+r+s=n}} a_i b_r c_s,$$

so $(fg)h=f(gh)$. For commutativity,

$$c_n(fg)=\sum_{k=0}^{n} a_k b_{n-k}=\sum_{k=0}^{n} b_k a_{n-k}=c_n(gf),$$

using commutativity of multiplication in R, hence $fg=gf$. We omit the proof of distributivity and the existence of a multiplicative identity.

(b) Let

$$u(x)=1+x+x^2+\cdots=\sum_{n=0}^{\infty} x^n \in R[[x]].$$

Compute $(1-x)u$ via coefficients. Write $1-x=1+(-1)x+0x^2+\cdots$. For $n=0$, the constant term of $(1-x)u$ is $1\cdot 1=1$. For $n\geq 1$, the coefficient of x^n is

$$(1)\cdot(1)+(-1)\cdot(1)=1-1=0,$$

since it comes only from $1\cdot x^n$ and $(-x)\cdot x^{n-1}$. Thus, $(1-x)u=1+0x+0x^2+\cdots=1$ and similarly $u(1-x)=1$ by commutativity, so $1-x$ is a unit with inverse u.

(c) For the forward direction, suppose $f(x)=\sum_{n=0}^{\infty} a_n x^n$ is a unit in $R[[x]]$, so there exists $g(x)=\sum_{n=0}^{\infty} b_n x^n$ with $f(x)g(x)=1$. Looking at constant terms,

$$a_0 b_0=1 \text{ in } R \quad \text{so} \quad a_0 \text{ is a unit in } R \text{ with inverse } b_0.$$

Conversely, assume a_0 is a unit in R. We construct $g(x)=\sum_{n=0}^{\infty} b_n x^n$ such that $f(x)g(x)=1$. The condition $fg=1$ is equivalent to the system of equations on coefficients:

$$\sum_{k=0}^{n} a_k b_{n-k}=\begin{cases}1 & \text{if } n=0;\\ 0 & \text{if } n\geq 1.\end{cases}$$

Set $b_0 = a_0^{-1}$. For $n \geq 1$, define b_n recursively by

$$b_n = -a_0^{-1} \sum_{k=1}^{n} a_k b_{n-k}.$$

This makes sense because the sum on the right involves only $b_0, \ldots, b_{n-1}$, already defined, and a_0^{-1} exists. Now, check the n^{th} coefficient of fg for $n \geq 1$, which is

$$\sum_{k=0}^{n} a_k b_{n-k} = a_0 b_n + \sum_{k=1}^{n} a_k b_{n-k} = a_0 \left(-a_0^{-1} \sum_{k=1}^{n} a_k b_{n-k} \right) + \sum_{k=1}^{n} a_k b_{n-k} = 0.$$

For $n = 0$, we have $a_0 b_0 = 1$ by construction. Hence $fg = 1$, so f is a unit.

Example 1.36 (Dummit and Foote p. 238 Question 5). Let R be a commutative ring with 1. Prove that if R is an integral domain then the ring of formal power series $R[[x]]$ (Definition 1.20) is also an integral domain.

Solution. Assume R is an integral domain. By Definition 1.9, we will show that $R[[x]]$ has no zero divisors. Let

$$f(x) = \sum_{n=0}^{\infty} a_n x^n \quad \text{and} \quad g(x) = \sum_{n=0}^{\infty} b_n x^n$$

be non-zero elements of $R[[x]]$. Since $f \neq 0$, there exists a least index $m = \min\{n \geq 0 : a_n \neq 0\}$ and since $g \neq 0$, there exists a least index $\ell = \min\{n \geq 0 : b_n \neq 0\}$. These minima exist because $\mathbb{Z}_{\geq 0}$ is well-ordered.

Consider the product

$$h(x) = f(x)g(x) = \sum_{n=0}^{\infty} c_n x^n \quad \text{where} \quad c_n = \sum_{k=0}^{n} a_k b_{n-k}.$$

We claim that $c_{m+\ell} \neq 0$. Indeed,

$$c_{m+\ell} = \sum_{k=0}^{m+\ell} a_k b_{m+\ell-k}.$$

If $k < m$, then $a_k = 0$ by minimality of m; if $k > m$ then $m + \ell - k < \ell$ so $b_{m+\ell-k} = 0$ by minimality of ℓ. Hence every term in the sum is 0 except possibly the term with $k = m$, and therefore $c_{m+\ell} = a_m b_\ell$. Since $a_m \neq 0$ and $b_\ell \neq 0$ and R is an integral domain, we have $a_m b_\ell \neq 0$. Thus $c_{m+\ell} \neq 0$, so $h(x) \neq 0$.

Therefore the product of two non-zero elements of $R[[x]]$ is non-zero, i.e. $R[[x]]$ has no zero divisors. Hence, $R[[x]]$ is an integral domain. □

1.3 Ring Homomorphisms and Quotient Rings

Definition 1.21 (ring homomorphism). Let R and S be rings. A ring homomorphism is a map $\varphi : R \to S$ such that for all $a, b \in R$, we have

$$\varphi(a+b) = \varphi(a) + \varphi(b) \quad \text{and} \quad \varphi(ab) = \varphi(a)\varphi(b).$$

Example 1.37 (Dummit and Foote p. 249 Question 25; binomial theorem). Assume R is a commutative ring with multiplicative identity 1. Prove that the binomial theorem

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} \tag{1.1}$$

holds in R, where the binomial coefficient $\binom{n}{k}$ is interpreted in R as the sum $1 + \ldots + 1$ of the identity 1 in R taken $\binom{n}{k}$ times.

Solution. One can use induction. Let $P(n)$ be the mentioned statement, where $n \in \mathbb{Z}^+$. When $n = 1$, we have $a+b$ on the left side of (1.1), whereas the right side is $\binom{n}{0}b + \binom{n}{1}a = a+b$ so $P(1)$ is true. Suppose that the proposition holds for some $m \in \mathbb{Z}^+$. Then, we consider $P(m+1)$, so

$$(a+b)^{m+1} = (a+b)^m (a+b) = \left[\sum_{k=0}^{m} \binom{m}{k} a^k b^{m-k}\right](a+b)$$

which is equal to

$$\left[\binom{m}{0} b^m + \binom{m}{1} a b^{m-1} + \ldots + \binom{m}{m} a^m\right](a+b) = \ldots = \sum_{k=0}^{m+1} \binom{m+1}{k} a^k b^{m+1-k}.$$

Here, we skipped many steps (routine exercise from MA1100 Basic Discrete Mathematics and/or MA2116 Probability) which uses Pascal's identity and some algebraic manipulation. □

Definition 1.22 (kernel). The kernel of a ring homomorphism φ, denoted by $\ker\varphi$, is the following set:

$$\ker\varphi = \{r \in R : \varphi(r) = 0_S\}$$

So, the kernel can be viewed as a homomorphism of additive groups.

Example 1.38 (quotient map and embedding). We discuss two simple examples.

(i) The quotient map $\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ is a ring homomorphism with kernel $n\mathbb{Z}$.

(ii) The embedding of the subring $n\mathbb{Z} \to \mathbb{Z}$ is a ring homomorphism with a trivial kernel.

Example 1.39. Consider the map

$$\varphi : \mathbb{C}[x] \to \mathbb{C} \quad \text{where} \quad f(x) \mapsto f(a).$$

The kernel is given by

$$\ker\varphi = \{f(x) \in \mathbb{C}[x] : f(a) = 0\} = \{(x-a)f(x) : f(x) \in \mathbb{C}[x]\}.$$

Example 1.40. Let

$$R = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} : a, b, d \in \mathbb{Z} \right\}$$

be the subring of $\mathcal{M}_{2\times 2}(\mathbb{Z})$ of 2×2 upper triangular matrices over the integers. One can prove that the map

$$\varphi : R \to \mathbb{Z} \times \mathbb{Z} \quad \text{where} \quad \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \mapsto (a, d)$$

is a surjective homomorphism and its kernel is $(a, b, d) = (0, b, 0)$, where $b \in \mathbb{Z}$.

Lemma 1.3. Let $\varphi : R \to S$ be a ring homomorphism. Then, $\operatorname{im}\varphi$ is a subring of S and $\ker\varphi$ is a subring of R.

The proof of Lemma 1.3 is similar to that in Group Theory. Recall from MA2202 Algebra I that the analogous result states that if $\varphi : G \to H$ is a group homomorphism, then $\operatorname{im}\varphi$ is a subring of H and $\ker\varphi$ is a subring of G.

Example 1.41 (Dummit and Foote p. 248 Question 5). Let R be a ring with identity $1 \neq 0$. Describe all ring homomorphisms from $\mathbb{Z} \times \mathbb{Z}$ to $\mathbb{Z}$. In each case, describe the kernel and the image.

Solution. Let $\varphi : \mathbb{Z} \times \mathbb{Z}$ where $(a, b) \mapsto c$ be a ring homomorphism. Just like in MA2001 Linear Algebra I, we can think of the images of the standard basis vectors. Let $\mathbf{e}_1 = (1, 0)$ and $\mathbf{e}_2 = (0, 1)$. Suppose $\varphi(\mathbf{e}_1) = u$ and $\varphi(\mathbf{e}_2) = v$. Since ring homomorphisms preserve products, then it is easy to deduce that[5] $u^2 = u$ and $v^2 = v$.

As the only idempotent elements in $\mathbb{Z}$ are 0 and 1, then $u, v \in \{0, 1\}$. Next, we have $uv = \varphi(\mathbf{e}_1)\varphi(\mathbf{e}_2) = \varphi(0, 0) = 0$ so it is impossible to have both u and v being equal to 1. We shall proceed with casework.

- **Case 1:** Suppose $u = 0$ and $v = 0$. Then, $(1, 0) \mapsto 0$ and $(0, 1) \mapsto 0$. Then, $(a, b) \mapsto 0$ for all $a, b \in \mathbb{Z}$ so the kernel of the homomorphism is $\mathbb{Z} \times \mathbb{Z}$ and the image is $\{0\}$.
- **Case 2:** Suppose $u = 1$ and $v = 0$. Then, $(a, b) \mapsto a$, which is known as the projection map where we project (a, b) onto the first coordinate a. The kernel is all $(a, b) \in \mathbb{Z} \times \mathbb{Z}$ such that $a = 0$, whereas the image is $\mathbb{Z}$.

Similarly, we have a third case $u = 0$ and $v = 1$, but note that this is similar to Case 2. Hence, the ring homomorphisms $\mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ are exactly the zero map and the two coordinate projections, with kernels and images as above. □

Definition 1.23 (ring isomorphism). A bijective ring homomorphism is an isomorphism. Two rings R and S are isomorphic if there exists an isomorphism between R and S and we write $R \cong S$.

[5] We say that u and v are idempotent in $\mathbb{Z}$.

Example 1.42 (Dummit and Foote p. 247 Question 1). Let R be a ring with identity $1 \neq 0$. Prove that $2\mathbb{Z}$ and $3\mathbb{Z}$ are not isomorphic as rings.

Solution. Suppose on the contrary that there exists a ring homomorphism $\varphi : 2\mathbb{Z} \to 3\mathbb{Z}$ as rings with identity. However, $2\mathbb{Z}$ and $3\mathbb{Z}$ do not have a multiplicative identity, but by the property of ring isomorphisms that they map the identity in $2\mathbb{Z}$ to the identity in $3\mathbb{Z}$, the result follows. □

Example 1.43 (Dummit and Foote p. 247 Question 2). Let R be a ring with identity $1 \neq 0$. Prove that $\mathbb{Z}[x]$ and $\mathbb{Q}[x]$ are not isomorphic as rings.

Solution. We note that the units in $\mathbb{Z}[x]$ are ± 1. To see why, suppose $f(x)$ is a unit in $\mathbb{Z}[x]$. Then, there exists $g(x) \in \mathbb{Z}[x]$ such that $f(x)g(x) = 1$. So, $\deg f + \deg g = 0$, which implies f, g must be constant polynomials. In other words, $f, g \in \mathbb{Z}$ and as the only invertible integers are ± 1, it follows that the units in $\mathbb{Z}[x]$ are ± 1. Next, it is clear that the units in $\mathbb{Q}^\times$ are all the non-zero constants, i.e. $\mathbb{Q}^\times$.

Since $\mathbb{Z}[x]$ has 2 units and $\mathbb{Q}[x]$ has infinitely many units, it follows that no such ring isomorphism exists since a ring isomorphism sends units to units. □

Definition 1.24 (ideal). Let R be a ring and $I \subseteq R$.

(i) I is a left ideal of R if I is an additive subgroup of R and $rI \subseteq I$ for any $r \in R$

(ii) I is a right ideal of R if I is an additive subgroup of R and $Ir \subseteq I$ for any $r \in R$

(iii) I is an ideal of R if I is both a left ideal and a right ideal of R

Example 1.44 (principal ideal). $n\mathbb{Z} = (n)$ is a principal ideal in $\mathbb{Z}$. Here, (n) denotes the ideal generated by n, i.e. the smallest ideal containing n. By the term 'principal', we mean that $n\mathbb{Z}$ is generated by only one element (Definition 1.30).

In general, we write rR for $\{rx : x \in R\}$, which is the right ideal of R generated by r. When R is commutative, we simply write (r).

Lemma 1.4. Let $\varphi : R \to S$ be a ring homomorphism. Then, $\ker \varphi$ is an ideal of R.

Proof. Let $x, y \in \ker \varphi$. Then,

$$\varphi(x-y) = \varphi(x) - \varphi(y) = 0 - 0 = 0$$

so $x - y \in \ker \varphi$ so $\ker \varphi$ is an additive subgroup of R. Then, let $r, r' \in R$ and $x \in \ker \varphi$. So,

$$\varphi\left(rxr'\right) = \varphi(r)\varphi(x)\varphi\left(r'\right) = \varphi(r) \cdot 0 \cdot \varphi\left(r'\right) = 0$$

so $rxr' \in \ker \varphi$. Here, we verified that $\ker \varphi$ is a two-sided ideal of R (or in short, just an ideal) by proving **(i)** and **(ii)** in Definition 1.24 concurrently. □

Example 1.45 (Dummit and Foote p. 267 Question 1). Let R be a ring with identity $1 \neq 0$. Assume e is an idempotent element[6] in R and $er = re$ for all $r \in R$.

(i) Prove that Re and $R(1-e)$ are two-sided ideals of R and that $R \cong Re \times R(1-e)$.

(ii) Show that e and $1-e$ are identities for the subrings Re and $R(1-e)$ respectively.

Solution.

(i) We first prove that Re is a two-sided ideal. Recall that $Re = \{re : r \in R\}$, which is an additive subgroup since

$$r_1e + r_2e = (r_1 + r_2)e \quad \text{and} \quad -(re) = (-r)e.$$

Let $x = re \in Re$ and $s \in R$. Then, $sx \in Re$ so Re is a left ideal (Definition 1.24). One can then show that $xs \in Re$, which shows that Re is a right ideal. Therefore, Re is a two-sided ideal.

Using the above idea, one can show that $R(1-e)$ is a two-sided ideal.

Lastly, we prove that $R \cong Re \times R(1-e)$. Consider the map

$$\Phi : R \to Re \times R(1-e) \quad \text{where} \quad \Phi(r) = (re, r(1-e)).$$

This is additive by construction. For multiplicativity, let $r, s \in R$. Then,

$$\Phi(rs) = ((rs)e, (rs)(1-e)).$$

On the other hand,

$$\Phi(r)\Phi(s) = (re, r(1-e)) \cdot (se, s(1-e)) = \big((re)(se), r(1-e)s(1-e)\big).$$

Using centrality of e and $1-e$ and idempotency, we have $(re)(se) = (rs)e$ and $r(1-e)s(1-e) = rs(1-e)$. So, Φ is a ring homomorphism. The map is surjective and $\ker\Phi = \{0\}$. So, Φ is an isomorphism and $R \cong Re \times R(1-e)$.

(ii) Let $x \in Re$, so $x = re$ for some $r \in R$. Then

$$ex = e(re) = (er)e = (re)e = re^2 = re = x,$$

and

$$xe = (re)e = re^2 = re = x.$$

Thus e is a (two-sided) identity element for the subring Re. Similarly, one can show that $1-e$ is the identity element for the subring $R(1-e)$. □

Example 1.46 (Dummit and Foote p. 267 Question 3). Let R and S be rings with identity $1 \neq 0$. Let R and S be rings with identities. Prove that every ideal of $R \times S$ is of the form $I \times J$ where I is an ideal of R and J is an ideal of S.

[6] Recall that an element $e \in R$ is called an idempotent if $e^2 = e$.

Solution. Let R and S be rings with identity. Let K be an ideal of the direct product ring $R \times S$. Consider the subsets I and J, which are the images of K under the coordinate projections. That is,

$$I = \{r \in R : \text{there exists } s \in S \text{ with } (r,s) \in K\}$$
$$J = \{s \in S : \text{there exists } r \in R \text{ with } (r,s) \in K\}$$

One can prove that I is an ideal of R and J is an ideal of S. Next, $K \subseteq I \times J$ and $I \times J \subseteq K$, so the result follows. □

Definition 1.25 (quotient ring). Let $I \subseteq R$ be an ideal. Define the quotient ring R/I as follows: we can view it as an Abelian group, where multiplication is defined as follows:

$$R/I \times R/I \to R/I \quad \text{where} \quad (a+I, b+I) \mapsto ab+I \quad \text{for all } a,b \in R.$$

The image of $a \in R$ in R/I is often denoted by $\overline{a}$.

Definition 1.26 (sum of ideals). Let I and J be ideals of R. Define

$$I+J = \{a+b : a \in I, b \in J\} \quad \text{which is an ideal of } R.$$

Definition 1.27 (product of ideals). Let I and J be ideals of R. Define

$$IJ = \left\{\sum ab : a \in I, b \in J\right\} \quad \text{which is an ideal of } R.$$

Consequently, for any $n \in \mathbb{N}$, I^n is an ideal of R.

Example 1.47 (union of ideals; Dummit and Foote p. 249 Question 19). Prove that if $I_1 \subseteq I_2 \subseteq \cdots$ are ideals of R, then the union $I_1 \cup I_2 \cup \cdots$ is an ideal of R.

Solution. Let I denote the union. We first show that $I \neq \emptyset$. Note that $0 \in I_1$ by definition of an ideal and because $I_1 \subseteq I$, then I is non-empty.

We then prove the additive closure property. Take $a,b \in I$. Then, there exist $m,n \in \mathbb{N}$ such that $a \in I_m$ and $b \in I_n$. Without loss of generality, suppose $n \geq m$ so $a \in I_n$. So, $a-b \in I_n \subseteq I$, which implies that I is closed under subtraction.

Next, suppose $r \in R$ and $a \in I$. Then, again there exists $m \in \mathbb{N}$ such that $a \in I_m$. By Definition 1.24, $ra \in I_m \subseteq I$, so the absorption by ring elements property holds. Hence, I is an ideal of R. □

Example 1.48 (intersection of ideals; Dummit and Foote p. 249 Question 20). Let I be an ideal of R and let S be a subring of R. Prove that $I \cap S$ is an ideal of S. Show by example that not every ideal of a subring S of a ring R need be of the form $I \cap S$ for some ideal I of R.

Solution. Let $J = I \cap S$. For the first part, since $0 \in I$ and $0 \in S$ by properties of ideals and subrings, then $0 \in J$ so $J \neq \emptyset$. Next, we prove that J is closed under subtraction. Take $a, b \in J$. Then, $a, b \in I$ and $a, b \in S$. Since I is an ideal of R, then $a - b \in I$; the *same* argument can be applied to S since it is a subring of R. Hence, J is closed under subtraction. We omit the proof that J is close under subtraction. Hence, J is an ideal of S.

For the second part, take $R = \mathbb{Z}$ and $S = 2\mathbb{Z}$. Then, clearly $2\mathbb{Z}$ is a subring of $\mathbb{Z}$. Note that $4\mathbb{Z}$ is an ideal of $2\mathbb{Z}$. Note that this ideal is not of the form $\mathbb{Z} \cap 2\mathbb{Z} = 2\mathbb{Z}$. For example, $6 \in 2\mathbb{Z}$ but 6 is not contained in $4\mathbb{Z}$. □

Theorem 1.1 (first isomorphism theorem). Let $\varphi : R \to S$ be a ring homomorphism. Then,

$$R/\ker\varphi \cong \varphi(R).$$

Also, for any ideal $I \subseteq R$, the quotient map

$$\pi : R \to R/I \quad \text{where} \quad a \mapsto a + I = \overline{a}$$

is a surjective ring homomorphism with kernel I.

From Theorem 1.2, we see that if $I \subseteq \ker\varphi$, then φ factors through R/I, i.e. we have the commutative diagram shown in Figure 1.2, where $\overline{\varphi}(\overline{a}) = \varphi(a)$ for $a \in R$. In fact, this is the universal property of the quotient ring.

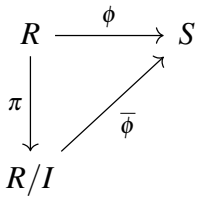


Figure 1.2: First isomorphism theorem for rings

Example 1.49. Let F be a field. Then, $F[x]/(x) \cong F$.

Theorem 1.2 (second isomorphism theorem). Let R be a ring. Let A be a subring and B be an ideal of R. Then, define

$$A + B = \{a + b : a \in A, b \in B\} \quad \text{to be a subring of } R.$$

We have the following isomorphism:

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

Next, let $I \subseteq J \subseteq R$ be ideals of R. Then, we have

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

See Figure 1.3 for the commutative diagram.

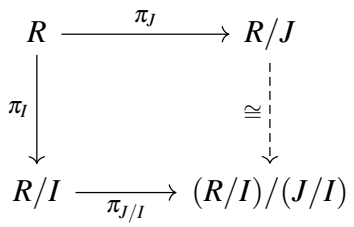


Figure 1.3: Second isomorphism theorem for rings

1.4 Properties of Ideals

Ideals are the basic congruence objects of ring theory: they are precisely the subsets of a ring that one may quotient by to form a new ring R/I, and they encode when two elements should be regarded as equivalent modulo I. This viewpoint makes ideals the ring-theoretic analogue of normal subgroups in group theory: just as normal subgroups are the kernels of group homomorphisms and quotient groups, ideals are the kernels of ring homomorphisms and quotient rings.

Definition 1.28 (smallest ideal). Let $A \subseteq R$ be a subset. Let (A) denote the smallest ideal in R generated by A, so

$$(A) = \bigcap_{\substack{A \subseteq J \\ J \text{ is an ideal of } R}} J = RAR = \left\{ \sum rar' : r, r' \in R, a \in A \right\}.$$

Definition 1.29 (left and right ideals). Let

$$RA = \left\{ \sum ra : r \in R, a \in A \right\} \quad \text{and} \quad AR = \left\{ \sum ar : r \in R, a \in A \right\}$$

to be the left ideal and right ideal generated by A respectively.

Definition 1.30 (principal ideal). An ideal generated by a single element x is called a principal ideal, denoted by (x).

Definition 1.31 (finitely generated ideal). An ideal that is generated by a finite set is a finitely generated ideal.

Example 1.50. $0 = (0)$ and $R = (1)$ are trivial examples of principal ideals.

Example 1.51. Recall Example 1.44 where we mentioned that (n) is a principal ideal.

Example 1.52. The ideal $(2, x) \subseteq \mathbb{Z}[x]$ is not principal.

Example 1.53. The ideal $\left(2x, 2x^2, 2x^3, \ldots\right) \subseteq 2\mathbb{Z}[x]$ is not finitely generated.

Lemma 1.5. Let R be a ring. For any ideal $I \subseteq R$,

$$I = R \quad \text{if and only if} \quad I \text{ contains a unit.}$$

Furthermore, if R is commutative, then

$$R \text{ is a field} \quad \text{if and only if} \quad R \text{ has only two ideals which are } 0 \text{ and } R.$$

Definition 1.32 (maximal ideal). Let $I \subseteq R$ be an ideal of R. I is maximal if

$$I \neq R \quad \text{and} \quad \text{for any ideal } J \text{ containing } I \text{ we have either } J = I \text{ or } J = R.$$

Lemma 1.6. Let R be a commutative ring. Then, the following hold:

(i) R is a field if and only if (0) is a maximal ideal

(ii) I is a maximal ideal of R if and only if R/I is a field

Note that **(i)** of Lemma 1.6 appears in Question 4 of p. 256 of Dummit and Foote [1]. We give a proof of this result.

Proof. We first prove the forward direction. Suppose R is a field. Then, as the only ideals of R are (0) and R, this implies there is no ideal strictly between (0) and R. The reverse direction follows from Lemma 1.5. □

Example 1.54. For any prime p, the ideal $(p) \subseteq \mathbb{Z}$ is maximal.

Example 1.55. The ideal $(x-a) \in \mathbb{C}[x]$ for any $a \in \mathbb{C}$ is a maximal ideal.

Proposition 1.7. Let R be a ring with multiplicative identity 1. Then, any ideal I is contained in a maximal ideal of R.

Proof. Use Zorn's lemma. □

Definition 1.33 (prime ideal). Let R be a commutative ring with multiplicative identity 1. An ideal $P \subseteq R$ is a prime ideal if

$$P \neq R \quad \text{and} \quad \text{for any } ab \in P \text{ either } a \in P \text{ or } b \in P.$$

Example 1.56. For any prime p, $(p) \subseteq \mathbb{Z}$ is a prime ideal.

Example 1.57. $(0) \subseteq \mathbb{Z}$ is a prime ideal.

Lemma 1.7. Let R be a commutative ring with multiplicative identity 1. Let $I \subseteq R$ be an ideal. Then,

$$I \text{ is a prime ideal} \quad \text{if and only if} \quad R/I \text{ is an integral domain.}$$

So, maximal ideals are also prime ideals.

Example 1.58 (Dummit and Foote p. 258 Question 23). Prove that in a Boolean ring (Definition 1.14), every prime ideal is a maximal ideal.

Solution. Recall from Definition 1.14 that in a Boolean ring R, $x^2 = x$ for all $x \in R$. Let $\mathfrak{p}$ be a prime ideal of R. We prove that $\mathfrak{p}$ is maximal. First, note that for all $x \in R$,

$$(x+\mathfrak{p})^2 = x+\mathfrak{p}$$

so every element of $R/\mathfrak{p}$ is idempotent. Hence, $R/\mathfrak{p}$ is a Boolean ring. Since $\mathfrak{p}$ is prime, by Lemma 1.7, $R/\mathfrak{p}$ is an integral domain.

Take $x \in R/\mathfrak{p}$. Since $R/\mathfrak{p}$ has no zero divisors (Definition 1.9), then $x = 0$ or $x = 1$, so $R/\mathfrak{p}$ has exactly two elements, implying that $R/\mathfrak{p} \cong \mathbb{F}_2$, the finite field with 2 elements. So, $R/\mathfrak{p}$ is a field. By Lemma 1.6, we conclude that $\mathfrak{p}$ is a maximal ideal. □

Definition 1.34 (nilradical). Let R be a commutative ring with multiplicative identity 1. Define the nilradical of R to be

$$\mathfrak{N}(R) = \text{set of nilpotent elements of } R = \{x \in r : x^n = 0 \text{ for some } n \in \mathbb{Z}_{\geq 0}\}.$$

Lemma 1.8. Let R be a commutative ring with multiplicative identity 1. Then, $\mathfrak{N}(R)$ is an ideal of R.

Proof. Let $x, y \in \mathfrak{N}(R)$. Then, there exist $n, m \in \mathbb{Z}_{\geq 0}$ such that $x^n = y^m = 0$. Note that for any $r \in R$, we have $(rx)^n = r^n x^n = 0$, where we used the fact that R is commutative, so $rx \in \mathfrak{N}(R)$. Now, let $l = 2\max\{m,n\}$, so

$$(x+y)^l = \sum_{i=0}^{l} \frac{l!}{i!\,(l-i)!} x^i y^{l-i} = 0$$

and it follows that $x+y \in \mathfrak{N}(R)$. As such, $\mathfrak{N}(R)$ is an ideal of R. □

1.5 Rings of Fractions

A recurring theme in algebra is that we often want to *divide* by non-zero elements, even when our ambient ring does not already allow this operation. For instance, in $\mathbb{Z}$, there is no element x satisfying $2x = 1$, yet in many computations it is natural to treat $\frac{1}{2}$ as a legitimate object. The construction of *rings of fractions* formalises this idea: we enlarge a ring by adjoining inverses of a chosen collection of elements, in a way that is canonical and respects the ring structure.

In the special case where R is an integral domain, we may invert every non-zero element. The resulting object is a field, called the field of fractions (or quotient field) of R. Conceptually, it is obtained by treating symbols $\frac{r}{d}$ with $r \in R$ and $d \in R \setminus \{0\}$ as fractions, and then imposing the expected identification

$$\frac{r}{d} = \frac{r'}{d'} \quad \text{if and only if} \quad rd' = r'd,$$

which captures the idea of cross-multiplying in a domain. After defining addition and multiplication in the familiar way, we obtain a field Q into which R embeds via $r \mapsto \frac{r}{1}$. Moreover, this field is universal among fields containing R: any field F that contains R must also contain a copy of Q. In this precise sense, Q is the *smallest* field in which the arithmetic of R can be carried out with division by non-zero elements.

Definition 1.35 (field of fractions). Let R be an integral domain. Let $D = R \setminus \{0\}$. Define

$$\widetilde{Q} = \{(r,d) : r \in R, d \in D\}.$$

Then, for any $d, d' \neq 0$, define an equivalence relation on $\widetilde{Q}$ via

$$(r,d) \sim (r',d') \quad \text{if and only if} \quad rd' = r'd.$$

Then, define $Q = \widetilde{Q}/\sim$ to be the field of fractions of R.

In Definition 1.35, we constructed the field of fractions of R, denoted by $Q = \widetilde{Q}/\sim$, where the equivalence relation $\sim$ means $(r,d) \sim (r',d')$ if and only if $rd' = r'd$. Intuitively, (r,d) is meant to represent the fraction r/d. However in Q, an element is an equivalence class so the same fraction has many representatives. To turn Q into a ring or a field, we must define addition and multiplication on these classes in a way that agrees with how fractions should add or multiply, and is well-defined. The latter means that addition and multiplication does not depend on which representatives we pick.

Theorem 1.3. Define addition on Q as follows:

$$\frac{r}{d} + \frac{s}{t} = \frac{rt + ds}{dt} \quad \text{which is well-defined} \tag{1.2}$$

Define multiplication on Q as follows:

$$\frac{r}{d} \cdot \frac{s}{t} = \frac{rs}{dt} \quad \text{which is well-defined} \tag{1.3}$$

Then, Q is a field.

Proof. We first prove (1.2). That is to say, addition is well-defined. Suppose

$$\frac{r}{d} = \frac{r'}{d'} \quad \text{and} \quad \frac{s}{t} = \frac{s'}{t'}. \tag{1.4}$$

As such, $rd' = r'd$ and $st' = s't$. We need to prove that

$$\frac{rt + ds}{dt} = \frac{r't' + d's'}{d't'}.$$

By definition of $\sim$ in Definition 1.35, it suffices to prove the cross multiplication identity $(rt + ds)(d't') = (r't' + d's')(dt)$. Expanding both sides and rewriting each term using the commutativity and and associativity of the elements in the integral domain (Definition 1.9), we have

$$rtd't' + dsd't' = r't'dt + d's'dt.$$

Thereafter, use the given equivalences in (1.4) to prove the mentioned equality.

Proving that multiplication is well-defined (1.3) is not that tedious. With the same assumptions $rd' = r'd$ and $st' = s't$, it is easy to prove that

$$\frac{rs}{dt} = \frac{r's'}{d't'}$$

via cross multiplication and commutativity of the elements in the integral domain. □

When we construct the field of fractions Q, we want Q to contain a copy of R so that elements of R can be viewed as fractions with denominator 1. Lemma 1.9 formalises this idea — it constructs a ring homomorphism ι and shows that it is injective (the hook arrow in Lemma 1.9 implies that the homomorphism is injective).

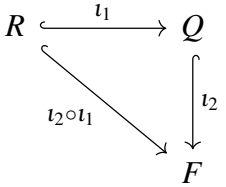
Lemma 1.9. There exists an embedding

$$\iota : R \hookrightarrow Q \quad \text{such that} \quad \iota(r) = \frac{r}{1} = \frac{rd}{d} \text{ for any } d \neq 0.$$

In Lemma 1.9, the equality $\frac{r}{1} = \frac{rd}{d}$ just expresses the familiar fact that multiplying the numerator and the denominator by the same non-zero element does not change a fraction.

Example 1.59. By applying Lemma 1.9 with $R = \mathbb{Z}$, one can easily deduce that $\mathbb{Q}$ is the field of fractions of $\mathbb{Z}$. By considering the equivalence relation $\frac{a}{b} \sim \frac{c}{d}$ if and only if $ad = bc$, the set of equivalence classes is $\mathbb{Q}$, and we define addition and multiplication as in (1.2) and (1.3).

Proposition 1.8. Let R be an integral domain with field of fractions Q. Let F be a field containing R. Then, F contains Q so Q is the smallest field containing R.



Example 1.60 (subfields of $\mathbb{R}$; Dummit and Foote p. 264 Question 4). Prove that any subfield of $\mathbb{R}$ must contain $\mathbb{Q}$.

Solution. Let $F \subseteq \mathbb{R}$ be a subfield. We shall prove that $\mathbb{Q} \subseteq F$. We first prove that $0, 1 \in F$. Since F is a field, then it must contain the additive identity 0 and the multiplicative identity 1, and by the subfield property, these are the same identity elements as in $\mathbb{R}$.

Next, we prove that all integers are in F. Since fields are closed under addition and additive inverses, as we know that $1 \in F$, then clearly $n \in F$ and $-n \in F$ for all $n \in \mathbb{N}$. As such, $\mathbb{Z} \subseteq F$.

Lastly, we prove that all rational numbers are in F, and consequently $\mathbb{Q} \subseteq F$. Take any $\frac{a}{b} \in \mathbb{Q}$, where $a, b \in \mathbb{Z}$ and $b \neq 0$. As we know that $a, b \in F$, by the property that multiplicative inverses exist, then $\frac{a}{b} \in F$ since $\frac{a}{b} = a \cdot b^{-1}$. □

In relation to Example 1.60, we shall briefly discuss some non-trivial examples of subfields of $\mathbb{R}$ that contain $\mathbb{Q}$. For example, we have the quadratic fields $\mathbb{Q}(\sqrt{2})$ which contain all numbers of the form $a + b\sqrt{2}$, where $a, b \in \mathbb{Q}$, and in general for any algebraic number $\alpha \in \mathbb{A}$, the set $\mathbb{Q}(\alpha)$ is a subfield of $\mathbb{R}$.

Example 1.61 (Dummit and Foote p. 264 Question 2). Let R be a commutative ring with identity $1 \neq 0$. Let R be an integral domain and let D be a non-empty subset of R that is closed under multiplication. Prove that the ring of fractions $D^{-1}R$ is isomorphic to a subring of the quotient field[7] of R (hence is also an integral domain).

Solution. We define

$$D^{-1}R = \{r/d : r \in R, d \in D\} / \sim \quad \text{where } \frac{r}{d} \sim \frac{r'}{d'} \text{ if and only if } rd' = r'd.$$

Let K denote the field of fractions of R. Consider the canonical inclusion map

$$\Phi : D^{-1}R \to K \quad \text{where} \quad \frac{r}{d} \mapsto \frac{r}{d} \text{ in } K$$

and it is clear that the map Φ is a well-defined injective ring homomorphism. As such, $D^{-1}R \cong \Phi\left(D^{-1}R\right) \subseteq K$, so $D^{-1}R$ is isomorphic to a subring of the field of fractions of R. □

1.6 The Chinese Remainder Theorem

We start with a motivating fact from MA1100 Basic Discrete Mathematics. Let $m, n \in \mathbb{Z}$ be coprime. By Bézout's lemma (and its converse), this is equivalent to saying that there exist $a, b \in \mathbb{Z}$ such that $am + bn = 1$. The new thing to take note of is that

$$(\mathbb{Z}/mn\mathbb{Z})^\times \cong (\mathbb{Z}/m\mathbb{Z})^\times \times (\mathbb{Z}/n\mathbb{Z})^\times$$

as an isomorphism of Abelian groups. This is precisely the Chinese remainder theorem. We will encounter another variant of the Chinese remainder theorem (for principal ideal domains) pretty soon in Theorem 2.5.

Definition 1.36 (coprime ideals). Let R be a commutative ring with multiplicative identity 1. Two ideals A and B of R are coprime if $A + B = R$.

One should note that Definition 1.36 is consistent with the definition in $\mathbb{Z}$.

Theorem 1.4 (Chinese remainder theorem). Let $A_1, \ldots, A_k$ be pairwise coprime

[7] Recall that this is the same as the field of fractions.

ideals of R (Definition 1.36). Then, we have the following isomorphism:

$$R/A_1 \dots A_k \to R/A_1 \times \dots \times R/A_k \quad \text{where} \quad r + A_1 \dots A_k \mapsto (r + A_1, \dots, r + A_k)$$

Proof. Let $A_1, \dots, A_k$ be ideals of R which are pairwise coprime. That is to say, $A_i + A_j = R$ for all $i \neq j$. Define the map

$$\varphi : R \to \prod_{i=1}^{k} R/A_i \quad \text{where} \quad r \mapsto r + A_1, \dots, r + A_k.$$

It is routine to check that φ is a ring homomorphism. We first identify the kernel. Note that

$$\ker \varphi = \{r \in R : r \in A_i \text{ for all } i\} = \bigcap_{i=1}^{k} A_i.$$

We claim that under the pairwise coprime hypothesis,

$$\bigcap_{i=1}^{k} A_i = \prod_{i=1}^{k} A_i = A_1 A_2 \cdots A_k.$$

The inclusion $A_1 \cdots A_k \subseteq \bigcap_{i=1}^{k} A_i$ always holds, since the product is contained in each factor. For the reverse inclusion, we first prove that for $k = 2$,

$$A_1 \cap A_2 = A_1 A_2 \quad \text{if } A_1 + A_2 = R.$$

Indeed, choose $x \in A_1$ and $y \in A_2$ with $x + y = 1$. If $r \in A_1 \cap A_2$, then $r = r \cdot 1 = r(x + y) = rx + ry$, where $rx \in A_2 A_1 \subseteq A_1 A_2$ (since $r \in A_2$ and $x \in A_1$) and $ry \in A_1 A_2$ (since $r \in A_1$ and $y \in A_2$). Hence $r \in A_1 A_2$, proving $A_1 \cap A_2 \subseteq A_1 A_2$.

Now proceed by induction on k. Let

$$B = A_1 A_2 \cdots A_{k-1}.$$

We claim $B + A_k = R$. For each $i < k$, pick $u_i \in A_i$ and $v_i \in A_k$ such that $u_i + v_i = 1$ (possible since $A_i + A_k = R$). Multiplying these identities gives

$$\prod_{i=1}^{k-1} (u_i + v_i) = 1.$$

Expanding the left hand side, the term $\prod_{i=1}^{k-1} u_i$ lies in $A_1 \cdots A_{k-1} = B$, and every other term contains at least one $v_i \in A_k$, hence lies in A_k. Therefore

$$1 \in B + A_k \quad \text{so} \quad B + A_k = R.$$

Using the $k = 2$ case for the coprime ideals B and A_k,

$$B \cap A_k = B A_k = A_1 \cdots A_{k-1} A_k.$$

Finally, by the induction hypothesis,

$$\bigcap_{i=1}^{k-1} A_i = A_1 \cdots A_{k-1} = B \quad \text{so} \quad \bigcap_{i=1}^{k} A_i = \left(\bigcap_{i=1}^{k-1} A_i \right) \cap A_k = B \cap A_k = A_1 \cdots A_k.$$

So, $\ker \varphi = A_1 \cdots A_k$.

We then prove that φ is surjective. Take some

$$(r_1 + A_1, \ldots, r_k + A_k) \in \prod_{i=1}^{k} R/A_i.$$

For each i, set

$$B_i = \prod_{\substack{1 \le j \le k \\ j \neq i}} A_j.$$

Since the A_i are pairwise coprime, one checks (as above) that $A_i + B_i = R$. Hence, there exist $e_i \in B_i$ and $f_i \in A_i$ with $e_i + f_i = 1$, so in particular

$$e_i \equiv 1 \pmod{A_i} \quad \text{and} \quad e_i \equiv 0 \pmod{A_j},$$

where $j \neq i$ because $e_i \in B_i \subseteq A_j$ for all $j \neq i$. Now, define

$$r = \sum_{i=1}^{k} r_i e_i \in R.$$

For a fixed index m, reducing modulo A_m gives

$$r \equiv r_m e_m + \sum_{i \neq m} r_i e_i \equiv r_m \cdot 1 + \sum_{i \neq m} r_i \cdot 0 \equiv r_m \pmod{A_m}.$$

So, the map is indeed surjective. By the first isomorphism theorem (Theorem 1.1), we indeed obtain the Chinese remainder theorem! □

Example 1.62 (Dummit and Foote p. 267 Question 2). Let R be a finite Boolean ring with identity $1 \neq 0$ (Definition 1.14). Use Example 1.45 to prove that

$$R \cong \mathbb{Z}/2\mathbb{Z} \times \cdots \times \mathbb{Z}/2\mathbb{Z}.$$

Solution. Let R be a finite Boolean ring with identity $1 \neq 0$, so $x^2 = x$ for all $x \in R$. We prove by induction on $|R|$ that

$$R \cong (\mathbb{Z}/2\mathbb{Z})^n \quad \text{for some } n \geq 1.$$

For the base case, if $|R| = 2$, then $R = \{0, 1\}$ as a set, and necessarily $1 + 1 = 0$ since the characteristic of R is $= 2$, so $R \cong \mathbb{Z}/2\mathbb{Z}$. As such, the base case holds.

Next, assume that the claim holds for all finite Boolean rings (with identity) of cardinality $< |R|$, and suppose $|R| > 2$. Then there exists an element $a \in R$ with $a \neq 0, 1$. Set

$e = a$. Since R is Boolean, $e^2 = e$, and since R is commutative, $er = re$ for all r. Thus, we may apply Example 1.45 to obtain a ring isomorphism

$$R \cong Re \times R(1-e) \quad \text{where} \quad r \mapsto (re, r(1-e)).$$

Moreover, by the same example, e is the identity of the subring Re and $1-e$ is the identity of the subring $R(1-e)$. We leave it to the reader that Re and $R(1-e)$ are finite Boolean rings (with identity). Finally, both factors are *proper* (hence strictly smaller) because $e \neq 0, 1$: indeed, Re has identity $e \neq 1$, so $Re \neq R$; similarly $R(1-e) \neq R$. Consequently,

$$|Re| < |R| \quad \text{and} \quad |R(1-e)| < |R|.$$

By the induction hypothesis, there exist integers $m, n \geq 1$ such that

$$Re \cong (\mathbb{Z}/2\mathbb{Z})^m \quad \text{and} \quad R(1-e) \cong (\mathbb{Z}/2\mathbb{Z})^n.$$

Therefore,

$$R \cong Re \times R(1-e) \cong (\mathbb{Z}/2\mathbb{Z})^m \times (\mathbb{Z}/2\mathbb{Z})^n \cong (\mathbb{Z}/2\mathbb{Z})^{m+n}.$$

This completes the induction, and hence every finite Boolean ring with identity is isomorphic to a finite direct product of copies of $\mathbb{Z}/2\mathbb{Z}$. □

A familiar version of the Chinese remainder theorem from Olympiad Mathematics states the following. Let $p_i \in \mathbb{Z}$ be distinct primes for $1 \leq i \leq k$. Then, there exists $x \in \mathbb{Z}$ unique mod $p_1 \dots p_k$ such that

$$x \cong x_i \pmod{p_i} \text{ for any } x_i \in \mathbb{Z}.$$

We will see how the Chinese remainder theorem in Theorem 1.4 is related to this version in Example 1.63.

Example 1.63 (Dummit and Foote p. 267 Question 5). Let $n_1, n_2, \dots, n_k$ be integers which are relatively prime in pairs: $(n_i, n_j) = 1$ for all $i \neq j$.

(a) Show that the Chinese remainder theorem (Theorem 1.4) implies that for any integers $a_1, \dots, a_k$, there is a solution $x \in \mathbb{Z}$ to the simultaneous congruences

$$x \equiv a_1 \pmod{n_1} \quad x \equiv a_2 \pmod{n_2} \quad \dots \quad x \equiv a_k \pmod{n_k},$$

and that the solution x is unique mod $n = n_1 n_2 \cdots n_k$.

(b) Let $n_i' = n/n_i$ be the quotient of n by n_i, which is relatively prime to n_i by assumption. Let t_i be the inverse of n_i' mod n_i. Prove that the solution x in **(a)** is given[8] by

$$x = a_1 t_1 n_1' + a_2 t_2 n_2' + \cdots + a_k t_k n_k' \pmod{n}.$$

[8]Note that the elements t_i can be quickly found by the Euclidean algorithm (writing $an_i + bn_i' = \gcd(n_i, n_i') = 1$ gives $t_i = b$) and that these then quickly give the solutions to the system of congruences above for any choice of $a_1, a_2, \dots, a_k$.

There was an original **(c)** in Example 1.63 of the book [1], but it is to solve two simultaneous systems of congruences. As this is something that has been discussed in Olympiad Mathematics, we will not delve into it here.

Solution.

(a) Since $\gcd(n_i, n_j) = 1$ for $i \neq j$, the ideals $(n_1), \ldots, (n_k)$ in $\mathbb{Z}$ are pairwise comaximal. That is, $(n_i) + (n_j) = \mathbb{Z}$ for $i \neq j$. By the Chinese remainder theorem (Theorem 1.4), we obtain an isomorphism

$$\mathbb{Z}/(n) \cong \mathbb{Z}/(n_1) \times \cdots \times \mathbb{Z}/(n_k) \quad \text{where } n = n_1 n_2 \cdots n_k.$$

Here, $(n) = (n_1) \cdots (n_k)$ because the ideals are pairwise comaximal in $\mathbb{Z}$.

Let $a_1, \ldots, a_k \in \mathbb{Z}$. Consider the element

$$(\overline{a_1}, \ldots, \overline{a_k}) \in \mathbb{Z}/(n_1) \times \cdots \times \mathbb{Z}/(n_k),$$

where $\overline{a_i}$ denotes the class of a_i modulo n_i. By surjectivity of the Chinese remainder theorem isomorphism, there exists $\bar{x} \in \mathbb{Z}/(n)$ mapping to $(\overline{a_1}, \ldots, \overline{a_k})$. Choosing a representative $x \in \mathbb{Z}$ of $\bar{x}$, this means

$$x \equiv a_i \pmod{n_i} \quad \text{for all } i = 1, \ldots, k,$$

so a solution exists. For uniqueness modulo n, suppose x and y are two integer solutions. Then for each i, $x \equiv a_i \equiv y \pmod{n_i}$ implies $n_i \mid (x - y)$. Since the n_i are pairwise coprime, their product divides $x - y$ so $x \equiv y \pmod{n}$.

(b) Fix $i \in \{1, \ldots, k\}$. Let $n_i' = n/n_i$. By assumption, $\gcd(n_i, n_i') = 1$, so n_i' is invertible modulo n_i. Let t_i be an inverse of n_i' modulo n_i, i.e. $t_i n_i' \equiv 1 \pmod{n_i}$. Define

$$x = \sum_{j=1}^{k} a_j t_j n_j' \in \mathbb{Z}.$$

We claim that x satisfies the system of congruences modulo each n_i. Reduce x modulo n_i. For $j \neq i$, we have $n_j' = \frac{n}{n_j}$ is divisible by n_i (since n_i is a factor of n and $n_i \neq n_j$), hence

$$n_i \mid n_j' \quad \text{so} \quad a_j t_j n_j' \equiv 0 \pmod{n_i} \text{ where } j \neq i.$$

Therefore,

$$x \equiv a_i t_i n_i' \pmod{n_i}.$$

Using $t_i n_i' \equiv 1 \pmod{n_i}$ gives $x \equiv a_i \pmod{n_i}$. Since this holds for each i, x is a simultaneous solution. Finally, any two solutions are congruent modulo n by **(a)**, so the congruence class of x modulo n is the unique solution as mentioned. □

Chapter 2

Euclidean Domains, Principal Ideal Domains, and Unique Factorization Domains

2.1 Euclidean Domains

A recurring theme in algebra is that many powerful theorems ultimately come from being able to divide with remainder. In the integers $\mathbb{Z}$, this is the familiar division algorithm: given integers a and $b \neq 0$, we can write $a = bq + r$ with $0 \leq r < |b|$. This simple statement implies the Euclidean algorithm, hence the existence of greatest common divisors, and even Bézout's identity.

Euclidean domains (Definition 2.2) are integral domains in which an analogous division algorithm is available, not necessarily using the usual absolute value, but using a norm that decreases on remainders. Once such a division is possible, the same mechanism that works in $\mathbb{Z}$ works again: we can run the Euclidean algorithm, compute greatest common divisors, and obtain nice consequences. In particular, Euclidean domains turn out to be principal ideal domains (Theorem 2.2), and hence enjoy unique factorisation properties (Theorem 2.4).

Definition 2.1 (norm). Let R be an integral domain. A norm on R is a function

$$N : R \to \mathbb{Z}_{\geq 0} \quad \text{such that} \quad N(0) = 0.$$

Definition 2.2 (Euclidean domain). Let R be an integral domain. We say that R is a Euclidean domain (ED) if we can perform the following division algorithm with respect to some norm N, which is for any $a, b \in R$ with $b \neq 0$, there exist $q, r \in R$ such that

$$a = bq + r \quad \text{with } r = 0 \text{ or } N(r) < N(b).$$

q is called the quotient and r is called the remainder.

Example 2.1. $\mathbb{Z}$ is a Euclidean domain with $N(r) = |r|$. Then, the division is exactly the division on $\mathbb{Z}$.

Example 2.2. The polynomial ring $\mathbb{R}[x]$ (or over any field) is a Euclidean domain with $N(f) = \deg(f)$.

Note that any field is a Euclidean domain with any norm, so the choice of the norm is not unique. The reader can attempt Examples 2.3 and 2.4, which are some simple questions involving the Euclidean algorithm and finding the inverse of a number modulo n.

Example 2.3 (Dummit and Foote p. 277 Question 1). For each of the following five pairs of integers a and b, determine their greatest common divisor d and write d as a linear combination $ax+by$ of a and b.

(a) $a = 20, b = 13$

(b) $a = 69, b = 372$

(c) $a = 11391, b = 5673$

(d) $a = 507885, b = 60808$

(e) $a = 91442056588823,\ b = 779086434385541$ (the Euclidean algorithm requires only 7 steps for these integers)

Example 2.4 (Dummit and Foote p. 277 Question 2). For each of the following pairs of integers a and n, show that a is relatively prime to n and determine the inverse of a mod n.

(a) $a = 13, n = 20$

(b) $a = 69, n = 89$

(c) $a = 1891, n = 3797$

(d) $a = 6003722857, n = 77695236973$ (the Euclidean algorithm requires only 3 steps for these integers)

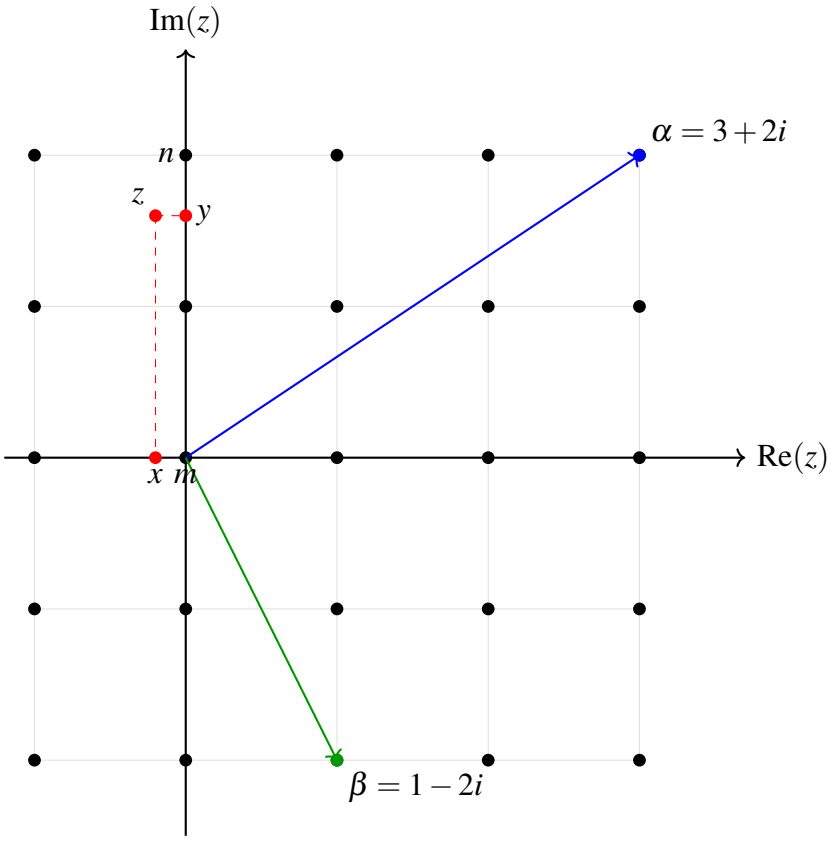
Proposition 2.1 (Gaussian integers forms a Euclidean domain). The ring of Gaussian integers $\mathbb{Z}[i] \subseteq \mathbb{C}$ equipped with the standard norm $N(a+bi) = a^2+b^2$ is a Euclidean domain.

Proof. We wish to show the division algorithm in $\mathbb{Z}[i]$ with the norm $N(a+bi) = a^2+b^2$ holds. This means that for any $\alpha, \beta \in \mathbb{Z}[i]$ with $\beta \neq 0$, we must find $q, r \in \mathbb{Z}[i]$ such that

$$\alpha = \beta q + r \quad \text{and} \quad r = 0 \text{ or } N(r) < N(\beta).$$

Since $\beta \neq 0$, we can form the quotient $z = \frac{\alpha}{\beta}$. Write $z = x+yi$, where $x, y \in \mathbb{Q}$. Choose $m, n \in \mathbb{Z}$ so that m is the nearest integer to x and n is the nearest integer to y. So, we need

to ensure that $|x-m| \leq \frac{1}{2}$ and $|y-n| \leq \frac{1}{2}$. Let $q = m+ni \in \mathbb{Z}[i]$.



Now, write z as $z = q + (z-q)$. Multiplying both sides by β yields $\alpha = \beta q + \beta(z-q)$. Setting $r = \beta(z-q)$, it just remains to show that $r \in \mathbb{Z}[i]$ and we need to bound $N(r)$ too. As $r = \alpha - \beta q$, where $\alpha, \beta, q \in \mathbb{Z}[i]$, then the first claim holds.

To bound $N(r)$, we recall that

$$\begin{aligned} z-q &= (x-m) + (y-n)\,i \\ |z-q|^2 &= (x-m)^2 + (y-n)^2 \end{aligned}$$

which is bounded above by $\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 = \frac{1}{2}$. Sine the standard norm on $\mathbb{Z}[i]$ is multiplicative, then

$$N(r) = N(\beta)N(z-q) = N(\beta)\,|z-q|^2 \leq \frac{1}{2}N(\beta) \leq N(\beta)$$

where in the last step, we used the fact that $N(\beta)$ being a non-negative integer and $\beta \neq 0$ implies $N(\beta) \geq 1$. So, the Euclidean condition as in Definition 2.2 holds. One notes that the quotient and remainder are not unique in this case. □

Example 2.5 (Dummit and Foote p. 278 Question 9). Prove that the ring of integers $\mathcal{O}$ in the quadratic integer ring $\mathbb{Q}(\sqrt{2})$ is a Euclidean domain with respect to the norm given by the absolute value of the field norm N as follows:

$$N : \mathbb{Q}(\sqrt{2}) \to \mathbb{Q} \quad \text{where} \quad N\left(a+b\sqrt{D}\right) = a^2 - 2b^2 \in \mathbb{Q}$$

Solution. The proof will be similar to that in Proposition 2.1. For any $\alpha, \beta \in \mathbb{Z}[\sqrt{2}]$ with $\beta \neq 0$, by the division algorithm, there exist $q, r \in \mathbb{Z}[\sqrt{2}]$ such that

$$\alpha = \beta q + r \quad \text{where} \quad r = 0 \text{ or } |N(r)| < |N(\beta)|.$$

This will prove that $\mathbb{Z}[\sqrt{2}]$ is Euclidean with Euclidean function $\delta(\gamma) = |N(\gamma)|$. Set $\gamma = \frac{\alpha}{\beta} \in \mathbb{Q}(\sqrt{2})$. Then, there exist $x, y \in \mathbb{Q}$ such that $\gamma = x + y\sqrt{2}$. Choose integers $m, n \in \mathbb{Z}$ such that $|x - m| \leq \frac{1}{2}$ and $|y - n| \leq \frac{1}{2}$. Then, we define

$$q = m + n\sqrt{2} \in \mathbb{Z}[\sqrt{2}] \quad \text{and} \quad \varepsilon = \gamma - q = (x - m) + (y - n)\sqrt{2}.$$

Let $u = x - m$ and $v = y - n$. Then, $|u| \leq \frac{1}{2}$ and $|v| \leq \frac{1}{2}$, and $\varepsilon = u + v\sqrt{2}$ so $N(\varepsilon) = u^2 - 2v^2$. As such, the absolute value of the norm of ε is bounded by 1. Since the norm function is multiplicative, one can then deduce that either $r = 0$ or $|N(r)| < |N(\beta)|$, which is exactly the Euclidean condition. □

2.2 Principal Ideal Domains

In Chapter 2.1, we isolated a strong computational property of certain integral domains: the existence of a division algorithm for Euclidean domains (Definition 2.2). The payoff of such a division algorithm is not merely that one can compute greatest common divisors efficiently; rather, it forces a remarkably rigid ideal structure. This motivates the next class of rings.

Recall that in $\mathbb{Z}$, every ideal is of the form (n) for some integer n. In other words, ideals in $\mathbb{Z}$ are generated by a single element, and the generator can be chosen in a canonical way (up to sign). This simple description underlies many familiar facts: gcds exist, gcds can be expressed as integer linear combinations (Bézout's lemma), and primes control factorisation. A principal ideal abstracts precisely this phenomenon: it is an integral domain in which every ideal is generated by one element (Definition 2.3).

> Definition 2.3 (principal ideal domain). A principal ideal domain (PID) is an integral domain in which every ideal is a principal ideal (Definition 1.30).

Example 2.6. For any field F, it is a PID. To see why, note that by Definition 2.3, F is an integral domain, so if $ab = 0$ in F and $a \neq 0$, then a has an inverse a^{-1} in F, so $ab = 0$ implies $b = 0$, hence F has no zero divisors.

Now, let I be an ideal of F. If $I = \{0\}$, then $I = (0)$ is principal since it is generated by a single element (Definition 1.30). Suppose otherwise, then we can pick $a \in I$ for some non-zero a. Since F is a field, then a is a unit so for any $x \in F$, we have $x = xa^{-1} \cdot a$. Since $a \in I$ and I is an ideal, then $x \in I$ so $I = F$. So, the only ideals of F are $\{0\}$ and F, and both are principal because $\{0\} = (0)$ and $F = (1)$.

Example 2.7 ($\mathbb{Z}$ is a PID). The ring of integers $\mathbb{Z}$ is a PID. It is also an ED from Example 2.1. We will appreciate a nice relation in Theorem 2.2 which states that every ED is a PID.

Example 2.8 (Dummit and Foote p. 283 Question 5). Let R be the quadratic integer ring $\mathbb{Z}[\sqrt{-5}]$. Define the ideals

$$I_2 = (2, 1+\sqrt{-5}) \quad \text{and} \quad I_3 = (3, 2+\sqrt{-5}) \quad \text{and} \quad I_3' = (3, 2-\sqrt{-5}).$$

(a) Prove that I_2, I_3, I_3' are non-principal ideals in R.

(b) Prove that the product of two non-principal ideals can be principal by showing that I_2^2 is the principal ideal generated by 2. That is, $I_2^2 = (2)$.

(c) Prove similarly that $I_2 I_3 = (1-\sqrt{-5})$ and $I_2 I_3' = (1+\sqrt{-5})$ are principal. Conclude that the principal ideal (6) is the product of 4 ideals. That is, $(6) = I_2^2 I_3 I_3'$.

Solution.

(a) We first prove that I_2 is non-principal. Suppose on the contrary that I_2 is principal. Then by Definition 1.30, there exists $a \in \mathbb{Z}[\sqrt{-5}]$ such that $I_2 = (a)$. Since $2 \in I_2$, then $a \mid 2$. A similar argument shows that $a \mid (1+\sqrt{-5})$. Let

$$N : R \to \mathbb{Z}_{\geq 0} \quad \text{where} \quad N(x+y\sqrt{-5}) = x^2 + 5y^2 \tag{2.1}$$

denote the norm function in $\mathbb{Z}[\sqrt{-5}]$. Then, $N(a) \mid N(2) = 4$ and $N(a) \mid N(1+\sqrt{-5}) = 6$.

As such, either $N(a) = 1$ or $N(a) = 2$. If $N(a) = 1$, then a is a unit so $(a) = R$, contradicting the fact that $I_2 \neq R$. On the other hand, if $N(a) = 2$, then there exist $x, y \in \mathbb{Z}$ such that $x^2 + 5y^2 = 2$, which is impossible. It follows that I_2 is not a principal ideal.

We then prove that I_3 is non-principal. Again, we proceed by contradiction. If I_3 was principal, then by Definition 1.30, there exists $b \in \mathbb{Z}[\sqrt{-5}]$ such that $I_3 = (b)$. Using the same argument as above, consider the norm function in (2.1). Then, $N(b) \mid 9$, so either $N(b) = 1, 3, 9$. As before, if $N(b) = 1$, it is impossible. If $N(b) = 3$, we consider the Diophantine equation $x^2 + 5y^2 = 3$, which has no integer solutions; if $N(b) = 9$, then again writing $b = x + y\sqrt{-5}$ yields $x^2 + 5y^2 = 9$. Reducing modulo 5, we get $x^2 \equiv 9 \equiv 4 \pmod 5$, hence $x \equiv \pm 2 \pmod 5$ (in particular, $x \neq \pm 3$). Moreover, if $y = 0$ then $x^2 = 9$, so $x = \pm 3$, contradicting $x \equiv \pm 2 \pmod 5$. If $y = \pm 1$, then $x^2 = 4$, so $x = \pm 2$, and then $b = \pm(2+\sqrt{-5})$, which forces

$$(b) = (2+\sqrt{-5}).$$

But $(2+\sqrt{-5})$ does not contain 3 (equivalently, $2+\sqrt{-5} \nmid 3$), contradicting $3 \in I_3 = (b)$. Therefore $N(b) \neq 9$ as well. This contradiction shows that I_3 is not principal.

In a similar fashion to proving I_3 is not a principal ideal, one can repeat it for I_3'.

(b) As mentioned, it suffices to prove that $I_2^2 = (2)$. By definition, I_2^2 is generated by all products of generators of I_2, so

$$I_2^2 = (4, 2(1+\sqrt{-5}), (1+\sqrt{-5})^2) = (4, 2+2\sqrt{-5}, -4+2\sqrt{-5}).$$

We can factor out 2 which yields $I_2^2 = (2)R$, where $R = \mathbb{Z}[\sqrt{-5}]$. By the absorption property for ideals (Definition 1.24), $I_2^2 = (2)$.

(c) We have

$$\begin{aligned} I_2 I_3 &= (2, 1+\sqrt{-5})(3, 2+\sqrt{-5}) \\ &= (6, 3+3\sqrt{-5}, 4+2\sqrt{-5}, -3+3\sqrt{-5}) \\ &= (3+3\sqrt{-5}, 4+2\sqrt{-5}, -3+3\sqrt{-5}) \end{aligned}$$

and the result follows. A similar computation shows that $I_2 I_3' = (1+\sqrt{-5})$. As such,

$$I_2^2 I_3 I_3' = I_2 I_3 I_2 I_3' = (1-\sqrt{-5})(1+\sqrt{-5}) = (6)$$

showing that (6) is the product of 4 non-principal ideals. □

Definition 2.4 (multiple and greatest common divisor). Let R be a commutative ring and $a, b \in R$ with $b \neq 0$.

(i) a is a multiple of b or b is a divisor of a if there exists $x \in R$ such that $a = xb$. We write $b \mid a$.

(ii) A greatest common divisor of a and b is a non-zero element d such that

$$d \mid a \text{ and } d \mid b \quad \text{and} \quad \text{for any } d' \text{ which divides both } a \text{ and } b \text{ we have } d' \mid d.$$

Lemma 2.1. Let R be an integral domain. If d, d' are both greatest common divisors of a and b, then $d' = ud$ for some unit u.

Proof. Since $d = \gcd(a,b)$ and $d' \mid a$ and $d' \mid b$, by Definition 2.4, $d' \mid d$. By a symmetric argument, $d \mid d'$. Hence, there exist $r, s \in R$ such that $d' = rd$ and $d = sd'$. As such, $d = (sr)d$, which implies $(1-sr)d = 0$. Since R is an integral domain and $d \neq 0$, it forces $sr = 1$. As such , r is a unit so $d' = rd = ud$ for some unit u. □

Proposition 2.2. Let R be a PID. Let a, b be non-zero such that $(a,b) = (d)$. Then, $d = \gcd(a,b)$. Hence, the gcd always exists and it is of the form $ax + by$.

Proposition 2.3. Every non-zero prime ideal in a PID is a maximal ideal.

Definition 2.5 (Noetherian ring). A commutative ring R is said to be Noetherian if it satisfies the following ascending chain condition on ideals. That is, if $I_1 \subseteq I_2 \subseteq \cdots$ is a chain of ideals of R, then there exists $m \in \mathbb{N}$ such that $I_k = I_m$ for all $k \geq m$. This is equivalent to saying that

$$\bigcup_{i=1}^{\infty} I_i = I_m.$$

Many rings that we have encountered are Noetherian. Take $\mathbb{Z}$ and $\mathbb{Z}[x]$ for example.

Example 2.9 ($\mathbb{Z}$ is Noetherian). Let $I_1 \subseteq I_2 \subseteq I_3 \subseteq \cdots$ be an ascending chain of ideals in $\mathbb{Z}$. If every $I_k = (0)$ then the chain stabilises, so assume some $I_k \neq (0)$. For each k, since every ideal of $\mathbb{Z}$ is principal (Example 2.7), write $I_k = (a_k)$ with $a_k \geq 0$. Then,

$$(a_k) \subseteq (a_{k+1}) \quad \text{if and only if} \quad a_{k+1} \mid a_k.$$

Hence, $(a_1) \subseteq (a_2) \subseteq \cdots$ implies the divisibility chain a_1 is divisible by a_2 is divisible by a_3 and so on with $a_k \geq 0$.

Now, consider the set $S = \{a_k : k \in \mathbb{N} \text{ and } a_k \neq 0\}$, which is $\subseteq \mathbb{N}$. It is a non-empty subset of $\mathbb{N}$ so by the well-ordering principle, it has a least element, say a_m. We claim the chain stabilises at m. For $k \geq m$, we have $(a_m) \subseteq (a_k)$, so $a_k \mid a_m$. But also $a_m \mid a_k$ (since a_m is minimal among the positive generators appearing in the chain, and a_k cannot be a smaller positive divisor of a_m). Thus every ascending chain of ideals in $\mathbb{Z}$ stabilises, so $\mathbb{Z}$ is Noetherian. For example,

$$(16) \subseteq (8) \subseteq (4) \subseteq (2) \subseteq (1) = \mathbb{Z}$$

is an ascending chain of ideals in $\mathbb{Z}$, which stabilises.

Example 2.10 ($\mathbb{Z}$ adjoin x is Noetherian). First, note that $\mathbb{Z}$ adjoin x is simply $\mathbb{Z}[x]$ — I could not write the adjoin symbol due to a technical error. It is a well-known fact by the Hilbert basis theorem (out of scope of this course) that $\mathbb{Z}[x]$ is Noetherian. In fact, in general, if R is a Noetherian ring, then $R[x]$ is Noetherian.

Theorem 2.1. Every PID is Noetherian.

Proof. Let $I_1 \subseteq I_2 \subseteq \cdots$ be an ascending chain of ideals of R, and let I denote their union. By Example 1.47, I is an ideal of R. Since R is a PID, I is principal, so by Definition 2.3, $I = (d)$ for some $d \in R$. As $d \in I$, then there exists $m \in \mathbb{N}$ such that $d \in I_m$. so, $(d) = I \subseteq I_m \subseteq I$. Note that the inclusion $(d) \subseteq I_m$ holds since I_m is an ideal containing d, so $I_m = I$.

For all $k \geq m$, we have $I_m \subseteq I_k \subseteq I$, but as $I = I_m$, then $I_k = I_m$ for all $k \geq m$. This shows that the chain stabilises and R satisfies the ascending chain condition on ideals[1], implying that R is Noetherian. □

[1] The proof is similar to our discussion in Example 2.9.

Theorem 2.2. Every ED is a PID.

Proof. Let R be an ED and I be a non-zero ideal of R. By Definition 2.3, we need to prove that R is generated by a single element. Among the non-zero elements of I, let b be such that $d(b)$ is minimum among all elements from I. Then, we shall prove that $I = (b)$. It is clear that $(b) \subseteq I$. For $a \in I$, by Definition 2.2,

$$\text{there exist } q, r \in R \text{ such that } a = bq + r \quad \text{where} \quad r = 0 \text{ or } d(r) < d(b).$$

Note that $r = a - bq \in I$. $d(r)$ cannot be less than $d(b)$, otherwise it would contradict the minimality of $d(b)$. So, $r = 0$. This shows that $a \in (b)$, so $I \in (b)$. Hence, $I = (b)$. □

Theorem 2.3 (Euclidean algorithm). Let R be a Euclidean domain and $a, b \in R$ such that both are non-zero. Then, one can use the Euclidean algorithm to compute $\gcd(a, b)$.

We will not discuss how the Euclidean algorithm works as the reader should have prior knowledge of it from MA1100 Basic Discrete Mathematics. Now, recall how factorisation works in $\mathbb{Z}$. We wish to generalise this idea to more general rings.

Definition 2.6 (irreducibles and units). Let R be an integral domain. Suppose $r \in R$ is non-zero and is not a unit. Then,

$$r \text{ is irreducible over } R \quad \text{if} \quad r = ab \text{ with } a, b \in R \text{ implies either } a \text{ or } b \text{ is a unit.}$$

Otherwise, r is reducible.

Example 2.11 (Dummit and Foote p. 278 Question 3). Let R be a Euclidean domain. Let m be the minimum integer in the set of norms of non-zero elements of R. Prove that every non-zero element of R of norm m is a unit. Deduce that a non-zero element of norm zero (if such an element exists) is a unit.

Solution. Let a be a non-zero element of R of norm m. Then, $a \neq 0$ such that $N(a) = m$. The trick is to apply the division algorithm to 1 by a: there exist $q, r \in R$ such that

$$1 = aq + r \quad \text{where} \quad r = 0 \text{ or } N(r) < N(a) = m.$$

If $r = 0$, then $aq = 1$ and we are done because a is a unit. Suppose otherwise. Then, we have $N(r) < N(a) = m$, then $N(r) \geq m$ by minimality of m among norms of non-zero elements, contradicting $N(r) < m$.

For the deduction: if there exists a non-zero element $u \in R$ with $N(u) = 0$, then 0 is the minimum norm among non-zero elements (since norms are ≥ 0), i.e. $m = 0$. By the first part, every non-zero element of norm $m = 0$ is a unit. In particular, u is a unit. □

Definition 2.7 (associate). Let R be an integral domain. If

$$a = ub \text{ where } u \text{ is a unit of } R \quad \text{then} \quad a \text{ and } b \text{ are associates.}$$

In $\mathbb{Z}$, the notion of irreducibility exactly matches the usual primes up to sign: an integer is irreducible in $\mathbb{Z}$ if and only if it is of the form $\pm p$ where $p > 1$ is a prime number. Also in $\mathbb{Z}$, the only units are ± 1. Thus a and b are associates in $\mathbb{Z}$ precisely when $a = \pm b$, i.e. when they differ only by multiplication by a unit. This explains why we often treat p and $-p$ as the same prime.

Example 2.12 (Dummit and Foote p. 282 Question 4). Let R be an integral domain. Prove that if the following two conditions hold then R is a PID:

(i) any two non-zero elements a and b in R have a greatest common divisor which can be written in the form $ra + sb$ for some $r, s \in R$, and

(ii) if $a_1, a_2, a_3, \ldots$ are non-zero elements of R such that $a_{i+1} \mid a_i$ for all i, then there exists $N \in \mathbb{N}$ such that a_n is a unit times a_N for all $n \geq N$.

Solution. Let I be a non-zero ideal of R. Choose $a_1 \in I$ with $a_1 \neq 0$. If $I = (a_1)$, then we are done. Otherwise choose $a_2 \in I \setminus (a_1)$. Let $d_2 = \gcd(a_1, a_2)$, and by **(i)** write $d_2 = r_2 a_1 + s_2 a_2$. Then $d_2 \in I$, and so $(a_1, a_2) = (d_2)$. Since $a_2 \notin (a_1)$, we have a strict containment

$$(a_1) \subsetneq (a_1, a_2) = (d_2).$$

Also $d_2 \mid a_1$, so $a_1 = d_2 c_1$ for some $c_1 \in R$. Proceed inductively. Suppose we have chosen $a_1, \ldots, a_k \in I$ and defined

$$(d_k) = (a_1, \ldots, a_k) \subseteq I$$

with $d_k \neq 0$. If $I = (d_k)$, we stop. Otherwise choose $a_{k+1} \in I \setminus (d_k)$. Let $d_{k+1} = \gcd(d_k, a_{k+1})$, and by **(i)** write

$$d_{k+1} = r_{k+1} d_k + s_{k+1} a_{k+1} \in I.$$

So, $(d_k, a_{k+1}) = (d_{k+1})$. Since $a_{k+1} \notin (d_k)$, we get strict containment

$$(d_k) \subsetneq (d_k, a_{k+1}) = (d_{k+1}).$$

Moreover, $d_{k+1} \mid d_k$. Thus, we have constructed a sequence of non-zero elements $d_1 = a_1, d_2, d_3, \ldots$ such that

$$d_{k+1} \mid d_k \quad \text{and} \quad (d_k) \subsetneq (d_{k+1}) \text{ for all } k \text{ as long as we do not stop.}$$

If this process never stopped, condition **(ii)** would apply to the divisibility chain $d_{k+1} \mid d_k$ and yield an N such that d_n is an associate of d_N for all $n \geq N$. But associates generate the same principal ideal, so $(d_n) = (d_N)$ for all $n \geq N$, contradicting the strict containments $(d_k) \subsetneq (d_{k+1})$. Hence the process *must* stop. That is, for some N we have $I = (d_N)$, so I is principal. □

Definition 2.8 (prime ideal). Let R be an integral domain. A non-zero element $p \in R$ is called a prime in R if (p) is a prime ideal. In other words,

$$p \mid ab \quad \text{implies} \quad p \mid a \text{ or } p \mid b$$

We know that $\mathbb{Z}$ is an integral domain by Example 1.12 so Definition 2.8 is in fact Euclid's lemma! Moreover, the remarkable fact here is that irreducible and prime *feels the same* but they are actually different. Without the abstractions coined in Definitions 2.6 and 2.8, this would be difficult to distinguish.

Example 2.13 (Dummit and Foote p. 282 Question 3). Prove that a quotient of a PID by a prime ideal is again a PID.

Solution. Let R be a PID. Since R is a PID, every ideal of R is principal, hence the prime ideal $\mathfrak{p}$ is of the form $\mathfrak{p} = (p)$ for some $p \in R$. We first prove that $R/\mathfrak{p}$ is an integral domain. Actually this follows from Lemma 1.7.

We then prove that every prime ideal of $R/\mathfrak{p}$ is principal. Let $\pi : R \to R/\mathfrak{p}$ be the canonical projection, where $\pi(r) = \overline{r} = r + \mathfrak{p}$. Let J be any ideal of $R/\mathfrak{p}$. Consider its preimage

$$I = \pi^{-1}(J) = \{r \in R : \overline{r} \in J\}.$$

Then I is an ideal of R and contains $\ker(\pi) = \mathfrak{p}$. Since R is a PID, I is principal so $I = (a)$ for some $a \in R$. We claim that J is generated by $\overline{a}$ in the quotient. First, $\overline{a} \in J$ because $a \in I = \pi^{-1}(J)$, hence $(\overline{a}) \subseteq J$. Conversely, take any $\overline{x} \in J$. Then $x \in I = (a)$, so $x = ar$ for some $r \in R$. Passing to the quotient gives

$$\overline{x} = \overline{ar} = \overline{a}\,\overline{r} \in (\overline{a}).$$

Thus $J \subseteq (\overline{a})$, hence $J = (\overline{a})$ is principal. □

Lemma 2.2. In an integral domain, a prime element is always irreducible.

Proof. Let p be a prime and suppose $p = ab$. Then, $pa' = a$ or $pb' = b$. Suppose the former holds. That is, $pa' = a$. Then, $a = pa' = aba'$, which implies $ba' = 1$, so b is a unit (Definition 1.8). □

Proposition 2.4. In a PID, a non-zero element is prime if and only if it is irreducible.

Example 2.14. We shall verify that $1 + \sqrt{-3}$ is irreducible but not prime in $\mathbb{Z}\left[\sqrt{-3}\right]$.

Solution. Note that $N\left(a + b\sqrt{-3}\right) = a^2 + 3b^2$. We first prove that $1 + \sqrt{-3}$ is irreducible in $\mathbb{Z}\left[\sqrt{-3}\right]$. Suppose $1 + \sqrt{-3}$ can be factorised as xy, where neither x nor y is a unit. Then,

$$N(xy) = N(x)N(y) = 4 \quad \text{which implies} \quad N(x) = N(y) = 2.$$

However, there are no integers x and y satisfying $a^2 + 3b^2 = 2$, which implies that either x or y is a unit and so $1 + \sqrt{-3}$ is an irreducible.

Now, we prove that $1+\sqrt{-3}$ is not a prime. We see that

$$\left(1+\sqrt{-3}\right)\left(1+\sqrt{-3}\right)=4=2\cdot 2$$

so $\left(1+\sqrt{-3}\right) \mid 4$. Suppose $\left(1+\sqrt{-3}\right) \mid 2$. Then, there exist $a,b\in\mathbb{Z}$ such that

$$\left(1+\sqrt{-3}\right)\left(a+b\sqrt{-3}\right)=2 \quad \text{so} \quad a-3b+(a+b)\sqrt{-3}=2.$$

Hence, $a-3b=2$ and $a+b=0$ but no integers satisfy these two equations. □

Example 2.15. Verify that 7 is irreducible over $\mathbb{Z}[\sqrt{5}]$.

Solution. Suppose on the contrary that $7=xy$, where neither x nor y is a unit. Then, $N(7)=N(x)N(y)=49$. As neither x nor y is a unit, then $N(x)=7$. Suppose $x=a+b\sqrt{5}$. Then, there exist $a,b\in\mathbb{Z}$ such that $\left|a^2-5b^2\right|=7$, i.e. $a^2-5b^2=7$ or $a^2-5b^2=-7$. Working in modulo 7, it is easy to see that the only solution is $(a,b)=(0,0)$. However, this implies that a and b are divisible by 7. We write $a=7r$ and $b=7s$, where $r,s\in\mathbb{Z}$. Then, $x=7\left(r+s\sqrt{5}\right)$, which implies that $N(x)$ is divisible by 49, which is a contradiction. □

2.3 Unique Factorisation Domains

A recurring theme in commutative algebra is that the arithmetic of a ring is encoded by how its elements factor. Over $\mathbb{Z}$, we have the familiar fundamental theorem of arithmetic: every non-zero non-unit factors into primes, and this factorisation is unique up to signs and ordering. The goal of this section is to isolate exactly the ring-theoretic hypotheses under which an analogous theory holds, and then to use these hypotheses as a tool for concrete computations.

The central notion is that of a *unique factorisation domain* (UFD): an integral domain in which every non-zero non-unit admits a factorisation into irreducibles, and where any two such factorizations agree up to associates and permutation. In a general integral domain, *irreducible* and *prime* are distinct concepts; a key structural fact is that in a UFD they coincide. This is precisely what makes the usual divisibility arguments work: once irreducibles are prime, one can prove cancellation-type statements (e.g. Euclid's lemma) and develop a clean gcd theory in terms of prime-power exponents.

Definition 2.9 (unique factorisation domain). A unique factorisation domain (UFD) is an integral domain R in which every non-zero element $r\in R$ which is not a unit has the following two properties:

(i) r can be written as a finite product of irreducibles p_i of R, i.e.

$$r=p_1\dots p_n$$

(ii) the decomposition $r = p_1 \dots p_n$ is unique up to multiplication by units and permutation, i.e. if there exist irreducibles q_i such that $r = q_1 \dots q_m$, then $m = n$ and $q_i = p_i$ after relabelling

Example 2.16. We can uniquely decompose $6 = 2 \cdot 3 = (-2) \cdot (-3)$ in $\mathbb{Z}$.

Proposition 2.5. In a UFD, a non-zero element is prime if and only if it is irreducible.

Proposition 2.6. Let a and b be two non-zero elements of a UFD R. Suppose

$$a = up_1^{a_1} \dots p_n^{a_n} \text{ and } a = vp_1^{b_1} \dots p_n^{b_n} \quad \text{are the prime factorisations of } a \text{ and } b,$$

where u and v are units, p_i are primes and the exponents $a_i, b_i \in \mathbb{Z}_{\geq 0}$. Take $p_i^0 = 1$. Then,

$$d = p_1^{d_1} \dots p_n^{d_n} \text{ where } d_i = \min\{a_i, b_i\} \quad \text{is a gcd of } a \text{ and } b.$$

Theorem 2.4. Every PID is a UFD.

Proof. Let R be a PID and a_0 be any non-zero non-unit in R. We need to prove the following:

- a_0 is a product of irreducibles (the product might consist of only one factor)
- the factorisation is unique up to associates and the order in which the factors appear

We address the first point. If a_0 is an irreducible, we are done. If not, write $a_0 = b_1 a_1$, where neither b_1 nor a_1 is a unit and $a_1 \neq 0$. If a_1 is not irreducible, write $a_1 = b_2 a_2$, where neither b_2 nor a_2 is a unit and $a_2 \neq 0$. In general, $a_n = b_{n+1} a_{n+1}$ for all $n \in \mathbb{N}$, where $b_1, b_2, \dots$ are not units in R and $a_0, a_1, a_2, \dots$ are non-zero elements of D.

So, $(a_0) \subseteq (a_1) \subseteq \dots$ is a strictly increasing chain of ideals. By the ascending chain condition for PIDs (by Theorem 2.1 since every PID is Noetherian and Definition 2.5 says that every Noetherian ring satisfies the ascending chain condition), this chain is finite, i.e. there exists $r \in \mathbb{N}$ such that $(a_r) = (a_{r+1}) = \dots$. In particular, a_r is an irreducible factor of a_0. Thus, every non-zero non-unit in R has at least one irreducible factor.

Now, write $a_0 = p_1 c_1$, where p_1 is irreducible and c_1 is not a unit. If c_1 is not irreducible, write $c_1 = p_2 c_2$, where p_2 is irreducible and c_2 is not a unit. Repeat to obtain the following strictly increasing chain of ideals: $(a_0) \subseteq (c_1) \subseteq (c_2) \subseteq \dots$, which terminates eventually. Suppose there exists $s \in \mathbb{N}$ such that $(c_s) = (c_{s+1})$. Then, c_s is irreducible and $a_0 = p_1 p_2 \dots p_s c_s$, where each p_i is irreducible. The first result follows.

For the second point, suppose some $a \in R$ has two different representations, i.e.

$$a = p_1 p_2 \dots p_r = q_1 q_2 \dots q_s,$$

where the p's and q's are irreducible and repetition is permitted. If $r = 1$, then a is irreducible, and consequently, $s = 1$ and $p_1 = q_1$, which supports the uniqueness of factorisation claim. Assume that a can be expressed as a product of fewer than r irreducible factors in only one way. Since $p_1 \mid q_1 q_2 \dots q_s$, then it must divide some q_i. Without a loss of generality, $p_1 \mid q_1$. Then, there exists $u \in R$ (u is a unit) such that $q_1 = up_1$.

So, $up_1 p_2 \dots p_r = uq_1 q_2 \dots q_s$, for which by cancellation, $p_2 \dots p_r = uq_2 \dots q_s$. The induction hypothesis tells us that these two factorisations are identical up to associates and the order in which the factors appear. We conclude that the same is true regarding the two factorisations of a. □

Corollary 2.1. Every ED is a UFD.

Proof. By Theorem 2.2, every ED is a PID and by Theorem 2.4, every PID is a UFD. The result follows. □

Example 2.17. $\mathbb{Z}$ is an ED, a PID, and a UFD (Definition 2.9).

Example 2.18. The ring $\mathbb{Z}[x]$ is a UFD but not a PID.

Example 2.19 (Motzkin). The ring

$$\mathbb{Z}\left[\frac{1+\sqrt{-19}}{2}\right]$$

is a PID but not an ED. This was the first example of a PID that is not an ED, which was given by Israeli-American mathematician Theodore Motzkin. See p. 282 of [1] for further discussion.

At this point, the following chain of inclusions should be quite obvious:

$$\text{fields} \subset \text{ED} \subset \text{PID} \subset \text{UFD} \subset \text{integral domains} \subset \text{commutative rings} \subset \text{rings}$$

Example 2.20. The ring $\mathbb{Z}\left[\sqrt{-5}\right]$ is an integral domain but not a UFD.

Theorem 2.5 (Chinese remainder theorem). Let R be a PID and $r = up_1^{a_1} \dots p_n^{a_n}$ be the prime factorisation of r. Then,

$$R/(r) \cong R/\left(p_1^{a_1}\right) \times R/\left(p_n^{a_n}\right).$$

We consider the ring $\mathbb{Z}[i]$ of Gaussian integers as an application and summary. With the standard complex norm, we know that $\mathbb{Z}[i]$ is an ED (Proposition 2.1), a PID (Theorem 2.2), and a UFD (Theorem 2.4). Note that the norm is always $\in \mathbb{Z}_{\geq 0}$. We have the following beautiful results (Theorems 2.6 and 2.7).

Theorem 2.6 (Fermat sum of two squares theorem). Let $p \in \mathbb{Z}$ be a positive prime. Then,

$$p = a^2 + b^2 \text{ for some } a, b \in \mathbb{Z} \quad \text{if and only if} \quad p = 2 \text{ or } p \equiv 1 \pmod 4.$$

The expression is unique up to multiplication by -1 and interchanging a and b.

Proof. The forward direction is easier so we will prove it first. Assume $p = a^2 + b^2$ for some integers a, b. If $p = 2$ we are done, so assume p is odd. Squares modulo 4 are 0 or 1, hence $a^2 + b^2 \equiv 0, 1, 2 \pmod 4$. Since p is an odd prime, $p \not\equiv 0, 2 \pmod 4$, so necessarily $p \equiv 1 \pmod 4$.

Recall that $N(\alpha\beta) = N(\alpha)N(\beta)$ and that $\mathbb{Z}[i]$ is a Euclidean domain with respect to N (Proposition 2.1). Assume $p \equiv 1 \pmod 4$. By Euler's criterion from MA3265 Number Theory,

$$\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}} = 1,$$

so -1 is a quadratic residue modulo p. Hence, there exists $t \in \mathbb{Z}$ such that $t^2 \equiv -1 \pmod p$. Equivalently, $p \mid \left(t^2 + 1\right) = N(t + i)$. In $\mathbb{Z}[i]$, we have $p \mid (t + i)(t - i)$.

We claim that p is not irreducible in $\mathbb{Z}[i]$. Indeed, if p were irreducible, then $p \mid (t + i)$ or $p \mid (t - i)$, i.e.

$$t \pm i = p\alpha \quad \text{for some } \alpha \in \mathbb{Z}[i].$$

Taking norms gives

$$t^2 + 1 = N(t \pm i) = N(p\alpha) = N(p)N(\alpha) = p^2 N(\alpha),$$

so $p^2 \mid (t^2 + 1)$, contradicting $t^2 \equiv -1 \pmod p$ (which implies $p \nmid t$ and hence $p^2 \nmid t^2 + 1$ for a suitable choice of t). Thus, p is reducible in $\mathbb{Z}[i]$.

Since $\mathbb{Z}[i]$ is a UFD, we can factor p non-trivially as $p = \pi\overline{\pi}$ for some non-unit $\pi \in \mathbb{Z}[i]$, where $\overline{\pi}$ is the complex conjugate of π. Taking norms,

$$p^2 = N(p) = N(\pi\overline{\pi}) = N(\pi)N(\overline{\pi}) = (N(\pi))^2.$$

Hence $N(\pi) = p$, so writing $\pi = a + ib$ with $a, b \in \mathbb{Z}$ gives $p = N(\pi) = a^2 + b^2$. This completes the proof of the equivalence.

We then prove that the expression is unique up to signs and swapping. Assume that

$$p = a^2 + b^2 = c^2 + d^2 \quad \text{where } a, b, c, d \in \mathbb{Z}.$$

In $\mathbb{Z}[i]$, this means

$$p = (a + ib)(a - ib) = (c + id)(c - id).$$

Since $\mathbb{Z}[i]$ is a UFD and p factors as $p = \pi\overline{\pi}$ with π irreducible (up to units), each of the factors $a + ib$ and $c + id$ must be associate to either π or $\overline{\pi}$. Thus

$$a + ib = u(c + id) \quad \text{or} \quad a + ib = u(c - id)$$

for some unit $u \in \{\pm 1, \pm i\}$. Multiplying by a unit corresponds exactly to changing signs and/or swapping real and imaginary parts. Therefore, the representation is unique up to multiplication by -1 and interchanging a and b. □

Theorem 2.7. The irreducible elements in $\mathbb{Z}[i]$, up to multiplication by units, are as follows:

(i) $1+i$, which is of norm 2

(ii) $p \in \mathbb{Z}$ with $p \equiv 3 \pmod 4$

(iii) $a \pm bi$, where

$$p = a^2 + b^2 = (a+bi)(a-bi) \text{ for } p \equiv 1 \pmod 4,$$

where p is a prime in $\mathbb{Z}$

Corollary 2.2. The irreducible elements in $\mathbb{Z}[i]$ are either up to some unit $u \in \mathbb{Z}[i]$ and a prime $p \in \mathbb{Z}$, or $a+bi$ with $a^2+b^2 = p$ for some prime $p \in \mathbb{Z}$. In particular, a prime $p \in \mathbb{Z}$ has at most 2 irreducible factors in $\mathbb{Z}[i]$.

Corollary 2.3. $1 \pm i$ are irreducible over $\mathbb{Z}[i]$.

So, it remains to determine whether we can factor an odd prime $p \in \mathbb{Z}$ in the bigger ring $\mathbb{Z}[i]$.

Lemma 2.3. If $p \equiv 3 \pmod 4$ is a prime, then p is irreducible over $\mathbb{Z}[i]$.

Proof. Consider $p = a^2 + b^2$ in $\mathbb{Z}/4\mathbb{Z}$, for which p cannot be 3 mod 4. □

Lemma 2.4. Let p be a prime, where $p \equiv 1 \pmod 4$. Then,

$$p \mid (n^2+1) \quad \text{for some } n \in \mathbb{Z}.$$

Proof. It suffices to show that the equation $x^2 + 1 = 0$ has a root in $\mathbb{Z}/p\mathbb{Z}$. Recall from MA2202 Algebra I that $(\mathbb{Z}/p\mathbb{Z})^\times$ is a cyclic group of order $p-1$ (we will state it again in Proposition 3.4). As $4 \mid (p-1)$, there exists $r \in (\mathbb{Z}/p\mathbb{Z})^\times$ of order 4, and the result follows. □

Corollary 2.4. Let p be a prime, where $p \equiv 1 \pmod 4$. Then,

$$p = a^2 + b^2 = (a+bi)(a-bi) \quad \text{for some } a, b \in \mathbb{Z}.$$

Chapter 3

Polynomial Rings

3.1 Definitions and Basic Properties

Polynomial rings are among the most fundamental constructions in algebra: they allow us to enlarge a ring R by adjoining new formal variables while keeping full control of arithmetic. Concretely, $R[x]$ consists of finite expressions $a_0 + a_1x + \cdots + a_nx^n$ with coefficients $a_i \in R$, and it serves simultaneously as an algebraic object in its own right and as a universal device for encoding substitution, factorisation, and algebraic relations. In this section, we only consider commutative rings R with multiplicative identity 1 for polynomial rings. Recall the polynomial ring $R[x]$ (Definition 1.16), so by repeatedly adjoining the elements $x_1, \ldots, x_n$, we define

$$R[x_1, \ldots, x_n] = R[x_1, \ldots, x_{n-1}][x_n].$$

This section develops the basic structural properties that make polynomial rings so useful. As a warmup, the reader can attempt Examples 3.1 and 3.2.

Example 3.1 (Dummit and Foote p. 298 Question 1). Let $p(x,y,z) = 2x^2y - 3xy^3z + 4y^2z^5$ and $q(x,y,z) = 7x^2 + 5x^2y^3z^4 - 3x^2z^3$ be polynomials in $\mathbb{Z}[x,y,z]$.

(a) Write each of p and q as a polynomial in x with coefficients in $\mathbb{Z}[y,z]$.

(b) Find the degree of each of p and q.

(c) Find the degree of p and q in each of the three variables x, y and z.

(d) Compute pq and find the degree of pq in each of the three variables x, y and z.

(e) Write pq as a polynomial in the variable z with coefficients in $\mathbb{Z}[x,y]$.

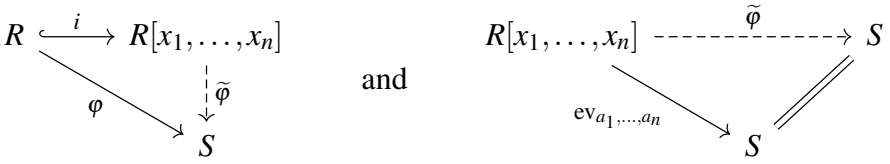
Example 3.2 (Dummit and Foote p. 298 Question 2). Repeat Example 3.1 under the assumption that the coefficients of p and q are in $\mathbb{Z}/3\mathbb{Z}$.

Proposition 3.1. Let S be a commutative ring with multiplicative identity 1. Let $\varphi : R \to S$ be a ring homomorphism. Then, for any $a_1, \ldots, a_n \in S$, there exists a

unique ring homomorphism

$$\widetilde{\varphi} : R[x_1,\dots,x_n] \to S \quad \text{such that} \quad \widetilde{\varphi}(r) = \varphi(r) \text{ and } \widetilde{\varphi}(x_i) = a_i.$$

In the following commutative diagram, i is the natural inclusion of constants, and $\mathrm{ev}_{a_1,\dots,a_n}$ denotes evaluation at $(a_1,\dots,a_n)$ (with coefficients mapped via φ).



Proof. Every element of $R[x_1,\dots,x_n]$ can be written uniquely as a finite sum

$$f = \sum_{\alpha\in\mathbb{N}^n} r_\alpha x^\alpha \quad \text{where} \quad x^\alpha = x_1^{\alpha_1}\cdots x_n^{\alpha_n},$$

where all but finitely many coefficients $r_\alpha \in R$ are zero. Define a map

$$\widetilde{\varphi} : R[x_1,\dots,x_n] \to S \quad \text{by} \quad \widetilde{\varphi}(f) = \sum_{\alpha\in\mathbb{N}^n} \varphi(r_\alpha)a^\alpha \text{ and } a^\alpha = a_1^{\alpha_1}\cdots a_n^{\alpha_n}.$$

This is well-defined because the sum is finite. We check that $\widetilde{\varphi}$ is a ring homomorphism. Additivity is immediate. That is, $\widetilde{\varphi}(f+g) = \widetilde{\varphi}(f) + \widetilde{\varphi}(g)$. For multiplicativity, we omit the details but anyway, one needs to make use of convolutions.

Also, $\widetilde{\varphi}(1) = 1$ since $1 \in R[x_1,\dots,x_n]$ corresponds to the constant polynomial 1_R and $\varphi(1_R) = 1_S$. Finally, the defining properties hold: for $r \in R$ viewed as a constant polynomial, $\widetilde{\varphi}(r) = \varphi(r)$; and for each i, since x_i has coefficient 1 on the monomial x_i and 0 elsewhere, $\widetilde{\varphi}(x_i) = a_i$.

We then prove uniqueness. Let $\psi : R[x_1,\dots,x_n] \to S$ be any ring homomorphism such that $\psi\,|_R = \varphi$ and $\psi(x_i) = a_i$ for all i. Then for every monomial $x^\alpha = x_1^{\alpha_1}\cdots x_n^{\alpha_n}$,

$$\psi(x^\alpha) = \psi(x_1)^{\alpha_1}\cdots\psi(x_n)^{\alpha_n} = a_1^{\alpha_1}\cdots a_n^{\alpha_n} = a^\alpha,$$

and for a general polynomial $f = \sum_\alpha r_\alpha x^\alpha$,

$$\psi(f) = \sum_\alpha \psi(r_\alpha)\psi(x^\alpha) = \sum_\alpha \varphi(r_\alpha)a^\alpha = \widetilde{\varphi}(f).$$

Hence, $\psi = \widetilde{\varphi}$, proving uniqueness. □

Example 3.3 (Dummit and Foote p. 298 Question 3). If R is a commutative ring and $x_1, x_2, \dots, x_n$ are independent variables over R, prove that $R[x_{\pi(1)}, x_{\pi(2)}, \dots, x_{\pi(n)}]$ is isomorphic to $R[x_1, x_2, \dots, x_n]$ for any permutation π of $\{1, 2, \dots, n\}$.

Solution. Let $\pi \in S_n$ be a permutation. Define a ring homomorphism

$$\varphi : R[x_1, x_2, \dots, x_n] \to R[x_{\pi(1)}, x_{\pi(2)}, \dots, x_{\pi(n)}]$$

by prescribing its values on generators, where $\varphi(r) = r$ for all $r \in R$ and $\varphi(x_i) = x_{\pi(i)}$ for all $1 \leq i \leq n$, and this extends multiplicatively and additively to all polynomials.

Note that every element of $R[x_1, \ldots, x_n]$ is a finite R-linear combination of monomials $x_1^{a_1} \cdots x_n^{a_n}$, so the above prescription forces

$$\varphi\left(x_1^{a_1} \cdots x_n^{a_n}\right) = x_{\pi(1)}^{a_1} \cdots x_{\pi(n)}^{a_n},$$

and hence determines φ uniquely. It is immediate from the definition that φ is additive, multiplicative, and restricts to the identity on R, so φ is an R-algebra homomorphism.

Let π^{-1} be the inverse permutation. Define similarly an R-algebra homomorphism

$$\psi : R[x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(n)}] \to R[x_1, x_2, \ldots, x_n]$$

by $\psi(r) = r$ for all $r \in R$ and $\psi(x_{\pi(i)}) = x_i$ for all $1 \leq i \leq n$, and extend to all polynomials.

For each $r \in R$ and each i,

$$(\psi \circ \varphi)(r) = \psi(r) = r \quad \text{and} \quad (\psi \circ \varphi)(x_i) = \psi(x_{\pi(i)}) = x_i,$$

so $\psi \circ \varphi = \mathrm{id}_{R[x_1, \ldots, x_n]}$ because $R[x_1, \ldots, x_n]$ is generated as an R-algebra by $\{x_1, \ldots, x_n\}$. Similarly,

$$(\varphi \circ \psi)(r) = r \quad \text{and} \quad (\varphi \circ \psi)(x_{\pi(i)}) = \varphi(x_i) = x_{\pi(i)},$$

so $\varphi \circ \psi = \mathrm{id}_{R[x_{\pi(1)}, \ldots, x_{\pi(n)}]}$. Therefore φ is an isomorphism of R-algebras, and the result follows. □

Example 3.4 (prime and maximal ideals in $\mathbb{Q}$ adjoin x, y; Dummit and Foote p. 298 Question 4). Prove that the ideals (x) and (x, y) are prime ideals in $\mathbb{Q}[x, y]$ but only the latter ideal is a maximal ideal.

Solution. We first prove that (x) is a prime ideal in $A = \mathbb{Q}[x, y]$. Consider the evaluation homomorphism at $x = 0$, which is

$$\varphi : A \to \mathbb{Q}[y] \quad \text{where} \quad \varphi(f(x, y)) = f(0, y).$$

This is a surjective ring homomorphism and $\ker \varphi = (x)$. By the first isomorphism theorem for rings (Theorem 1.1), $A/(x) \sim \mathbb{Q}[y]$. Since $\mathbb{Q}[y]$ is an integral domain, (x) is a prime ideal by Lemma 1.7.

We then prove that (x, y) is a maximal ideal which is a stronger result than just saying (x, y) is a prime ideal. Again, we consider some evaluation homomorphism, but this time at $(x, y) = (0, 0)$, so

$$\psi : A \to \mathbb{Q} \quad \text{where} \quad \psi(f(x, y)) = f(0, 0).$$

This map is surjective, and its kernel is precisely the set of polynomials with zero constant term, i.e. $\ker \psi = (x, y)$. By the first isomorphism theorem (Theorem 1.1), $A/(x, y) \cong \mathbb{Q}$.

Since $\mathbb{Q}$ is a field, by Lemma 1.6, (x,y) is a maximal ideal, and in particular it is prime.

Lastly, we prove that (x) is not a maximal ideal. Earlier, we proved that $\mathbb{Q}[x,y]/(x) \cong \mathbb{Q}[y]$ and $\mathbb{Q}[y]$ is clearly not a field (for instance, y is not a unit). The result follows by Lemma 1.6. □

One can use the technique in Example 3.4 to solve Example 3.5. It is interesting to note that when using the first isomorphism theorem (Theorem 1.1) to prove that $(2,x,y)$ is a maximal ideal in $\mathbb{Z}[x,y]$, the idea is to show that $\mathbb{Z}[x,y]/(2,x,y) \cong \mathbb{F}_2$, where $\mathbb{F}_2$ denotes the finite field with 2 elements. Thereafter, one makes use of Lemma 1.6.

Example 3.5 (ideals in $\mathbb{Z}$ adjoin x,y; Dummit and Foote p. 298 Question 5). Prove that (x,y) and $(2,x,y)$ are prime ideals in $\mathbb{Z}[x,y]$ but only the latter ideal is a maximal ideal.

Example 3.6 (Dummit and Foote p. 298 Question 7). Let R be a commutative ring with 1. Prove that a polynomial ring in more than one variable over R is not a PID.

Solution. Let R be a commutative ring with 1 and let $n \geq 2$. We prove that $R[x_1,\ldots,x_n]$ is not a PID. Consider the ideal

$$I = (x_1,x_2) \subseteq R[x_1,\ldots,x_n].$$

We claim that I is not principal[1]. Suppose on the contrary that $I = (f)$ for some $f \in R[x_1,\ldots,x_n]$. Then, $f \mid x_1$ and $f \mid x_2$.

Now use the evaluation homomorphism at $x_2 = 0$, so

$$\varphi : R[x_1,\ldots,x_n] \to R[x_1,x_3,\ldots,x_n] \quad \text{where} \quad \varphi(g) = g(x_1,0,x_3,\ldots,x_n).$$

Applying φ to the divisibility $f \mid x_1$ gives $\varphi(f) \mid x_1$ in $R[x_1,x_3,\ldots,x_n]$, so $\varphi(f)$ must be of the form $\varphi(f) = ux_1$ for some unit $u \in R[x_1,x_3,\ldots,x_n]$. But the only units in a polynomial ring over R are the constant units (if a polynomial has positive total degree, it cannot have a multiplicative inverse), hence $u \in R^\times$. In particular, $\varphi(f)$ is associate to x_1, so $x_1 \mid \varphi(f)$ and $\varphi(f) \mid x_1$.

On the other hand, apply φ to $f \mid x_2$. Since $\varphi(x_2) = 0$, we get $\varphi(f) \mid 0$, which is automatic and gives no information. So instead evaluate at $x_1 = 0$ to obtain the map

$$\psi : R[x_1,\ldots,x_n] \to R[x_2,x_3,\ldots,x_n] \quad \text{where} \quad \psi(g) = g(0,x_2,x_3,\ldots,x_n).$$

From $f \mid x_2$, we obtain $\psi(f) \mid x_2$, hence similarly $\psi(f) = vx_2$ for some unit $v \in R^\times$.

Write f as a polynomial in x_2 with coefficients in $R[x_1,x_3,\ldots,x_n]$. That is,

$$f = a_0 + a_1x_2 + \cdots + a_mx_2^m \quad \text{where } a_i \in R[x_1,x_3,\ldots,x_n].$$

[1] One can see this as a generalisation of Examples 3.4 and 3.5.

Then, $\varphi(f) = a_0$ (since setting $x_2 = 0$ kills all higher terms), so $a_0 = ux_1$. In particular, a_0 has no term independent of x_1. Similarly, writing f as a polynomial in x_1 with coefficients in $R[x_2, x_3, \ldots, x_n]$, and setting $x_1 = 0$ shows that $\psi(f)$ is the constant term in x_1, so that constant term equals vx_2. In particular, f has no term independent of x_2.

Combining these, we see that every term of f is divisible by x_1 (because the x_2-constant term is ux_1 and the other terms all contain x_2, hence (by the second condition) must also contain x_1), and likewise every term of f is divisible by x_2. Concretely, $f \in (x_1) \cap (x_2)$. In a polynomial ring, $(x_1) \cap (x_2) = (x_1 x_2)$: indeed, if a polynomial is divisible by both x_1 and x_2, then every monomial is divisible by $x_1 x_2$, hence the polynomial is divisible by $x_1 x_2$. Thus $f \in (x_1 x_2)$, so

$$f = x_1 x_2 \cdot h \quad \text{for some } h \in R[x_1, \ldots, x_n].$$

However, f cannot divide x_1: if $x_1 = f \cdot g = (x_1 x_2 h) g$, the right-hand side is divisible by x_2, whereas the left-hand side x_1 is not divisible by x_2 in $R[x_1, \ldots, x_n]$. This contradiction shows that $I = (x_1, x_2)$ is not principal. Therefore $R[x_1, \ldots, x_n]$ is not a PID for every $n \geq 2$. □

Proposition 3.2. Let I be an ideal of R. We consider I as a subset of $R[x]$ and denote $(I) = I[x]$ to be the ideal generated by I in $R[x]$. We have the following isomorphism:

$$(R/I)[x] \cong R[x]/(I).$$

So, if I is prime in R, then (I) is prime in $R[x]$.

Proposition 3.3. Let F be a field. Then, $F[x]$ is an ED with the norm $N(f) = \deg f$.

Proof. Let $f, g \in F[x]$ with $g \neq 0$. By Definition 2.2, we wish prove that there exist $q, r \in F[x]$ such that

$$f = gq + r \quad \text{and} \quad r = 0 \text{ or } \deg r < \deg g.$$

This is exactly the Euclidean division property for the Euclidean function $N(h) = \deg h$ (with the convention $N(0) = -\infty$, so that $N(r) < N(g)$ when $r = 0$ or $\deg r < \deg g$). We shall write

$$f = a_m x^m + a_{m-1} x^{m-1} + \cdots + a_0 \quad \text{and} \quad g = b_n x^n + b_{n-1} x^{n-1} + \cdots + b_0$$

where $a_m \neq 0$ and $b_n \neq 0$, so $\deg f = m$ and $\deg g = n$. If $m < n$, take $q = 0$ and $r = f$, and we are done.

Assume now that $m \geq n$. Since F is a field, b_n is invertible, so define $c = a_m b_n^{-1} \in F$ and $h = c x^{m-n} \in F[x]$. Then, $\deg(hg) = m$ and the leading term of hg is $a_m x^m$, so the polynomial $f_1 = f - hg$ either is 0 or satisfies $\deg f_1 < m = \deg f$.

If $\deg f_1 < n$, set $q = h$ and $r = f_1$ and we are done. Otherwise repeat the same procedure with (f_1, g) in place of (f, g). Each repetition strictly decreases the degree of the current remainder, and degrees are non-negative integers, so the process terminates after finitely many steps. Hence we obtain polynomials $h_1, \ldots, h_k \in F[x]$ and a final remainder r with $r = 0$ or $\deg r < \deg g$ such that

$$f = g(h_1 + \cdots + h_k) + r.$$

Let $q = h_1 + \cdots + h_k$. Then, $f = gq + r$ with $r = 0$ or $\deg r < \deg g$, as required. Therefore, $F[x]$ is a Euclidean domain with Euclidean norm $N(f) = \deg f$. □

Corollary 3.1. Let F be a field. Then, the following hold:

(i) $F[x]$ is a PID, hence a UFD

(ii) Let $f(x) \in F[x]$. Then,

$$f(a) = 0 \text{ for } a \in F \quad \text{if and only if} \quad (x-a) \mid f(x)$$

(iii) Let $f(x) \in F[x]$ be of degree n. Then, $f(x)$ has at most n roots in F counting multiplicity

(iv) Let $f(x) \in F[x]$. Then,

$$F[x]/(f) \text{ is a field} \quad \text{if and only if} \quad f \text{ is prime.}$$

(v) If $p(x), q(x) \in F[x]$ are distinct irreducible polynomials, i.e. $(p(x)) \neq (q(x))$, then they are coprime

(vi) Let

$$f(x) = p_1^{a_1}(x) \ldots p_n^{a_n}(x) \quad \text{be} \quad \text{an irreducible factorisation of } f(x).$$

Then,

$$F[x]/(f(x)) \cong F[x]/\left(p_1^{a_1}(x)\right) \times \ldots \times F[x]/(p_n^{a_n}(x))$$

Proposition 3.4. Let F be a field and $G \subseteq F^\times$ be a finite subgroup. Then, G is cyclic.

Example 3.7. For any prime p, $\mathbb{Z}/p\mathbb{Z}$ is a field. Then, by Proposition 3.4, $(\mathbb{Z}/p\mathbb{Z})^\times \cong \mathbb{Z}/(p-1)\mathbb{Z}$ is a cyclic group.

Proposition 3.4 was used in the proof of Lemma 2.4. Moreover, the proof of this proposition requires a well-known result known as the fundamental theorem of finitely generated Abelian groups.

Definition 3.1 (polynomial functions). Let F be a field. Then, define the set of polynomial functions on F, denoted by $\mathcal{P}$, to be functions from F to F of the form

$$F \to F \quad \text{where} \quad x \mapsto f(x) = \sum_{i=0}^{n} a_i x^i.$$

Note that the set of polynomial functions on F, denoted by $\mathcal{P}$, is obviously a ring under pointwise addition and multiplication. As such, we have an obvious ring homomorphism $F[x] \to \mathcal{P}$ (Example 3.8).

Example 3.8. Consider the ring of polynomial functions on $\mathbb{Z}/2\mathbb{Z}$.

Proposition 3.5. The ring homomorphism

$$\varphi : F[x] \to \mathcal{P} \text{ is an isomorphism} \quad \text{if and only if} \quad F \text{ is infinite.}$$

Proof. Clearly, φ is a surjective ring homomorphism. As such, it suffices to prove that

$$\ker \varphi = \{e\} \quad \text{if and only if} \quad F \text{ is infinite.}$$

We first prove the forward direction by contraposition. Suppose F is a finite set. Then, $F = \{a_1, \ldots, a_n\}$. So, the image of the non-zero polynomial $(x-a_1)\ldots(x-a_n)$ is the zero function, so $\ker \varphi \neq \{e\}$.

We then prove the reverse direction. Suppose F is an infinite set. Suppose

$$f(x) = \sum_{i=1}^{n} a_i x^i \in \ker \varphi.$$

Then, $f(a) = 0$ for all $a \in F$. By **(ii)** of Corollary 3.1, the result follows. □

Theorem 3.1. R is a UFD if and only if $R[x]$ is a UFD.

Corollary 3.2. For any field F, $F[x_1, \ldots, x_n]$ is a UFD.

Example 3.9 (Dummit and Foote p. 298 Question 6). Prove that (x, y) is not a principal ideal in $\mathbb{Q}[x, y]$.

Solution. Suppose on the contrary that (x, y) is principal in $\mathbb{Q}[x, y]$. By Definition 1.30, we can write the ideal as (a) for some $a \in \mathbb{Q}[x, y]$. Since $x \in (x, y) = (a)$, there exists $f \in \mathbb{Q}[x, y]$ such that $x = af$. Similarly, there exists $g \in \mathbb{Q}[x, y]$ such that $y = ag$.

In Example 3.4, we showed that (x, y) is maximal in $\mathbb{Q}[x, y]$. Note that $\mathbb{Q}[x, y]$ is a UFD by Theorem 3.1, so x and y are irreducible: viewing $\mathbb{Q}[x, y] = \mathbb{Q}[y][x]$, the polynomial x has degree 1 in x and hence cannot factor non-trivially; similarly for y. Therefore, from $a \mid x$ we obtain that either a is a unit or a is an associate of x. The first case is impossible

because if a is a unit then $(a) = \mathbb{Q}[x,y] \neq (x,y)$. Hence a must be an associate of x, and so $(a) = (x)$. But then $y \in (a) = (x)$, so there exists $h \in \mathbb{Q}[x,y]$ such that

$$y = xh.$$

Taking degrees in x gives a contradiction: $\deg_x(y) = 0$ whereas $\deg_x(xh) \geq 1$ for any non-zero h. □

Now, we shall consider some irreducible polynomials in $F[x]$ and construct some interesting fields.

Lemma 3.1. Let F be a field and $f(x) \in F[x]$ be of degree 2 or 3. Then,

$$f(x) \text{ is reducible} \quad \text{if and only if} \quad f(x) \text{ has a root in } F.$$

Example 3.10. We have the factorisation

$$x^2 + 3x + 4 = (x-3)(x-5) \quad \text{in } \mathbb{Z}/11\mathbb{Z}.$$

To see why, working from right to left,

$$(x-3)(x-5) = x^2 - 8x + 15 = x^2 + 3x + 4.$$

Example 3.11. Suppose $F = \mathbb{Z}/2\mathbb{Z}$ and $f(x) = x^2 + x + 1 \in F[x]$. Since $\deg f = 2$, by Lemma 3.1, f is irreducible so $F[x]/(f(x))$ is a field. In fact, $F[x]/(f(x))$ is an F-vector space with basis $\{\overline{1}, \overline{x}\}$ and the vector space has cardinality 4. So, F is a field with 4 elements.

3.2 Irreducibility Criteria

A recurring theme in algebra is that complicated rings become manageable once we understand how their elements factor. For polynomial rings, the basic question is given an integral domain R, when is a polynomial $f(x) \in R[x]$ irreducible?

Over a field K, this is already subtle, but we at least have a clean definition: f is irreducible in $K[x]$ if it has no non-trivial factorisation into lower-degree polynomials. When R is not a field (e.g. $R = \mathbb{Z}$), new arithmetic phenomena appear: coefficients may share common factors. This section develops a toolkit for handling irreducibility in $R[x]$, with $\mathbb{Z}[x]$ as the main case of interest.

Definition 3.2 (content and primitive polynomial). Let

$$f(x) = a_0 + a_1 x + \ldots + a_n x^n \quad \text{be} \quad \text{a non-zero polynomial in } R[x].$$

Then, $\gcd(a_0, a_1, \ldots, a_n)$ is known as the content of $f(x)$, denoted by $\operatorname{cont} f$. We say that

$$f(x) \in R[x] \text{ is a primitive polynomial} \quad \text{if} \quad \operatorname{cont} f \text{ is a unit in } R.$$

Example 3.12. Let $f(x) = 6x^4 - 9x^3 + 3x - 12 \in \mathbb{Z}[x]$. Then, $\operatorname{cont} f = 3$ so f is not a primitive polynomial because 3 is not a unit in $\mathbb{Z}$.

Example 3.13 (Dummit and Foote p. 311 Question 3). Show that the polynomial[2]

$$(x-1)(x-2)\cdots(x-n)-1 \text{ is irreducible over } \mathbb{Z} \quad \text{for all } n \geq 1. \tag{3.1}$$

Solution. Let $f_n(x)$ denote the polynomial in (3.1). We prove that f_n is irreducible over $\mathbb{Z}$ for every $n \geq 1$. Suppose on the contrary that f_n factors non-trivially in $\mathbb{Z}[x]$ as follows:

$$f_n(x) = g(x)\,h(x) \quad \text{where } g(x), h(x) \in \mathbb{Z}[x] \text{ are non-constant.} \tag{3.2}$$

For each $1 \leq k \leq n$, we have $f_n(k) = -1$ so $g(k) = -h(k)$. Define $p(x) = g(x) + h(x) \in \mathbb{Z}[x]$. Then, $h(k) = 0$ for all $1 \leq k \leq n$ so h has at least n distinct integer roots. By considering the degrees of polynomials in (3.1) and (3.2), we obtain a contradiction. □

Example 3.14 (Dummit and Foote p. 311 Question 4). Show that the polynomial

$$(x-1)(x-2)\cdots(x-n)+1 \text{ is irreducible over } \mathbb{Z} \quad \text{for all } n \geq 1, n \neq 4. \tag{3.3}$$

Solution. In a similar fashion to Example 3.13, let $f_n(x)$ denote the polynomial in (3.3). We prove that f_n is irreducible over $\mathbb{Z}$ for every $n \geq 1$. Suppose on the contrary that f_n factors non-trivially in $\mathbb{Z}[x]$ as follows:

$$f_n(x) = g(x)\,h(x) \quad \text{where } g(x), h(x) \in \mathbb{Z}[x] \text{ are non-constant.}$$

For each $1 \leq k \leq n$, we have $f_n(k) = 1$ so $g(k)\,h(k) = 1$. First, suppose $g(k) = h(k) = 1$ for all $1 \leq k \leq n$. Define $u(x) = g(x) - h(x) \in \mathbb{Z}[x]$. Then, $u(k) = 0$ for all $1 \leq k \leq n$. However, $\deg u \leq n-1$ so a non-zero polynomial of degree at most $n-1$ cannot have n distinct roots, which is a contradiction. Then, repeat the argument for the case when $g(k) = h(k) = -1$.

Having said that, when $n = 4$, one can verify that $f_n(x)$ is reducible. □

Lemma 3.2. Let R be a UFD. If

$$g(x), h(x) \text{ are primitive in } R[x] \quad \text{then} \quad f(x) = g(x)\,h(x) \text{ is primitive in } R[x].$$

Proof. Let

$$g(x) = \sum_{i=0}^{m} b_i x^i \quad \text{and} \quad h(x) = \sum_{j=0}^{n} c_j x^j$$

with $b_i, c_j \in R$. Since g and h are primitive, no irreducible element of R divides all the b_i, and likewise no irreducible divides all the c_j. Set

$$f(x) = g(x)h(x) = \sum_{k=0}^{m+n} a_k x^k \quad \text{where } a_k = \sum_{i+j=k} b_i c_j.$$

[2] A hint is as follows: if the polynomial factors, consider the values of the factors at $x = 1, 2, \ldots, n$.

We prove that f is primitive by contradiction. Suppose on the contrary that f is not primitive. Then by Definition 3.2, there exists an irreducible element $\pi \in R$ such that

$$\pi \mid a_k \quad \text{for every } k = 0, 1, \ldots, m+n.$$

Equivalently, in the quotient ring $\overline{R} = R/(\pi)$ we have $\overline{a_k} = 0$ for all k, so

$$\overline{f}(x) = \sum_{k=0}^{m+n} \overline{a_k} x^k = 0 \in \overline{R}[x].$$

But $\overline{f} = \overline{g} \cdot \overline{h}$ in $\overline{R}[x]$, hence $\overline{g}(x)\overline{h}(x) = 0$ in $\overline{R}[x]$. Now, we use the fact that in a UFD every irreducible element is prime (Proposition 2.5), so (π) is a prime ideal and $\overline{R} = R/(\pi)$ is an integral domain (Lemma 1.7). Therefore, $\overline{R}[x]$ is also an integral domain, so it has no zero divisors. Hence

$$\overline{g}(x) = 0 \quad \text{or} \quad \overline{h}(x) = 0.$$

If $\overline{g} = 0$, then every coefficient $\overline{b_i} = 0$, i.e. $\pi \mid b_i$ for all i, contradicting that g is primitive. Similarly, $\overline{h} = 0$ contradicts primitivity of h. Thus no irreducible π can divide all coefficients of f, so $\operatorname{cont}(f)$ is a unit in R, i.e. f is primitive. □

Lemma 3.3 (Gauss' lemma). If R be a UFD with field of fractions Q and let $f(x) \in R[x]$ be primitive. If

$$f(x) \text{ is reducible in } Q[x] \quad \text{then} \quad f(x) \text{ is reducible in } R[x].$$

Corollary 3.3. Let R be a UFD with field of fractions Q. Then,

$$f(x) \in R[x] \text{ is reducible in } Q[x] \quad \text{if and only if} \quad f(x) \text{ is reducible in } R[x].$$

Example 3.15 (Dummit and Foote p. 312 Question 11). Prove that $x^2 + y^2 - 1$ is irreducible in $\mathbb{Q}[x, y]$.

Solution. We view $\mathbb{Q}[x, y]$ as the field adjunction $\mathbb{Q}[y][x]$. So, we have

$$f(x, y) = x^2 + \left(y^2 - 1\right) \in \mathbb{Q}[y][x].$$

If the polynomial were reducible in $\mathbb{Q}[y][x]$, then we can factor it as

$$f(x, y) = (x + a(y))(x + b(y)) \quad \text{where } a(y), b(y) \in \mathbb{Q}[y].$$

Expanding and comparing coefficients of powers yields

$$x^2 + (a(y) + b(y))x + a(y)b(y) = x^2 + 0 \cdot x + (y^2 - 1)$$

so $a(y) + b(y) = 0$ and $a(y)\, b(y) = y^2 - 1$. This implies that $1 - y^2$ must be a square in $\mathbb{Q}[y]$. We show this is impossible.

Write $a(y) = \alpha y + \beta$ with $\alpha, \beta \in \mathbb{Q}$. Then

$$(a(y))^2 = (\alpha y + \beta)^2 = \alpha^2 y^2 + 2\alpha\beta\, y + \beta^2.$$

Comparing with $1-y^2$ yields $\alpha^2=-1$, $2\alpha\beta=0$ and $\beta^2=1$, which is impossible over $\mathbb{Q}$ since $\alpha^2=-1$ has no solution in $\mathbb{Q}$. Hence $1-y^2$ is not a square in $\mathbb{Q}[y]$, so $f(x,y)$ has no root in $\mathbb{Q}(y)$ and cannot factor into linear factors in $\mathbb{Q}[y][x]$. Therefore f is irreducible in $\mathbb{Q}[y][x]$. Finally, since $\mathbb{Q}[y][x]=\mathbb{Q}[x,y]$, the result follows. □

Lemma 3.4. Let R be a UFD. Then, $R[x]$ is a UFD.

Example 3.16 (Dummit and Foote p. 311 Question 9). Prove that the polynomial $x^2-\sqrt{2}$ is irreducible over $\mathbb{Z}[\sqrt{2}]$.[3]

Solution. Suppose on the contrary that x^2-2 is reducible over $\mathbb{Z}[\sqrt{2}]$. Then, we can write

$$x^2-2=(x-(a+b\sqrt{2}))(x-(c+d\sqrt{2})) \quad \text{where } a,b,c,d\in\mathbb{Z}.$$

Upon expansion and comparing coefficients, $a=-c$, $b=-d$, and so $a^2+2b^2=2$. However, there are no integer solutions to this equation, and the result follows. □

Corollary 3.4. If R is a UFD, then $R[x_1,\ldots,x_n]$ is a UFD. In particular, $\mathbb{Z}[x]$ is a UFD.

We consider irreducible polynomials in $R[x]$ for an arbitrary integral domain R. We are mainly interested in $\mathbb{Z}[x]$ actually.

Lemma 3.5 (irreducibility modulo ideal). Let R be an integral domain with a proper ideal I. Let $p(x)\in R[x]$ be monic and non-constant. Then,

$$p(x) \text{ is irreducible over } R/I[x] \quad \text{implies} \quad p(x) \text{ is irreducible over } R[x].$$

Example 3.17. We give some applications of Lemma 3.5.

(i) The polynomial x^2+x+1 is irreducible over $\mathbb{Z}[x]$ as it is irreducible over $\mathbb{Z}/2\mathbb{Z}[x]$.

(ii) The polynomial x^2+1 is irreducible over $\mathbb{Z}[x]$ as it is irreducible over $\mathbb{Z}/3\mathbb{Z}[x]$.

Example 3.18 (Dummit and Foote p. 311 Question 1). Determine whether the following polynomials are irreducible in the rings indicated. For those that are reducible, determine their factorisation into irreducibles. The notation $\mathbb{F}_p$ denotes the finite field $\mathbb{Z}/p\mathbb{Z}$, where p is prime.

(a) x^2+x+1 in $\mathbb{F}_2[x]$

(b) x^3+x+1 in $\mathbb{F}_3[x]$

(c) x^4+1 in $\mathbb{F}_5[x]$

(d) x^4+10x^2+1 in $\mathbb{Z}[x]$

Solution.

[3] One can also use the fact that $\mathbb{Z}[\sqrt{2}]$ is a UFD to prove this statement.

(a) Let $f(x) = x^2 + x + 1$. Then, $f(0) \neq 0$ and $f(1) \neq 0$ which has no root in $\mathbb{F}_2$. So, f is irreducible in $\mathbb{F}_2[x]$.

(b) Let $f(x) = x^3 + x + 1$. Then, note that $f(1) = 0$ so we can write f as $(x-1)\left(x^2 + x + 2\right)$, which implies f is reducible in $\mathbb{F}_3[x]$.

(c) Observe that

$$x^4 + 1 = x^4 - 4 = \left(x^2 - 2\right)\left(x^2 + 2\right)$$

where $x^2 - 2$ and $x^2 + 2$ are irreducible in $\mathbb{F}_5[x]$.

(d) By the rational root test, any rational root of f must be ± 1. One checks that f has no linear factor in $\mathbb{Z}[x]$. Since f is monic of degree 4, if it is reducible in $\mathbb{Z}[x]$ then it must factor as a product of two monic quadratics. That is,

$$f(x) = (x^2 + ax + b)(x^2 + cx + d)$$

for some $a, b, c, d \in \mathbb{Z}$ with $bd = 1$ (comparing constant terms). Hence either

$$(b,d) = (1,1) \quad \text{or} \quad (b,d) = (-1,-1).$$

One can consider these cases and eventually reach a contradiction, so $f(x)$ cannot factor into quadratics in $\mathbb{Z}[x]$. Since it also has no linear factor, it follows that $f(x)$ is irreducible in $\mathbb{Z}[x]$. □

Example 3.19 (Dummit and Foote p. 311 Question 5). Find all the monic irreducible polynomials of degree ≤ 3 in $\mathbb{F}_2[x]$, and the same in $\mathbb{F}_3[x]$.

Solution. Note that $\mathbb{F}_2[x] = \mathbb{Z}/2\mathbb{Z}$ and $\mathbb{F}_3[x] = \mathbb{Z}/3\mathbb{Z}$. We first find all irreducible monic polynomials of degree ≤ 3 in $\mathbb{Z}/2\mathbb{Z}$. By enumeration, we have

$$x, x+1, x^2+x+1, x^3+x+1, x^3+x^2+1.$$

We then see which polynomials are also irreducible in $\mathbb{Z}/3\mathbb{Z}$. The polynomials are x and $x+1$. The learning point here is that higher-degree polynomials that are irreducible over $\mathbb{Z}/2\mathbb{Z}$ may become reducible over $\mathbb{Z}/3\mathbb{Z}$, and vice versa because $\mathbb{Z}/3\mathbb{Z}$ has different roots and different possible factorisations. □

Lemma 3.6. Let P be a prime ideal in an integral domain R. Let

$$f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_0 \in R[x] \quad \text{be} \quad \text{monic and non-constant.}$$

Suppose $a_{n-1}, \ldots, a_0$ are all in P but a_0 is not in P^2. Then, $f(x)$ is irreducible over $R[x]$.

Corollary 3.5 (Eisenstein's cirterion). Let

$$f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_0 \in \mathbb{Z}[x] \quad \text{be} \quad \text{monic and non-constant.}$$

Suppose $p \mid a_{n-1}, \ldots, a_0$ but p^2 does not divide a_0 for some prime p. Then, $f(x)$ is

irreducible over $\mathbb{Z}[x]$.

Example 3.20. The polynomial x^2+4x+2 is irreducible over $\mathbb{Z}[x]$. To see why, suppose

$$\text{there exist polynomials } f(x), g(x) \in \mathbb{Z}[x] \quad \text{such that} \quad x^2+4x+2 = f(x)\,g(x).$$

This forces f and g to be monic linear polynomials. In particular, set $f(x) = x+c$ and $g(x) = x+d$ for some $c, d \in \mathbb{Z}$. Then,

$$f(x)\,g(x) = x^2+(c+d)x+cd \quad \text{so} \quad c+d = 4 \text{ and } cd = 2.$$

This pair of equations does not have integer solutions, resulting in a contradiction.

Example 3.21 (Dummit and Foote p. 311 Question 2). Prove that the following polynomials are irreducible in $\mathbb{Z}[x]$:

(a) x^4-4x^3+6

(b) $x^6+30x^5-15x^3+6x-120$

(c) $x^4+4x^3+6x^2+2x+1$ (hint is to substitute $x-1$ for x)

(d) $\frac{(x+2)^p-2^p}{x}$, where p is an odd prime.

Solution.

(a) Consider the prime $p = 2$ and Eisenstein's criterion (Corollary 3.5).

(b) Consider the prime $p = 3$ and Eisenstein's criterion (Corollary 3.5).

(c) Replace x with $x-1$ to obtain the polynomial x^4-2x+2. Then, consider the prime $p = 2$ and Eisenstein's criterion (Corollary 3.5)[4]. This shows that sometimes, we may not be able to directly apply Eisenstein's criterion.

(d) By the binomial theorem, we can write the expression as

$$x^{p-1}+\cdots+\binom{p}{3}2^{p-3}x^2+\binom{p}{2}2^{p-2}x+p2^{p-1}.$$

Consider the prime p and Eisenstein's criterion (Corollary 3.5) and the result follows immediately. □

Example 3.22 (Dummit and Foote p. 312 Question 12). Prove that

$$x^{n-1}+x^{n-2}+\cdots+x+1 \text{ is irreducible over } \mathbb{Z} \quad \text{if and only if} \quad n \text{ is prime.}$$

Solution. Let

$$\Phi_n(x) = x^{n-1}+x^{n-2}+\cdots+x+1 = \frac{x^n-1}{x-1} \in \mathbb{Z}[x] \quad \text{where } n \geq 2.$$

[4]Since the substitution map $\varphi : \mathbb{Q}[x] \to \mathbb{Q}[x]$ given by $\varphi(h(x)) = h(x-1)$ is a ring automorphism, irreducibility is preserved under translation: if f were reducible, then $g = f(x-1)$ would also be reducible. Hence $f(x)$ is irreducible in $\mathbb{Q}[x]$ (and thus in $\mathbb{Z}[x]$ by Gauss' lemma in Lemma 3.3).

Since Φ_n is primitive, Gauss' lemma (Lemma 3.3) implies that Φ_n is irreducible over $\mathbb{Z}$ if and only if it is irreducible over $\mathbb{Q}$.

We first prove the forward direction. Suppose on the contrary that n is composite. Then, we can write $n = ab$ with $a, b > 1$. As such,

$$x^n - 1 = x^{ab} - 1 = (x^a - 1)\left(x^{a(b-1)} + x^{a(b-2)} + \cdots + x^a + 1\right).$$

Dividing by $x-1$ and using $x^a - 1 = (x-1)\left(x^{a-1} + \cdots + x + 1\right)$ yields

$$\Phi_{ab}(x) = \frac{x^{ab}-1}{x-1} = \left(\frac{x^a-1}{x-1}\right)\left(x^{a(b-1)} + x^{a(b-2)} + \cdots + x^a + 1\right).$$

In other words,

$$\Phi_n(x) = \Phi_a(x) \cdot \left(x^{a(b-1)} + x^{a(b-2)} + \cdots + x^a + 1\right).$$

Both factors lie in $\mathbb{Z}[x]$ and have positive degree (since $a, b > 1$), hence Φ_n is reducible in $\mathbb{Z}[x]$. By contraposition, if Φ_n is irreducible, then n cannot be composite, so n is prime.

To prove the reverse direction, suppose n is prime, then we wish to prove that Φ_n is irreducible over $\mathbb{Z}$. So, let $n = p$ be a prime and set

$$f(x) = \Phi_p(x) = x^{p-1} + x^{p-2} + \cdots + x + 1.$$

Consider the translate $g(x) = f(x+1)$ (same trick as **(iii)** of Example 3.21). We compute using the identity

$$f(x) = \frac{x^p - 1}{x-1} \quad \text{which implies} \quad g(x) = f(x+1) = \frac{(x+1)^p - 1}{x}.$$

By the binomial theorem,

$$(x+1)^p - 1 = \sum_{k=1}^{p} \binom{p}{k} x^k,$$

hence

$$g(x) = \sum_{k=1}^{p} \binom{p}{k} x^{k-1} = \binom{p}{1} + \binom{p}{2} x + \cdots + \binom{p}{p-1} x^{p-2} + \binom{p}{p} x^{p-1}.$$

Then $g(x)$ satisfies Eisenstein's criterion with the prime p (Corollary 3.5), so g is irreducible in $\mathbb{Z}[x]$. Finally, irreducibility is preserved under the change of variables $x \mapsto x+1$ (an automorphism of $\mathbb{Q}[x]$ with inverse $x \mapsto x-1$). Hence $f(x) = g(x-1)$ is irreducible over $\mathbb{Q}$, and by Gauss' lemma f is irreducible in $\mathbb{Z}[x]$ (Lemma 3.3). The result follows. □

Example 3.23. Let $p \in \mathbb{Z}$ be a prime. Then,

$$x^n - p \quad \text{is} \quad \text{irreducible over } \mathbb{Z}[x].$$

Example 3.24. Define $R = \mathbb{R}[x][y] = \mathbb{R}[x,y]$. Then, $y^n - x$ is irreducible over R. As such, noting that

$$R/(x) \cong \mathbb{R}[y] \quad \text{is} \quad \text{an integral domain,}$$

it follows that x is prime. Knowing that $\mathbb{R}[x,y]$ is a UFD (use Corollary 3.4), where we note that $\mathbb{R}$ is a UFD), then $y^n - x$ is a prime, so $R/(y^n - x)$ is an integral domain.

Lemma 3.7. Let R be a commutative ring with multiplicative identity 1. Then,

$$R \text{ is Noetherian} \quad \text{if and only if} \quad \text{every ideal is finitely generated.}$$

Theorem 3.2. Let R be a commutative ring with multiplicative identity 1. Then,

$$R \text{ is Noetherian} \quad \text{implies} \quad R[x] \text{ is Noetherian.}$$

A standard way to adjoin new elements to a ring is to take a polynomial ring and then impose a relation by taking the quotient with an ideal. Here we adjoin an element x to $\mathbb{R}$ subject to the relation $x^2+1=0$, i.e. we formally force $x^2 = -1$. Since the complex numbers are precisely obtained from $\mathbb{R}$ by adjoining an element i with $i^2 = -1$, it is natural to expect that

$$\mathbb{R}[x]/(x^2+1)$$

is a concrete algebraic model for $\mathbb{C}$ (Example 3.25.

Example 3.25 (Dummit and Foote p. 311 Question 7). Prove that $\mathbb{R}[x]/\left(x^2+1\right)$ is a field which is isomorphic to the complex numbers $\mathbb{C}$.

Solution. Let $I = (x^2+1) \subseteq \mathbb{R}[x]$ and write $\overline{f(x)}$ for the residue class of $f(x)$ modulo I. Since $\mathbb{R}[x]$ is an ED (hence a PID by Theorem 2.2), an ideal generated by an irreducible polynomial is maximal. Thus it suffices to show that x^2+1 is irreducible in $\mathbb{R}[x]$.

Because x^2+1 has degree 2, it is reducible over $\mathbb{R}$ if and only if it has a real root. But $x^2+1=0$ would imply $x^2 = -1$, which has no solution in $\mathbb{R}$. Hence, x^2+1 is irreducible in $\mathbb{R}[x]$, so (x^2+1) is maximal, and therefore $\mathbb{R}[x]/(x^2+1)$ is a field.

By the division algorithm in $\mathbb{R}[x]$, for each $f(x) \in \mathbb{R}[x]$ there exist $q(x), r(x) \in \mathbb{R}[x]$ with $\deg r < 2$ such that

$$f(x) = q(x)(x^2+1) + r(x).$$

Write $r(x) = a + bx$ for some $a, b \in \mathbb{R}$. Passing to the quotient gives

$$\overline{f(x)} = \overline{r(x)} = \overline{a+bx} = a + b\bar{x}.$$

In particular, every element of $\mathbb{R}[x]/(x^2+1)$ can be written uniquely as $a + b\bar{x}$ with $a, b \in \mathbb{R}$, and moreover,

$$\bar{x}^2 = \overline{x^2} = \overline{-1} \quad \text{in } \mathbb{R}[x]/(x^2+1).$$

Now, we construct the isomorphism to $\mathbb{C}$. Define a map

$$\varphi : \mathbb{R}[x]/(x^2+1) \to \mathbb{C} \quad \text{where} \quad \overline{f(x)} \mapsto f(i).$$

Here, $i \in \mathbb{C}$ satisfies $i^2 = -1$. Note that the map is a well-defined homomorphism, which is injective and surjective. So, φ is a bijective ring homomorphism and hence it is a field isomorphism. As such,

$$\mathbb{R}[x]/(x^2+1) \cong \mathbb{C}$$

as fields. □

Chapter 4

Introduction to Module Theory

4.1 Basic Definitions and Examples

We first collect the basic definitions and first examples of modules, the natural generalisation of vector spaces in which the scalars come from an arbitrary ring rather than a field. Concretely, an R-module is an Abelian group $(M,+)$ together with an action of R on M that is compatible with the ring operations; this recovers familiar linear algebra when R is a field, but also allows genuinely new phenomena (e.g. torsion, non-trivial annihilators, and asymmetric left/right actions when R is non-commutative).

After introducing left and right modules (Definitions 4.1 and 4.2) and the notion of a submodule (Definition 4.3), we record a few foundational identities (such as $0m = 0$ and $(-1)m = -m$), explain how ring homomorphisms allow restriction of scalars, and construct the submodule generated by a subset. These results will be used repeatedly: they supply the basic closure properties needed to build new modules from old ones (submodules, quotients, sums etc.) and set up the ideal-module dictionary that underlies much of the structure theory to come.

Definition 4.1 (left R-module). Let R be a ring. A left R-module is an Abelian group M equipped with a map (known as the action of R), defined by

$$R \times M \to M \quad \text{where} \quad (r,m) \mapsto rm = r \cdot m$$

such that for any $r,s \in R$ and $m,n \in M$, the following hold:

(i) **Distributivity:** $(r+s)\,m = rm + sm$

(ii) **Associativity:** $(rs)\cdot m = r\cdot(sm)$

(iii) **Distributivity:** $r\,(m+n) = rm + rs$

(iv) $1 \cdot m = m$ if R has a multiplicative identity

Recall Definition 1.6. Note that an R-module structure on M is equivalent to a ring

homomorphism $R \to \mathrm{End}_{\mathrm{Ab}}(M)$, where Ab is some Abelian group — $\mathrm{End}_{\mathrm{Ab}}(M)$ denotes the ring of all group endomorphisms of M (i.e. homomorphisms from M to itself). We also require 1_R to map to the identity map on M if 1_R exists.

Example 4.1 (Dummit and Foote p. 343 Question 1). Let R be a ring with 1 and M be a left R-module. Prove that $0m = 0$ and $(-1)m = -m$ for all $m \in M$.

Solution. Since $0 = 0+0$ in R, by distributivity of the scalar action over addition in R, we have $0m = (0+0)m = 0m+0m$. Add the additive inverse of $0m$ (in the abelian group $(M,+)$) to both sides to get $0m = 0$.

We then prove that $(-1)m = -m$. Since $1+(-1) = 0$ in R, then by distributivity, we have $(1+(-1))m = 1m+(-1)m$ so $0m = 1m+(-1)m$. Substituting $0m = 0$ and $1m = m$ gives $0 = m+(-1)m$. Thus, $(-1)m$ is the additive inverse of m in M, i.e. $(-1)m = -m$. □

Example 4.2 (Dummit and Foote p. 343 Question 3). Let R be a ring with 1 and M be a left R-module. Assume that $rm = 0$ for some $r \in R$ and some $m \in M$ with $m \neq 0$. Prove that r does not have a left inverse (i.e. there is no $s \in R$ such that $sr = 1$).

Solution. Suppose on the contrary that r has a left inverse. Then, there exists $s \in R$ such that $sr = 1$. Right multiplying both sides by m, we have $srm = m$, so $m = 0$, but this is a contradiction. □

Definition 4.2 (right R-module). Let R be a ring with 1 and M be a left R-module. A right R-module is an Abelian group M equipped with a map (known as the action of R), defined by

$$M \times R \to M \quad \text{where} \quad (m,r) \mapsto mr = m \cdot r$$

such that for any $r,s \in R$ and $m,n \in M$, the following hold:

(i) **Distributivity:** $m(r+s) = mr+ms$

(ii) **Associativity:** $m \cdot (rs) = (mr) \cdot s$

(iii) **Distributivity:** $(m+n)r = mr+nr$

(iv) $m \cdot 1 = m$ if R has a multiplicative identity

Although Definitions 4.1 and 4.2 hold, we often only consider left actions, or left modules over R (Definition 4.1). We just refer to them as R-actions or R-modules. Note that if R is commutative, then the left action is the same as the right action.

Example 4.3 (the trivial R-module). We have the trivial 0 module for any ring R. To see why, consider the Abelian group $\{0\}$, known as the *zero group*. Define an action of R on 0 by

$$R \times 0 \to 0 \quad \text{where} \quad (r,0) \mapsto r \cdot 0 = 0.$$

This is the only possible action since 0 has only one element. We verify the module axioms. For all $r,s \in R$, we have

$$(r+s)\cdot 0 = 0 = r\cdot 0 + s\cdot 0 \quad \text{and} \quad (rs)\cdot 0 = 0 = r\cdot(s\cdot 0).$$

Moreover, since $0+0=0$ in the group 0, then

$$r\cdot(0+0) = r\cdot 0 = 0 = r\cdot 0 + r\cdot 0.$$

If R has a multiplicative identity 1, then $1\cdot 0 = 0$. Hence, 0 is an R-module for every ring R. This module is called the zero module or the trivial module.

Example 4.4 (F-module). For any field F, the F-modules are just the F-vector spaces.

Example 4.5. For any Abelian group M, we say that M is a $\mathbb{Z}$-module.

Example 4.6. Let R be a ring. Then, R is a left R-module under left multiplication and a right R-module under right multiplication.

Example 4.7. Let I be a left ideal of R. Then, I is a left R-module. Actually, I is a left R-submodule of R.

Example 4.8. Let I be a left ideal of R. Then, R/I (the quotient Abelian group) is a left R-module.

Example 4.9 (F adjoin x module). Let F be a field. An $F[x]$-module structure on an Abelian group V is a map

$$F[x]\times V \to V \quad \text{where} \quad (f,v)\mapsto f\cdot v$$

such that for all $f,g \in F[x]$, $v,w \in V$, and $a \in F$, the following hold:

$$\begin{aligned}(f+g)\cdot v &= f\cdot v + g\cdot v\\ f\cdot(v+w) &= f\cdot v + f\cdot w\\ (fg)\cdot v &= f\cdot(g\cdot v)\\ 1\cdot v &= v\end{aligned}$$

and (since $F \subseteq F[x]$) scalar multiplication by $a \in F$ agrees with the given F-vector space structure. Since $F[x]$ is generated (as an F-algebra) by the single element x, specifying an $F[x]$-module structure is essentially the same as specifying how x acts on V.

More precisely, define a map

$$T : V \to V \quad \text{where} \quad T(v) = x\cdot v.$$

Then, the module axioms force T to be F-linear:

$$T(v+w) = x\cdot(v+w) = x\cdot v + x\cdot w = T(v)+T(w),$$

and for $a \in F$,

$$T(av) = x\cdot(av) = (xa)\cdot v = (ax)\cdot v = a(x\cdot v) = aT(v),$$

using commutativity of $F[x]$ and the requirement that scalars in $F \subseteq F[x]$ act as the usual scalar multiplication. Conversely, given any F-linear map $T : V \to V$, we can define an $F[x]$-action on V by

$$\left(\sum_{k=0}^{m} a_k x^k\right) \cdot v = \sum_{k=0}^{m} a_k T^k(v),$$

where $T^0 = \mathrm{id}_V$ and $T^{k+1} = T \circ T^k$. One checks directly that this satisfies the module axioms. Hence, giving an $F[x]$-module structure on an F-vector space V is equivalent to giving an F-linear operator $T : V \to V$, with x acting as T.

With $T(v) = x \cdot v$, the action of a polynomial $f(x) = a_0 + a_1 x + \cdots + a_m x^m$ is

$$f \cdot v = a_0 v + a_1 T(v) + a_2 T^2(v) + \cdots + a_m T^m(v) = f(T)(v).$$

So saying that f acts on V literally means to evaluate the polynomial f at the operator T and apply it to v.

Definition 4.3 (submodule). Let M be an R-module. Then, a subgroup $N \subseteq M$ is an R-submodule of M if N is closed under the R-action.

Example 4.10. The zero module is a trivial submodule of any R-module.

Example 4.11. Let $S \subseteq R$ be a subring. Then, S is an S-submodule of R.

Example 4.12. Let M be an Abelian group or equivalently a $\mathbb{Z}$-module (Example 4.5). Then, a subgroup of N is the same as a $\mathbb{Z}$-submodule of M.

Example 4.13 (Dummit and Foote p. 343 Question 4). Let M be the module R^n as follows:

$$M = R^n = \{(a_1, \ldots, a_n) : a_i \in R\}$$

and let $I_1, I_2, \ldots, I_n$ be left ideals of R. In due course, you will learn that M is the free left R-module of rank n. Prove that the following are submodules of M:

(a) $\{(x_1, x_2, \ldots, x_n) \mid x_i \in I_i\}$

(b) $\{(x_1, x_2, \ldots, x_n) \mid x_i \in R \text{ and } x_1 + x_2 + \cdots + x_n = 0\}$

Solution.

(a) Let N_1 denote the mentioned set. We wish to prove that N_1 is a submodule of M. First, $N_1 \neq \emptyset$ since $(0, \ldots, 0) \in N_1$ (each I_i contains 0). Next, if $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ lie in N_1, then $x_i, y_i \in I_i$ for each i, hence $x_i + y_i \in I_i$ because I_i is an additive subgroup of R. Therefore,

$$x + y = (x_1 + y_1, \ldots, x_n + y_n) \in N_1.$$

Finally, if $r \in R$ and $x = (x_1, \ldots, x_n) \in N_1$, then $x_i \in I_i$ for each i and since I_i is a *left* ideal we have $rx_i \in I_i$ for each i. Hence

$$rx = (rx_1, \ldots, rx_n) \in N_1.$$

Thus N_1 is non-empty and closed under addition and left scalar multiplication, so N_1 is a submodule of M.

(b) Let N_2 denote the mentioned set. Again, $N_2 \neq \emptyset$ since $(0,\dots,0) \in N_2$. If $x = (x_1,\dots,x_n)$ and $y = (y_1,\dots,y_n)$ lie in N_2, then

$$x_1 + \cdots + x_n = 0 \quad \text{and} \quad y_1 + \cdots + y_n = 0.$$

Hence,

$$(x_1+y_1) + \cdots + (x_n+y_n) = (x_1 + \cdots + x_n) + (y_1 + \cdots + y_n) = 0 + 0 = 0,$$

so $x+y \in N_2$. Finally, if $r \in R$ and $x = (x_1,\dots,x_n) \in N_2$, then $x_1 + \cdots + x_n = 0$, so by left distributivity,

$$(rx_1) + \cdots + (rx_n) = r(x_1 + \cdots + x_n) = r \cdot 0 = 0,$$

and therefore $rx \in N_2$. Thus N_2 is non-empty and closed under addition and left scalar multiplication, hence it is a submodule of M. □

Example 4.14 (Dummit and Foote p. 344 Question 17). For any field F, let T be the shift operator on the vector space $V = F^n$ and let $\mathbf{e}_1,\dots,\mathbf{e}_n$ be the standard basis vectors from the usual example of $F[x]$-modules. The shift operator is defined to be

$$T(\mathbf{e}_i) = \begin{cases} \mathbf{e}_{i+1} & \text{if } 1 \le i \le n-1; \\ \mathbf{0} & \text{if } i = n. \end{cases}$$

For $m \geq n$, compute

$$\left(a_m x^m + a_{m-1}x^{m-1} + \cdots + a_0\right)\mathbf{e}_n.$$

Solution. Note that the action of

$$f(x) = \sum_{k\geq 0} a_k x^k \in F[x] \quad \text{is} \quad \text{given by } f(x)\cdot \mathbf{v} = \sum_{k\geq 0} a_k T^k(\mathbf{v}).$$

Let

$$f(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_0 \quad \text{where } m \geq n.$$

Then,

$$f(x)\mathbf{e}_n = \sum_{k=0}^{m} a_k\, T^k(\mathbf{e}_n).$$

But $T(\mathbf{e}_n) = \mathbf{0}$, hence $T^k(\mathbf{e}_n) = \mathbf{0}$ for every $k \geq 1$. Therefore every term with $k \geq 1$ vanishes, and we get $\left(a_m x^m + a_{m-1}x^{m-1} + \cdots + a_0\right)\mathbf{e}_n = a_0\mathbf{e}_n$. □

Example 4.15 (Dummit and Foote p. 344 Question 18). Let $F = \mathbb{R}$ denote the field of real numbers, $V = \mathbb{R}^2$, and let $T : V \to V$ be the linear map given by clockwise rotation about the origin through angle $\pi/2$. Prove that the only $F[x]$-submodules (i.e. T-invariant subspaces) are the zero subspace and V.

Solution. Let $F = \mathbb{R}$, $V = \mathbb{R}^2$, and let $T : V \to V$ be clockwise rotation by $\pi/2$ about the origin. Explicitly, with respect to the standard basis, we have

$$T(x,y) = (y,-x) \quad \text{and} \quad [T] = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Let W be an $\mathbb{R}$-subspace that is T-invariant. We show $W = \{0\}$ or $W = V$.

If $W = \{0\}$, there is nothing to prove. Otherwise pick a non-zero vector $w \in W$. Since W is T-invariant, we have $T(w) \in W$ as well. We claim that w and $T(w)$ are linearly independent over $\mathbb{R}$. Indeed, write $w = (a,b) \neq (0,0)$ so $T(w) = (b,-a)$. If $T(w) = \lambda w$ for some $\lambda \in \mathbb{R}$, then

$$(b,-a) = \lambda(a,b),$$

so $b = \lambda a$ and $-a = \lambda b$. Substituting $b = \lambda a$ into the second equation gives

$$-a = \lambda(\lambda a) = \lambda^2 a \quad \text{so} \quad (\lambda^2+1)a = 0.$$

If $a \neq 0$ this forces $\lambda^2 = -1$, which is impossible over $\mathbb{R}$. If $a = 0$, then $w = (0,b)$ with $b \neq 0$, and the first equation $b = \lambda a = 0$ is a contradiction. Hence no such real λ exists, so $T(w)$ is not a scalar multiple of w, and therefore $\{w, T(w)\}$ is linearly independent.

Consequently, $\dim_{\mathbb{R}}(\operatorname{span}\{w,T(w)\}) = 2$, so $V = \operatorname{span}\{w,T(w)\} \subseteq W$. Thus $W = V$. Therefore the only T-invariant subspaces of V are $\{0\}$ and V. □

Proposition 4.1. Let R be a commutative ring with multiplicative identity 1. Let M be an R-module.

(i) We have

$$0 \cdot m = 0 \text{ and } -1 \cdot m = -m \quad \text{for any } m \in M$$

The intersection of any non-empty collection of submodules of M is also an R-submodule

(ii) The annihilator of M in R,

$$\operatorname{Ann}_M(R) = \{r \in R : rm = 0 \text{ for any } m \in M\} \quad \text{is an ideal of } R$$

(iii) Let $z \in Z(R)$ (center of R). Then,

$$zM = \{zm : m \in M\} \quad \text{is an } R\text{-submodule of } M.$$

Corollary 4.1. For any commutative ring with multiplicative identity 1, let M be an R-module. Let $I \subseteq \operatorname{Ann}_M(R)$ be an ideal. Then, M is an R/I-module.

Proposition 4.2. Let R be a commutative ring with multiplicative identity 1. Let M

be an R-module and $N \subseteq M$. Define

$$RN = \left\{ \sum_{\text{finite}} an : a \in R, n \in N \right\},$$

which is the set of finite R-linear combinations. This is an R-submodule of M generated by N.

Proof. We first prove that RN is an R-submodule of M. We check the submodule axioms (Definition 4.3). Firstly, the sum is non-empty since taking the empty sum shows that $0 \in RN$. Next, we prove that RN is closed under addition. Let $x, y \in RN$. Then, we can write

$$x = \sum_{i=1}^{k} a_i n_i \quad \text{and} \quad y = \sum_{j=1}^{\ell} b_j m_j$$

where $a_i, b_j \in R$ and $n_i, m_j \in N$. It is then clear that the sum $x + y \in RN$. Lastly, one can prove that if $x \in RN$, then $-x \in RN$. Closure under scalar multiplication is clear too.

We then prove that RN is the submodule generated by N. First, $N \subseteq RN$ because for each $n \in N$ we have $n = 1 \cdot n \in RN$. Now, let $L \leq M$ be any R-submodule with $N \subseteq L$. Then for any finite sum

$$x = \sum_{i=1}^{k} a_i n_i \quad \text{where } n_i \in N \subseteq L,$$

we have $a_i n_i \in L$ (since L is closed under scalar multiplication) and hence $x \in L$ (since L is closed under addition). Thus $RN \subseteq L$. Therefore RN is the smallest R-submodule of M containing N, i.e. the R-submodule of M generated by N. □

Proposition 4.3. Let R be a commutative ring with multiplicative identity 1. Let $M_1, \ldots, M_n$ be R-submodules of an R-module M. We define the R-submodule of M as follows:

$$M_1 + \ldots + M_n = \{m_1 + \ldots + m_n \in M : m_i \in M_i, n \in M\}$$

Compare this with **(a)** of Example 4.13.

Proof. Set $S = M_1 + \cdots + M_n$. Since each M_i is a submodule, $0 \in M_i$ for all i. Hence,

$$0 = 0 + \cdots + 0 \in S,$$

so $S \neq \emptyset$.

We then prove that S is closed under addition. Let $x, y \in S$. Then, there exist $m_i, m_i' \in M_i$ such that

$$x = m_1 + \cdots + m_n \quad \text{and} \quad y = m_1' + \cdots + m_n'.$$

Therefore,

$$x+y=(m_1+m_1')+\cdots+(m_n+m_n').$$

Since each M_i is a submodule, it is closed under addition, so $m_i+m_i'\in M_i$ for all i. Hence, $x+y\in S$.

Lastly, we prove that S is closed under scalar multiplication. Let $r\in R$ and $x\in S$, say $x=m_1+\cdots+m_n$ with $m_i\in M_i$. Then by distributivity in the R-module M,

$$rx=r(m_1+\cdots+m_n)=rm_1+\cdots+rm_n.$$

Since each M_i is a submodule, it is closed under scalar multiplication, so $rm_i\in M_i$ for all i. Hence $rx\in S$.

Thus S is non-empty and closed under addition and scalar multiplication, so S is an R-submodule of M. □

Example 4.16 (Dummit and Foote p. 343 Question 5). Let R be a ring with 1 and M be a left R-module. For any left ideal I of R define

$$IM=\left\{\sum_{\text{finite}} a_im_i \mid a_i\in I, m_i\in M\right\}$$

to be the collection of all finite sums of elements of the form am where $a\in I$ and $m\in M$. Prove that IM is a submodule of M. This is somewhat similar to Definition 1.26 on the sum of ideals.

Solution. Let

$$IM=\left\{\sum_{i=1}^{t} a_im_i \;\middle|\; t\in\mathbb{N}, a_i\in I, m_i\in M\right\}\subseteq M,$$

where I is a left ideal of R and M is a left R-module. We verify that IM is a submodule of M (Definition 4.3).

We first prove that IM is non-empty. Since I is a left ideal, $0\in I$, and since M is an R-module, then $0\in M$. As such, $0\cdot 0=0\in M$, hence $0=0\cdot 0\in IM$ as a finite sum with one term. Thus $IM\neq\emptyset$.

We then prove that IM is closed under addition. Let $x,y\in IM$. Then, there exist $t,u\in\mathbb{N}$ with

$$x=\sum_{i=1}^{t} a_im_i \quad\text{and}\quad y=\sum_{j=1}^{u} b_jn_j,$$

where $a_i,b_j\in I$ and $m_i,n_j\in M$. Then

$$x+y=\sum_{i=1}^{t} a_im_i+\sum_{j=1}^{u} b_jn_j=\sum_{k=1}^{t+u} c_k\ell_k,$$

where $(c_1,\dots,c_{t+u}) = (a_1,\dots,a_t,b_1,\dots,b_u)$ and $(\ell_1,\dots,\ell_{t+u}) = (m_1,\dots,m_t,n_1,\dots,n_u)$. Each $c_k \in I$ and each $\ell_k \in M$, so $x+y \in IM$.

Closure under additive inverse is easy. In other words, given $x \in IM$, one can easily show that $-x \in IM$. To show closure under scalar multiplication by R, let $r \in R$ and $x \in IM$, where $x = \sum_{i=1}^{t} a_i m_i$ with $a_i \in I$. Using distributivity and associativity of the module action, we have

$$rx = r\left(\sum_{i=1}^{t} a_i m_i\right) = \sum_{i=1}^{t} r(a_i m_i) = \sum_{i=1}^{t} (ra_i)m_i.$$

Because I is a *left* ideal, $ra_i \in I$ for each i. Therefore $rx \in IM$. So, IM is an R-submodule of M. □

Example 4.17 (Dummit and Foote p. 343 Question 6). Let R be a ring with 1. Show that the intersection of any non-empty collection of submodules of an R-module is a submodule. This is similar to Example 1.48 on the intersection of ideals.

Solution. Let M be a left R-module, and let $\{N_\lambda\}_{\lambda\in\Lambda}$ be a non-empty collection of R-submodules of M. As such, $\Lambda \neq \emptyset$. We verify that N is an R-submodule of M.

First, we prove that N is non-empty. Since each N_λ is a submodule, then $0 \in N_\lambda$ for every $\lambda \in \Lambda$. Hence, 0 is contained in the intersection, so $N \neq \emptyset$. We then prove that N is closed under addition. Let $x,y \in N$. Then, $x,y \in N_\lambda$ for every $\lambda \in \Lambda$. Since each N_λ is a submodule, then $x+y \in N_\lambda$ for all λ. So, the sum $x+y \in N$. Closure under additive inverses and scalar multiplication is straightforward — it is just the same old steps as before. □

Example 4.18 (Dummit and Foote p. 344 Question 7). Let R be a ring with 1 and M be a left R-module. Let $N_1 \subseteq N_2 \subseteq \cdots$ be an ascending chain of submodules of M. Prove that the union $N_1 \cup N_2 \cup \cdots$ is a submodule of M.

We omit the solution to this example. Recall Example 1.47 for a similar result.

Definition 4.4 (torsion). Let R be a ring with 1 and M be a left R-module. An element m of M is called a torsion element if $rm = 0$ for some non-zero $r \in R$. The set of torsion elements is denoted by

$$\mathrm{Tor}(M) = \{m \in M : rm = 0 \text{ for some non-zero } r \in R\}.$$

Definition 4.5 (annihilator). Let R be a ring with 1 and M be a left R-module.

(i) If N is a submodule of M, the annihilator of N in R is defined to be

$$\{r \in R : rn = 0 \text{ for all } n \in N\}.$$

(ii) If I is a right ideal of R, the annihilator of I in M is defined to be

$$\{m \in M : am = 0 \text{ for all } a \in I\}.$$

Example 4.19 (Dummit and Foote p. 344 Question 9). Let R be a ring with 1 and M be a left R-module. Suppose N is a submodule of M. Recall Definition 4.5 on the annihilator of N in R. Prove that this annihilator is a 2-sided ideal of R.

Solution. Let the annihilator of N in R be denoted by $\mathrm{Ann}_N(R)$. We prove that it is a two-sided ideal of R.

First, we prove that the annihilator is non-empty. Note that for all $n \in N$, we have $0 \cdot n = 0$ so $0 \in \mathrm{Ann}_R(N)$. We then prove closure under addition. Suppose $r, s \in \mathrm{Ann}_R(N)$. Then, for every $n \in N$, we have

$$(r+s)n = rn + sn = 0 + 0 = 0 \quad \text{so} \quad r+s \in \mathrm{Ann}_R(N).$$

One then needs to show that the annihilator is closed under additive inverses, left and right multiplication by arbitrary elements, so by Definition 1.24, the result follows. □

Example 4.20 (Dummit and Foote p. 344 Question 10). Let R be a ring with 1 and M be a left R-module. Suppose I is a right ideal of R. Recall Definition 4.5 on the annihilator of I in M. Prove that this annihilator is a submodule of M.

Solution. Let the annihilator be denoted by $\mathrm{Ann}_M(I)$. We need to show that the annihilator is a submodule of M (Definition 4.3). So, it suffices to prove that it is non-empty, closed under addition, closed under additive inverses, and scalar multiplication.

We will only prove closure under scalar multiplication and leave the rest as simple exercises for the reader. Suppose $a \in R$ and $m \in \mathrm{Ann}_M(I)$. We must show that $am \in \mathrm{Ann}_M(I)$, i.e. for any $i \in I$, we have $i(am) = 0$. This follows by the associativity of the module action as we have $i(am) = (ia)m$. Since I is a right ideal, then $ia \in I$. Next, since $m \in \mathrm{Ann}_M(I)$, then $(ia)m = 0$ and the result follows. □

Example 4.21 (Dummit and Foote p. 344 Question 8). Let R be a ring with 1 and M be a left R-module. Recall Definition 4.4.

(a) Prove that if R is an integral domain then $\mathrm{Tor}(M)$ is a submodule of M (known as the torsion submodule of M)

(b) Give an example of a ring R and an R-module M such that $\mathrm{Tor}(M)$ is not a submodule[1].

(c) If R has zero divisors, show that every non-zero R-module has non-zero torsion elements.

Solution.

(a) Suppose R is an integral domain. We show that $\mathrm{Tor}(M)$ is an R-submodule of M (Definition 4.3). First, $\mathrm{Tor}(M)$ is non-empty because $r \cdot 0_M = 0$ for all non-zero

[1] Consider the torsion elements in the R-module R.

$r \in R$ so $0_M \in \mathrm{Tor}(M)$. Next, closure under addition is clear too. Let $m, n \in \mathrm{Tor}(M)$. Then by Definition 4.4, there exist non-zero $r, s \in R$ such that

$$rm = 0 \quad \text{and} \quad sn = 0.$$

Now, consider $rs \in R$. Since R is an integral domain and $r, s \neq 0$, then $rs \neq 0$. So,

$$(rs)(m+n) = (rs)m + (rs)n = s(rm) + r(sn) = s \cdot 0 + r \cdot 0 = 0.$$

This shows that $m+n \in \mathrm{Tor}(M)$. As before, showing closure under additive inverses and scalar multiplication is straightforward so we omit the solution.

(b) Let $R = \mathbb{Z}/6\mathbb{Z}$ which is a ring. We then view $M = R$ as a left R-module over itself. Take $2, 3 \in \mathrm{Tor}(R)$ since

$$3 \cdot 2 = 6 \equiv 0 \pmod{6} \quad \text{and} \quad 2 \cdot 3 = 6 \equiv 0 \pmod{6}.$$

Clearly, $2 \neq 0$, $3 \neq 0$ in R, with $3 \neq 0$, $2 \neq 0$ as annihilators (Definition 4.5). However, $2 + 3 = 5 \notin \mathrm{Tor}(R)$, since no non-zero $r \in R$ kills 5. As $\mathrm{Tor}(R)$ is not closed under addition, it is not a submodule.

(c) Assume R has zero divisors, and let M be a non-zero left R-module. Choose $0 \neq m \in M$. Since R has zero divisors, there exist $a, b \in R$ with $a \neq 0$, $b \neq 0$ and $ab = 0$. Consider $bm \in M$.

If $bm \neq 0$, then bm is a non-zero torsion element because $a(bm) = (ab)m = 0 \cdot m = 0$ with $a \neq 0$. On the other hand, if $bm = 0$, then m itself is a non-zero torsion element (killed by the non-zero scalar b). In either case, M has a non-zero torsion element, i.e. $\mathrm{Tor}(M) \neq 0$. □

Example 4.22 (Dummit and Foote p. 344 Question 11). Let M be the abelian group (i.e. $\mathbb{Z}$-module) $\mathbb{Z}/24\mathbb{Z} \times \mathbb{Z}/15\mathbb{Z} \times \mathbb{Z}/50\mathbb{Z}$.

(a) Find the annihilator of M in $\mathbb{Z}$ (i.e. a generator for this principal ideal).

(b) Let $I = 2\mathbb{Z}$. Describe the annihilator of I in M as a direct product of cyclic groups.

Solution.

(a) Note that the prime factorisations of 24, 15, and 50 are $24 = 2^3 \cdot 3$, $15 = 3 \cdot 5$ and $50 = 2 \cdot 5^2$. Note that elements in m are of the form

$$(a, b, c) \quad \text{where } a \in \mathbb{Z}/24\mathbb{Z}, b \in \mathbb{Z}/15\mathbb{Z}, c \in \mathbb{Z}/50\mathbb{Z}.$$

For a direct product of Abelian groups, the action is component-wise, so n annihilates M if and only if it annihilates each factor. That is, to consider

$$n \cdot (a, b, c) = (na, nb, nc) = (0, 0, 0).$$

Thus, n must be divisible by 24, 15, and 50. Hence, by considering the lowest common multiple of 24, 15, an 50, the annihilator is $600\mathbb{Z}$.

(b) By Definition 4.5, the annihilator of $I = 2\mathbb{Z}$ in M is

$$\begin{aligned}\mathrm{Ann}_M(2\mathbb{Z}) &= \{m \in M : im = 0 \text{ for all } i \in 2\mathbb{Z}\} \\ &= \{m \in \mathbb{Z}/24\mathbb{Z} \times \mathbb{Z}/15\mathbb{Z} \times \mathbb{Z}/50\mathbb{Z} : 2km = 0 \text{ for all } k \in \mathbb{Z}\}\end{aligned}$$

As such, we need

$$2ka = 0, 2kb = 0, 2kc = 0 \quad \text{where } a \in \mathbb{Z}/24\mathbb{Z}, b \in \mathbb{Z}/15\mathbb{Z}, c \in \mathbb{Z}/50\mathbb{Z}.$$

It is easy to deduce that $\mathrm{Ann}_M(2\mathbb{Z}) \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.[2] □

Example 4.23 (Dummit and Foote p. 344 Question 12). Prove the following facts about annihilators.

(a) Let N be a submodule of M and let I be its annihilator in R. Prove that the annihilator of I in M contains N. Give an example where the annihilator of I in M does not equal N.

(b) Let I be a right ideal of R and let N be its annihilator in M. Prove that the annihilator of N in R contains I. Give an example where the annihilator of N in R does not equal I.

Solution. Throughout, R is a ring with 1 and M is a left R-module. For a subset $X \subseteq M$ and a subset $J \subseteq R$, we use the notations

$$\begin{aligned}\mathrm{Ann}_R(X) &= \{r \in R : rx = 0 \text{ for all } x \in X\} \\ \mathrm{Ann}_M(J) &= \{m \in M : jm = 0 \text{ for all } j \in J\}\end{aligned}$$

(a) Let N be a submodule of M and let $I = \mathrm{Ann}_R(N)$. We claim that $N \subseteq \mathrm{Ann}_M(I)$. Indeed, take $n \in N$ and $i \in I$. Since $i \in \mathrm{Ann}_R(N)$, we have $in = 0$. Thus $in = 0$ for all $i \in I$, so $n \in \mathrm{Ann}_M(I)$, proving $N \subseteq \mathrm{Ann}_M(I)$.

We give an example where strict containment occurs. Let $R = \mathbb{Z}$ and $M = \mathbb{Z}$ as a $\mathbb{Z}$-module. Take $N = 2\mathbb{Z}$ which is a submodule of M. Then,

$$I = \mathrm{Ann}_{\mathbb{Z}}(2\mathbb{Z}) = \{a \in \mathbb{Z} : a(2k) = 0 \text{ for all } k \in \mathbb{Z}\} = \{0\}.$$

Hence,

$$\mathrm{Ann}_M(I) = \mathrm{Ann}_{\mathbb{Z}}(\{0\}) = \{m \in \mathbb{Z} : 0 \cdot m = 0\} = \mathbb{Z},$$

so $\mathrm{Ann}_M(I) = \mathbb{Z} \supsetneq 2\mathbb{Z} = N$.

(b) Let I be a right ideal of R and let $N = \mathrm{Ann}_M(I)$. We claim that $I \subseteq \mathrm{Ann}_R(N)$. Indeed, take $i \in I$. For any $m \in N = \mathrm{Ann}_M(I)$ we have $im = 0$. Since this holds for all $m \in N$, it follows that $i \in \mathrm{Ann}_R(N)$. Thus $I \subseteq \mathrm{Ann}_R(N)$.

[2] This is implicitly $\mathrm{Ann}_M(2\mathbb{Z}) \cong \mathbb{Z}/2\mathbb{Z} \times \{0\} \times \mathbb{Z}/2\mathbb{Z}$.

We give an example where strict containment occurs. Again let $R = \mathbb{Z}$ and $M = \mathbb{Z}$, and take the (two-sided, hence right) ideal $I = 2\mathbb{Z}$. Then,

$$N = \mathrm{Ann}_{\mathbb{Z}}(2\mathbb{Z}) = \{m \in \mathbb{Z} : (2k)m = 0 \text{ for all } k \in \mathbb{Z}\} = \{0\}.$$

Therefore

$$\mathrm{Ann}_R(N) = \mathrm{Ann}_{\mathbb{Z}}(\{0\}) = \{a \in \mathbb{Z} : a \cdot 0 = 0\} = \mathbb{Z},$$

so $\mathrm{Ann}_R(N) = \mathbb{Z} \supsetneq 2\mathbb{Z} = I$. □

4.2 Quotient Modules and Module Homomorphisms

One reason modules are so useful is that they support the same two organising ideas that run throughout group theory and ring theory: homomorphisms and quotients.

Definition 4.6 (module homomorphism). Let M and N be R-modules. An R-module homomorphism is a map $\varphi : M \to N$ satisfying the following properties:

$$\varphi(x+y) = \varphi(x) + \varphi(y) \text{ for } x, y \in M \quad \text{and} \quad \varphi(rx) = r\varphi(x) \text{ for } r \in R, x \in M.$$

Definition 4.7 (module isomorphism). An R-module homomorphism $\varphi : M \to N$ is an isomorphism if it is a bijection, i.e. there also exists an R-module homomorphism $\psi : N \to M$ such that

$$\varphi \circ \psi = \mathrm{id}_N \quad \text{and} \quad \psi \circ \varphi = \mathrm{id}_M.$$

On p. 350 of Dummit and Foote [1], Question 2 asks readers to prove that the relation 'is R-module isomorphic to' is an equivalence relation on any set of R-modules. As expected, the notion is very similar to proving that two groups/rings are isomorphic and one can use Definition 4.7.

Example 4.24 (Dummit and Foote p. 350 Question 3). Give an explicit example of a map from one R-module to another which is a group homomorphism but not an R-module homomorphism.

Solution. We want a map $f : M \to N$ between R-modules such that

$$f(m+m') = f(m) + f(m') \quad \text{for all } m, m' \in M$$

which signifies a group homomorphism, but there exists some $r \in R$ and $m \in M$ with $f(rm) \neq rf(m)$ so f fails to be an R-module homomorphism.

To come up with a counterexample, a standard choice is $R = \mathbb{Z}$ because $\mathbb{Z}$-modules are exactly abelian groups. But if we choose $R = \mathbb{Z}$, then group homomorphism and $\mathbb{Z}$-module homomorphism are the same thing, so we cannot get a counterexample. As such, we must choose a ring R larger than $\mathbb{Z}$, for example $R = \mathbb{Z}[i]$.

Then, take $M = N = R$ itself, viewed as a left R-module via multiplication. A very common trick is to use a ring automorphism that is additive but does not commute with the scalar action you want. For $R = \mathbb{Z}[i]$, complex conjugation $a + bi \mapsto a - bi$ is additive (hence a group homomorphism), but it is not R-linear because it sends i to $-i$, so it does not respect multiplication by i.

Often the scalar $r = i$ and element $m = 1$ already works. As a concrete example, let $R = \mathbb{Z}[i]$ and let $M = N = R$ as left R-modules (scalar action is multiplication in R). Define

$$f : R \to R \quad \text{where} \quad f(a+bi) = a - bi.$$

This denotes complex conjugation. Note that f is a group homomorphism because for any $z_1, z_2 \in R$, we have

$$f(z_1 + z_2) = \overline{z_1 + z_2} = \overline{z_1} + \overline{z_2} = f(z_1) + f(z_2),$$

so f is additive, hence a homomorphism of abelian groups. However, f is not an R-module homomorphism. To be R-linear, we would need $f(rz) = rf(z)$ for all $r, z \in R$. Take $r = i$ and $z = 1$. Then

$$f(i \cdot 1) = f(i) = \bar{i} = -i \quad \text{but} \quad if(1) = i \cdot \overline{1} = i \cdot 1 = i.$$

So, we have $f(i \cdot 1) \neq if(1)$, which implies f is not R-linear. □

Definition 4.8 (kernel and image). Let $\varphi : M \to N$ be an R-module homomorphism. Then, define

$$\ker \varphi = \{m \in M : \varphi(m) = 0 \text{ in } N\}$$
$$\operatorname{im} \varphi = \varphi(M) = \{\varphi(m) \text{ in } N : m \in M\}$$

Proposition 4.4 (Dummit and Foote p. 350 Question 1). Let $\varphi : M \to N$ be an R-module homomorphism. Then, $\ker \varphi$ is an R-submodule of M and $\operatorname{im} \varphi$ is an R-submodule of N.

The proof is straightforward and one can use Definition 4.3.

Definition 4.9 (endomorphism ring). Define

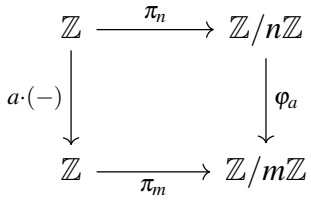
$$\operatorname{Hom}_R(M,N) \quad \text{to be} \quad \text{the set of } R\text{-module homomorphisms from } M \text{ to } N.$$

We often write $\operatorname{End}_R(M) = \operatorname{Hom}_R(M,M)$.

Example 4.25 (Dummit and Foote p. 350 Question 6). Prove that

$$\operatorname{Hom}_{\mathbb{Z}}(\mathbb{Z}/n\mathbb{Z}, \mathbb{Z}/m\mathbb{Z}) \cong \mathbb{Z}/\gcd(n,m)\mathbb{Z}.$$

The commutative diagram is as follows.



Solution. Let $d = \gcd(n,m)$. Any group homomorphism (equivalently $\mathbb{Z}$-module homomorphism) $\varphi : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}$ is determined uniquely by the image of $\overline{1} \in \mathbb{Z}/n\mathbb{Z}$: if $\varphi(\overline{1}) = \overline{a} \in \mathbb{Z}/m\mathbb{Z}$, then

$$\varphi(\overline{k}) = \varphi(k\overline{1}) = k\,\varphi(\overline{1}) = \overline{ka} \quad \text{where } \overline{k} \in \mathbb{Z}/n\mathbb{Z}.$$

The map is well-defined if and only if $\overline{n} = \overline{0}$ in the domain maps to 0 in the codomain, i.e.

$$0 = \varphi(\overline{0}) = \varphi(\overline{n}) = n\,\varphi(\overline{1}) = n\overline{a} = \overline{na} \in \mathbb{Z}/m\mathbb{Z},$$

equivalently $m \mid na$. Thus, homomorphisms correspond bijectively to residue classes $\overline{a} \in \mathbb{Z}/m\mathbb{Z}$ satisfying $m \mid na$. Let

$$H = \{\overline{a} \in \mathbb{Z}/m\mathbb{Z} : m \mid na\}.$$

Then H is a subgroup of the cyclic group $\mathbb{Z}/m\mathbb{Z}$, and the map $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/n\mathbb{Z}, \mathbb{Z}/m\mathbb{Z}) \to H$, where $\varphi \mapsto \varphi(\overline{1})$ is an isomorphism of abelian groups (pointwise addition of homomorphisms corresponds to addition in H).

It remains to identify H. Write $n = dn'$ and $m = dm'$ with $\gcd(n',m') = 1$. The condition $m \mid na$ becomes $dm' \mid dn'a$ if and only if $m' \mid n'a$. Since $\gcd(n',m') = 1$, this is equivalent to $m' \mid a$. Hence,

$$H = \{\overline{a} \in \mathbb{Z}/m\mathbb{Z} : a \equiv 0 \pmod{m'}\} = \langle \overline{m'} \rangle \leq \mathbb{Z}/m\mathbb{Z}.$$

The element $\overline{m'}$ has order

$$\frac{m}{\gcd(m,m')} = \frac{m}{m'} = d,$$

so H is a cyclic subgroup of order d. The result follows. □

We then collect some basic results about module homomorphisms.

Proposition 4.5. Let M and N be R-modules for a commutative ring R with multiplicative identity 1. If $\varphi, \psi \in \mathrm{Hom}_R(M,N)$, then

$$\varphi + \psi : M \to N \quad \text{defined by} \quad (\varphi + \psi)(m) = \varphi(m) + \psi(m)$$

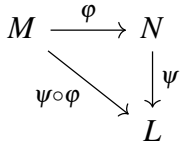
is also an R-module homomorphism.

Proposition 4.6 (composition). Let M, N, L be R-modules for a commutative ring R with multiplicative identity 1. Suppose

$$\varphi \in \mathrm{Hom}_R(M,N) \quad \text{and} \quad \psi \in \mathrm{Hom}_R(N,L).$$

Then,

$$\psi \circ \varphi \quad \text{is an } R\text{-module homomorphism from } M \text{ to } L.$$



Proposition 4.7. Let M be an R-module for a commutative ring R with multiplicative identity 1. Then, the map

$$\varphi : M \to M \quad \text{where} \quad m \mapsto rm \text{ for any } r \in R$$

is an R-module homomorphism.

Example 4.26. Let M and N be $\mathbb{Z}$-modules. Then, any homomorphism $M \to N$ is the same as a homomorphism of Abelian groups.

Example 4.27. For any field F, let V and W be F-vector spaces. Then, an F-module map from V to W is just a linear map from V to W.

Example 4.28. For any R-modules M and N, we can define the product module $M \oplus N = M \times N$ as the expected one (this is known as the direct sum of M and N which we will formally introduce in Definition 4.14).

Example 4.29 (Dummit and Foote p. 350 Question 5). Exhibit all $\mathbb{Z}$-module homomorphisms from $\mathbb{Z}/30\mathbb{Z}$ to $\mathbb{Z}/12\mathbb{Z}$.

Solution. This is similar to what was covered in MA2202 Algebra I. Since $\mathbb{Z}$-modules are the same as abelian groups, a $\mathbb{Z}$-module homomorphism $\varphi : \mathbb{Z}/30\mathbb{Z} \to \mathbb{Z}/12\mathbb{Z}$ is determined by the image of the generator $\overline{1} \in \mathbb{Z}/30\mathbb{Z}$. For any $\overline{k} \in \mathbb{Z}/30\mathbb{Z}$, let

$$\varphi_a(\overline{k}) = \overline{ak} \in \mathbb{Z}/12\mathbb{Z}.$$

We leave it to the reader to deduce that

$$\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/30\mathbb{Z}, \mathbb{Z}/12\mathbb{Z}) = \{\varphi_0, \varphi_2, \varphi_4, \varphi_6, \varphi_8, \varphi_{10}\},$$

so there are 6 such homomorphisms. □

Lemma 4.1. Let R and S be commutative rings with multiplicative identity 1. Let $\varphi : R \to S$ be a ring homomorphism such that $1_R \mapsto 1_S$. Let M be an S-module.

Then,

$$M \text{ is an } R\text{-module} \quad \text{via the action} \quad r \cdot m = \varphi(r) \cdot m.$$

Definition 4.10 (quotient R-module). Let R be a ring. Let M be an R-module with a submodule N. Define the quotient R-module M/N as follows:

(i) $M/N = M/N$ as Abelian groups

(ii) the R-action is given by

$$r(m+N) = rm+N \quad \text{for any } m \in M, r \in R.$$

Moreover, the natural quotient map $\pi : M \to M/N$ (Figure 4.1) is an R-module homomorphism.

Lemma 4.2 (universal property of the quotient). Let R be a ring. Let M be an R-module with a submodule N. Let L be another R-module. Then, for any R-module homomorphism $\varphi : M \to L$ such that $\varphi(N) = 0$, there exists a unique R-module homomorphism $\varphi : M/N \to L$ such that Figure 4.1 commutes.

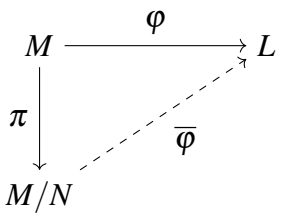


Figure 4.1: Universal property of the quotient

Theorem 4.1 (isomorphism theorems). Let R be a ring with multiplicative identity 1. Then, the following hold:

(i) Let M and N be R-modules and $\varphi : M \to N$ be an R-module homomorphism. Then,

$$M/\ker M \cong \varphi(M).$$

(ii) Let M, N be submodules of L. Then,

$$(M+N)/M \cong N/(M \cap N).$$

Here, $M+N$ is an R-submodule of L (Proposition 4.3.

(iii) Let M, N be submodules of L such that $N \subseteq M$. Then,

$$L/M \cong (L/N)/(M/N).$$

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

Definition 4.11 (R-algebra). Let R be a commutative ring with multiplicative identity 1. Then, an R-algebra A is a ring with multiplicative identity 1 equipped with

$$\text{a ring homomorphism } f : R \to A \quad 1 \mapsto 1 \quad \text{such that} f(R) \in Z(A).$$

Example 4.30. A ring with multiplicative identity 1 is just a $\mathbb{Z}$-algebra.

Definition 4.12. Let F be a field. An F-algebra R is finite-dimensional if R is a finite-dimensional F-vector space.

Example 4.31. Let F be a field. The polynomial ring $F[x]$ is an F-algebra.

Example 4.32. Let G be a finite group. Then, $R[G]$ is an R-algebra. Furthermore, if R is a field, then $R[G]$ is a finite-dimensional F-algebra.

Example 4.33. Let R be a commutative ring with multiplicative identity 1. Let M be an R-module. Then, $\mathrm{End}_R(M)$ is an R-algebra.

Definition 4.13 (simple module). Let R be a ring with multiplicative identity 1. Let M be an R-module. Then,

$$M \text{ is simple} \quad \text{if} \quad M \neq 0 \text{ such that } 0 \text{ and } M \text{ are the only submodules of } M.$$

Example 4.34. All 1-dimensional vector spaces over a field F are simple modules. Recall that an F-module is simply an F-vector space. We claim that

$$\text{a non-zero vector space } V \text{ over } F \text{ is simple} \quad \text{if and only if} \quad \dim(V) = 1.$$

So, if $\dim(V) = 1$, then there are no proper non-zero subspaces. Hence, V is a simple F-module. On the other hand, if $\dim(V) > 1$, then any non-zero vector generates a 1-dimensional subspace, which is a proper non-zero submodule.

Example 4.35. Take $R = \mathbb{Z}$. Then, the simple $\mathbb{Z}$-modules are $\mathbb{Z}/p\mathbb{Z}$ for p prime. Then, we have

$$\mathrm{End}_{\mathbb{Z}}(\mathbb{Z}/p\mathbb{Z}) \cong \mathbb{Z}/p\mathbb{Z} \quad \text{as rings.}$$

Example 4.36. Let F be a field and $R = \mathcal{M}_{n\times n}(F)$ be the matrix ring. Then, F^n is a simple R-module via matrix multiplication.

Lemma 4.3 (Schur's lemma). Let R be a ring with multiplicative identity 1. Let M and N be simple R-modules. Then,

$$\text{any } R\text{-module homomorphism } \varphi : M \to N \quad \text{is} \quad \text{either 0 or an isomorphism.}$$

In particular, $\mathrm{End}_R(M)$ is a division ring.

4.3 Generation of Modules, Direct Sums, and Free Modules

Definition 4.14 (direct sum). Let R be a ring with multiplicative identity 1. Let $M_1, \ldots, M_n$ be R-modules. Define their direct sum

$$M = M_1 \oplus \ldots \oplus M_n = \bigoplus_{i=1}^{n} M_i$$

as follows:

$M = M_1 \times \ldots \times M_n$ as sets and $r(m_1, \ldots, m_n) = (rm_1, \ldots, rm_n)$ for the R-action.

In relation to Definition 4.14, the more precise definition of direct sums and direct products involve categories. Finite direct sums and finite direct products are the same for R-modules.

Example 4.37. $R^n = R \oplus \ldots \oplus R$ is called the free R-module (simply said, a free module is one which has a basis) of rank n. We often write

$$\mathbf{e}_i = (0, \ldots, 1, \ldots, 0) \quad \text{where} \quad 1 \text{ is in the } i^{\text{th}} \text{ component.}$$

Example 4.38. By the Chinese remainder theorem (Theorem 2.5), we have

$$\mathbb{Z}/6\mathbb{Z} \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \quad \text{as } \mathbb{Z}\text{-modules.}$$

Lemma 4.4. Let R be a ring with multiplicative identity 1. Let M be an R-module and $m_1, \ldots, m_n \in M$. Then,

there exists a unique R-module homomorphism $\varphi : R^n \to M$ such that $\mathbf{e}_i \mapsto m_i$.

We call R^n the free R-module of rank n.

Lemma 4.5 (universal property of direct sum). Let R be a ring with multiplicative identity 1. Let $N, M_1, \ldots, M_n$ be R-modules. Then, for any R-module map $\varphi_i : M_i \to N$, there exists

a unique R-module map $\varphi : \bigoplus_{i=1}^{n} M_i \to N$ such that Figure 4.2 commutes.

Proposition 4.8. Let R be a ring with multiplicative identity 1. Let M be an R-module with submodules $N_1, \ldots, N_k$. The following are equivalent:

(i) The natural map induced by the embeddings

$$\bigoplus_{i=1}^{k} N_i \to \sum_{i=1}^{k} N_i \quad \text{is an isomorphism}$$

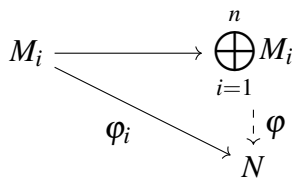


Figure 4.2: Universal property of direct sum

(ii) Any $x \in N_1 + \ldots + N_k$ can be uniquely written as

$$x = a_1 + \ldots + a_k \quad \text{where } a_i \in N_i$$

(iii) For any j,

$$N_j \cap \sum_{i \neq j} N_i = 0$$

Theorem 4.2 (universal property for free modules). Let R be a ring with multiplicative identity 1. Let $A = \{a_1, \ldots, a_n\}$ be a finite set. The free R-module over A is an R-module $F(A)$ together with a map of sets $\iota : A \to F(A)$ such thart for any R-module M and a map of sets $\varphi_A : A \to M$, we have

a unique R-module homomorphism $\varphi : F(A) \to M$ such that Figure 4.3 commutes.

The free module exists and is unique up to isomorphism. Actually,

$$F(A) \cong \bigoplus_{i=1}^{n} R.$$

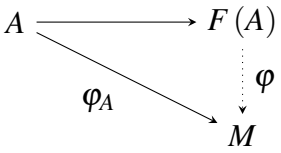


Figure 4.3: Universal property for free modules

Lemma 4.6. Let R be a ring with multiplicative identity 1. Also, let

$$A_1, \ldots, A_n \text{ be } R\text{-modules} \quad \text{with} \quad \text{respective submodules } B_1, \ldots, B_n.$$

Then, we have the following isomorphism:

$$\bigoplus_{i=1}^{n} A_i / \bigoplus_{i=1}^{n} B_i \cong \bigoplus A_i / B_i$$

Proof. We will only prove the case where $n = 2$. We have the following commutative diagram:

$$\begin{array}{ccccc} A_1 & \longrightarrow & A_1 \oplus A_2 & \longrightarrow & A_2 \\ \downarrow & & \vdots\,\varphi & & \downarrow \\ A_1/B_1 & \longrightarrow & A_1/B_1 \oplus A_2/B_2 & \longrightarrow & A_2/B_2 \end{array}$$

Here, φ is surjective with kernel $B_1 \oplus B_2$. The result follows by the first isomorphism theorem (**(i)** of Theorem 4.1). □

Corollary 4.2. Let R be a ring with multiplicative identity 1 and a left ideal I. Then, we have the following isomorphism of R-modules:

$$R^n/IR^n \cong R/I \oplus \ldots R/I \quad \text{which consists of } n \text{ copies.}$$

Theorem 4.3. Let R be a commutative ring with multiplicative identity 1. Then,

$$R^n \cong R^m \quad \text{if and only if} \quad n = m.$$

Theorem 4.4 (universal property for arbitrary direct sum of modules). Let R be a ring with multiplicative identity 1. Let M_c be a collection of R-modules for, i.e. $c \in I$ for some index set I. Their direct sum is an R-module

$$\bigoplus_{c\in I} M_c \quad \text{together with} \quad \iota_c : M_c \to \bigoplus_{c\in I} M_c$$

such that for any R-module N and R-module map $\varphi_c : M_c \to N$, there exists

a unique R-module homomorphism $\varphi : \bigoplus_{c\in I} M_c \to N$ such that Figure 4.4 commutes.

The aforementioned direct sum exists and is unique up to isomorphism.

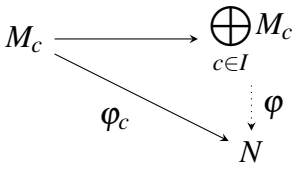


Figure 4.4: Universal property for arbitrary direct sum of modules

Theorem 4.5 (universal property for arbitrary direct product of modules). Let R be a ring with multiplicative identity 1. Let M_c be a collection of R-modules for, i.e. $c \in I$ for some index set I. Their direct product is an R-module

$$\pi_{c\in I} M_c \quad \text{together with} \quad \pi_c : \prod_{c\in I} M_c \to M_c$$

such that for any R-module N and R-module map $\varphi_c : N \to M_c$, there exists

a unique R-module homomorphism $\varphi : N \to \prod_{c\in I} M_c$ such that Figure 4.5 commutes.

The aforementioned direct product exists and is unique up to isomorphism.

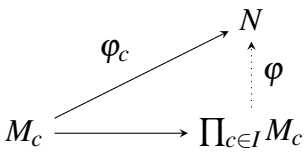


Figure 4.5: Universal property for arbitrary direct product of modules

4.4 Tensor Product of Modules

In this section, we assume that our rings R have a multiplicative identity 1. We often consider both left modules and right modules. We start with the motivation of the tensor product $\otimes$ with two examples (Examples 4.39 and 4.40).

Example 4.39. Let $V = \mathbb{R}^3$ and $W = \mathbb{R}^3$ be two 3-dimensional vector spaces over $\mathbb{R}$. We shall write elements of V and W as row matrices, i.e. (v_1, v_2, v_3) and (w_1, w_2, w_3) respectively. Then, we know that $V \oplus W$ is also an $\mathbb{R}$-vector space of dimension $3+3=6$. Recall that

$$V \oplus W = \{(v,w) : v \in V, w \in W\} = \left\{(v_1, v_2, v_3, w_1, w_2, w_3) \in \mathbb{R}^6\right\}.$$

We now define a new vector space $V \otimes_{\mathbb{R}} W$ as follows:

$$V \otimes_{\mathbb{R}} W = \left\{ \sum_{\text{finite}} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \cdot (w_1, w_2, w_3) = \sum_{\text{finite}} \begin{bmatrix} v_1w_1 & v_1w_2 & v_1w_3 \\ v_2w_1 & v_2w_2 & v_2w_3 \\ v_3w_1 & v_3w_2 & v_3w_3 \end{bmatrix} \right\}.$$

Here, $(v_1, v_2, v_3) \in V$ and $(w_1, w_2, w_3) \in W$.

As vector spaces, we have

$$V \otimes_{\mathbb{R}} W \cong \mathcal{M}_{3\times 3}(\mathbb{R}) \quad \text{which is of dimension } 3 \times 3 = 9.$$

For $\mathbf{v} = (v_1, v_2, v_3)$ and $\mathbf{w} = (w_1, w_2, w_3)$, if we write

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \cdot (w_1, w_2, w_3) \quad \text{as} \quad \mathbf{v} \otimes_{\mathbb{R}} \mathbf{w} \in V \otimes_{\mathbb{R}} W,$$

we have the identities

$$\left(\mathbf{v} + \mathbf{v}'\right) \otimes_{\mathbb{R}} \mathbf{w} = \mathbf{v} \otimes_{\mathbb{R}} \mathbf{w} + \mathbf{v}' \otimes_{\mathbb{R}} \mathbf{w} \quad \text{and} \quad \mathbf{v} \otimes_{\mathbb{R}} \left(\mathbf{w} + \mathbf{w}'\right) = \mathbf{v} \otimes_{\mathbb{R}} \mathbf{w} + \mathbf{v} \otimes_{\mathbb{R}} \mathbf{w}'.$$

Also, for any $k \in \mathbb{R}$, we have

$$(k\mathbf{v}) \otimes_{\mathbb{R}} \mathbf{w} = \mathbf{w} \otimes_{\mathbb{R}} (k\mathbf{w})$$

Proposition 4.9 (tensor product). Let $F(V \times W)$ be the free Abelian group of infinite rank over the set $V \times W$. Let A be the subgroup of $F(V \times W)$ generated by the following:

(i) $(\mathbf{v}+\mathbf{v}', \mathbf{w}) - (\mathbf{v}, \mathbf{w}) - (\mathbf{v}', \mathbf{w})$

(ii) $(\mathbf{v}, \mathbf{w}+\mathbf{w}') - (\mathbf{v}-\mathbf{w}) - (\mathbf{v}-\mathbf{w}')$

(iii) $(k\mathbf{v}, \mathbf{w}) - (\mathbf{v}, k\mathbf{w})$, where $k \in \mathbb{R}$

Then, the natural map

$$\pi : F(V \times W) \to V \otimes_{\mathbb{R}} W \quad \text{where} \quad (\mathbf{v}, \mathbf{w}) \mapsto \mathbf{v} \otimes_{\mathbb{R}} \mathbf{w}$$

has kernel A. Very often, we write $\otimes$ for $\otimes_{\mathbb{R}}$ (or $\otimes_R$ over an arbitrary ring with multiplicative identity).

We shall look at another example (Example 4.40) for a more conceptual perspective. Let $S \subseteq R$ be a subring of R containing the multiplicative identity 1 (note that both rings share the same multiplicative identity). Then, any R-module M is naturally an S-module via restriction. We would like to reserve the process to equip an S-module N with a natural R-module structure.

Example 4.40. Let $R = \mathbb{Q}$ and $S = \mathbb{Z}$. One checks that $\mathbb{Z} \subseteq \mathbb{Q}$ and $\mathbb{Z}$ is a subring of R. Consider the $\mathbb{Z}$-module $\mathbb{Z}/2\mathbb{Z}$. There is no $\mathbb{Q}$-module M such that $M \cong \mathbb{Z}/2\mathbb{Z}$ as $\mathbb{Z}$-modules via restriction.

Proposition 4.10 (universal property of the tensor product of modules). Let $S \subseteq R$ be a subring of R that contains the same multiplicative identity 1. Let N be an S-module. We define an R-module $R \otimes_S N$ via the following universal property: for any R-module M, considered as an S-module via restriction, and any S-module homomorphism $\varphi : N \to M$,

there exists a unique R-module homomorphism $\overline{\varphi} : R \otimes_S N \to M$ such that Figure 4.6 commutes.

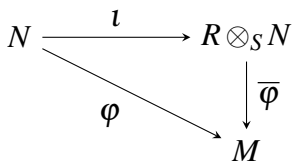


Figure 4.6: Universal property of the tensor product of modules

Theorem 4.6. The tensor product $R \otimes_S N$ exists and is unique up to isomorphism of R-modules.

Example 4.41. One can check that

$$\mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Z}/4\mathbb{Z} \quad \text{as } \mathbb{Q}\text{-modules}$$

from the concrete construction as well as the universal properties.

Example 4.42. For any R-module N, we have $R \otimes_R N \cong N$.

Example 4.43. For any $\mathbb{R}$-vector space $V \cong \mathbb{R}^n$, we have

$$\mathbb{C} \otimes_{\mathbb{R}} V \cong \mathbb{C}^n.$$

Proposition 4.11 (universal property of the tensor product of modules). Let M be a right R-module and N be a left R-module. Their tensor product $M \otimes_R N$ over R is an Abelian group together with a map of Abelian groups

$$\iota : M \times N \to M \otimes_R N \quad \text{where} \quad \iota(mr, n) = \iota(m, rn) = m \otimes_R rn = mr \otimes_R n$$

such that for any map of Abelian groups

$$\varphi : M \times N \to L \quad \text{where} \quad \varphi(mr, n) = \varphi(m, rn),$$

there exists a unique

$$\text{map of Abelian groups } \overline{\varphi} : M \otimes_R N \to L \quad \text{such that} \quad \text{Figure 4.7 commutes.}$$

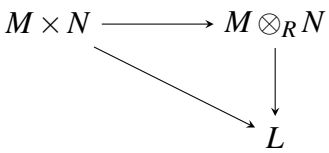


Figure 4.7: Universal property of the tensor product of modules

Example 4.44. From Example 4.39, if $V = W = \mathbb{R}^3$, then we know that

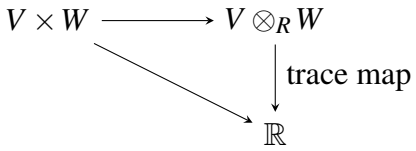
$$V \otimes_{\mathbb{R}} W \cong \mathcal{M}_{3\times 3}(\mathbb{R}).$$

It is equipped with the following map:

$$((v_1, v_2, v_3), (w_1, w_2, w_3)) \mapsto \begin{bmatrix} v_1w_1 & v_1w_2 & v_1w_3 \\ v_2w_1 & v_2w_2 & v_2w_3 \\ v_3w_1 & v_3w_2 & v_3w_3 \end{bmatrix}$$

As such, this induces the trace map on $\mathcal{M}_{3\times 3}(\mathbb{R})$, i.e. the following diagram commutes:

A very important case of the previous propositions is the case when R is a commutative ring. Here, the left modules can also be regarded as the right modules.



Definition 4.15 (bilinearity). Let R be a commutative ring with multiplicative identity 1. Let M, N, L be R-modules. A map of Abelian groups $\varphi : M \times N \to L$ is r-bilinear if

$$\varphi(r_1 m_1 + r_2 m_2, n) = r_1 \varphi(m_1, n) + r_2 \varphi(m_2, n)$$
$$\varphi(m, r_1 n_1 + r_2 n_2) = r_1 \varphi(m, n_1) + r_2 \varphi(m, n_2)$$

where $m_1, m_2 \in M, n_1, n_2 \in N$ and $r_1, r_2 \in R$.

Lemma 4.7. There is a bijection of the following sets:

$$\mathrm{Hom}_R(N, \mathrm{Hom}_R(M, L)) \quad \text{and} \quad \text{the set of } R\text{-bilinear maps from } M \times N \text{ to } L$$

Definition 4.16 (bimodule). Let R and S be rings. An R-S-bimodule is an Abelian group M that is simultaneously a left R-module and a right S-module such that

$$(rm)s = r(ms) = rms \quad \text{where } r \in R, s \in S, m \in M.$$

Example 4.45. Let M be a right R-module and N be a left R-module. Then, we have

$$m \otimes 0 = 0 \times n = 0.$$

Example 4.46. Consider the $\mathbb{Z}$-modules $\mathbb{Z}/2\mathbb{Z}$ and $\mathbb{Z}/3\mathbb{Z}$. We claim that $\mathbb{Z}/2\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/3\mathbb{Z}$ is trivial. Note that 4 acts as the identity in $\mathbb{Z}/3\mathbb{Z}$. Then,

$$m \otimes n = m \otimes (4n) = (4m) \otimes n = 0 \otimes n = 0.$$

Example 4.47. We consider the $\mathbb{Z}$-modules $\mathbb{Z}/2\mathbb{Z}$ and $\mathbb{Z}/2\mathbb{Z}$. We claim that

$$\mathbb{Z}/2\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/2\mathbb{Z} \cong \mathbb{Z}/2\mathbb{Z}.$$

Consider the map

$$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \quad \text{where} \quad (a, b) \mapsto ab.$$

Using Definition 4.15, one checks that the map is bilinear. On the other hand, we see that $1 \otimes 0 = 0 \times 0 = 0 \times 1$, so the induced map is bijective and hence, an isomorphism.

Example 4.48. Let $V = \mathbb{C}$. As a $\mathbb{C}$-vector space, this is is 1-dimensional. However, V is 2-dimensional as an $\mathbb{R}$-vector space. One checks that

$$V \otimes_{\mathbb{C}} V \cong \mathbb{C} \quad \text{and} \quad V \otimes_{\mathbb{R}} V \cong \mathbb{R}^4.$$

Proposition 4.12. Let M_1, M_2, M_3 be right R-modules and N_1, N_2, N_3 be left R-modules. Let

$$f : M_1 \to M_2 \quad \text{and} \quad g : N_1 \to N_2, \quad h : M_2 \to M_3, \quad k : N_2 \to N_3$$

be (left or right, respectively) R-module maps. Then, the following hold:

(i) There is a unique map of abelian groups

$$f \times g : M_1 \otimes_R N_1 \to M_2 \otimes_R N_2$$

such that

$$(f \otimes_R g)(m \otimes_R n) = f(m) \otimes_R g(n) \quad \text{with } m \in M_1 \text{ and } n \in N_1$$

(ii) We have

$$(h \otimes_R k) \circ (f \otimes_R g) = (h \circ f) \otimes (k \circ g) : M_1 \otimes_R N_1 \to M_3 \otimes_R N_3$$

(iii) If M_1 and M_2 are both S-R-bimodules with $f : M_1 \to M_2$ being a morphism of bimodules, then $f \otimes_R g$ is also a map of S-modules

Theorem 4.7. Let M be a right R-module, N be an R-S-bimodule, and L be a left S-module. Then, we have the following isomorphism of Abelian groups:

$$M \otimes_R (N \otimes_S L) \cong (M \otimes_R N) \otimes_S L$$

Chapter 5

Modules over Principal Ideal Domains

5.1 The Basic Theory

This section develops the structural backbone of finitely generated modules over a PID. In MA2001 Linear Algebra I, much of the theory rests on two foundational facts: every subspace of a finite-dimensional vector space admits a basis, and linear maps can be simplified (diagonalised, put into row-reduced form, etc.) by performing invertible changes of basis. Over a *PID* R, finitely generated R-modules behave in many ways like finite-dimensional vector spaces, but there is a crucial new phenomenon: torsion.

Once torsion is introduced, freeness over a PID admits an intrinsic characterisation: a finitely generated R-module is free precisely when it has no non-zero torsion. This should be compared with vector spaces over a field, where torsion cannot occur at all (since $rv = 0$ with $r \neq 0$ forces $v = 0$). From this point of view, vector space is the torsion-free part of the theory, and the genuinely new content over a PID is that torsion modules exist and are controlled by ideals of R.

Theorem 5.1 (submodule of free module is also free). Let R be a PID and M be a free R-module of rank n and N be an R-submodule of M. Then,

$$N \text{ is also free} \quad \text{with} \quad \operatorname{rank}(N) \leq \operatorname{rank}(M)$$

Compare and contrast Theorem 5.1 with a known result from MA2001 Linear Algebra — say we have a finite-dimensional vector space V of rank n (which means any basis of V has n linearly independent vectors). If U is a subspace of V, then U also has a basis[1] such that $\operatorname{rank}(U) \leq \operatorname{rank}(V) = n$.

Recall the definition of a torsion element in a module (Definition 4.4). We state it again

[1] The proof of this is merely a consequence of Zorn's lemma.

here. Let M be an R-module over an integral domain R. An element $m \in M$ is called torsion if

$$rm = 0 \quad \text{for some } r \in R \text{ where } r \neq 0.$$

The set of torsion elements is denoted by $\mathrm{Tor}(M)$.

(i) An R-module M is torsion-free if $\mathrm{Tor}(M) = 0$

(ii) An R-module M is a torsion module if $\mathrm{Tor}(M) = M$

Theorem 5.2. Let R be a PID and M be a finitely generated R module. Then,

$$M \text{ is free} \quad \text{if and only if} \quad M \text{ is torsion-free.}$$

Let R be a PID. Recall from Theorem 2.4 that R is also a UFD. We wish to study the diagonalisation of matrices with entries in R, or equivalently,

$$R\text{-module homomorphisms} \quad \varphi = (a_{ij}) \in \mathcal{M}_{n\times m}(R) : R^{\oplus m} \to R^{\oplus n}.$$

Example 5.1. Let $R = \mathbb{Q}$. Then, recall the usual elementary row operations (which are invertible) from MA2001 Linear Algebra 1. Say we have

$$\begin{bmatrix} 5 & 7 & 9 \\ 2 & 3 & 5 \end{bmatrix} \longrightarrow \cdots \longrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/5 & 0 \end{bmatrix}.$$

Here, one performed some elementary row operations. We need more complicated operations for a general PID R.

We have the following interesting fact about PIDs (Lemma 5.1).

Lemma 5.1 (recursively appending gcd). For any PID R, let $a_1, \ldots, a_n \in R$ such that $(\alpha) = (a_1, \ldots, a_n)$. Then, up to multiplication by units, we have

$$\alpha = \gcd(a_1, \ldots, a_n) = \gcd(a_1, \gcd(a_2, \ldots, a_n)). \tag{5.1}$$

Example 5.2. Take $R = \mathbb{Z}$ and $(a_1, a_2, a_3, a_4) = (18, 30, 42, 105)$. Then, $(18, 30, 42, 105) = (3)$ so we may take $\alpha = 3$. Indeed, $\gcd(18, 30, 42, 105) = 3$ and recursively,

$$\gcd(30, 42, 105) = \gcd(30, \gcd(42, 105)) = \gcd(30, 21) = 3,$$

hence

$$\gcd(18, 30, 42, 105) = \gcd(18, \gcd(30, 42, 105)) = \gcd(18, 3) = 3.$$

All equalities are understood up to multiplication by units (± 1 in $\mathbb{Z}$).

We now prove Lemma 5.1.

Proof. Since R is a PID, the ideal generated by $a_1, \ldots, a_n$ is principal, so $(a_1, \ldots, a_n) = (\alpha)$. By definition, a (chosen) gcd of $a_1, \ldots, a_n$ is any element $\delta \in R$ such that

$$(\delta) = (a_1, \ldots, a_n),$$

and such a δ is unique up to multiplication by a unit. Hence α is a gcd of $a_1,\dots,a_n$, i.e. $\alpha = \gcd(a_1,\dots,a_n)$ up to a unit.

For the recursive identity (5.1), set β such that $(a_2,\dots,a_n) = (\beta)$ so that $\beta = \gcd(a_2,\dots,a_n)$ up to a unit. Then,

$$(a_1,\beta) = (a_1) + (\beta) = (a_1) + (a_2,\dots,a_n) = (a_1,\dots,a_n) = (\alpha).$$

Therefore, α is also a gcd of a_1 and β, i.e.

$$\alpha = \gcd(a_1,\beta) = \gcd\left(a_1, \gcd(a_2,\dots,a_n)\right)$$

up to multiplication by a unit, as required. □

Lemma 5.2. Let R be a PID. Suppose $a,b \in R$ such that $\gcd(a,b) = 1$. By Bézout's identity (works for arbitrary PID), there exist $c,d \in R$ such that $ac + bd = 1$. Then,

$$\begin{bmatrix} a & b \\ -d & c \end{bmatrix}, \begin{bmatrix} a & -d \\ b & c \end{bmatrix} \in \mathcal{M}_{2\times 2}(R) \quad \text{are invertible} \quad \text{with inverses} \quad \begin{bmatrix} c & -b \\ d & a \end{bmatrix}, \begin{bmatrix} c & d \\ -b & a \end{bmatrix} \text{ respectively.}$$

Corollary 5.1. Let $a,b \in R$ and define $\alpha = \gcd(a,b)$. Then, there exist $\mathbf{S},\mathbf{T} \in \mathrm{GL}_2(R)$ such that

$$\begin{bmatrix} a & b \\ \star & \star \end{bmatrix} \mathbf{S} = \begin{bmatrix} \alpha & 0 \\ \star & \star \end{bmatrix} \quad \text{and} \quad \mathbf{T} \begin{bmatrix} a & \star \\ b & \star \end{bmatrix} = \begin{bmatrix} \alpha & \star \\ 0 & \star \end{bmatrix}.$$

Here, $\star$ denotes an arbitrary element in R.

Lemma 5.3. Let R be a PID. Let $a,b \in R$. Then, there exist $\mathbf{S},\mathbf{T} \in \mathrm{GL}_2(R)$ such that

$$\mathbf{T} \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \mathbf{S} = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix} \quad \text{where} \quad \alpha \mid \beta.$$

Actually,

$$\alpha = \gcd(a,b) \quad \text{and} \quad \alpha\beta = ab = \det \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$$

uniquely up to multiplication by units.

Theorem 5.3 (Smith normal form). There exist $\mathbf{T} \in \mathcal{M}_{n\times n}(R)$ and $\mathbf{S} \in \mathcal{M}_{m\times m}(R)$ such that

$$\mathbf{TAS} = \mathrm{diag}(\alpha_1,\dots,\alpha_r,0,\dots,0) \quad \text{such that} \quad \alpha_i \mid \alpha_{i+1}.$$

The α_i's are the invariant factors of $\mathbf{A}$ and the RHS is the Smith normal form of $\mathbf{A}$.

Theorem 5.4 (structure theorem for finitely generated modules over a PID). Let M be a finitely generated module over a PID R. We know that

$$M \cong \operatorname{coker}(\varphi : R^m \to R^n) \quad \text{since} \quad M \text{ is finitely generated and finitely presented.}$$

Let $S : R^n \to R^n$ and $T : R^m \to R^m$ be isomorphisms of R-modules. Then, we have the following isomorphism:

$$\operatorname{coker} \varphi \cong M \cong \operatorname{coker}(S \circ \varphi \circ T).$$

Theorem 5.5 (invariant factor). Let M be a finitely generated module over a PID R. Then,

$$M \cong R^k \oplus R/(\alpha_1) \oplus \ldots \oplus R(\alpha_r) \quad \text{where } \alpha_i \mid \alpha_{i+1}.$$

The elements α_i are the invariant factors of M.

Theorem 5.6 (elementary factor). Let M be a finitely generated module over a PID R. Then,

$$M \cong R^k \oplus R/\left(p_1^{a_1}\right) \oplus \ldots \oplus R/\left(p_s^{a_s}\right)$$

for not necessarily distinct prime elements p_i in R. The elements $p_i^{a_i}$ are called the elementary factors of M.

Lemma 5.4. Let $\varphi : M \to N$ be an R-module homomorphism. Then,

$$\varphi(\operatorname{Tor} M) \subseteq \operatorname{Tor} N.$$

Lemma 5.5. Let R be a PID. Let M be a finitely generated R-module. Then,

$$M \text{ is free} \quad \text{if and only if} \quad M \text{ is torsion free.}$$

Example 5.3. Let M be a finitely generated Abelian group. Then,

$$M \cong \mathbb{Z}^k \oplus \mathbb{Z}/a_1\mathbb{Z} \oplus \ldots \oplus \mathbb{Z}/a_r\mathbb{Z} \quad \text{where } a_i \mid a_{i+1} \text{ in } \mathbb{Z}.$$

Example 5.4. Let F be a field. Let $G \subseteq F^\star$ be a finite subgroup. Then, G is cyclic.

Example 5.5. The Abelian group of rational numbers $\mathbb{Q}$ is not finitely generated as a $\mathbb{Z}$-module. It is torsion free but not free.

Example 5.6. Let M be an Abelian group generated by x and y subjected to the relation $2x+5y=0$ and $3x+7y=0$. Then, we wish to determine the structure of M. Consider the map

$$R^2 \to M \quad \text{where} \quad (a,b) \mapsto ax+by.$$

The kernel is generated by $(2,5)$ and $(3,7)$. We consider another map

$$\varphi : R^2 \to R^2 \quad \text{where} \quad (a,b) \mapsto a(2,5)+b(3,7) = (2a+3b, 5a+7b).$$

Then, we have $M \cong \operatorname{coker} \varphi$. The smith normal form of φ is

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \text{so} \quad M \cong 0.$$

Lemma 5.6. Let p and q be distinct primes in R. Let $F = R/(p)$ be a field. Then, the following hold:

(i) Let $M = R/(q^a)$ for $a \geq 1$. As F-modules, we have

$$p^M/p^{t+1}M \cong 0.$$

(ii) Let $M = R/(p^a)$ for $a \geq 1$. Then, as F-modules,

$$p^t M/p^{t+1}M \cong \begin{cases} F & \text{if } t < a; \\ 0 & \text{if } t \geq a. \end{cases}$$

(iii) Let $M = R/(q^a) \oplus R/(p^b)$ for $a, b \geq 1$. Then, as F-modules,

$$p^t M/p^{t+1}M \cong \begin{cases} F & \text{if } t < a; \\ 0 & \text{if } t \geq a. \end{cases}$$

Theorem 5.7. Let

$$M \cong R/(p_1^{a_1}) \oplus \ldots \oplus R/(p_s^{a_s}) \cong R/(q_1^{b_1}) \oplus \ldots \oplus R/(q_r^{b_r})$$

as modules over a PID R for non-necessarily distinct primes p_i and q_i. Then, the elementary factors are unique up to permutation and multiplication by units.

Corollary 5.2. Let $\varphi \in \mathcal{M}_{m\times n}(R)$ for a PID R. Then, the invariant factors of the Smith normal form of φ is unique up to permutation and multiplication by units. In other words, it is independent of the invertible matrices $\mathbf{S}$ and $\mathbf{T}$ we used.

5.2 The Rational Canonical Form

We first recall some concepts from Linear Algebra. Let F be a field and let V be a finite-dimensional $F[x]$-module. Suppose x acts on V via a linear transformation T. Assume that $\dim_F V = n$ and fix an F-basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ of V. We then have $T = (a_{ij}) \in \mathcal{M}_{n\times n}(F)$. We write $\mathbf{I} \in \mathcal{M}_{n\times n}(F)$ for the identity matrix.

Note that the set $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ forms a basis for the F-module V. However, it is only a set of generators for the $F[x]$-module V. Hence, we have the following isomorphism of

$F[x]$-modules:

$$V \cong F[x]/(f_1(x)) \oplus \ldots \oplus F[x]/(f_r(x)) \quad \text{where } f_i \mid f_{i+1}$$

Let

$$V \cong F[x]/(f(x)) \quad \text{where} \quad f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_0.$$

We consider the basis $\{1, x, x^2, \ldots, x^{n-1}\}$ under the isomorphism. With respect to this basis, x acts via the matrix

$$\mathbf{T} = \begin{bmatrix} 0 & 0 & \ldots & \ldots & -a_0 \\ 1 & 0 & \ldots & \ldots & -a_1 \\ 0 & 1 & \ldots & \ldots & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -a_{n-1} \end{bmatrix}.$$

By direct computation, we have $\det(x\mathbf{I} - \mathbf{T}) = f(x)$.

Definition 5.1 (rational canonical form). The rational canonical form (RCF) of a matrix $\mathbf{T}$ is the matrix

$$\begin{bmatrix} T_1 & & \\ & T_2 & \\ & & \ddots \end{bmatrix}$$

where each T_i is of the form as mentioned earlier with respect to the $F[x]$-submodule $F[x]/(f_i(x))$.

Proposition 5.1. Let $\mathbf{A}, \mathbf{B} \in \mathcal{M}_{n\times n}(F)$.

(i) The RCF of $\mathbf{A}$ is unique

(ii) $\mathbf{A}$ is similar to its RCF

(iii) $\mathbf{A}$ is similar to $\mathbf{B}$ if and only if they have the same RCF

Example 5.7. We shall compute the RCF of the matirx

$$\mathbf{A} = \begin{bmatrix} 2 & -2 & 14 \\ 0 & 3 & -7 \\ 0 & 0 & 2 \end{bmatrix} \quad \text{in } \mathcal{M}_{3\times 3}(\mathbb{Q}).$$

Solution. We have

$$x\mathbf{I} - \mathbf{A} = \begin{bmatrix} x-2 & 2 & -14 \\ 0 & x-3 & 7 \\ 0 & 0 & x-2 \end{bmatrix}.$$

Let T_1 denote the gcd of all 1×1 minors, T_1T_2 denote the gcd of all 2×2 minors and $T_1T_2T_3$ denote $\det(x\mathbf{I} - \mathbf{A})$.

We first compute T_1. Note that all entries of $x\mathbf{I}-\mathbf{A}$ are

$$x-2, 2, -14, x-3, 7, x-2 \quad \text{whose gcd is } 1.$$

So, $T_1 = 1$. We then look at all 2×2 sub-determinants. One is able to determine that $T_1 T_2 = x-2$, so $T_2 = x-2$. Lastly, $T_3 = (x-2)(x-3)$.

The Smith normal form of $x\mathbf{I}-\mathbf{A}$ is the diagonal matrix $\operatorname{diag}(1, x-2, (x-2)(x-3))$. The RCF of $\mathbf{A}$ is a block-diagonal matrix whose blocks are the companion matrices of the invariant factors. As our invariant factors are $1, x-2, (x-2)(x-3)$, the companion blocks we need in the RCF are for the polynomials $x-2$ and $(x-2)(x-3)$.

Hence, the RCF is a direct sum of

$$\text{a } 1\times 1 \text{ companion block for } x-2 \quad \text{and} \quad \text{a } 2\times 2 \text{ companion block for } x^2-5x+6.$$

Note that

$$\text{for a quadratic polynomial } x^2+ax+b \quad \text{its companion matrix is} \quad \begin{bmatrix} 0 & -b \\ 1 & -a \end{bmatrix}.$$

As such, the RCF of $\mathbf{A}$ is

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & -6 \\ 0 & 1 & 5 \end{bmatrix}.$$

□

5.3 The Jordan Canonical Form

Now, we talk about the characteristic polynomial and minimal polynomial of a square matrix.

Definition 5.2 (minimal polynomial). Let $T \in \mathcal{M}_{n\times n}(F)$. We consider the $F[x]$-module $V = F^n$, where x acts as T. Let $\operatorname{Ann}_{F[x]}(V) = (p(x))$. Then, $p(x)$ is the minimal polynomial (often assumed to be monic) of T.

In fact, a number of properties of the characteristic polynomial and minimal polynomial have already been covered in MA2101.

Definition 5.3 (algberaically closed field). Let F be a field. F is algebraically closed if every non-constant polynomial in $F[x]$ has a root in F.

Example 5.8. $\mathbb{C}$ is algebraically closed, while $\mathbb{Q}$ and $\mathbb{R}$ are not.

Lemma 5.7. Let F be an algebraically closed field. Then, the following hold:

(i) F is infinite

(ii) If $f(x) \in F[x]$ is irreducible, then

$$f(x) = k(x-a) \quad \text{for some } a \in F \text{ and } k \in F^\times.$$

Lemma 5.8. Let F be an algebraically closed field. Let $T \in \mathcal{M}_{n\times n}(F)$. Then, T has an eigenvalue.

Example 5.9. The matrix

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \in \mathcal{M}_{2\times 2}(\mathbb{R}) \quad \text{has no eigenvalue in } \mathbb{R}.$$

Example 5.10. Let $V \cong F[x]/(x-a)^n$ for some $a \in F$. Then, we consider the following F-basis of V via the isomorphism:

$$1, x-\lambda, (x-\lambda)^2, \ldots, (x-\lambda)^{n-1}$$

Then, x acts on the matrix

$$\begin{bmatrix} \lambda & 1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & \lambda \end{bmatrix}$$

Let $T \in \mathcal{M}_{n\times n}(F)$ for some algebraically closed field F. We consider $V = F^n$ as an $F[x]$-module where x acts on T. Then, we have

$$V \cong F[x]/((x-\lambda_1)^{a_1}) \oplus \ldots \oplus F[x]/((x-\lambda_r)^{a_r})$$

Definition 5.4 (Jordan canonical form). The Jordan canonical form (JCF) of a matrix is

$$\mathbf{J} = \begin{bmatrix} T_1 & & \\ & T_2 & \\ & & \ddots \end{bmatrix}$$

where each T_i is of the form in the example with respect to the $F[x]$-submodule $F[x]/((-\lambda_i)^{a_i})$.

Proposition 5.2. Let $\mathbf{A}, \mathbf{B} \in \mathcal{M}_{n\times n}(F)$ for an algebraically closed field F.

(i) The JCF of $\mathbf{A}$ is unique up to permutation

(ii) $\mathbf{A}$ is similar to its JCF

(iii) $\mathbf{A}$ is similar to $\mathbf{B}$ if and only if they have the same JCF up to permutation

Example 5.11. One can compute the JCF of the matrix

$$\mathbf{A} = \begin{bmatrix} 2 & -2 & 14 \\ 0 & 3 & -7 \\ 0 & 0 & 2 \end{bmatrix} \quad \text{in } \mathcal{M}_{3\times 3}(\mathbb{C}).$$

The JCF is

$$\mathbf{J} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Since each Jordan block is of size 1, $\mathbf{A}$ is said to be diagonalisable over $\mathbb{C}$.

Example 5.12. The JCF of

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \in \mathcal{M}_{2\times 2}(\mathbb{C}) \quad \text{is} \quad \mathbf{J} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$$

which is diagonalisable over $\mathbb{C}$ as each Jordan block is of size 1.

Example 5.13. The JCF of

$$\mathbf{A} = \begin{bmatrix} 0 & -1 \\ 1 & 2 \end{bmatrix} \in \mathcal{M}_{2\times 2}(\mathbb{C}) \quad \text{is} \quad \mathbf{J} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

which contains one Jordan block of size 2. Hence, $\mathbf{A}$ is not diagonalisable over $\mathbb{C}$.

Corollary 5.3. Let F be an algebraically closed field. Let $T \in \mathcal{M}_{n\times n}(F)$ with a Jordan canonical form $\mathbf{J}$.

(i) T is diagonalisable over F if and only if $\mathbf{J}$ is a diagonal matrix

(ii) T is diagonalizable over F if and only if its minimal polynomial does not have multiple roots

Proposition 5.3. Let $\mathbf{A}, \mathbf{B} \in \mathcal{M}_{n\times n}(F)$. Let V and W be $F[x]$-modules, where x acts on $\mathbf{A}$ and $\mathbf{B}$, respectively. Then, the following are equivalent:

(i) $\mathbf{A}$ and $\mathbf{B}$ are similar

(ii) $\mathbf{A}$ and $\mathbf{B}$ have the same RCF

(iii) $V \cong W$

(iv) V and W have the same invariant factors

(v) V and W have the same elementary factors

(vi) If F is algebraically closed, then $\mathbf{A}$ and $\mathbf{B}$ have the same JCF

Definition 5.5 (nilpotent matrix). Let F be a field. A matrix $\mathbf{A} \in \mathcal{M}_{n\times n}(F)$ is nilpotent if

$$\text{there exists } k \in \mathbb{N} \quad \text{such that} \quad \mathbf{A}^k = \mathbf{0}.$$

We wish to classify the *nilpotent orbits*, i.e. orbits consisting of nilpotent matrices.

Lemma 5.9. Let $\mathbf{A} \in \mathcal{M}_{n\times n}(F)$ be nilpotent. The minimal polynomial $m_{\mathbf{A}}(x) = x^k$ for some $k \in \mathbb{N}$ and the characteristic polynomial is $c_{\mathbf{A}}(x) = x^n$.

Corollary 5.4. The set of nilpotent orbits on $M_{n\times n}(F)$ is in bijection with the set of partitions of n.

We then consider the conjugacy classes in $\mathrm{GL}_2(F_2)$, where F_2 is the finite field with 2 elements. We wish to determine the number of conjugacy classes and find a representative of each class.

Proposition 5.4. There are 3 conjugacy classes in $\mathrm{GL}_2(F_2)$. Their elementary factors are

$$\left\{(x-1)^2\right\}, \{x-1, x-1\}, \left\{x^2+x+1\right\}.$$

We can find the representative using either the JCF or the RCF.

Next, we consider orbits of $\mathrm{GL}_3(\mathbb{Q})$ on the set

$$S = \left\{\mathbf{A} \in \mathrm{GL}_3(\mathbb{Q}) : \mathbf{A}^6 = \mathbf{I}\right\}.$$

The elements in this set are of order $2, 3, 6$. We consider the factorisation

$$\begin{aligned} x^6 - 1 &= \left(x^3+1\right)\left(x^3-1\right) \\ &= (x+1)\left(x^2-x+1\right)(x-1)\left(x^2+x+1\right) \\ &= (x-1)(x+1)\left(x^2-x+1\right)\left(x^2+x+1\right) \end{aligned}$$

Let $\mathbf{A} \in S$. Then, the minimal polynomial $m_{\mathbf{A}}(x)$ of $\mathbf{A}$ must divide $x^6 - 1$. We also know that $\deg m_{\mathbf{A}}(x) \leq 3$. As such, we have the following possibilities (at first glance) for $m_{\mathbf{A}}(x)$:

(1) $x-1$
(2) $x+1$
(3) x^2-x+1
(4) x^2+x+1
(5) $(x-1)(x+1)$
(6) $(x-1)\left(x^2-x+1\right)$
(7) $(x-1)\left(x^2+x+1\right)$
(8) $(x+1)\left(X^2-x+1\right)$
(9) $(x+1)\left(x^2+x+1\right)$

By Proposition 5.4, we shall classify the $F[x]$-modules of F-dimension 3 with the largest invariant factors being $m_{\mathbf{A}}(x)$. They are classified by the invariant factors. Consequently, the invariant factors corresponding to each case are as follows:

(1) $x-1, x-1, x-1$
(2) $x+1, x+1, x+1$
(3) Impossible
(4) Impossible
(5) $x-1, (x-1)(x+1)$ or $1, (x-1)(x+1)$
(6) $(x-1)\left(x^2-x+1\right)$
(7) $(x-1)\left(x^2+x+1\right)$
(8) $(x+1)\left(x^2-x+1\right)$
(9) $(x+1)\left(x^2+x+1\right)$

We know that the orbits of $\mathbf{A}$ are uniquely determined by the invariant factors. So, there are 7 orbits.

Bibliography

[1] Dummit, D. S., and Foote, R. M. *Abstract Algebra* (3rd ed.). Wiley, Hoboken, New Jersey, 2003. ISBN: 9780471433347.

[2] Gallian, J. A. *Contemporary Abstract Algebra* (7th ed.). Brooks/Cole, Cengage Learning, Belmont, California, 2009. ISBN: 9780547165097.