# MATH2027 Rings and Polynomials

Cheatsheet

2022/23

This document collects together the important definitions and results presented throughout the lecture notes. The numbering used throughout will be consistent with that in the lecture notes.

## Contents


**1 Introduction** **2**

**2 Ideals and Factor Rings** **6**

**3 Homomorphisms** **10**

**4 Fields and Integral Domains** **13**

**5 Classes of Integral Domains** **16**

**6 Elements in Integral Domains** **19**

**7 Prime and Irreducible Elements** **22**

**8 Irreducible Polynomials** **27**

# 1 Introduction

**Definition 1.1** Let $A$ and $B$ be sets. The Cartesian product of $A$ and $B$ is the set of all pairs of elements the first of which is from $A$ and the second of which is from $B$, that is

$$A \times B := \{(a, b) : a \in A \text{ and } b \in B\}.$$

**Definition 1.2** Let $G$ be a set. A binary operation on $G$ is a function $G \times G \to G$.

**Note:** In other words, it is a function taking two elements of a set and spitting out another element which also lives inside the same set. We do **not** assume it is associative (see below).

**Definition 1.4** A group $(G, *)$ is a pair consisting of a non-empty set $G$ and a binary operation $* : G \times G \to G$ on $G$ which satisfying the following axioms:
(G1) For all $g, h \in G$, we have $g * h \in G$. (**Closed**)
(G2) For all $g, h, k \in G$, we have $(g * h) * k = g * (h * k)$. (**Associativity**)
(G3) There exists $e \in G$ such that $g * e = g = e * g$ for all $g \in G$. (**Identity**)
(G4) For all $g \in G$, we have $h \in G$ such that $g * h = e = h * g$. (**Existence of Inverses**)

**Remark** By stipulating the operation $*$ is binary, we automatically get that $g * h \in G$ for all $g, h \in G$. The only reason we write out the closure rule is to make us remember to check that the operation is indeed a binary operation (i.e. a valid function whose output lives in the set $G$).

**Note:** Often, we refer to a group by the underlying set $G$ and don't explicitly mention $*$.

**Lemma** *Let $G$ be a group. Then, the identity is unique.*

*Proof*: Suppose $e, f \in G$ are two identities. Then, we have the following:

(i) $e * f = e$, because $f$ is an identity.

(ii) $e * f = f$, because $e$ is an identity.

But clearly, $e * f = e * f$, so it follows that $e = f$. □

**Lemma** *Let $G$ be a group. Then, the inverse of $g \in G$ is unique.*

*Proof*: Suppose $h, k \in G$ are two inverses of the element $g$. Then, we have the following:

(i) $g * h = e = h * g$, by definition.

(ii) $g * k = e = k * g$, by definition.

Therefore, we see that $h = h * e = h * (g * k) = (h * g) * k = e * k = k$ by associativity. □

**Note:** Per this lemma, we henceforth denote the inverse of $g \in G$ by the symbol $g^{-1}$.

**Definition** A group $G$ is Abelian if the operation $*$ is commutative, that is for all $g, h \in G$,

$$g * h = h * g.$$

**Remark 1.6** Addition is clearly Abelian. Therefore, we use this notation for any Abelian group:

(i) The operation $*$ is denoted $+$.

(ii) The identity $e$ is denoted $0$.

(iii) The inverse $g^{-1}$ is denoted $-g$.

**Definition 1.7** A ring $(R, +, \times)$ is a triple consisting of a non-empty set $R$ and two binary operations $+ : R \times R \to R$ and $\times : R \times R \to R$ satisfying the following axioms:

(R1) The pair $(R, +)$ is an Abelian group.

(R2) For all $r, s \in R$, we have $r \times s \in R$. (**Closure of** $\times$)

(R3) For all $r, s, t \in R$, we have $(r \times s) \times t = r \times (s \times t)$. (**Associativity of** $\times$)

(R4) For all $r, s, t \in R$, we have each of these: (**Distributivity of** $\times$ **over** $+$)

- (i) $r \times (s + t) = (r \times s) + (r \times t)$.
- (ii) $(r + s) \times t = (r \times t) + (s \times t)$.

**Remark 1.8** For the sake of nicer notation, we often write $rs := r \times s$ and $r - s := r + (-s)$.

**Definition** Let $R$ be a ring. A multiplicative identity is some $1_R \in R$ where for all $r \in R$,

$$1_R \times r = r = r \times 1_R.$$

**Note:** We do **not** assume that every ring has a multiplicative identity (note there is no mention of this in Definition 1.7). However, those that do we herein call rings with one.

**Lemma** *Let $R$ be a ring. If it exists, the multiplicative identity $1_R$ is unique.*

*Sketch of Proof*: This is the same proof as the uniqueness of the identity of a group. □

**Proposition** *Let $R$ be any ring. Then, the following are also rings:*

(i) *The matrix ring $M_n(R)$ of $n \times n$ matrices with entries in $R$.*

(ii) *The polynomial ring $R[x]$ of polynomials in one variable $x$ with coefficients in $R$.*

(iii) *The polynomial ring $R[x_1, ..., x_k]$ of polynomials in $k$ variables with coefficients in $R$.*

(iv) *The Gaussian integers $\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\} \subseteq \mathbb{C}$.*

**Lemma 1.11** *Let $R$ be a ring and $r, s, t \in R$ be any elements. Then, the following are true:*

(i) *The additive identity $0_R$ is unique.*
(ii) *The additive inverse $-r$ of $r$ is unique.*
(iii) *If $r + t = s + t$, then $r = s$.*
(iv) *We have $-(r + s) = (-r) + (-s)$.*
(v) *We have $-(-r) = r$.*
(vi) *We have $r0_R = 0_R = 0_R r$.*
(vii) *We have $(-r)s = -(rs) = r(-s)$.*

*Proof*: By Axiom (R1), we know $(R, +)$ is a group. So we've already proved (i) and (ii) earlier.

(iii) Suppose $r + t = s + t$. Because $t \in R$ and $R$ is a group, it is closed under forming inverses, that is there exists an element $-t \in R$ such that $t + (-t) = 0_R$. Thus, adding this element to both sides of the equation tells us that $r + t + (-t) = s + t + (-t)$, but this is nothing other than $r + 0_R = s + 0_R$ which is the same as $r = s$.

(iv) Well, we can see that

$$\begin{aligned}
\big((-r) + (-s)\big) + (r + s) &= \big((-s) + (-r)\big) + (r + s), && \text{as } + \text{ is commutative,} \\
&= (-s) + \big((-r) + (r + s)\big), && \text{as } + \text{ is associative,} \\
&= (-s) + \big(((-r) + r) + s\big), && \text{as } + \text{ is associative,} \\
&= (-s) + (0_R + s), && \text{as } -r \text{ is the additive inverse of } r, \\
&= (-s) + s, && \text{as } 0_R \text{ is the additive identity,} \\
&= 0_R, && \text{as } -s \text{ is the additive inverse of } s.
\end{aligned}$$

Doing a similar argument, we conclude that $(r + s) + \big((-r) + (-s)\big) = 0_R$. Therefore, we see that the inverse of $(r + s)$ is $(-r) + (-s)$, which is as written in the statement of the lemma; this uses the uniqueness we know from (ii).

(v) This is immediate from the fact that $(-r) + r = 0_R$; the inverse of $(-r)$ is $r$.

(vi) We can write $0_R = 0_R + 0_R$. Thus, we see that

$$\begin{aligned}
r0_R &= r(0_R + 0_R) \\
&= (r0_R) + (r0_R), && \text{as } \times \text{ distributes over } +.
\end{aligned}$$

But by the existence of additive inverses, we know that there exists $-(r0_R) \in R$. Therefore,

$$\begin{aligned}
r0_R + (-r0_R) &= (r0_R + r0_R) + (-r0_R), && \text{as additive inverses exist,} \\
&= r0_R + \big(r0_R + (-r0_R)\big), && \text{as } + \text{ is associative.}
\end{aligned}$$

However, this just tells us that $0_R = r0_R + 0_R$, so we conclude that $r0_R = 0_R$. We can proceed similarly in the other order to get the result.

(vii) Continuing on from (vi), we see that $0_R = 0_R s = \big((-r) + r\big)\, s = (-r)s + rs$ by distributivity of multiplication over addition. Therefore, adding $-(rs)$ to both sides gives the result. Similarly, one can do it the other way around. $\square$

**Remark 1.12** Since addition $+$ is associative, it is common to not write brackets, e.g. $r + s + t$.

**Definition 1.13** A ring $R$ is commutative if the operation $\times$ is commutative: for all $r, s \in R$,

$$rs = sr.$$

**Definition 1.14** Let $R$ be a ring. A subring is a subset $S \subseteq R$ where the following hold:

(S1) It contains the additive identity, that is $0_R \in S$.

(S2) For all $r \in S$, we have $-r \in S$.

(S3) For all $r, s \in S$, we have $r + s \in S$.

(S4) For all $r, s \in S$, we have $rs \in S$.

**Note:** A subring $S \subseteq R$ is a ring in its own right, whose operations are the same as those for $R$ but restricted onto $S$ and whose additive identity $0_S = 0_R$.

**Proposition** *Let $R$ be any ring. Then, $\{0_R\} \subseteq R$ and $R \subseteq R$ are subrings automatically.*

*Sketch of Proof*: Simply show that each of the axioms in Definition 1.14 is satisfied. $\square$

## 2 Ideals and Factor Rings

**Reminder:** Let $G$ be a group. We call the subgroup $N \leq G$ a normal subgroup if for all $n \in N$ and $g \in G$, we have $gng^{-1} \in N$. This is then denoted $N \trianglelefteq G$. For a normal subgroup, we can define the quotient group $G/N = \{gN : g \in G\}$ under coset addition and multiplication; a coset is the set $gH := \{gh : h \in H\}$ for **any** subgroup $H \leq G$.

**Definition 2.1** Let $R$ be a ring. An ideal (of $R$) is a subset $I \subseteq R$ satisfying the following:

(I1) It contains the additive identity, that is $0_R \in I$.

(I2) For all $x \in I$, we have $-x \in I$. (**Closed under Negation**)

(I3) For all $x, y \in I$, we have $x + y \in I$. (**Closed under Addition**)

(I4) For all $x \in I$ and $r \in R$, we have $rx \in I$ and $xr \in I$. (**Absorbing Property**)

**Note:** An ideal of a ring is automatically a subring of said ring; compare (I4) with (S4).

**Lemma 2.3** *Let $R$ be a commutative ring and $a \in R$. Then, the* ideal generated by $a$

$$(a) := \{ar : r \in R\}$$

*is indeed an ideal of $R$.*

*Proof*: One need only verify the axioms of an ideal written in Definition 2.1.

- Clearly, $0_R = a0_R$ which means that $0_R \in (a)$.
- Suppose $x \in I$, that is $x = ar$ for some $r \in R$. Then, the inverse $-x = -(ar) = a(-r) \in (a)$ because $-r \in R$ since $R$ is a ring and it is closed under additive inverses.
- Suppose $x, y \in I$, that is $x = ar$ and $y = as$ for some $r, s \in R$. Then, their sum is $x + y = ar + as = a(r + s) \in (a)$ because $r + s \in R$ since $R$ is a ring and it is closed under addition.
- Suppose $x \in I$, that is $x = as$ for some $s \in R$, and $r \in R$. Then, $xr = asr = a(sr) \in (a)$ because $sr \in R$ since $R$ is a ring and it is closed under multiplication. □

**Note:** An ideal $(a)$ generated by a **single** element $a \in R$ is called a principal ideal of $R$.

**Lemma 2.4** *Let $R$ be a commutative ring with $1_R \in R$ and $a \in R$. Then, $a \in (a)$ and any ideal of $R$ that contains the element $a$ also contains the entire ideal $(a)$.*

*Proof*: Well, $a = a1_R \in (a)$ is pretty clear. Next, let $I \subseteq R$ be an ideal with $a \in I$ by Axiom (I4), we know that $ar \in I$ for any $r \in R$. Consequently, $\{ar : r \in R\} = (a) \subseteq I$. □

**Definition 2.5** Let $R$ be a ring and $I, J \subseteq R$ be ideals. Then, the sum of ideals is

$$I + J \coloneqq \{x + y : x \in I \text{ and } y \in J\}.$$

**Lemma 2.6** *Let $R$ be a ring and $I, J \subseteq R$ be ideals. Then, we have the following:*
(i) *The set $I + J$ is an ideal of $R$.*
(ii) *The set $I \cap J$ is an ideal of $R$.*

*Proof*: (i) This very much hinges on the fact that $I$ and $J$ are ideals.

- Clearly, $0_R = 0_R + 0_R \in I + J$ since $0_R \in I$ and $0_R \in J$.
- Now then, let $a \in I + J$, which means that $a = x + y$ where $x \in I$ and $y \in J$. Then, $-a = -(x + y) = (-x) + (-y) \in I + J$ because $-x \in I$ and $-y \in J$.
- Let $a, b \in I + J$, which means $a = x + y$ and $b = s + t$ for $x, s \in I$ and $y, t \in J$. Then, $a + b = (x + y) + (s + t) = (x + s) + (y + t) \in I + J$ because addition is associative and commutative and $x + s \in I$ and $y + t \in J$.
- Finally, let $a \in I + J$, which means $a = x + y$ for $x \in I$ and $y \in J$, and $r \in R$. Then, $ar = (x + y)r = (xr) + (yr) \in I + J$ because $xr \in I$ and $yr \in J$.

(ii) This again rests on the fact that $I$ and $J$ are ideals.

- Clearly, $0_R \in I \cap J$ because $0_R \in I$ and $0_R \in J$.
- Now, let $a \in I \cap J$, meaning that $a \in I$ and $a \in J$. Because $I$ and $J$ are ideals, it follows that $-a \in I$ and $-a \in J$, which is to say that $-a \in I \cap J$.
- Let $a, b \in I \cap J$, meaning that $a, b \in I$ and $a, b \in J$. Because $I$ and $J$ are ideals, we know $a + b \in I$ and $a + b \in J$, which therefore means $a + b \in I \cap J$.
- Last, let $a \in I \cap J$, meaning that $a \in I$ and $a \in J$, and $r \in R$. Because $I$ and $J$ are ideals, we conclude $ar \in I$ and $ar \in J$; it immediately follows that $ar \in I \cap J$. □

**Note:** Because $(R, +)$ is an Abelian group and Axioms (I1), (I2) and (I3) imply that $I \leq R$ is a subgroup, we know that $I \trianglelefteq R$ is normal (true of any subgroup of an Abelian group).

**Definition 2.7** Let $R$ be a ring and $I \subseteq R$ an ideal. A coset of $I$ is a subset of the form

$$r + I \coloneqq \{r + x : x \in I\} \subseteq R.$$

**Lemma 2.8** *Let $R$ be a ring and $I \subseteq R$ an ideal, with $r, s \in R$. Then, $r + I = s + I$ if and only if $r - s \in I$.*

*Proof*: ($\Rightarrow$) Suppose $r + I = s + I$. Then, $r + 0_R \in r + I = s + I$, because ideals contain zero. Therefore, $r + 0_R = s + x$ for some element $x \in I$, but the left-hand side is just $r$. Therefore, this rearranges to say that $r - s = x \in I$.

($\Leftarrow$) Suppose $r - s \in I$ and define $x \coloneqq r - s$ (which means that $r = x + s$ and $s = r - x$). We show the cosets $r + I$ and $s + I$ are equal by demonstrating that they are subsets of one another.

- Let $a \in r + I$, which means that $a = r + y$ for some $y \in I$. Therefore, we see that $a = (x + s) + y = s + (x + y) \in s + I$ because ideals are closed under addition and so $x + y \in I$. Because any element of $r + I$ also appears in $s + I$, we know that $r + I \subseteq s + I$.
- Let $b \in s + I$, which means that $b = s + z$ for some $z \in I$. Therefore, we see that $b = (r - x) + z = r + (z - x) \in r + I$ because ideals are closed under addition and negation and so $z - x \in I$. Because any element of $s + I$ also appears in $r + I$, we get $s + I \subseteq r + I$.

Therefore, having both subset inclusions implies that $r + I = s + I$. □

**Lemma 2.10** *Let $R$ be a ring and $I \subseteq R$ an ideal. If $X_1 = a_1 + I, ..., X_n = a_n + I$ are cosets of $I$ in $R$ whose union $\bigcup_{i=1}^{n} X_i = R$, then every coset of $I$ is equal to some $X_i$.*

*Proof*: Let $r \in R$, meaning $r \in X_i = a_i + I$ for some $i$ since $R$ is the union of the $X_i$. Therefore, $r - a_i \in I$ which is equivalent to saying that $r + I = a_i + I = X_i$ by Lemma 2.8. □

**Definition** Let $R$ be a ring and $I \subseteq R$ an ideal. The set of cosets of $I$ in $R$ is

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

**Reminder:** An operation is well-defined if it doesn't depend on the representative taken.

**Lemma 2.12** *The following binary operations defined on $R/I$ are well-defined:*
(i) *The coset addition operation $(r + I) + (s + I) \coloneqq (r + s) + I$.*
(ii) *The coset multiplication operation $(r + I)(s + I) \coloneqq rs + I$.*

*Proof*: (i) To show that coset addition is well-defined, suppose $r_1, r_2, s_1, s_2 \in R$ are such that $r_1 + I = r_2 + I$ and $s_1 + I = s_2 + I$. By Lemma 2.8, this means $r_1 - r_2 \in I$ and $s_1 - s_2 \in I$. Hence, we see that $(r_1 + s_1) - (r_2 + s_2) = (r_1 - r_2) + (s_1 - s_2) \in I$ because ideals are closed under addition and negation. Therefore, again applying Lemma 2.8, we conclude that $(r_1 + s_1) + I = (r_2 + s_2) + I$. Thus, picking different representatives for the left-hand side of the coset addition operation doesn't change what we get in the output, so it is well-defined.

(ii) To show that coset multiplication is well-defined, suppose $r_1, r_2, s_1, s_2 \in R$ are such that $r_1 + I = r_2 + I$ and $s_1 + I = s_2 + I$. By Lemma 2.8, this means $r_1 - r_2 \in I$ and $s_1 - s_2 \in I$. Hence, we see that $r_1 s_1 - r_2 s_2 = (r_1 - r_2)s_1 + r_2(s_1 - s_2) \in I$ because ideals are closed under addition and negation. Therefore, again applying Lemma 2.8, we conclude that $r_1 s_1 + I = r_2 s_2 + I$. □

**Theorem 2.13** *Let $R$ be a ring and $I \subseteq R$ an ideal. Then, $R/I$ together with the coset addition and multiplication operations from* Lemma 2.12 *is a ring with additive identity $0_R + I$. We call $R/I$ a* quotient ring *or* factor ring. *Moreover, if $R$ is a ring with one whose multiplicative identity is $1_R$, then so too is $R/I$, with multiplicative identity $1_R + I$.*

*Proof*: One need only show that the axioms in Definition 1.7 are satisfied.

- Closure under coset addition is immediate from its definition. Now, $0_R + I$ is the additive identity: $(r + I) + (0_R + I) = (r + 0_R) + I = r + I$. Finally, if we continue to assume that $r + I \in R/I$, then $(-r) + I \in R/I$ is the additive inverse. Indeed, we see that $(r+I) + \big((-r) + I\big) = (r + (-r)) + I = 0_R + I$. Hence, $R/I$ is closed under taking additive inverses. This shows that $(R/I, +)$ is an Abelian group.
- Closure under coset multiplication is immediate from its definition.
- Let $r + I, s + I, t + I \in R/I$. Then, we see that

$$\begin{aligned}\big((r+I)(s+I)\big)(t+I) &= (rs+I)(t+I)\\ &= (rs)t + I\\ &= r(st) + I\\ &= (r+I)(st+I)\\ &= (r+I)\big((s+I)(t+I)\big),\end{aligned}$$

  which demonstrates associativity of coset multiplication.
- Let $r + I, s + I, t + I \in R/I$. Then, we see that

$$\begin{aligned}(r+I)\big((s+I)+(t+I)\big) &= (r+I)\big((s+t)+I\big)\\ &= r(s+t) + I\\ &= (rs+rt) + I\\ &= (rs+I) + (rt+I)\\ &= (r+I)(s+I) + (r+I)(t+I)\end{aligned}$$

  and

$$\begin{aligned}\big((r+I)+(s+I)\big)(t+I) &= \big((r+s)+I\big)(t+I)\\ &= (r+s)t + I\\ &= (rt+st) + I\\ &= (rt+I) + (st+I)\\ &= (r+I)(t+I) + (s+I)(t+I),\end{aligned}$$

  which demonstrates distributivity of coset multiplication over coset addition. □

## 3 Homomorphisms

**Definition 3.1** A ring homomorphism is a map $\varphi : R \to S$ between rings satisfying these:
(H1) For all $r_1, r_2 \in R$, we have $\varphi(r_1 + r_2) = \varphi(r_1) + \varphi(r_2)$.
(H2) For all $r_1, r_2 \in R$, we have $\varphi(r_1 r_2) = \varphi(r_1)\varphi(r_2)$.

**Note:** If $\varphi$ is a bijective ring homomorphism, we call it a ring isomorphism and write $R \cong S$.

**Remark 3.3** We always have the following for any ring homomorphism $\varphi : R \to S$:

(i) $\varphi(0_R) = 0_S$. Indeed, $0_S + \varphi(0_R) = \varphi(0_R) = \varphi(0_R + 0_R) = \varphi(0_R) + \varphi(0_R)$ by Axiom (H1). But now, Lemma 1.11(iii) means we cancel one of the $\varphi(0_R)$ to get that $0_S = \varphi(0_R)$.

(ii) $\varphi(-r) = -\varphi(r)$ for all $r \in R$. Indeed, $\varphi(-r) + \varphi(r) = \varphi(-r + r) = \varphi(0_R) = 0_S$ by Axiom (H1) and by (i) above. This shows the inverse of $\varphi(r)$ is $\varphi(-r)$, exactly what we wanted.

**Definition 3.4** Let $R$ and $S$ be rings and $\varphi : R \to S$ be a ring homomorphism.
(i) The kernel of $\varphi$ is $\ker(\varphi) := \{r \in R : \varphi(r) = 0_S\}$.
(ii) The image of $\varphi$ is $\operatorname{im}(\varphi) := \{\varphi(r) : r \in R\}$.

**Proposition 3.6** *Let $R$ and $S$ be rings and $\varphi : R \to S$ be a ring homomorphism.*
(i) *The kernel $\ker(\varphi) \subseteq R$ is an ideal of $R$.*
(ii) *The image $\operatorname{im}(\varphi) \subseteq S$ is a subring of $S$.*

**Note:** Be aware of the fact that the kernel is an ideal but the image is only a subring.

*Proof*: (i) We show the axioms from Definition 2.1.

- Per Remark 3.3, we see that $\varphi(0_R) = 0_S$, so $0_R \in \ker(\varphi)$.
- Let $x \in \ker(\varphi)$. Then, again by Remark 3.3, $\varphi(-x) = -\varphi(x) = -0_S = 0_S$, so $-x \in \ker(\varphi)$.
- Let $x, y \in \ker(\varphi)$. Then, $\varphi(x + y) = \varphi(x) + \varphi(y) = 0_S + 0_S = 0_S$, so $x + y \in \ker(\varphi)$.
- Let $x \in \ker(\varphi)$ and $r \in R$. Then, $\varphi(xr) = \varphi(x)\varphi(r) = 0_S\varphi(r) = 0_S$, so $xr \in \ker(\varphi)$.

(ii) We show the axioms from Definition 1.14.

- Per Remark 3.3, we see that $\varphi(0_R) = 0_S$, so $0_S \in \operatorname{im}(\varphi)$.
- Let $s \in \operatorname{im}(\varphi)$, meaning $s = \varphi(r)$ for some $r \in R$. Again by Remark 3.3, we see that $-s = -\varphi(r) = \varphi(-r)$, so $-s \in \operatorname{im}(\varphi)$.
- Let $s_1, s_2 \in \operatorname{im}(\varphi)$, meaning $s_1 = \varphi(r_1)$ and $s_2 = \varphi(r_2)$ for some $r_1, r_2 \in R$. Then, $s_1 + s_2 = \varphi(r_1) + \varphi(r_2) = \varphi(r_1 + r_2)$, so $s_1 + s_2 \in \operatorname{im}(\varphi)$.
- Let $s_1, s_2 \in \operatorname{im}(\varphi)$, meaning $s_1 = \varphi(r_1)$ and $s_2 = \varphi(r_2)$ for some $r_1, r_2 \in R$. Then, $s_1 s_2 = \varphi(r_1)\varphi(r_2) = \varphi(r_1 r_2)$, so $s_1 s_2 \in \operatorname{im}(\varphi)$. □

**Reminder:** Let $f : A \to B$ be an arbitrary function.

(i) $f$ is injective (or one-to-one) if for every $a_1, a_2 \in A$, $f(a_1) = f(a_2)$ implies $a_1 = a_2$.

(ii) $f$ is surjective (or onto) if for every $b \in B$, there exists $a \in A$ such that $f(a) = b$.

**Lemma 3.7** *A ring homomorphism* $\varphi : R \to S$ *is injective if and only if* $\ker(\varphi) = \{0_R\}$.

*Proof*: ($\Rightarrow$) Let $\varphi$ be injective. We know that $0_R \in \ker(\varphi)$ from Remark 3.3, so the kernel is non-empty. Suppose $x \in \ker(\varphi)$. Then, $\varphi(x) = 0_S = \varphi(0_R)$. But injectivity then allows us to conclude that $x = 0_R$, so in fact the kernel consists only of the zero of $R$.

($\Leftarrow$) Let $\ker(\varphi) = \{0_R\}$ and assume $x, y \in R$ with $\varphi(x) = \varphi(y)$. We conclude from Axiom (H1) that $\varphi(x - y) = \varphi(x) - \varphi(y) = 0_S$, so it follows that $x - y \in \ker(\varphi)$. But this means that $x - y = 0_R$, which is to say $x = y$. □

**Theorem 3.9** *Let $R$ be a ring and $I \subseteq R$ an ideal. Then, the* quotient map $\varphi : R \to R/I$ *given by* $\varphi(r) = r + I$ *is a ring homomorphism. Furthermore,* $\ker(\varphi) = I$ *and* $\operatorname{im}(\varphi) = R/I$.

*Proof*: The first part of the proof concerns showing the axioms in Definition 3.1.

- Let $r_1, r_2 \in R$. Then, $\varphi(r_1 + r_2) = (r_1 + r_2) + I = (r_1 + I) + (r_2 + I) = \varphi(r_1) + \varphi(r_2)$.
- Let $r_1, r_2 \in R$. Then, $\varphi(r_1 r_2) = r_1 r_2 + I = (r_1 + I)(r_2 + I) = \varphi(r_1)\varphi(r_2)$.

Furthermore, we see that

$$\begin{aligned}\ker(\varphi) &= \{r \in R : \varphi(r) = 0_{R/I}\} \\ &= \{r \in R : r + I = 0_R + I\} \\ &= \{r \in R : r - 0_R \in I\} \\ &= \{r \in R : r \in I\} \\ &= I\end{aligned}$$

and

$$\begin{aligned}\operatorname{im}(\varphi) &= \{\varphi(r) : r \in R\} \\ &= \{r + I : r \in R\} \\ &= R/I.\end{aligned}$$

□

**Note:** The quotient map in Theorem 3.9 is surjective as its image is the whole codomain.

**Theorem 3.10** (First Isomorphism Theorem) *Let* $\varphi : R \to S$ *be a ring homomorphism. Then, there exists an* induced ring isomorphism $\overline{\varphi} : R/\ker(\varphi) \to \operatorname{im}(\varphi)$ *given by*

$$\overline{\varphi}\left(r + \ker(\varphi)\right) = \varphi(r).$$

*Proof*: There are a few things to prove about the induced map $\overline{\varphi}$, namely that it is well-defined, it is a ring homomorphism and that it is bijective (i.e. has trivial kernel and has full image).

- Let $r_1 + \ker(\varphi) = r_2 + \ker(\varphi)$ for $r_1, r_2 \in R$. By Lemma 2.8, we know that $r_1 - r_2 \in \ker(\varphi)$. But this is to say $\varphi(r_1 - r_2) = 0_S$; applying Axiom (H1) to the left-hand side results in $\varphi(r_1) - \varphi(r_2) = 0_S$, which is equivalent to $\varphi(r_1) = \varphi(r_2)$. Therefore, $\overline{\varphi}$ is well-defined.
- Let $r + \ker(\varphi), s + \ker(\varphi) \in R/\ker(\varphi)$. Then, we see that

  $$\begin{aligned} \overline{\varphi}\left((r+\ker(\varphi)) + (s+\ker(\varphi))\right) &= \overline{\varphi}\left((r+s)+\ker(\varphi)\right) \\ &= \varphi(r+s) \\ &= \varphi(r)+\varphi(s) \\ &= \overline{\varphi}(r+\ker(\varphi)) + \overline{\varphi}(s+\ker(\varphi)) \end{aligned}$$

  and

  $$\begin{aligned} \overline{\varphi}\left((r+\ker(\varphi))(s+\ker(\varphi))\right) &= \overline{\varphi}(rs+\ker(\varphi)) \\ &= \varphi(rs) \\ &= \varphi(r)\varphi(s) \\ &= \overline{\varphi}(r+\ker(\varphi))\overline{\varphi}(s+\ker(\varphi)). \end{aligned}$$

  Hence, we know that $\overline{\varphi}$ is a ring homomorphism.
- To show that $\overline{\varphi}$ is injective, we will use Lemma 3.7. Indeed, let $r + \ker(\varphi) \in \ker(\overline{\varphi})$, which means that $\overline{\varphi}(r + \ker(\varphi)) = 0_S$. By the definition of $\overline{\varphi}$, this is equivalent to $\varphi(r) = 0_S$, meaning $r \in \ker(\varphi)$. Therefore, Lemma 2.8 tells us that $r + \ker(\varphi) = 0_R + \ker(\varphi) = 0_{R/I}$. In other words, anything in the kernel is always $0_{R/I}$, so we indeed get injectivity.
- Finally, to show that $\overline{\varphi}$ is surjective, suppose that $s \in \operatorname{im}(\varphi)$, meaning that $s = \varphi(r)$ for some $r \in R$. But by definition of $\overline{\varphi}$, this means that $s = \overline{\varphi}(r + \ker(\varphi))$, so we indeed get surjectivity. □

**Definition 3.12** Let $R$ and $S$ be rings. The direct product of these rings is defined as

$$R \times S := \{(r, s) : r \in R \text{ and } s \in S\}.$$

**Proposition** *Let $R$ and $S$ be rings. Then, $R \times S$ is a ring with the following operations:*
(i) *The* pointwise addition *operation* $(r_1, s_1) + (r_2, s_2) := (r_1 + r_2, s_1 + s_2)$.
(ii) *The* pointwise multiplication *operation* $(r_1, s_1)(r_2, s_2) := (r_1 r_2, s_1 s_2)$.

*Sketch of Proof*: Simply show that each of the axioms in Definition 1.7 is satisfied. □

# 4 Fields and Integral Domains

**Definition 4.1** Let $R$ be a ring with one. An element $a \in R$ is called a unit (or invertible) if there exists an element $b \in R$ such that $ab = 1_R = ba$. The set of units is denoted $U(R)$.

**Reminder:** We call two integers $a, b \in \mathbb{Z}$ coprime (or relatively prime) if $\gcd(a, b) = 1$.

**Definition 4.2** A ring $R$ is called a field if it satisfies the following axioms:
(F1) $R$ is a ring with one, namely $1_R$.
(F2) The identities are distinct, that is $1_R \neq 0_R$.
(F3) $R$ is commutative.
(F4) Every non-zero element of $R$ is a unit, that is $U(R) = R \setminus \{0_R\}$.

Henceforth, we use the blackboard font to denote arbitrary fields, in particular $\mathbb{K}$.

**Definition 4.3** Let $\mathbb{K}$ be a field. A subfield is a subset $\mathbb{F} \subseteq \mathbb{K}$ where the following hold:
(SF1) It contains the identities, that is $0_{\mathbb{K}}, 1_{\mathbb{K}} \in \mathbb{F}$.
(SF2) For all $r \in \mathbb{F}$, we have $-r \in \mathbb{F}$.
(SF3) For all $r, s \in \mathbb{F}$, we have $r + s \in \mathbb{F}$ and $rs \in \mathbb{F}$.
(SF4) For all $r \in \mathbb{F} \setminus \{0_{\mathbb{K}}\}$, we have $r^{-1} \in \mathbb{F}$.

**Note:** Much like a subring, a subfield is a field in its own right. Also, a subfield is simply a subring containing the multiplicative identity and whose non-zero elements are units.

**Reminder:** Let $\mathbb{K}$ be a field. A $\mathbb{K}$-vector space is a set $V$ satisfying the following axioms:
(V1) $V$ is an Abelian group under addition.
(V2) For all $v \in V$ and $k_1, k_2 \in \mathbb{K}$, we have $k_1(k_2 v) = (k_1 k_2)v$.
(V3) For all $v \in V$, we have $1_{\mathbb{K}} v = v$.
(V4) For all $v \in V$ and $k_1, k_2 \in \mathbb{K}$, we have $(k_1 + k_2)v = k_1 v + k_2 v$.
(V5) For all $v_1, v_2 \in V$ and $k \in \mathbb{K}$, we have $k(v_1 + v_2) = kv_1 + kv_2$.

**Theorem 4.5** *Let $\mathbb{K}$ be a field and $\mathbb{F} \subseteq \mathbb{K}$ a subfield. Then, $\mathbb{K}$ is an $\mathbb{F}$-vector space with addition being the usual addition on $\mathbb{K}$ and scalar multiplication defined by $\lambda \cdot r := \lambda r$, where $\lambda \in \mathbb{F}$ and $r \in \mathbb{K}$ and the right-hand side is the usual multiplication in $\mathbb{K}$.*

*Sketch of Proof*: Simply check the vector space axioms written above. □

**Definition 4.6** Let $R$ be a ring and $r \in R$. For $n \in \mathbb{Z}$, we define the product notation

$$nr := \begin{cases} \overbrace{r + \cdots + r}^{n \text{ copies}} & \text{if } n > 0 \\ 0 & \text{if } n = 0 \\ \underbrace{(-r) + \cdots + (-r)}_{n \text{ copies}} & \text{if } n < 0 \end{cases}.$$

**Note:** In general, $n \notin R$ so $nr$ as defined above is **not** just multiplication in the ring $R$.

**Remark** In fact, the ring $R$ is behaving analogously to a vector space where $\mathbb{Z}$ is acting as the scalars. However, a vector space uses a field as scalars and $\mathbb{Z}$ is **not** a field. What we are touching on here is a slight generalisation of the notion of a vector space over a field, that being a so-called *module* over a ring (more on this in MATH3195/5195M).

**Lemma 4.7** *Let $R$ be a ring with $r, s \in R$ and $n, m \in \mathbb{Z}$. Then, we have the following:*

(i) $mr + nr = (m + n)r$.
(ii) $(-n)r = -(nr)$.
(iii) $n(-r) = -(nr)$.
(iv) $m(r + s) = mr + ms$.
(v) $m(nr) = (mn)r$.
(vi) $(mr)(ns) = (mn)rs = (nr)(ms)$.

*Proof*: This is an exercise in using Definition 4.6 in conjunction with previously-seen axioms. □

**Definition 4.8** Let $\mathbb{K}$ be a field. The characteristic of $\mathbb{K}$ is the least positive integer $n \in \mathbb{Z}^+$ such that $n1_\mathbb{K} = 0_\mathbb{K}$ (if such $n$ exists, otherwise we define it to be zero), denoted $\text{char}(\mathbb{K})$.

**Lemma 4.9** *Let $\mathbb{K}$ be a field. Then,* $\text{char}(\mathbb{K})$ *is either zero or a prime number.*

*Proof*: Assume that $\text{char}(\mathbb{K}) = n \neq 0$. Suppose for a contradiction that $n = ab$ where $a, b \in \mathbb{Z}^+$ such that $1 < a, b < n$. Then, we see that

$$0_\mathbb{K} = n1_\mathbb{K} = (ab)1_\mathbb{K} = (a1_\mathbb{K})(b1_\mathbb{K}).$$

Since $1 < a, b < n$, we must have that $a1_\mathbb{K} \neq 0$ and $b1_\mathbb{K} \neq 0$ (because Definition 4.8 defines the characteristic to be the **least** positive integer and we are assuming this to be $n$, so anything less than it cannot multiply $1_\mathbb{K}$ to get $0_\mathbb{K}$). Note that these are non-zero elements of a field, so their inverses exist. As such, multiplying the above equation on the left by $(a1_\mathbb{K})^{-1}$ tells us that $b1_\mathbb{K} = 0$, a contradiction. Therefore, we cannot write $n = ab$ with $1 < a, b < n$ so it must be that $n$ is prime. □

**Definition 4.11** Let $R$ be a commutative ring. We call a non-zero element $r \in R \setminus \{0_R\}$ a non-zero zero divisor if there exists an element $s \in R \setminus \{0_R\}$ such that $rs = 0_R$.

**Remark** Most people simply call them *zero divisors*, omitting the "non-zero" for brevity.

**Definition 4.11** Let $R$ be a ring. It is an integral domain (ID) if it satisfies the following:
(ID1) $R$ is a ring with one, namely $1_R$.
(ID2) The identities are distinct, that is $1_R \neq 0_R$.
(ID3) $R$ is commutative.
(ID4) $R$ has **no** non-zero zero divisors.

**Note:** We can restate Axiom (ID4) in the following alternative-yet-equivalent ways:
(i) For all $r, s \in R \setminus \{0_R\}$, we have $rs \neq 0_R$.
(ii) For all $r, s \in R$, $rs = 0_R$ implies that $r = 0_R$ or $s = 0_R$.

**Lemma 4.12** *Every field is an integral domain.*

*Proof*: Let $\mathbb{K}$ be a field. Compare Definitions 4.2 and 4.11 to see Axioms (ID1), (ID2) and (ID3) hold automatically. It only remains to show the final integral domain axiom. Indeed, let $r, s \in \mathbb{K}$ and suppose that $rs = 0_{\mathbb{K}}$. If $r \neq 0_{\mathbb{K}}$, then $r^{-1} \in \mathbb{K}$ exists and we can consider $r^{-1}rs = s = 0_{\mathbb{K}}$, so $s = 0_{\mathbb{K}}$. So, $rs = 0_{\mathbb{K}}$ implies that $r = 0_{\mathbb{K}}$ or $s = 0_{\mathbb{K}}$, which is the alternative form of (ID4). □

**Definition 4.13** Let $R$ be a ring and $f \in R[x]$ be a non-zero polynomial. Then, we write $f(x) = a_0 + a_1x + \cdots + a_nx^n$ where $n \in \mathbb{N}$ and each $a_i \in R$ with $a_n \neq 0_R$.
(i) The degree of $f$ is the integer $n$, denoted $\deg(f)$.
(ii) The leading term of $f$ is the term $a_nx^n$.
(iii) The leading coefficient of $f$ is the element $a_n$.

**Proposition 4.14** *Let $R$ be an integral domain. Then, $R[x]$ is also an integral domain.*

*Proof*: We just need to show the integral domain axioms.

- $R[x]$ is a ring with one where $1_{R[x]} = 1_R$, regarded as a constant polynomial.
- Because $R$ is an integral domain, $1_{R[x]} = 1_R \neq 0_R = 0_{R[x]}$.
- Because $R$ is commutative, so too is $R[x]$.
- Let $f, g \in \mathbb{R}[x] \setminus \{0_{R[x]}\}$ where $f = a_0 + a_1x + \cdots + a_nx^n$ and $g = b_0 + b_1x + \cdots + b_mx^m$ where $a_n, b_m \neq 0_R$. Then, their product is $fg = a_nb_mx^nx^m + \cdots = a_nb_mx^{n+m} + \cdots$. Since $R$ is an integral domain, we know that $a_nb_m \neq 0$, which means that $fg \neq 0$. □

**Remark 4.15** We have actually also shown $\deg(fg) = \deg(f) + \deg(g)$ for $f, g \in R[x] \setminus \{0_{R[x]}\}$.

## 5 Classes of Integral Domains

**Theorem 5.1** (Division Algorithm for $\mathbb{Z}$) *For every $a, b \in \mathbb{Z}$ with $b \neq 0$, there exist unique $q, r \in \mathbb{Z}$ such that $0 \leq r < |b|$ and $a = qb + r$.*

*Proof*: Omitted. □

**Theorem 5.2** *Every ideal of $\mathbb{Z}$ is principal, that is generated by a single element.*

*Proof*: Let $I \subseteq \mathbb{Z}$ be an ideal. If $I = \{0\}$, we are done since $\{0\} = (0)$, so we henceforth assume that $I \neq \{0\}$ is a non-zero ideal. By Axiom (I2), we know that $I$ contains positive elements (since both $\pm x \in I$ for any $x \in I$). As such, let $a \in I$ be the smallest positive element and take some $x \in I$. By the Division Algorithm for $\mathbb{Z}$, we can write $x = qa + r$ for $q, r \in \mathbb{Z}$ and $0 \leq r < |a| = a$. By Axiom (I4), the absorbing property, we know that $qa \in I$. Therefore, since ideals are closed under addition and negation, $r = x - qa \in I$ . Now, $r > 0$ contradicts the minimality of $a$, so we must have that $r = 0$. In other words, $x = qa \in (a)$. This shows the inclusion $I \subseteq (a)$. Conversely, we assumed that $a \in I$ so we immediately have $(a) \subseteq I$ from Lemma 2.4. Consequently, $I = (a)$. □

**Definition 5.3** A principal ideal domain (PID) is an integral domain in which every ideal is principal. In other words, for each ideal, there exists a single element generating it.

**Lemma 5.4** *Every field is a principal ideal domain.*

*Proof*: Let $\mathbb{K}$ be a field. Per Lemma 4.12, we know that $\mathbb{K}$ is an integral domain. Suppose that $I \subseteq \mathbb{K}$ is an ideal. If it is zero, it is principal, so assume $I \neq \{0_{\mathbb{K}}\}$. Thus, it contains a non-zero element $a \in I$. But if $x \in \mathbb{K}$, we can write $x = (xa^{-1})a \in I$ by Axiom (I4). Thus, $I = \mathbb{K}$, so we can write $I = (1_{\mathbb{K}})$. □

**Note:** The proof of Lemma 5.4 shows the only ideals of a field $\mathbb{K}$ are $\{0_{\mathbb{K}}\}$ and $\mathbb{K}$ itself.

**Definition 5.5** A Euclidean domain is an integral domain $R$ with a map $V : R \setminus \{0_R\} \to \mathbb{N}$ called the valuation satisfying the following axioms:

(ED1) For all $a, b \in R \setminus \{0_R\}$, we have $V(a) \leq V(ab)$.

(ED2) For every $a, b \in R$ with $b \neq 0_R$, there exist $q, r \in R$ such that $a = qb + r$ and **one** of the following occurs: (i) $r = 0_R$ **or** (ii) $r \neq 0_R$ and $V(r) < V(b)$.

**Remark 5.6** Comparing Definition 5.5 to the Division Algorithm for $\mathbb{Z}$, notice (ED2) is almost the same **except** we don't insist that $q, r \in R$ are unique which we did do for $\mathbb{Z}$. Furthermore, we see that the valuation isn't defined on $0_R$.

**Note:** It is enough to have (ED2) only. Indeed, if $R$ is an integral domain with a valuation $\mathcal{V}$ satisfying only (ED2), we can define a new valuation $V$ which satisfies (ED1) and (ED2):

$$V : R \setminus \{0_R\} \to \mathbb{N}, \qquad V(a) = \min\{\mathcal{V}(ra) : r \in R \setminus \{0_R\}\}.$$

In words, $V(a)$ is the minimum value attained by $\mathcal{V}$ on non-zero elements of the ideal $(a)$.

**Lemma** *Every field is a Euclidean domain.*

*Proof*: Let $\mathbb{K}$ be a field. Per Lemma 4.12, we know that $\mathbb{K}$ is an integral domain; it remains to define a valuation map. Indeed, let $V : \mathbb{K} \setminus \{0_\mathbb{K}\} \to \mathbb{N}$ be given by $V(a) = 1$, that is it always outputs the integer one. The fact that Axiom (ED1) holds is trivial. Next, let $a, b \in \mathbb{K}$ with $b \neq 0_\mathbb{K}$. Then, we can always write $a = ab^{-1}b + 0_\mathbb{K}$, that is $q := ab^{-1}$ and $r = 0_\mathbb{K}$. This shows that Axiom (ED2) is satisfied. □

**Theorem 5.8** *Every Euclidean domain is a principal ideal domain.*

*Proof*: Let $R$ be a Euclidean domain with valuation map $V$ and let $I \subseteq R$ be an ideal. Again, $I$ being the zero ideal is nothing special because we know it is generated by $0_R$ and we are done; assume therefore that $I \neq \{0_R\}$. As such, we can choose a non-zero element $a \in I \setminus \{0_R\}$ for which $V(a)$ is minimal. The goal is to establish $I = (a)$ by showing each inclusion.

- If $x \in (a)$, then $x = ra$ for some $r \in R$. By Axiom (I4), since $a$ is an element of the ideal, absorption means that $x \in I$. This shows that $(a) \subseteq I$.
- If $x \in I$, then we can write $x = qa + r$ where either (i) $r = 0$ or (ii) $r \neq 0$ but $V(r) < V(a)$ by Axiom (ED2). But if (ii) is true, then $r = x - qa \in I$ but this contradicts the minimality of $V(a)$. The only situation that can occur is (i), so $x = qa$ and thus $x \in (a)$. This shows that $I \subseteq (a)$. □

**Remark 5.9** The converse of Theorem 5.8 is **not** true; there exist principal ideal domains that are not Euclidean domains, e.g. the (sub)ring $\{a + b\sqrt{-19} : a, b \in \mathbb{Z} \text{ with } a \equiv b \pmod 2\} \subseteq \mathbb{C}$.

**Proposition 5.10** (Division Algorithm for $\mathbb{K}[x]$) *Let $\mathbb{K}$ be a field. For every $f, g \in \mathbb{K}[x]$ with $g \neq 0$, there exist unique $q, r \in \mathbb{K}[x]$ such that $f = qg + r$ and either* (i) $r = 0$ *or* (ii) $r \neq 0$ *and* $\deg(r) < \deg(g)$.

*Proof*: Omitted. □

**Corollary 5.12** *For $\mathbb{K}$ a field, $\mathbb{K}[x]$ is a Euclidean domain, and a principal ideal domain.*

*Proof*: Combining Lemma 4.12 and Proposition 4.14, $\mathbb{K}[x]$ is an integral domain. It remains to exhibit a valuation map satisfying the axioms in Definition 5.5. Indeed, let $V : \mathbb{K}[x] \setminus \{0_{\mathbb{K}[x]}\} \to \mathbb{N}$ be given by $V(f) = \deg(f)$. If $f, g \in \mathbb{K}[x] \setminus \{0_{\mathbb{K}[x]}\}$, then $V(fg) = \deg(fg) = \deg(f) + \deg(g)$ by Remark 4.15. Because $\deg(g) \geq 0$, this tells us that $V(f) \leq V(fg)$, so Axiom (ED1) is satisfied. Finally, Axiom (ED2) is an immediate consequence of the Division Algorithm for $\mathbb{K}[x]$. □

**Reminder:** The Gaussian integers is the ring $\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\}$, a subring of $\mathbb{C}$.

**Lemma 5.13** *The ring $\mathbb{Z}[i]$ is an integral domain.*

*Proof*: As usual, it suffices to show each of the axioms in Definition 4.11.

- It is certainly a ring with multiplicative identity $1_{\mathbb{Z}[i]} = 1_{\mathbb{C}} = 1$.
- Clearly, $1_{\mathbb{Z}[i]} = 1 \neq 0 = 0_{\mathbb{Z}[i]}$.
- Because $\mathbb{C}$ is commutative, so too is $\mathbb{Z}[i]$.
- Let $a, b \in \mathbb{Z}[i]$ with $ab = 0$. If $a \neq 0$, then $b = a^{-1}ab = a^{-1}0 = 0$; either $a = 0$ or $b = 0$. □

**Definition 5.14** The norm on $\mathbb{Z}[i]$ is $N : \mathbb{Z}[i] \setminus \{0\} \to \mathbb{N}$ with $N(a+bi) = |a + bi|^2 = a^2 + b^2$.

**Proposition 5.15** *The ring $\mathbb{Z}[i]$ is a Euclidean domain, and a principal ideal domain.*

*Proof*: We know from Lemma 5.13 that $\mathbb{Z}[i]$ is an integral domain. It remains to show that there exists a valuation satisfying the relevant axioms. Indeed, we claim that the norm $N$ is such a map. We first notice that $N(x) \geq 1$ for all $x \in \mathbb{Z}[i] \setminus \{0\}$, from which it follows that $N(xy) = |xy|^2 = |x|^2|y|^2 \geq |x|^2 = N(a)$, so Axiom (ED1) is satisfied.

As for Axiom (ED2), consider $x, y \in \mathbb{Z}[i]$ with $y \neq 0$ and write them as $x = s + ti$ and $y = u + vi$ for $s, t, u, v \in \mathbb{Z}$. We can form the quotient $\frac{a}{b} = l + mi \in \mathbb{C}$ where $l, m \in \mathbb{R}$. However, for (ED2) to be satisfied, we want to use $m + ni$ to define $L + Mi$ where now $L, M \in \mathbb{Z}$. Indeed, let $L, M \in \mathbb{Z}$ be such that $|l - L| \leq \frac{1}{2}$ and $|m - M| \leq \frac{1}{2}$. Then, we can write

$$\frac{a}{b} = L + Mi + (l - L) + (m - M)i \qquad \Rightarrow \qquad a = (L + Mi)b + \big((l - L) + (m - M)i\big)\, b..$$

Because $a - (L + Mi)b \in \mathbb{Z}[i]$, the term $\big((l - L) + (m - M)i\big)\, b \in \mathbb{Z}[i]$ also. If this is zero, we are done. Hence, assume $\big((l - L) + (m - M)i\big)\, b \neq 0$, in particular $(l - L) + (m - M)i \neq 0$ since we already assume $b \neq 0$. Thus, we use the Triangle Inequality to see that the norm satisfies

$$N\Big(\big((l - L) + (m - M)i\big)\, b\Big) = \big|(l - L) + (m - M)i\big|^2 |b|^2 \leq \left(\frac{1}{4} + \frac{1}{4}\right) |b|^2 = \frac{1}{2}N(b) < N(b).$$

For $q = L + Mi$ and $r = \big((l - L) + (m - M)i\big)\, b$, we conclude that Axiom (ED2) is satisfied. □

# 6 Elements in Integral Domains

**Definition 6.1** Let $R$ be an integral domain and $a, b \in R$. We say that $a$ divides $b$ (or $a$ is a divisor of $b$) if there exists $d \in R$ with $da = b$; we write $a \mid b$. Otherwise, we write $a \nmid b$.

**Definition 6.3** Let $R$ be an integral domain and $a, b \in R$. Then, $b$ is an associate of $a$ if there exists a unit $u \in U(R)$ such that $ua = b$.

**Note:** The notion of being associate is symmetric: if $b$ is an associate of $a$, then $ua = b$ for some $u \in U(R)$. But because $u^{-1} \in U(R)$ automatically, we can also write that $u^{-1}b = a$, so $a$ is an associate of $b$. Consequently, we may just say that they are associates in $R$.

**Remark** Recall that units are invertible elements (Definition 4.1) and that a field is a commutative ring where every non-zero element is invertible (Definition 4.2). Therefore, for $\mathbb{K}$ a field, we have $U(\mathbb{K}) = \mathbb{K} \setminus \{0_{\mathbb{K}}\}$ – this is Axiom (F4) – and thus all non-zero elements are associates.

**Lemma 6.5** *Let $R$ be an integral domain and $a, b \in R$. Then, $a$ and $b$ are associates in $R$ if and only if* **both** *$a \mid b$ and $b \mid a$.*

*Proof*: ($\Rightarrow$) Let $a$ and $b$ be associates. Then, $ua = b$ for some $u \in U(R)$, which is precisely to say that $a \mid b$. But we can equally write $u^{-1}b = a$, which is precisely to say that $b \mid a$.

($\Leftarrow$) Suppose $a \mid b$ and $b \mid a$. Per Definition 6.1, this means there exist $d, e \in R$ such that $da = b$ and $eb = a$. Therefore, we can substitute the first into the second to get $a = eb = eda$, which is equivalent to $(1_R - ed)a = 0$. Because $R$ is an integral domain, there are no non-zero zero divisors, so either $a = 0$ or $1_R - ed = 0$.

- If $a = 0$, then $b = 0$ and they are trivially associates.
- If $1_R - ed = 0$, then $1_R = ed = de$, so $d, e \in U(R)$ because they are inverse to each other. Therefore, $a$ and $b$ are associates once again. $\square$

**Proposition 6.7** *Let $R$ be an integral domain and $a, b \in R$. Then, $a$ and $b$ are associates in $R$ if and only if $(a) = (b)$.*

*Proof*: ($\Rightarrow$) Let $a$ and $b$ be associates. Then, $ua = b$ and $vb = a$ for some $u, v \in U(R)$. We now show both inclusions of the ideals each of the associate elements generate.

- If $x \in (a)$, then $a = ra = rvb$ for some $r \in R$, so $x \in (b)$. Consequently, $(a) \subseteq (b)$.
- If $y \in (b)$, then $y = sb = sua$ for some $s \in R$, so $y \in (a)$. Consequently, $(b) \subseteq (a)$.

($\Leftarrow$) Suppose $(a) = (b)$. Clearly, $a = 1_R a \in (a)$ which means $a \in (b)$, that is $a = rb$ for some $r \in R$. This is to say that $b \mid a$. Similarly, $b = 1_R b \in (b)$ which means $b \in (a)$, that is $b = sa$ for some $s \in R$. This is to say that $a \mid b$. By Lemma 6.5, we know that $a$ and $b$ are associates. $\square$

**Definition 6.9** Let $R$ be an integral domain and $a, b \in R$ **not both** zero. Then, an element $d \in R$ is a greatest common divisor (GCD) of $a$ and $b$, denoted $d = \gcd(a, b)$, if these hold:

(i) Both $d \mid a$ and $d \mid b$.

(ii) If $c \in R$ such that $c \mid a$ and $c \mid b$, then $c \mid d$.

**Note:** A greatest common divisor is **not** unique, but any two are related; see Lemma 6.12.

**Remark 6.11** Let $R$ be an integral domain with $a \in R \setminus \{0_R\}$. Then, $\gcd(a, 0) = \gcd(0, a) = a$.

**Lemma 6.12** *Let $R$ be an integral domain and $a, b \in R$* **not both** *zero. If $d_1$ and $d_2$ are greatest common divisors of $a$ and $b$, then $d_1$ and $d_2$ are associates.*

*Proof*: By Definition 6.9(i), we know that $d_1 \mid a$ and $d_2 \mid b$. But using the fact that $d_2$ is also a greatest common divisor (in particular that it divides both $a$ and $b$), Definition 6.9(ii) tells us that $d_1 \mid d_2$. However, we can exchange the roles of $d_1$ and $d_2$ above and run the same logic to conclude that $d_2 \mid d_1$. Therefore, Lemma 6.5 tells us that $d_1$ and $d_2$ are associates. $\square$

**Remark 6.13** Let $R$ be an integral domain and $a, b \in R$ **not both** zero. If $d$ is a greatest common divisor of $a$ and $b$, then so too is any associate of $d$. Indeed, let $d = \gcd(a, b)$ have an associate $\delta = ud$ for some $u \in U(R)$. We now show that Definition 6.9 is satisfied by the element $\delta$.

(i) Because $d \mid a$ and $d \mid b$, we see that $a = rd$ and $b = sd$ for some $r, s \in R$. But using the fact that $d = u^{-1}\delta$, this tells us $a = ru^{-1}\delta$ and $b = su^{-1}\delta$; we therefore have $\delta \mid a$ and $\delta \mid b$.

(ii) Let $c \in R$ such that $c \mid a$ and $c \mid b$. As $d$ is a greatest common divisor, $c \mid d$ which means $d = tc$ for some $t \in R$. Again, this tells us $u^{-1}\delta = tc \Leftrightarrow \delta = utc$; we therefore have $c \mid \delta$.

**Reminder:** The Euclidean Algorithm in $\mathbb{Z}$ is a method for computing a greatest common divisor of two integers. Indeed, let $a, b \in \mathbb{Z}$ with $b \neq 0$. We proceed as follows for $q_i, r_i \in \mathbb{Z}$:

$$\begin{aligned}
a &= q_1 b + r_1, && \text{for } 0 \leq r_1 < |b_1|, \\
b &= q_2 r_1 + r_2, && \text{for } 0 \leq r_2 < r_1, \\
r_1 &= q_3 r_2 + r_3, && \text{for } 0 \leq r_3 < r_2, \\
&\ \vdots \\
r_{k-3} &= q_{k-1} r_{k-2} + r_{k-1}, && \text{for } 0 \leq r_{k-1} < r_{k-2}, \\
r_{k-2} &= q_k r_{k-1} + 0.
\end{aligned}$$

The algorithm terminates when $r_k = 0$ for some $k \in \mathbb{Z}^+$ and we obtain $\gcd(a, b) = r_{k-1}$.

**Note:** In general, the greatest common divisor does **not** exist in integral domains (that is, being an ID isn't sufficient to guarantee GCDs are well-defined). An example is $\mathbb{Z}[\sqrt{-3}]$; this is an integral domain but $2 + 2\sqrt{-3}$ and 4 do not have a greatest common divisor.

**Theorem 6.16** *Let $R$ be a principal ideal domain and $a, b \in R$* **not both** *zero. Then, $a$ and $b$ have a greatest common divisor $d$. Moreover, there exist $s, t \in R$ such that $sa + tb = d$.*

*Proof*: Let $I \coloneqq \{ua + vb : u, v \in R\}$; this is an ideal of $R$ (one can prove this by showing the usual axioms are satisfied). Because $R$ is a principal ideal domain, there exists an element that generates this ideal, say $d \in R$ where $I = (d)$. In particular, we have $d \in I$ so there exist $s, t \in R$ with $d = sa + tb$. It remains to show that $d$ is a greatest common divisor of $a$ and $b$.

(i) Because $a, b \in I = (d)$, we have $a = xd$ and $b = yd$ for $x, y \in R$; this says $d \mid a$ and $d \mid b$.

(ii) Let $c \in R$ such that $c \mid a$ and $c \mid b$. This means that $a = mc$ and $b = nc$ for some $m, n \in R$. Substituting, we see that $d = sa + tb = smc + tnc = (sm + tn)c$ so $c \mid d$. □

**Note:** Writing a greatest common divisor in the form $sa + tb$ is called Bézout's Lemma.

**Remark 6.17** Recall that any Euclidean domain is automatically a principal ideal domain by Theorem 5.8. Hence, Theorem 6.16 implies that Euclidean domains also have greatest common divisors. In fact, we can use a corresponding Euclidean Algorithm to compute greatest common divisors (it will be a slight adaptation of the Euclidean Algorithm for $\mathbb{Z}$ in the previous reminder).

**Definition 6.19** Let $R$ be an integral domain and $a, b \in R$ **not both** zero. We say that $a$ and $b$ are coprime (or relatively prime) if $\gcd(a, b) = 1_R$.

**Note:** In the case of coprime elements, the greatest common divisors are precisely $U(R)$; this is a consequence of Lemma 6.12 and Remark 6.13. In particular, if $a$ and $b$ do **not** have a greatest common divisor, then they are not coprime with each other.

**Remark 6.18** Recall that the Fibonacci numbers are the sequence defined by $F_0 = F_1 = 1$ and

$$F_n = F_{n-1} + F_{n-2}.$$

If we apply the Euclidean Algorithm to consecutive Fibonacci numbers, we should see that they are coprime. Indeed, let $F_{n+1}$ and $F_{n+2}$ be two consecutive Fibonacci numbers. Then, we have

$$\begin{aligned} F_{n+2} &= 1F_{n+1} + F_n, \\ F_{n+1} &= 1F_n + F_{n-1}, \\ F_n &= 1F_{n-1} + F_{n-2}, \\ &\vdots \\ F_4 &= 1F_3 + F_2 \\ F_3 &= 2F_2 + 0. \end{aligned}$$

The algorithm terminates and we can read from it that $\gcd(F_{n+1}, F_{n+2}) = F_2 = 1$.

## 7 Prime and Irreducible Elements

**Reminder:** An integer $p \in \mathbb{Z}$ is prime if it has two distinct positive divisors, namely 1 and $p$ itself (the fact that we declare these to be distinct excludes calling the number 1 a prime number, which is the normal thing to do). An important property of a prime $p$ is this:

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

**Definition 7.1** Let $R$ be an integral domain and $a \in R$.

(a) We call $a \in R$ prime if these hold:
   (i) Both $a \neq 0_R$ and $a \notin U(R)$.
   (ii) For all $b, c \in R$, we have $a \mid bc$ implies either $a \mid b$ or $a \mid c$.

(b) We call $a \in R$ irreducible if these hold:
   (i) Both $a \neq 0_R$ and $a \notin U(R)$.
   (ii) If $a = bc$ for some $b, c \in R$, then $b \in U(R)$ or $c \in U(R)$.

**Note:** Any associate of a prime/irreducible element is itself a prime/irreducible element.

**Proposition 7.3** *Let $R$ be an integral domain. Then, any prime element is irreducible.*

*Proof*: Let $a \in R$ be prime. By Definition 7.1(a)(i), we know that $a \neq 0_R$ and that $a$ is **not** a unit. This automatically satisfies Definition 7.1(b)(i), so it remains to show Definition 7.1(b)(ii). Indeed, let $a = bc$ for some $b, c \in R$. This clearly tells us that $a \mid bc$. By Definition 7.1(a)(ii), we know therefore that either $a \mid b$ or $a \mid c$.

- If $a \mid b$, then $b = da$ for some $d \in R$. Therefore, $a = bc = dac = adc$, the last equality coming from Axiom (ID3) which says $R$ is commutative. We can re-write this equation as $a(1_R - dc) = 0_R$. We already know that $a \neq 0_R$, so it must follow that $1_R - dc = 0_R$ because Axiom (ID4) tells us $R$ has **no** non-zero zero divisors. But this equation is the same as $dc = 1_R$ so $c$ is a unit.

- If $a \mid c$, a near-identical argument works to imply that $b$ is a unit. □

**Note:** The converse is **not** true in general, but it is in some broad cases; see Theorem 7.4.

**Theorem 7.4** *Let $R$ be a principal ideal domain. Then, any irreducible element is prime.*

*Proof*: Let $a \in R$ be irreducible. By Definition 7.1(b)(i), we know that $a \neq 0_R$ and that $a$ is **not** a unit. This automatically satisfies Definition 7.1(a)(i), so it remains to show Definition 7.1(a)(ii). Indeed, let $a \mid bc$ for some $b, c \in R$. Per Theorem 6.16, there exists a greatest common divisor, $d$ say, of $a$ and $b$. We know therefore that $d \mid a$ and $d \mid b$. In particular, $a = ed$ for some $e \in R$. By Definition 7.1(b)(ii), we know $e$ is a unit or $d$ is a unit.

- If $e$ is a unit, then $e^{-1}a = d$ which tells us that $a \mid d$. But because $d \mid b$, transitivity of division implies $a \mid b$.
- If $d$ is a unit, then $d$ is associate to $1_R$. Consequently, Remark 6.13 implies $\gcd(a, b) = 1_R$. We can use a result from Question Sheet 4 to conclude straight away that $a \mid c$. □

**Corollary 7.5** *Let $R$ be a principal ideal domain. Then, primes and irreducibles coincide.*

*Proof*: This is a direct application of Proposition 7.3 and Theorem 7.4. □

The goal is that we want to write any element as a product of irreducibles; this mimics how any integer can be written as a product of prime numbers. The idea is that any $r \in R$ is either irreducible (so we are done) or it is not, and we can factorise it; we then repeat this with the factors until the process terminates. Said process does terminate if $R$ is a PID, but also if it is another class of rings that we next introduce.

**Definition 7.7** Let $R$ be an integral domain. It is a unique factorisation domain (UFD) if it satisfies the following, where $a \in R \setminus \{0_R\}$ is **not** a unit:
(UFD1) We can write $a = p_1 \cdots p_n$ where each $p_i \in R$ is irreducible.
(UFD2) If $a = p_1 \cdots p_n = q_1 \cdots q_m$ where the $p_i, q_j \in R$ are irreducible, then $n = m$ and $p_i$ is associate with $q_i$ (after reordering if necessary).

**Theorem 7.8** *Every principal ideal domain is a unique factorisation domain.*

*Proof*: Omitted. □

**Note:** The converse to Theorem 7.8 is not true, e.g. the ring $\mathbb{Q}[x, y]$ of polynomials in two indeterminates with rational coefficients is a unique factorisation domain but is **not** a principal ideal domain since the ideal $(x, y)$ cannot be generated by a single element.

**Corollary** *Every Euclidean domain is a unique factorisation domain.*

*Proof*: This is an immediate consequence of the fact that every Euclidean domain is a principal ideal domain (Theorem 5.8) in conjunction with Theorem 7.8. □

**Definition 7.10** Let $d \in \mathbb{Z}$. The ring of square root-adjoined integers is a ring on the set $\mathbb{Z}[\sqrt{d}] := \{a + b\sqrt{d} : a, b \in \mathbb{Z}\}$, where $\sqrt{d}$ is as usual for $d \geq 0$ and $\sqrt{d} = i\sqrt{-d}$ for $d < 0$.

**Lemma 7.11** *For $d \in \mathbb{Z}$, the ring $\mathbb{Z}[\sqrt{d}]$ is an integral domain.*

*Sketch of Proof*: One can show $\mathbb{Z}[\sqrt{d}] \subseteq \mathbb{C}$ is a subring in a similar way as for the Gaussian integers in Question Sheet 1. Showing the integral domain axioms is similar to Lemma 5.13. □

**Definition 7.12** A non-zero $d \in \mathbb{Z} \setminus \{0\}$ is called square-free if it has **no** repeated prime factors, that is $a^2 \mid d$ for some $a \in \mathbb{Z}$ implies that $a^2 = 1$.

**Lemma 7.13** *If $d \in \mathbb{Z} \setminus \{0, 1\}$ is square-free, then the square root $\sqrt{d} \notin \mathbb{Q}$.*

*Proof*: If $d < 0$, then $\sqrt{d} = i\sqrt{-d} \notin \mathbb{Q}$ because it isn't even in the real numbers. It remains to consider $d > 1$. Assume to the contrary that $\sqrt{d} \in \mathbb{Q}$, so there exist integers $a, b \in \mathbb{Z}$ with $b \neq 0$ such that $\sqrt{d} = a/b$. Without loss of generality, suppose $\gcd(a, b) = 1$. This equation implies that $a^2 = db^2$. If $p$ is a prime factor of $a$, then $p^2 \mid a^2$, which implies that $p^2 \mid db^2$. Because $\gcd(p, b) = 1$, it follows that $\gcd(p^2, b^2) = 1$ also. Therefore, $p^2 \mid d$ by a result from Question Sheet 4, but this contradicts the fact that $d$ is square-free. Therefore, $p$ is **not** a prime factor of $a$, so $a = \pm 1$. However, because $d \mid a$, this means that $d = \pm 1$, which is again a contradiction. □

**Corollary 7.14** *If $d \in \mathbb{Z} \setminus \{0, 1\}$ is square-free, then $a + b\sqrt{d} \in \mathbb{Z}[\sqrt{d}]$ is zero if and only if $a = b = 0$.*

*Proof*: If $a = b = 0$, $a + b\sqrt{d} = 0$. Conversely, let $a + b\sqrt{d} = 0$. If $b \neq 0$, then $\sqrt{d} = -a/b \in \mathbb{Q}$, contradicting Lemma 7.13. Hence, $b = 0$ and $a = 0$ follows immediately. □

**Definition 7.15** Let $d \in \mathbb{Z} \setminus \{0, 1\}$. The norm on $\mathbb{Z}[\sqrt{d}]$ is the map $N : \mathbb{Z}[\sqrt{d}] \to \mathbb{N}$ where

$$N(a + b\sqrt{d}) = \left|a^2 - db^2\right|.$$

**Note:** In other words, then above norm $N$ is such that $a + b\sqrt{d} \mapsto (a + b\sqrt{d})(a - b\sqrt{d})$.

**Lemma** *The norm $N$ in* Definition 7.15 *is well-defined.*

*Proof*: Suppose that $a + b\sqrt{d} = s + t\sqrt{d}$. Then, we see that $(a - s) + (b - t)\sqrt{d} = 0$. By Corollary 7.14, this is true if and only if $a - s = 0$ and $b - t = 0$; this means that $a = s$ and $b = t$. In particular, $N(a + b\sqrt{d}) = N(s + t\sqrt{d})$ which is to say that $N$ is well-defined. □

**Lemma 7.16** *Let $d \in \mathbb{Z} \setminus \{0, 1\}$ be square-free. The norm satisfies the following:*

(i) *For $x, y \in \mathbb{Z}[\sqrt{d}] \setminus \{0\}$, we have $N(xy) = N(x)N(y)$.*

(ii) *For $x \in \mathbb{Z}[\sqrt{d}]$, $x$ is a unit if and only if $N(x) = 1$.*

*Proof*: (i) Let $x = a + b\sqrt{d}$ and $y = s + t\sqrt{d}$ be non-zero where $a, b, s, t \in \mathbb{Z}$. Then,

$$\begin{aligned} N(xy) &= N\left((a + b\sqrt{d})(s + t\sqrt{d})\right) \\ &= N\left(as + btd + (at + bs)\sqrt{d}\right) \\ &= \left|(as + btd)^2 - d(at + bs)^2\right| \\ &= \left|a^2s^2 + 2asbtd + b^2t^2d^2 - da^2t^2 - 2asbtd - db^2s^2\right| \\ &= \left|a^2s^2 + b^2t^2d^2 - da^2t^2 - db^2s^2\right| \\ &= \left|(a^2 - db^2)(s^2 - dt^2)\right| \\ &= N(x)N(y). \end{aligned}$$

(ii) Let $x \in \mathbb{Z}[\sqrt{d}]$ be a unit. Then, there exists $y \in \mathbb{Z}[\sqrt{d}]$ such that $xy = yx = 1$. Clearly we have that $x \neq 0$ and $y \neq 0$. Thus, we conclude from (i) above that $1 = N(xy) = N(x)N(y)$. Because $N(x)$ and $N(y)$ are non-negative integers, it must be that $N(x) = 1$ and $N(y) = 1$. Conversely, suppose that $x = a + b\sqrt{d} \neq 0$ and that $N(x) = 1$. This means that

$$(a + b\sqrt{d})(a - b\sqrt{d}) = a^2 - db^2 = \pm 1,$$

from which we conclude that either $a - b\sqrt{d}$ or $-(a - b\sqrt{d})$ is an inverse for $x = a + b\sqrt{d}$. This is equivalent to saying that $x \in U(\mathbb{Z}[\sqrt{d}])$. □

**Lemma 7.18** *Let $d \in \mathbb{Z} \setminus \{0, 1\}$ be square-free and $x \in \mathbb{Z}[\sqrt{d}] \setminus \{0\}$ such that $N(x)$ is prime. Then, $x$ is an irreducible element of $\mathbb{Z}[\sqrt{d}]$.*

*Proof*: We know that $x \neq 0$ and since $N(x) \neq 1$, we know that $x$ is not a unit via Lemma 7.16(ii). Suppose $x = yz$ where $y, z \in \mathbb{Z}[\sqrt{d}]$. Taking norms tells us that $N(x) = N(yz) = N(y)N(z)$. However, we are assuming that $N(x)$ is prime so one of $N(y) = 1$ and $N(z) = 1$ is true. Thus, either $y$ is a unit or $z$ is a unit. Hence, the irreducibility conditions are satisfied. □

**Theorem 7.21** *Let $R$ be a principal ideal domain and $p \in R$ be irreducible. Then, the quotient ring $R/(p)$ is a field.*

*Proof*: This amounts to showing the field axioms from Definition 4.2.

- Because $R$ is a principal ideal domain, it has a one; Theorem 2.13 tells us that the quotient ring also has a one, namely $1_{R/(p)} = 1_R + (p)$.

- If $1_{R/(p)} = 0_{R/(p)}$, then we would have $1_R + (p) = 0_R + (p)$, which implies that $1_R \in (p)$ by Lemma 2.8. Hence, $1_R = rp$ for some $r \in R$, so $p$ is a unit; this is a contradiction. Therefore, we must have $1_{R/(p)} \neq 0_{R/(p)}$.
- Because $R$ is commutative, so too is $R/(p)$.
- Let $r \in R$ and suppose that $r + (p) \neq 0_{R/(p)}$. The aim is to show that this is a unit. Well, Lemma 2.8 again applies to reveal that $r \notin (p)$. Since $r \neq 0_R$, it follows from Theorem 6.16 that $d := \gcd(r, p)$ exists. In particular, $d \mid p$ which means $p = cd$ for some $c \in R$. Because $p$ is assumed irreducible, it must be that $c$ is a unit or $d$ is a unit. Note that $c$ being a unit means $c^{-1}p = d$, so $p \mid d$. But because $d \mid r$, transitivity implies that $p \mid r$, which contradicts $r \notin (p)$. Thus, $c$ is **not** a unit but $d$ **is** a unit. By Theorem 6.16, specifically Bézout's Lemma, we can find $s, t \in R$ such that $d = sr + tp$. This implies that $1_R = d^{-1}sr + d^{-1}tp$. Consequently, $1 - d^{-1}sr \in (p)$ and we again use Lemma 2.8 to conclude that $1_R + (p) = d^{-1}sr + (p)$; this final equation can be re-written as $1_{R/(p)} = (d^{-1}s + (p))(r + (p))$ using coset multiplication. But this tells us that $r + (p)$ has an inverse, so it is a unit. □

**Note:** We have this chain of class inclusions for the different types of rings we encountered:

$$\text{rings} \supseteq \text{rings } \textbf{with} \text{ multiplicative identity} \supseteq \text{commutative rings} \supseteq \text{integral domains} \supseteq \text{unique factorisation domains} \supseteq \text{principal ideal domains} \supseteq \text{Euclidean domains} \supseteq \text{fields.}$$

## 8 Irreducible Polynomials

**Definition** Let $R$ be a ring. We call $f \in R[x]$ a constant polynomial if $\deg(f) = 0$.

**Lemma 8.1** *Let $R$ be an integral domain. Then, $U(R[x]) = U(R)$.*

*Proof*: We show both inclusions. Indeed, if $f \in U(R)$, then we regard it as a constant polynomial (a polynomial of degree zero). Because $f$ is a unit of $R$, there exists an inverse $g \in R$ which is also a constant polynomial. Therefore, $f \in U(R[x])$; this shows $U(R) \subseteq U(R[x])$. Conversely, if $f \in U(R[x])$, then there exists $g \in R[x]$ such that $fg = 1_{R[x]} = 1_R$ (regarded as the constant polynomial). In particular, we know that $f$ and $g$ are non-zero. Remark 4.15 readily implies that $\deg(fg) = \deg(f) + \deg(g) = \deg(1_R) = 0$. Because the degree is non-negative, it must be that $\deg(f) = \deg(g) = 0$, so they are both constant polynomials. In particular, $f \in U(R)$; this shows $U(R[x]) \subseteq U(R)$. $\square$

**Note:** If $R$ is **not** an integral domain, Lemma 8.1 can fail, e.g. $U(\mathbb{Z}_4) \not\ni 1 + 2x \in U(\mathbb{Z}_4[x])$.

**Lemma 8.3** *An element $f \in \mathbb{Z}[x] \setminus \{0\}$ is irreducible in $\mathbb{Z}[x]$ if and only if*
- (i) *$f \neq \pm 1$; and*
- (ii) *$f = gh$ where $g, h \in \mathbb{Z} \setminus \{0\}$ implies that $g = \pm 1$ or $h = \pm 1$.*

*Proof*: Clear from Definition 7.1(b) and Lemma 8.1, which says $U(\mathbb{Z}[x]) = U(\mathbb{Z}) = \{\pm 1\}$. $\square$

**Lemma 8.4** *For $\mathbb{K}$ a field, an element $f \in \mathbb{K}[x] \setminus \{0_{\mathbb{K}}\}$ is irreducible in $\mathbb{K}[x]$ if and only if*
- (i) *$f$ is **not** a constant polynomial; and*
- (ii) *$f = gh$ where $g, h \in \mathbb{K}[x] \setminus \{0_{\mathbb{K}}\}$ implies that $g \in \mathbb{K} \setminus \{0_{\mathbb{K}}\}$ or $h \in \mathbb{K} \setminus \{0_{\mathbb{K}}\}$.*

*Proof*: Clear from Definition 7.1(b) and Lemma 8.1, which says $U(\mathbb{K}[x]) = U(\mathbb{K}) = \mathbb{K} \setminus \{0_{\mathbb{K}}\}$. $\square$

**Note:** Condition (ii) in Lemma 8.4 can be altered to the following similar statement:
- (ii) $f = gh$ where $g, h \in \mathbb{K}[x] \setminus \{0_{\mathbb{K}}\}$ implies that $g \in \mathbb{K}$ or $h \in \mathbb{K}$.

This is because we assume that $f$ is non-zero, so automatically $g$ and $h$ must be non-zero.

**Lemma 8.5** *Let $\mathbb{K}$ be a field. Any degree one polynomial in $\mathbb{K}[x]$ is irreducible.*

*Proof*: Let $f \in \mathbb{K}[x]$ have degree one; so $f$ is non-constant. Assume $f = gh$ for some non-zero $g, h \in \mathbb{K}[x]$. Then, Remark 4.15 tells us $\deg(g) + \deg(h) = \deg(f) = 1$. Hence, either $\deg(g) = 0$ or $\deg(h) = 0$, which is to say $g \in \mathbb{K} \setminus \{0_{\mathbb{K}}\}$ or $h \in \mathbb{K} \setminus \{0_{\mathbb{K}}\}$; this demonstrates irreducibility. $\square$

**Note:** Recall Corollary 5.12 says $\mathbb{K}[x]$ is a principal ideal domain, so Theorem 7.8 implies any non-constant polynomial in $\mathbb{K}[x]$ can be uniquely written as a product of irreducible polynomials, up to reordering and multiplication by non-zero scalars (i.e. the units).

**Reminder:** A root of a polynomial $f \in \mathbb{K}[x]$ is an element $a \in \mathbb{K}$ such that $f(a) = 0$.

**Lemma 8.6** *Let $\mathbb{K}$ be a field and $f \in \mathbb{K}[x]$. Then, $a \in \mathbb{K}$ is a root if and only if $(x-a) \mid f$.*

*Proof*: ($\Rightarrow$) Assume $f(a) = 0$, i.e. $a$ is a root of $f$. Then, the Division Algorithm for $\mathbb{K}[x]$ (Proposition 5.10) allows us to write $f = q(x-a)+r$, where (i) $r = 0_{\mathbb{K}}$ or (ii) $\deg(r) < \deg(x-a)$, but $\deg(x-a) = 1$ so this forces $\deg(r) = 0$. Either way, we see that $r$ is a constant polynomial. Because $f(a) = 0$, this necessarily means that $r = 0_{\mathbb{K}}$ and so $(x-a) \mid f$.

($\Leftarrow$) Assume $(x-a) \mid f$. Then, $f = (x-a)g$ for some $f \in \mathbb{K}[x]$. But clearly $f(a) = 0$. ☐

**Corollary 8.7** *Let $\mathbb{K}$ be a field. Any polynomial in $\mathbb{K}[x]$ with degree at least two that also has a root in $\mathbb{K}$ is* **not** *irreducible.*

*Proof*: By Lemma 8.6, such a polynomial $f$ has a degree one factor, so $f = gh$ where $\deg(g) = 1$ and $\deg(h) \geq 1$; this means neither $g$ nor $h$ is constant and thus $f$ is not irreducible. ☐

**Method – Non-Irreducibility:** Suppose we have a polynomial $f \in \mathbb{K}[x]$ where $\deg(f) \geq 2$. Then, we can immediately show that it is **not** irreducible by finding a root $a \in \mathbb{K}$.

**Corollary 8.8** *Let $\mathbb{K}$ be a field. Any polynomial in $\mathbb{K}[x]$ with degree two or three that has* **no** *root in $\mathbb{K}$ is irreducible.*

*Proof*: Let $f$ be such a polynomial; in particular, it is non-constant. Furthermore, Lemma 8.6 implies that it has no degree one factor, so any factorisation $f = gh$ must be such that $\deg(g)$ and $\deg(h)$ are not one and sum to either two or three; at least one has to be zero degree. ☐

**Theorem 8.9** (Fundamental Theorem of Algebra) *Any non-constant polynomial $f \in \mathbb{C}[x]$ has a root in $\mathbb{C}$.*

*Proof*: Omitted. ☐

**Proposition 8.10** *Let $\mathbb{K}$ be a field and consider the polynomial ring $\mathbb{K}[x]$.*

(i) *If $\mathbb{K} = \mathbb{C}$, the irreducible polynomials are the linear polynomials.*

(ii) *If $\mathbb{K} = \mathbb{R}$, the irreducible polynomials are the linear polynomials and the quadratic polynomials with* **no** *real roots.*

*Proof*: (i) Lemma 8.5 says precisely that linear polynomials are irreducible. Next, let $f \in \mathbb{C}[x]$ with $\deg(f) > 1$; the Fundamental Theorem of Algebra implies $f$ has a root in $\mathbb{C}$, so Corollary 8.7 tells us $f$ is **not** irreducible. In other words, linear polynomials are the only irreducibles.

(ii) Omitted. □

**Theorem 8.11** (Rational Root Test) *Let $f = a_0 + a_1x + \cdots + a_nx^n \in \mathbb{Z}[x]$. If $a \in \mathbb{Q}$ is a rational root of $f$ of the form $a = p/q$ with $q \neq 0$ and $\gcd(p,q) = 1$, then $p \mid a_0$ and $q \mid a_n$.*

*Proof*: Let $a = p/q \in \mathbb{Q}$ with $q \neq 0$ and $\gcd(p,q) = 1$ (in particular, if $a = 0$, take $p = 0$ and $q = 1$). Because $a$ is a root of $f$, we know that $f(a) = 0$; this can be written fully as

$$a_0 + a_1\left(\frac{p}{q}\right) + a_2\left(\frac{p}{q}\right)^2 + \cdots + a_n\left(\frac{p}{q}\right)^n = 0.$$

Multiplying both sides by $q^n$ tells us that

$$a_0q^n + a_1pq^{n-1} + a_2p^2q^{n-2} + \cdots + a_np^n = 0.$$

In other words, we have

$$a_0q^n = -p(a_1q^{n-1} + a_2pq^{n-2} + \cdots + a_np^{n-1}),$$

so we conclude $p \mid a_0q^n$. Because $\gcd(p,q) = 1$, it follows also that $\gcd(p,q^n) = 1$ and thus $p \mid a_0$. On the other hand, we could rewrite the root equation as

$$a_np^n = -q(a_0q^{n-1} + a_1pq^{n-2} + \cdots + a_{n-1}p^{n-1}),$$

from which we conclude $q \mid a_np^n$. A similar argument to the above means we also have $q \mid a_n$. □

**Method – Rational Root Test:** Suppose we have a polynomial $f$ of degree two or three.

(i) Check that the coefficients of $f$ are integers.

(ii) Let $a = p/q \in \mathbb{Q}$ be a root. Write the possible values of $p$ and $q$ using Theorem 8.11.

(iii) Use Step (ii) to find a list of candidates for $a$.

(iv) Check the values of $f(a)$ for each candidate from Step (iii).

If $f(a) \neq 0$ for each candidate from Step (iii), then Corollary 8.8 tells us $f$ is irreducible.

**Definition 8.13** Let $a_1, ..., a_n \in \mathbb{Z}$ **not all** zero. A greatest common divisor of $a_1, ..., a_n$ is an integer $d \in \mathbb{Z}$ such that the following are satisfied:

(i) $d \mid a_i$ for all $i$ .

(ii) If $c \in \mathbb{Z}$ such that $c \mid a_i$ for all $i$, then $c \mid d$.

**Definition 8.14** A non-zero polynomial $f \in \mathbb{Z}[x]$ is primitive if its coefficients are coprime.

**Lemma 8.15** *Let $f \in \mathbb{Q}[x] \setminus \{0\}$. Then, $f$ can be written uniquely in the form*

$$f = c_f f_0,$$

*where $c_f \in \mathbb{Q}^+$ is a positive rational, the so-called* content of $f$, *and $f_0 \in \mathbb{Z}[x]$ is primitive. Moreover, if $f \in \mathbb{Z}[x]\setminus\{0\}$, then $c_f$ is a positive greatest common divisor of the coefficients.*

*Proof*: Let $b \in \mathbb{Z}^+$ such that $bf \in \mathbb{Z}[x]$; one way to choose $b$ is to take the absolute value of the product of the denominators of the coefficients of $f$. Let $a$ be a positive greatest common divisor of the coefficients of $bf$ and let $f_0$ be the element of $\mathbb{Q}[x]$ satisfying $bf = af_0$. Such an $f_0$ is primitive. Then, we see that

$$f = \frac{a}{b} f_0 \qquad \Rightarrow \qquad c_f = \frac{a}{b}.$$

If $f \in \mathbb{Z}[x]$, we take $b = 1$ and this means $c_f = a$ as required. Uniqueness is omitted. $\square$

**Reminder:** The map $\varphi_n : \mathbb{Z} \to \mathbb{Z}_n$ given by $\varphi_n(a) = a \pmod{n}$ is a ring homomorphism.

**Definition** We can extend $\varphi_n$ from above by defining the map $\psi_n : \mathbb{Z}[x] \to \mathbb{Z}_n[x]$ as follows:

$$\psi_n(a_0 + a_1 x + \cdots + a_k x^k) = \varphi_n(a_0) + \varphi_n(a_1)x + \cdots + \varphi_n(a_k)x^k,$$

that is we apply the map $\varphi_n$ to the coefficients of the polynomial we input into $\psi_n$.

**Lemma 8.17** *The map $\psi_n : \mathbb{Z}[x] \to \mathbb{Z}_n[x]$ from above is a ring homomorphism.*

*Sketch of Proof*: We must show the axioms from Definition 3.1. To this end, let $f, g \in \mathbb{Z}[x]$ be given by $f = a_0 + a_1 x + \cdots + a_k x^k$ and $g = b_0 + b_1 x + \cdots + b_m x^m$. Without loss of generality, let $n \leq m$. Therefore, we see that

$$\begin{aligned}
\psi_n(f+g) &= \psi_n(a_0 + a_1 x + \cdots + a_k x^k + b_0 + b_1 x + \cdots + b_m x^m) \\
&= \psi_n\Big((a_0 + b_0) + \cdots + (a_n + b_n)x^n + b_{n+1}x^{n+1} + \cdots + b_m x^m\Big) \\
&= \varphi_n(a_0 + b_0) + \cdots + \varphi_n(a_n + b_n)x^n + \varphi_n(b_{n+1})x^{n+1} + \cdots + \varphi_n(b_m)x^m \\
&= \varphi_n(a_0) + \varphi_n(b_0) + \cdots + \varphi_n(a_n)x^n + \varphi_n(b_n)x^n + \varphi_n(b_{n+1})x^{n+1} + \cdots + \varphi_n(b_m)x^m \\
&= \big(\varphi_n(a_0) + \cdots + \varphi_n(a_n)x^n\big) + \big(\varphi_n(b_0) + \cdots + \varphi_n(b_m)x^m\big) \\
&= \psi_n(f) + \psi_n(g).
\end{aligned}$$

Similarly, we can show $\psi_n(fg) = \psi_n(f)\psi_n(g)$; this again relies on the fact that $\phi_n$ is itself a ring homomorphism (which we used in the fourth equality above). $\square$

**Lemma 8.18** (Gauss' Lemma) *Let $f, g \in \mathbb{Z}[x]$ be primitive. Then, $fg \in \mathbb{Z}[x]$ is primitive.*

*Proof*: Suppose to the contrary that $f$ and $g$ are primitive but that $fg$ is **not**. Then, the positive greatest common divisor of the coefficients of $fg$ is more than one (if it was one, they are all coprime and it is primitive). Hence, there is a prime number $p$ which divides every coefficient of $fg$. Therefore, $\psi_p(f)\psi_p(g) = \psi_p(fg) = 0 \in \mathbb{Z}_p[x]$, using Lemma 8.17 to get the left-hand equality. But $\mathbb{Z}_p$ is a field by Theorem 7.21, so it is an integral domain by Proposition 4.12. But Proposition 4.14 implies that $\mathbb{Z}_p[x]$ is therefore also an integral domain. As there are no non-zero zero divisors, we have $\psi_p(f) = 0$ or $\psi_p(g) = 0$. We now consider these (identical) cases below:

- If $\psi_p(f) = 0$, then $p$ divides every coefficient of $f$, contradicting $f$ being primitive.
- If $\psi_p(g) = 0$, then $p$ divides every coefficient of $g$, contradicting $g$ being primitive.

Either way, we achieve a contradiction; it must be that $fg$ is primitive. □

**Corollary 8.19** *Let $f, g \in \mathbb{Z}[x]\backslash\{0\}$. In* Lemma 8.15 *notation, $c_{fg} = c_f c_g$ and $(fg)_0 = f_0 g_0$.*

*Proof*: Let $f, g \in \mathbb{Z}[x]$; we can write $f = c_f f_0$ and $g = c_g g_0$ and $fg = c_{fg}(fg)_0$ via Lemma 8.15, where the polynomials $f_0, g_0, (fg)_0 \in \mathbb{Z}[x]$ are primitive and the contents $c_f, c_g, c_{fg} \in \mathbb{Q}^+$ are positive rationals. But we can also write the product as

$$fg = c_f c_g f_0 g_0.$$

We know from Gauss' Lemma that $f_0 g_0$ is primitive. But Lemma 8.15 also tells us that the expressions are unique, so we must have that $c_{fg} = c_f c_g$ and $(fg)_0 = f_0 g_0$. □

**Theorem 8.20** (Gauss' Theorem) *Let $f \in \mathbb{Z}[x] \setminus \mathbb{Z}$ be a non-constant polynomial. If $f$ is* **not** *a product of two non-constant polynomials in $\mathbb{Z}[x]$, then $f$ is irreducible in $\mathbb{Q}[x]$.*

*Proof*: Suppose $f = gh$ where $g, h \in \mathbb{Q}[x] \setminus \{0\}$. By Lemma 8.15, we can write the following:

$$\begin{aligned} f &= c_f f_0, && \text{with } c_f \in \mathbb{Q}^+ \text{ and } f_0 \text{ primitive,} \\ g &= c_g g_0, && \text{with } c_g \in \mathbb{Q}^+ \text{ and } g_0 \text{ primitive,} \\ h &= c_h h_0, && \text{with } c_h \in \mathbb{Q}^+ \text{ and } h_0 \text{ primitive.} \end{aligned}$$

Because $f = gh$, we must have that $c_f = c_g c_h$ and $f_0 = g_0 h_0$ by the uniqueness part of Lemma 8.15. Consequently, we see that $f = c_f f_0 = c_f g_0 h_0$. Now, $f \in \mathbb{Z}[x]$ which means that $c_f \in \mathbb{Z}$ by Corollary 8.19. Therefore, $c_f g_0 \in \mathbb{Z}[x]$ and $h_0 \in \mathbb{Z}[x]$. But $f$ is **not** a product of non-constant polynomials by assumption, so either $c_f g_0$ (and therefore $g$) is constant or $h_0$ (and therefore $h$) is constant. □

**Method – Irreducibility via Gauss' Theorem:** Let $f \in \mathbb{Q}[x]$ be some polynomial.

(i) Check that $f \in \mathbb{Q}[x] \setminus \mathbb{Z}$.

(ii) Apply the Rational Root Test and Lemma 8.6 to conclude that $f$ has **no** linear factors in $\mathbb{Q}[x]$, and hence $\mathbb{Z}[x]$.

(iii) Writing $f$ as a product of integer polynomials of degree at least two, show that this is not possible by expanding and comparing coefficients.

(iv) Use Gauss' Theorem to conclude that $f$ is irreducible.

**Theorem 8.22** (Eisenstein's Irreducibility Criterion) *Let $f = a_0 + a_1 x + \cdots + a_n x^n \in \mathbb{Z}[x] \setminus \mathbb{Z}$ be non-constant and assume there exists a prime $p \in \mathbb{Z}$ satisfying the following:*

(i) $p \mid a_0, \ldots, p \mid a_{n-1}$.

(ii) $p \nmid a_n$.

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

*Then, $f$ is irreducible in $\mathbb{Q}[x]$.*

*Proof*: Suppose $f = gh$ is a product of non-constant polynomials $g, h \in \mathbb{Z}[x] \setminus \mathbb{Z}$ of this form:

$$f = (b_0 + b_1 x + \cdots + b_r x^r)(c_0 + c_1 x + \cdots + c_s x^s),$$

where $b_r \neq 0$ and $c_s \neq 0$. Then, $r + s = n$ for $0 < r < n$ and $0 < s < n$ by comparing degrees. Moreover, we have that the constant part $a_0 = b_0 c_0$. We also have that $a_n = b_r c_s$. Because $p \mid a_0$ and $p^2 \nmid a_0$, there are two cases to consider.

- Assume $p \mid b_0$ but $p \nmid c_0$. Because we also assume $p \nmid a_n$, it must be that $p \nmid b_r$ and $p \nmid c_s$. Suppose that $b_m$ is the first coefficient in $g$ such that $p \nmid b_m$. Notice that we can write

$$a_m = b_0 c_m + b_1 c_{m-1} + \cdots + b_m c_0,$$

  with $c_i = 0$ for all $i > s$. Then, $p$ divides every term in this sum **except** the last one (because this is the case where $p \nmid c_0$). Hence, it follows that $p \nmid a_m$. By (i) in the statement of the theorem, this means $m = n$; this is a contradiction as $m \leq r < n$, so we end up with $n < n$.

- Assume $p \nmid b_0$ but $p \mid c_0$. A near-identical argument will also yield a contradiction

Either way, we have a contradiction so $f$ is **not** the product of two non-constant polynomials with integer coefficients. Thus, Gauss' Theorem tells us $f$ is irreducible in $\mathbb{Q}[x]$. □

**Method – Irreducibility via Eisenstein's Criterion:** Let $f \in \mathbb{Q}[x]$ be some polynomial.

(i) Check that $f \in \mathbb{Z}[x] \setminus \mathbb{Z}$.

(ii) Find a prime $p \in \mathbb{Z}$ satisfying the conditions of Theorem 8.22.