A **Ring** (with 1) is a set \( R \) with two binary operations \(+\) and \( \times \) such that

- **R1.** \( (R, +) \) is an Abelian group under \( + \).
- **R2.** \( (R, \times) \) is a Monoid under \( \times \), (so \( \times \) is associative and has an identity 1).
- **R3.** The distributive laws hold: \( a(b + c) = ab + ac, \ (b + c)a = ba + ca \).

Many of the standard facts from algebra follow from these axioms. In particular, \( 0a = a0 = 0, \ a(−b) = (−1)a, \ (\sum a_i)(\sum b_j) = \sum a_i b_j \).

The ring \( R \) is **commutative** if \( \times \) is commutative.

An element of \( R \) is a **unit** if it has a (2-sided) multiplicative inverse.

The set of units \( R \times \) (or \( U(R) \)) is a group under \( \times \).

**Examples**

1. \( \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C} \) are all rings under the usual \(+\) and \( \times \). \( \mathbb{Q}, \mathbb{R}, \mathbb{C} \) are fields. \( \mathbb{Z} \) is an ID.
2. \( \mathbb{Z}/n\mathbb{Z} \) is a ring under \(+\) and \( \times \) mod \( n \). This ring is an ID iff \( n \) is prime. In fact, if \( n \) is prime then \( \mathbb{Z}/n\mathbb{Z} \) is a field.
3. If \( R \) is a ring then the set \( M_n(R) \) of \( n \times n \) matrices with entries in \( R \) is a ring under matrix addition and multiplication. \( M_n(R) \) is non-commutative in general.
4. Let \( (A, +) \) be an abelian group and let \( \text{End}(A) \) be the set of group homomorphisms \( A \to A \). Define addition pointwise, \( (f + g)(a) = f(a) + g(a) \), and multiplication by composition, \( fg(a) = f(g(a)) \). Then \( \text{End}(A) \) is a (usually non-commutative) ring.
5. If \( A = \prod_{i \in \mathbb{N}} \mathbb{Z} = \{(a_0, a_1, \ldots) : a_i \in \mathbb{Z} \} \) then the maps \( R((a_0, \ldots)) = (0, a_0, a_1, \ldots) \) and \( L((a_0, a_1, \ldots)) = (a_1, a_2, \ldots) \) lie in \( \text{End}(A) \) and \( LR = 1 \neq RL \). Hence \( R \) has a left, but not a right inverse. [Recall that left and right inverses must be equal if they both exist.]
6. Let \( C[0, 1] \) be the set of continuous functions from \( [0, 1] \) to \( \mathbb{R} \) with addition and multiplication defined pointwise. Then \( C[0, 1] \) is a ring. It is not an ID (why?).

A subset \( S \) of \( R \) is a **subring** iff \( (S, +) \) is a subgroup of \( (R, +) \) and \( (S, \times) \) is a submonoid of \( (R, \times) \). Equivalently, \( 1_R \in S \) and \( a, b \in S \) implies \( a - b, ab \in S \).

A subset \( I \) of \( R \) is a **left ideal** iff \( (I, +) \) is a subgroup of \( (I, +) \) such that for all \( r \in R, \ a \in I \), we have \( ra \in I \). A subset \( I \) of \( R \) is a **right ideal** iff \( (I, +) \) is a subgroup of \( (I, +) \)
such that for all \( r \in R, a \in I \), we have \( ar \in I \). An **ideal** is a subset that is both a left ideal and a right ideal. Equivalently, \( I \neq \emptyset \) and \( a, b \in I, r \in R \), implies \( a - b, ra, ar \in I \). The sets \( \{0\} \) and \( R \) are ideals of \( R \). An ideal \( I \) is **proper** if \( I \neq R \), and **non-trivial** if \( I \neq \{0\} \).

**Examples**

1. \( n\mathbb{Z} \) is an ideal of \( \mathbb{Z} \) but not a subring (unless \( n = \pm 1 \)).
2. \( \mathbb{Z} \) is a subring of \( \mathbb{R} \) but not an ideal.
3. The set of the matrices \( I = \left\{ \begin{pmatrix} 0 & b \\ 0 & d \end{pmatrix} \mid b, d \in \mathbb{R} \right\} \) is a left ideal, but not a right ideal of \( M_2(\mathbb{R}) \). But \( I \) is a 2-sided ideal of the subring \( T = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \mid a, b, d \in \mathbb{R} \right\} \) of \( M_2(\mathbb{R}) \).
4. The **quaternions** \( \mathbb{H} = \left\{ \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \mid \alpha, \beta \in \mathbb{C} \right\} \) form a subring of \( M_2(\mathbb{C}) \). Any \( x \in \mathbb{H} \) can be written uniquely as \( x = a + bi + cj + dk \) where \( i = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, j = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, k = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \). Then \( i^2 = j^2 = k^2 = -1, ij = k, ji = -k, \) and \( (a + bi + cj + dk)^{-1} = (a/r) - (b/r)i - (c/r)j - (d/r)k \) where \( r = a^2 + b^2 + c^2 + d^2 \). Thus \( \mathbb{H} \) is a noncommutative division ring.

**Lemma 1.1** If \( S_\alpha, \alpha \in A \), are subrings of \( R \) then \( \bigcap_{\alpha \in A} S_\alpha \) is a subring of \( R \).

If \( I_\alpha \) are ideals of \( R \) then \( \bigcap_{\alpha \in A} I_\alpha \) is an ideal of \( R \).

The ideal \( (S) \) generated by a subset \( S \subseteq R \) is the smallest ideal of \( R \) containing \( S \). It can be defined as the intersection \( \bigcap_{J \supseteq S} J \) of all ideals containing \( S \).

An ideal \( I \) is **principal** if it is generated by a single element, \( I = (a) \) for some \( a \in R \). An ideal is **finitely generated** if it is generated by a finite set, \( I = (S), |S| < \infty \).

We can also define the subring generated by a subset. More generally, if \( R \) is a subring of \( R' \) and \( S \subseteq R' \), then \( R[S] \) is the smallest subring of \( R' \) containing \( R \) and \( S \) ( = the intersection of all subrings of \( R' \) containing \( R \) and \( S \)).

**Exercises**

1. Show that an ideal is proper iff it does not contain a unit.
2. Show that \( (S) = \{ \sum_{i=1}^n r_is_ir'_i \mid r_i, r'_i \in R, s_i \in S, n \in \mathbb{N} \} \).
3. Show that if \( R \) is commutative then the principal ideal \( (a) \) is \( \{ ra \mid r \in R \} \).
4. Show that \( R[\alpha] \) is the set of all polynomial expressions \( \sum_{i=0}^n a_i \alpha^i \) with coefficients \( a_i \in R \).
5. Deduce that \( \mathbb{Z}[i] = \{ a + bi \mid a, b \in \mathbb{Z} \} \) as a subring of \( \mathbb{C} \) and \( \mathbb{Q}[\sqrt{2}] = \{ a + b\sqrt{2} + c\sqrt{4} \mid a, b, c \in \mathbb{Q} \} \) as a subring of \( \mathbb{R} \).
6. Describe \( \mathbb{Z}[1/2] \) as a subring of \( \mathbb{Q} \).
7. Let \( I \) be the set of continuous functions \( f \in C[0, 1] \) such that \( f(0.5) = 0 \). Show that \( I \) is an ideal of \( C[0, 1] \) that is not principal (or even finitely generated).
A (ring) homomorphism from the ring $R$ to the ring $S$ is a function $f : R \to S$ that is a group homomorphism $(R, +) \to (S, +)$ and a monoid homomorphism $(R, \times) \to (S, \times)$. Equivalently $f(a + b) = f(a) + f(b), f(ab) = f(a)f(b), f(1_R) = 1_S$.

Examples

1. The map $f : \mathbb{T} \to \mathbb{R}$ given by $f(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = a$ where $T = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, b, c, d \in \mathbb{R} \right\}$.

2. If $S$ is a subring of $R$ then the inclusion map $i : S \to R, i(r) = r$, is a homomorphism.

A (ring) isomorphism is a homomorphism $R \to S$ that has a 2-sided inverse map $g : S \to R$ which is also a homomorphism. It is sufficient for $f$ to be a bijective homomorphism.

If $I$ is an ideal of $R$ then the quotient ring $R/I$ is the quotient group $(R/I, +)$ with multiplication defined by $(a + I)(b + I) = ab + I$.

Lemma 2.1 The quotient ring $R/I$ is indeed a ring and the projection map $\pi : R \to R/I$ given by $\pi(a) = a + I$ is a surjective ring homomorphism.

Example $R = \mathbb{Z}$, $I = (n)$, then $R/I = \mathbb{Z}/n\mathbb{Z}$ is the integers mod $n$ with addition and multiplication mod $n$.

Theorem (1st Isomorphism Theorem) If $f : R \to S$ then $\text{Ker} \, f = \{ r \mid f(r) = 0 \}$ is an ideal of $R$, $\text{Im} \, f = \{ f(r) \mid r \in R \}$ is a subring of $S$ and $f = i \circ \tilde{f} \circ \pi$ where

- $\pi : R \to R/\text{Ker} \, f$ is the (surjective) projection homomorphism.
- $\tilde{f} : R/\text{Ker} \, f \to \text{Im} \, f$ is a (bijective) ring isomorphism.
- $i : \text{Im} \, f \to S$ is the (injective) inclusion homomorphism.

Theorem (2nd Isomorphism Theorem) If $I$ is an ideal of $R$ then there is a bijection

> $\{\text{subgroups } H \text{ of } (R, +) \text{ with } I \leq H \leq R\} \leftrightarrow \{\text{subgroups of } (R/I, +)\}$,

where $H$ corresponds to $H/I$. In this correspondence subrings correspond to subrings and ideals correspond to ideals. Moreover, if $J$ is an ideal with $I \leq J \leq R$ then there is an isomorphism $R/J \cong (R/I)/(J/I)$.

Theorem (3rd Isomorphism Theorem) If $I$ is an ideal of $R$ and $S$ is a subring of $R$ then $S + I$ is a subring of $R$, $S \cap I$ is an ideal of $S$, and $(S + I)/I \cong S/(S \cap I)$.

Example For any ring $R$ define $f : \mathbb{Z} \to R$ by $f(n) = n.1_R$ ($1_R + \cdots + 1_R$ defined as for additive groups). Then $f$ is a ring homomorphism. The kernel is a subgroup of $(\mathbb{Z}, +)$ so is $n\mathbb{Z}$ for some $n \geq 0$. The image $S = \{ n.1_R \mid n \in \mathbb{Z} \}$ is called the prime subring of $R$ and is isomorphic to $\mathbb{Z}/n\mathbb{Z}$. The characteristic of $R$, $\text{char}(R)$, is the integer $n$. E.g., $\text{char}(\mathbb{R}) = 0$, $\text{char}(\mathbb{Z}/n\mathbb{Z}) = n$, $\text{char}(\{0\}) = 1$. 

3
A **maximal ideal** is a proper ideal $M$ of $R$ such that for any ideal $I$, $M \subseteq I \subseteq R$ implies $I = M$ or $I = R$.

**Example**  The ideal $(n)$ is a maximal ideal of $\mathbb{Z}$ iff $n$ is prime.

A non-trivial ring is **simple** if the only ideals of $R$ are $(0)$ and $R$. Equivalently, $(0)$ is maximal.

**Lemma 2.2** Let $R$ be a commutative ring. Then $R$ is simple iff $R$ is a field.

**Proof.** If $R$ is a field and $I \neq (0)$ is an ideal then $u \in I$ for some $u \neq 0$. But $u$ is a unit so $(ru^{-1})u = r \in I$ for all $r \in R$. Thus $I = R$. Conversely, if $a \neq 0$ and $a$ is not a unit then $(a) = \{ra \mid r \in R\}$ is a non-trivial proper ideal of $R$. \hfill $\Box$

Note that if $R$ is a division ring then $R$ is simple. However the converse fails:

**Lemma 2.3** Let $D$ be a division ring. Then $M_n(D)$ is a simple ring for any $n \geq 1$.

**Proof.** Let $I$ be a non-zero ideal of $M_n(D)$ and let $A = (a_{ij}) \in I$, $A \neq 0$. In particular $a_{kl} \neq 0$ for some $k$, $l$. Let $E_{ij}$ be the matrix with 1 in entry $(i,j)$ and zeros elsewhere. Then $E_{ik}AE_{lj} = a_{kl}E_{ij} \in I$. Since $a_{kl} \in D$ and $D$ is a division ring, $a_{kl}^{-1} \in D$, so $a_{kl}^{-1}I \subseteq M_n(D)$. Now $(a_{kl}^{-1}I)(a_{kl}E_{ij}) = E_{ij} \in I$. But any matrix $B = (b_{ij})$ is a linear combination $\sum(b_{ij}I)E_{ij}$, so $B \in I$ and $I = M_n(D)$. \hfill $\Box$

So by the 2nd Isomorphism Theorem, for commutative $R$, $M$ is maximal iff $R/M$ is a field, but for non-commutative $R$, $M$ may be maximal without $R/M$ being a division ring.

**Exercises**

1. Show that any finite ID is a field.

2. An element $a$ of a ring is **nilpotent** if $a^n = 0$ for some $n \in \mathbb{N}$. Show that if $a$ is nilpotent then $1 + a$ is a unit.

3. Show that if $R$ is commutative then the set of nilpotent elements forms an ideal of $R$. [Hint: make sure you check that $a$, $b$ nilpotent implies $a - b$ is nilpotent.]

4. Show that if $r \in R$ lies in the intersection of all maximal ideals of $R$ then $1 + r$ is a unit.

5. Show that any homomorphism $f: F \rightarrow R$ from a field $F$ to a non-trivial ring $R$ is injective, so in particular $R$ contains a subring isomorphic to $F$. 
A partial ordering on a set $\mathcal{X}$ is a relation $\leq$ satisfying the properties:

O1. $\forall x: x \leq x$,
O2. $\forall x, y$: if $x \leq y$ and $y \leq x$ then $x = y$,
O3. $\forall x, y, z$: if $x \leq y$ and $y \leq z$ then $x \leq z$.

A total ordering is a partial ordering which also satisfies:

O4. $\forall x, y$: either $x \leq y$ or $y \leq x$.

**Example** Any collection of sets with $\subseteq$ as the ordering forms a partially ordered set that is not in general totally ordered.

If $(\mathcal{X}, \leq)$ is a partially ordered set, a **chain** in $\mathcal{X}$ is a non-empty subset $\mathcal{C} \subseteq \mathcal{X}$ that is totally ordered by $\leq$.

If $\mathcal{S} \subseteq \mathcal{X}$, and $x \in \mathcal{X}$, we say $x$ is an **upper bound** for $\mathcal{S}$ if $y \leq x$ for all $y \in \mathcal{S}$. [Note that we do not require $x$ to be an element of $\mathcal{S}$.]

A maximal element of $\mathcal{X}$ is an element $x$ such that for any $y \in \mathcal{X}$, $x \leq y$ implies $x = y$. [Note: This does not imply that $y \leq x$ for all $y$ since $\leq$ is only a partial order. In particular there may be many maximal elements.]

**Theorem (Zorn’s Lemma)** If $(\mathcal{X}, \leq)$ is a non-empty partially ordered set for which every chain has an upper bound then $\mathcal{X}$ has a maximal element.

This result follows from (and is equivalent to) the Axiom of choice, which states that if $X_i$ are non-empty sets then $\prod_{i \in I} X_i$ is non-empty. [I will not give the proof here as it is rather long.]

Note: If we had defined things so that $\emptyset$ were a chain, we would not need the condition that $\mathcal{X} \neq \emptyset$ in Zorn’s Lemma since the existence of an upper bound for $\emptyset$ is just the condition that an element of $\mathcal{X}$ exists. However, in practice it is easier to check $\mathcal{X} \neq \emptyset$ and then check separately that each non-empty totally ordered subset has an upper bound.

**Theorem 3.1** If $I$ is a proper ideal of a ring $R$ (with 1) then there exists a maximal ideal $M$ such that $I \subseteq M$.

**Proof.** If an ideal $J$ contains 1 then $J = R$, so an ideal is proper iff it does not contain 1. Let $\mathcal{X}$ be the set of proper ideals $J$ of $R$ with $I \subseteq J$. The partial order on $\mathcal{X}$ will be $\subseteq$. Since $I \in \mathcal{X}$, $\mathcal{X} \neq \emptyset$. Now let $\mathcal{C}$ be a chain in $\mathcal{X}$, i.e., a set of ideals $\{J_\alpha\}$ such that for every $J_\alpha, J_\beta \in \mathcal{C}$ either $J_\alpha \subseteq J_\beta$ or $J_\beta \subseteq J_\alpha$. Let $K = \bigcup_{J_\alpha \in \mathcal{C}} J_\alpha$. We shall show that $K$ is an upper bound for $\mathcal{C}$.
Firstly $C \neq \emptyset$, so some ideal $J_\alpha$ lies in $C$ and $I \subseteq J_\alpha \subseteq K$. In particular $K \neq \emptyset$. If $x, y \in K$ then $x \in J_\alpha$, $y \in J_\beta$, say. Since $C$ is totally ordered, we can assume without loss of generality that $J_\alpha \subseteq J_\beta$. Thus $x, y \in J_\beta$, and $x - y \in J_\beta \subseteq K$. If $x \in K$, $r \in R$, then $x \in J_\alpha$, say, so $xr, rx \in J_\alpha \subseteq K$. Hence $K$ is an ideal with $I \subseteq K$. However $1 \notin J_\alpha$ for each $J_\alpha \in C$, so $1 \notin K$. Hence $K$ is proper. Therefore $K \in \mathcal{X}$ and is clearly an upper bound for $C$.

The conditions of Zorn’s Lemma apply, so $\mathcal{X}$ has a maximal element $M$, say. Now $M$ is a proper ideal containing $I$ and is maximal, since if $M \subset J \subset R$ then $J \in \mathcal{X}$ and $M$ would not be maximal in $\mathcal{X}$. \qed

We now give an example from linear algebra. Let $V$ be a vector space (possibly infinite dimensional).

A set $S \subseteq V$ is called linearly independent if there are no non-trivial finite linear combinations that give 0. In other words if $\sum_{i=1}^{n} \lambda_i s_i = 0$ and the $s_i$ are distinct elements of $S$ then $\lambda_i = 0$ for each $i$.

A set $S \subseteq V$ is called spanning if every element $v \in V$ can be written as a finite linear combinations of elements of $S$, $v = \sum_{i=1}^{n} \lambda_i s_i$.

A set $S \subseteq V$ is called a basis if it is a linearly independent spanning set. Note that every element $v \in V$ can be written as a linear combination of elements of a basis in a unique way. [Spanning implies existence, linear independence implies uniqueness.]

**Theorem 3.2** Every vector space has a basis.

**Proof.** Let $\mathcal{X}$ be the set of all linearly independent sets in $V$ partially ordered by $\subseteq$. Since $\emptyset$ is linearly independent, $\mathcal{X} \neq \emptyset$. Let $\mathcal{C}$ be a chain in $\mathcal{X}$ and let $S = \bigcup_{S_\alpha \in \mathcal{C}} S_\alpha$. We shall show that $S$ is linearly independent.

Suppose $\sum_{i=1}^{n} \lambda_i s_i = 0$ and $s_i \in S_{\alpha_i} \in C$ (the $s_i$ are distinct but the $\alpha_i$ need not be). Then by total ordering of the $S_{\alpha_i}$, there must be one $S_{\alpha_j}$ that contains all the others (use induction on $n$). But then $\sum_{i=1}^{n} \lambda_i s_i = 0$ is a linear relation in $S_{\alpha_j}$ which is linearly independent. Thus $\lambda_i = 0$ for all $i$. Hence $S$ is linearly independent, so $S \in \mathcal{X}$ and is an upper bound for $\mathcal{C}$.

Now apply Zorn’s Lemma to give a maximal linearly independent set $M$. We shall show that $M$ spans $V$ and so is a basis. Clearly any element of $M$ is a linear combination of elements of $M$, so pick any $v \notin M$ and consider $M \cup \{v\}$. By maximality of $M$ this cannot be linearly independent. Hence there is a linear combination $\lambda v + \sum_{i=1}^{n} \lambda_i s_i = 0$, $s_i \in M$, with not all the $\lambda$’s zero. If $\lambda = 0$ this gives a linear relation in $M$, contradicting linear independence of $M$. Hence $\lambda \neq 0$ and $v = \sum_{i=1}^{n} (-\lambda_i / \lambda) s_i$ is a linear combination of elements of $M$. \qed
Anti-isomorphisms

An anti-homomorphism is a map \( f : R \rightarrow S \) such that \( f(a + b) = f(a) + f(b) \), \( f(1) = 1 \), and \( f(ab) = f(b)f(a) \). An anti-isomorphism is an invertible anti-homomorphism.

Examples The transpose map \( T : M_n(\mathbb{R}) \rightarrow M_n(\mathbb{R}) \).
The map \( \mathbb{H} \rightarrow \mathbb{H} \) given by \( f(a + bi + cj + dk) = a - bi - cj - dk \).

The opposite ring \( R^o \) of \( R \) is the ring \( R \) with multiplication defined by \( a \times_{R^o} b = b \times_R a \).

Note that \( R^{oo} = R \).

Lemma 4.1 A map \( f : R \rightarrow S \) is an anti-homomorphism iff it is a homomorphism viewed as a map \( R \rightarrow S^o \) (or \( R^o \rightarrow S \)).

Example \( M_n(\mathbb{R})^o \) is isomorphic to \( M_n(\mathbb{R}) \), one isomorphism being the transpose map \( T \).

Rngs (Rings without 1s)

A Rng (or “ring which does not necessarily have a 1”) is a set \( R \) with + and \( \times \) defined so that \((R,+)\) is an abelian group, \((R,\times)\) is a semigroup (\( \times \) is associative), and the distributive laws hold. However, \( R \) need not contain a multiplicative identity.

Subrings, rng-homomorphisms etc., can be defined without the conditions involving 1. The definition of an ideal is the same, and an ideal is a special case of a subring. The theory of rngs is similar to that of rings, although they are more awkward to deal with later on. The following lemma shows that we can regard a rng as an ideal of a bigger ring.

Lemma 4.2 Let \( R \) be a rng and define \( R_1 = \mathbb{Z} \times R \) with addition \((n,r) + (m,s) = (n+m,r+s)\) and multiplication \((n,r)(m,s) = (nm,n.s+m.r+rs)\), where \( n.s = s + \cdots + s \) etc.. Then \( R_1 \) is a ring containing an ideal \( \{0\} \times R \) isomorphic to \( R \).

Direct sums and the Chinese Remainder Theorem

If \( R_1 \) and \( R_2 \) are rings, define the ring \( R_1 \oplus R_2 \) as the set \( R_1 \times R_2 \) with addition \((a_1,a_2) + (b_1,b_2) = (a_1+b_1,a_2+b_2)\) and multiplication \((a_1,a_2)(b_1,b_2) = (a_1b_1,a_2b_2)\). The identity is \((1,1)\). The direct sum \( R_1 \oplus \cdots \oplus R_n \) is defined similarly. Note that even if \( R_1 \) and \( R_2 \) are IDS, \( R_1 \oplus R_2 \) will not be since \((1,0)(0,1) = (0,0)\).

If \( R \) is a ring and \( I \) and \( J \) are ideals of \( R \), we can define the following ideals.

- \( I + J = \{a + b \mid a \in I, \ b \in J\} \)
- \( I \cap J = \{c \mid c \in I, \ c \in J\} \)
- \( IJ = \{\sum_{i=1}^{n} a_ib_i \mid a_i \in I, \ b_i \in J, \ n \in \mathbb{N}\} \)
It is easily checked that each of these is indeed an ideal. Note that in general $IJ \neq \{ab \mid a \in I, b \in J\}$, but $IJ$ is the ideal generated by all the products $ab, a \in I, b \in J$.

**Example** For $R = \mathbb{Z}$, $I = (x) = \{ax \mid a \in \mathbb{Z}\}$, $J = (y) = \{by \mid b \in \mathbb{Z}\}$

1. $I + J = (\gcd(x, y))$.
   Note $\gcd(x, y) = ax + by$ for some $a, b \in \mathbb{Z}$, so $\gcd(x, y) \in I + J$. Conversely $I + J = \{ax + by \mid a, b \in \mathbb{Z}\}$ and $ax + by$ is always a multiple of $\gcd(x, y)$.

2. $I \cap J = (\lcm(x, y))$.
   $m \in I \iff x \mid m$ and $m \in J \iff y \mid m$. Hence if $m \in I \cap J$ then $m$ must be a common multiple of $x$ and $y$. Thus $m \in (\lcm(x, y))$ Conversely $\lcm(x, y)$ is a common multiple of $x$ and $y$ so lies in $I \cap J$. Hence $I \cap J = (\lcm(x, y))$.

3. $IJ = (xy)$.
   $IJ = \{\sum a_i x b_i y \mid a_i, b_i \in \mathbb{Z}\} \subseteq (xy)$. Conversely $xy \in IJ$, so $(xy) \subseteq IJ$.

Ideals $I$ and $J$ are relatively prime if $I + J = R$. Equivalently $\exists a \in I, b \in J$: $a + b = 1$ (recall that an ideal equals $R$ iff it contains 1).

**Lemma 4.3** $IJ \subseteq I \cap J$. Moreover, if $R$ is commutative and $I + J = R$ then $IJ = I \cap J$.

**Proof** If $a_i \in I$ then $\sum a_ib_i \in I$. If $b_i \in J$ then $\sum a_i b_i \in J$. Hence $IJ \subseteq I \cap J$. Now let $I + J = R$ so that $a + b = 1$ for some $a \in I, b \in J$. Then if $c \in I \cap J, ac + cb \in IJ$. But $ac + cb = c(a + b) = c, so c \in IJ$. Thus $I \cap J \subseteq IJ$ and so $IJ = I \cap J$.

**Theorem (Chinese Remainder Theorem)** If $I$ and $J$ are ideals of a commutative ring $R$ and $I + J = R$ then $R/IJ \cong R/I \oplus R/J$.

**Proof** Let $f: R \to R/I \oplus R/J$ be defined by $f(r) = (r + I, r + J)$. Then $f(r + s) = (r + s + I, r + s + J) = (r + I, r + J) + (s + I, s + J) = f(r) + f(s), f(rs) = (rs + I, rs + J) = (r + I, r + J)(s + I, s + J) = f(r)f(s)$, and $f(1) = (1 + I, 1 + J)$ is the identity in $R/I \oplus R/J$. Now Ker $f = \{r \mid r + I = r + J = 0\} = I \cap J$ so Ker $f = IJ$ by Lemma 4.3. For the image of $f$, write $1 = a + b$ with $a \in I, b \in J$. Then $f(sa + rb) = (sa + r(1-a) + I, s(1-b) + rb + J) = (r + I, s + J)$. Thus $f$ is surjective. Hence $R/IJ \cong R/I \oplus R/J$.

**Example** If $\gcd(n, m) = 1$ then $\mathbb{Z}/nm\mathbb{Z} \cong \mathbb{Z}/n\mathbb{Z} \oplus \mathbb{Z}/m\mathbb{Z}$.

**Exercises**

1. Show that composing two anti-homomorphisms gives a homomorphism and composing an anti-homomorphism with a homomorphism gives an anti-homomorphism.

2. Define $\phi(n) = |(\mathbb{Z}/n\mathbb{Z})^\times|$. Show that if $\gcd(n, m) = 1$ then $\phi(nm) = \phi(n)\phi(m)$. If $n = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$ is the prime factorization of $n$, deduce that $\phi(n) = \prod_i p_i^{\alpha_i-1}(p_i - 1)$.

3. Generalize the CRT: if $I_1, \ldots, I_n$ are ideals of a commutative ring $R$ and for each $i$ and $j, I_i + I_j = R$, show that $R/I_1I_2 \cdots I_n \cong I_1 \oplus I_2 \oplus \cdots \oplus I_n$.
Throughout this section we shall assume $R$ is a commutative ring.

Recall: An Integral Domain (ID) is a non-trivial ring in which $ab = 0$ implies either $a = 0$ or $b = 0$.

A prime ideal of a commutative ring $R$ is a proper ideal such that $ab \in P$ implies either $a \in P$ or $b \in P$.

**Lemma 5.1** An ideal $P$ is prime iff $R/P$ is an ID.

**Proof.** Assume $P$ is prime. Then $R/P$ is non-trivial since $P$ is proper. If $(a+P)(b+P) = 0+P$ then $ab+P = P$ and so $ab \in P$. Thus either $a \in P$ or $b \in P$, so either $a+P = P$ or $b+P = P$. Thus $R/P$ is an ID. Conversely, if $R/P$ is an ID then $P$ is proper since $R/P$ is non-trivial. If $a,b \notin P$, then $a+P,b+P \neq 0+P$, so $(a+P)(b+P) = ab+P \neq 0+P$, so $ab \notin P$. Thus $P$ is a prime ideal. \qed

**Corollary 5.2** Any maximal ideal of a commutative ring is also a prime ideal.

**Proof.** $M$ maximal $\Rightarrow R/M$ is a field $\Rightarrow R/M$ is an ID $\Rightarrow M$ is prime. \qed

The converse does not hold: $(0)$ is prime but not maximal in $\mathbb{Z}$.

Examples of prime ideals: $(p)$ in $\mathbb{Z}$, $(0)$ in any ID. The ideal $(X)$ in the ring $\mathbb{Z}[X]$ of polynomials in $X$ with coefficients in $\mathbb{Z}$. This last example is also not maximal.

Every field is an ID. Furthermore, every subring of a field is an ID (e.g., $\mathbb{Z} \subseteq \mathbb{Q}$). We shall show that conversely, every ID can be embedded as a subring of a field.

Assume $R$ is a commutative ring and $S \subseteq R$ is a submonoid of $(R, \times)$. In other words, $1 \in S$ and $a,b \in S$ implies $ab \in S$. For example, set $S = R \setminus P$ for any prime $P$. One particularly important case is when $R$ is an ID and $S = R \setminus \{0\}$.

**Definition** $S^{-1}R$ is defined as $(R \times S)/\sim$, where $(r,s) \sim (r',s')$ iff $\exists u \in S$: $urs' = ur's$. We write $r/s$ for the equivalence class $(r,s) \in S^{-1}R$.

Note: if $S$ contains no zero-divisors then $(r,s) \sim (r',s')$ iff $rs' = r's$.

**Lemma 5.3** The relation $\sim$ defined above is an equivalence relation and $S^{-1}R$ can be made into a ring so that the map $i: R \to S^{-1}R$, $i(r) = r/1$ is a homomorphism. Also $i(S) \subseteq (S^{-1}R)^\times$ and the map $i$ is injective iff $S$ contains no zero-divisors.

**Proof.** Reflexivity and symmetry of $\sim$ are immediate. For transitivity, if $(r,s) \sim (r',s') \sim (r'',s'')$ then $\exists u,u': urs' = ur's$, $u'r's'' = u'r''s'$. Hence $(uu's')(rs'') = u's''u'sr = u's''u'sr' = usu'r's'' = usu'rs'' = (uu's')(r''s)$. But $uu's' \in S$, so $(r,s) \sim (r'',s'')$.

Define addition by $r_1/s_1 + r_2/s_2 = (r_1s_2 + r_2s_1)/(s_1s_2)$ and multiplication by $(r_1/s_1)(r_2/s_2) = (r_1r_2)/(s_1s_2)$. A long and rather tedious check shows that under these operations $S^{-1}R$
becomes a commutative ring with identity $1/1$. The map $i(r) = r/1$ is a ring homomorphism since $i(r) + i(r') = r/1 + r'/1 = (r + r')/1 = i(r + r')$, $i(r)i(r') = (r/1)(r'/1) = (rr')/1 = i(rr')$, and $i(1) = 1/1$. The element $1/s \in S^{-1}R$ is the inverse of $i(s) = s/1$, so $i(S) \subseteq (S^{-1}R)\times$.
The kernel of $i$ is $\{r \in R : r/1 = 0/1\} = \{r \in R : \exists u \in S : ur = 0\}$. Thus $\text{Ker } i = \{0\}$ iff $S$ contains no zero-divisors.

**Lemma 5.4** $S^{-1}R$ satisfies the following universal property: If $f : R \rightarrow R'$ is a homomorphism with $f(S) \subseteq (R')\times$ then $f$ factors uniquely as $f = h \circ i$ where $h : S^{-1}R \rightarrow R'$ is a homomorphism.

**Proof.** Any such $\tilde{f}$ must satisfy $\tilde{f}(r/s)\tilde{f}(s/1) = \tilde{f}(r/1)$ and $\tilde{f}(t/1) = f(t)$. Hence $\tilde{f}(r/s)f(s) = f(r)$ and $\tilde{f}(r/s) = f(r)f(s)^{-1}$. Conversely, defining $\tilde{f}(r/s) = f(r)f(s)^{-1}$ gives a homomorphism $S^{-1}R \rightarrow R'$ (check this!).

**Notation:** If $S = R \setminus P$ for some prime ideal $P$, we also write $S^{-1}R$ as $R_P$ and call it the **localization of $R$ at $P$**.

**Lemma 5.5** If $R$ is an ID then $(R \setminus \{0\})^{-1}R = R_{(0)}$ is a field containing a subring isomorphic to $R$.

**Proof.** Let $S = R \setminus \{0\}$. If $r/s \neq 0/1$ then $r \neq 0$, so $s/r \in S^{-1}R$ and $(s/r)(r/s) = 1/1$. Hence any non-zero element of $S^{-1}R$ is invertible. The map $i$ is injective, so $\text{Im } i$ is a subring of $S^{-1}R$ isomorphic to $R$.

In this case we call $R_{(0)} = S^{-1}R$ the **field of fractions** of $R$, or Frac $R$. For example Frac$\{\mathbb{Z}\} = \mathbb{Q}$.

**Exercises**

1. Show that the units of $R_P$ consists of the elements $r/s$ where $r \notin P$ and there is a unique maximal ideal of $R_P$ consisting of all the non-unit elements. [Rings that have a unique maximal ideal are called **local rings**.]

2. Show that if $R$ is an ID, then for any prime ideal $P$, $R_P$ is isomorphic to a subring of Frac $R$.

3. Describe $\mathbb{Z}_{(2)}$ explicitly as a subring of $\mathbb{Q}$.

4. What is the field of fractions of a field?

5. What is the field of fractions of the ring of entire functions (holomorphic functions $f : \mathbb{C} \rightarrow \mathbb{C}$).

6. What is the field of fraction of the ring of polynomial functions $\mathbb{C}[X] = \{\sum_{i=0}^{n} a_i X^i : a_i \in \mathbb{C}, \ n \in \mathbb{N}\}$.
Assume that $R$ is a commutative ring. We wish to construct the ring $R[X]$ of polynomials in $X$ with coefficients in $R$.

Define $R[X]$ as the set of sequences $(a_0, a_1, \ldots)$ with the property that all but finitely many of the $a_i$s are zero. Define $(a_0, \ldots) + (b_0, \ldots) = (a_0 + b_0, a_1 + b_1, \ldots)$ (so $R[X] = \bigoplus_{i \in \mathbb{N}} R$ as group under $+$) and define $R[X] \cdot R[X] = (c_0, c_1, \ldots)$ where $c_i = \sum_{0 \leq j \leq i} a_j b_{i-j}$. We call $R[X]$ the ring of polynomials in $X$ over $R$. Let $i: R \rightarrow R[X]$ be defined by $i(a) = (a, 0, 0, \ldots)$ and let $X \in R[X]$ be the element $X = (0, 1, 0, 0, \ldots)$. Note that $X(a_0, a_1, \ldots) = (a_0, a_0, a_1, \ldots)$ and $i(a)(a_0, a_1, \ldots) = (aa_0, aa_1, \ldots)$.

**Lemma 6.1** $R[X]$ is a ring, $i: R \rightarrow R[X]$ is an injective ring homomorphism, and if $a_i = 0$ for all $i > n$ then $(a_0, a_1, \ldots) = \sum_{i=0}^n a_i X^i$.

We shall normally identify $i(a)$ with $a$ and write polynomials $f(X) \in R[X]$ in the form $\sum_{i=0}^n a_i X^i$. The degree $\deg f(X)$ of a polynomial is the largest $n$ such that $a_n \neq 0$, (or $-\infty$ if $f = 0$). The leading coefficient of $f(X)$ is $a_n$ where $n = \deg f$, (or $0$ if $f = 0$). A polynomial is monic if the leading coefficient is $1$.

**Lemma 6.2** If $f, g \in R[X]$ then

1. $\deg(f + g) \leq \max\{\deg f, \deg g\}$,
2. $\deg(fg) \leq \deg f + \deg g$, with equality holding if $R$ is an ID.

**Lemma 6.3** If $R$ is an ID then $R[X]$ is an ID and $(R[X])^\times = R^\times$.

**Proof.** If $f, g \in R[X]$ and $f, g \neq 0$ then $\deg(fg) = \deg f + \deg g \geq 0$, so $fg \neq 0$. If $fg = 1$ then $0 = \deg(fg) = \deg f + \deg g$ so $\deg f = \deg g = 0$ and $f, g \in R$. Hence $f \in (R[X])^\times$ implies $f \in R^\times$. Conversely $f \in R^\times$ clearly implies $f \in (R[X])^\times$. \qed

**Theorem (Universal property of polynomial rings)** If $\phi: R \rightarrow R'$ is a ring homomorphism and $\alpha \in R'$ then there exists a unique homomorphism $ev_{\phi, \alpha} : R[X] \rightarrow R'$ such that $ev_{\phi, \alpha}(a) \quad (= ev_{\phi, \alpha}(i(a))) = \phi(a)$ for all $a \in R$ and $ev_{\phi, \alpha}(X) = \alpha$.

If $R$ is a subring of $R'$ and $\phi$ is the inclusion map we write $f(\alpha)$ for $ev_{\phi, \alpha}(f)$. More generally, if just $R$ is a subring of $R'$ we write $\phi(f)(\alpha)$ for $ev_{\phi, \alpha}(f)$.

**Lemma 6.4** If $R$ is a subring of $R'$ and $\alpha \in R'$ then $R[\alpha]$ is isomorphic to a quotient $R[X]/I$ where $I$ is an ideal of $R[X]$ containing no non-zero constants: $I \cap R = \{0\}$.

**Proof.** Apply 1st Isomorphism Theorem to $ev_{\alpha}: R[X] \rightarrow R'$. \qed

We say $\alpha \in R'$ is transcendental over $R \subseteq R'$ if the map $ev_{\alpha}$ is injective. In other words, if $f(\alpha) = 0$ implies $f(X) = 0$. Otherwise we say that $\alpha$ is algebraic over $R$. 

Examples  The element $\pi \in \mathbb{R}$ is transcendental over $\mathbb{Z}$, so $\mathbb{Z}[\pi] \cong \mathbb{Z}[X]$. The elements $i, \sqrt{2}, \sqrt{3} \in \mathbb{C}$ are all algebraic over $\mathbb{Z}$. However $\pi$ is algebraic over $\mathbb{R}$ (since it is a root of $X - \pi \in \mathbb{R}[X]$).

Theorem (Division Algorithm)  If $f, g \in R[X]$ and the leading coefficient of $g$ is a unit in $R$, then there exist unique $q, r \in R[X]$ such that $f = qg + r$ and $\deg r < \deg g$ (or $r = 0$).

If $a, b \in R$, we say $a$ divides $b$, $a \mid b$, if there exists $c \in R$ such that $b = ca$.

Examples  In any ring, $u \mid 1$ if and only if $u \in R^\times$, $a \mid 0$ for all $a$. In $\mathbb{Z}$, $7 \mid 21$. In $\mathbb{Q}$, $21 \mid 7$.

Lemma 6.5  If $\alpha \in R$ and $f \in R[X]$ then $f(X) = (X - \alpha)q(X) + f(\alpha)$ for some $q \in R[X]$. In particular, $X - \alpha \mid f$ if and only if $f(\alpha) = 0$.

Lemma 6.6  If $R$ is an integral domain and $f \in R[X]$, $f \neq 0$, then $|\{\alpha \in R : f(\alpha) = 0\}| \leq \deg f$.

Lemma 6.7  If $R$ is an integral domain and $G$ is a finite subgroup of $R^\times$ then $G$ is cyclic.

Proof.  $G$ is a finite abelian group, so $G \cong C_{d_1} \times \cdots \times C_{d_r}$. But then $x^{d_1} = 1$ for all $x \in G$. Thus the polynomial $X^{d_1} - 1$ has $|G|$ zeros. Thus $|G| = d_1d_2\ldots d_r \leq d_1$, so $d_2 = \cdots = d_r = 1$ and $G \cong C_{d_1}$ is cyclic.

We can generalize polynomial rings to polynomials in many variables. If $\{X_i\}_{i \in I}$ is a set (possibly infinite) of indeterminates, define a term $t$ to be a function $I \to \mathbb{N}$ which is non-zero for only finitely many $i \in I$. We think of $t$ as corresponding to a finite product $\prod_{i \in I} X_i^{t_i}$. Let $T$ be the set of terms. Now define the ring

$$R[\{X_i\}_{i \in I}] = \bigoplus_{t \in T} R = \{(a_t)_{t \in T} \mid a_t = 0 \text{ for all but finitely many } t\},$$

with addition of coefficients componentwise $(a_t) + (b_t) = (a_t + b_t)$ and multiplication defined by $(a_t)(b_t) = (c_t)$ where $c_t = \sum_{r+s=t} a_rb_s$ (note that this is a finite sum). As for $R[X]$, we can identify $R$ as a subring of $R[\{X_i\}_{i \in I}]$ and define elements $X_i$ so that $(a_t)_{t \in T}$ is equal to the (finite) sum $\sum_{t \in T} a_t \prod_{i \in I} X_i^{t_i}$.

Theorem (Universal property of polynomial rings)  If $\phi: R \to R'$ is a ring homomorphism and $\alpha_i \in R'$ for all $i \in I$ then there exists a unique homomorphism $ev_{\phi, (\alpha_i)}: R[\{X_i\}_{i \in I}] \to R'$ such that $ev_{\phi, (\alpha_i)}(a) = \phi(a)$ for all $a \in R$ and $ev_{\phi, (\alpha_i)}(X_i) = \alpha_i$ for all $i \in I$.

If $I$ is finite then we can also identify $R[X_1, \ldots, X_n]$ with $R[X_1, \ldots, X_{n-1}][X_n]$ (use universal properties to define the isomorphism).
A Euclidean Domain (ED) is an ID for which there is a function \( d: R \setminus \{0\} \to \mathbb{N} \) such that if \( a, b \in R, b \neq 0 \) then there exists \( q, r \in R \) such that \( a = qb + r \) with either \( d(r) < d(b) \) or \( r = 0 \).

**Examples**

1. \( \mathbb{Z} \) with \( d(a) = |a| \).
2. \( F[X] \), where \( F \) is a field, \( d(f) = \deg f \).
3. \( \mathbb{F} \), where \( F \) is a field, \( d(a) = 0 \).
4. \( \mathbb{Z}[i] \), with \( d(a + ib) = |a + ib|^2 = a^2 + b^2 \). Write \( a/b = x + iy \) and let \( q = x' + iy' \) with \( |x-x'|, |y-y'| \leq \frac{1}{2} \). Then \( d(r) = |q - a| = |q - a/b|^2|b|^2 = ((x-x')^2 + (y-y')^2)d(b) \leq \frac{1}{2}d(b) \).

A Principal Ideal Domain (PID) is an ID in which every ideal \( I \) is principal, i.e., \( I = (a) \) for some \( a \in R \).

**Theorem 7.1** Every ED is a PID.

*Proof.* If \( R \) is an ED then \( R \) is an ID, so it is enough to show that any ideal \( I \) is principal. Let \( I \) be an ideal of \( R \) and assume \( I \neq \{0\} \). Pick \( b \in I \setminus \{0\} \) with minimal value of \( d(b) \) (by well ordering of \( \mathbb{N} \)). If \( a \in I \) then \( a = qb + r \) with \( d(r) < d(b) \) or \( r = 0 \). But \( r = a - qb \in I \), so by choice of \( b \) we must have \( r = 0 \). Thus \( a = qb \in (b) \). Thus \( I \subseteq (b) \). But \( b \in I \), so \((b) \subseteq I \). Thus \( I = (b) \) is principal. \( \square \)

Note: PID \( \neq \) ED.

If \( I = (a) \) is a principal ideal then \( b \in I \) implies there exists \( a \in R \) with \( b = ca \). Thus \( b \in I \) is equivalent to \( a \mid b \). In particular \((b) \subseteq (a) \iff a \mid b \). If \( (a) = (b) \) then \( b = ua \) and \( a = vb \). Thus either \( a = b = 0 \) or \( uv = 1 \) and \( u, v \in R^\times \). Conversely, if \( a = ub \) with \( u \in R^\times \) then \((a) = (b) \).

The elements \( a, b \in R \) are called **associates** if \( b = ua \) for some \( u \in R^\times \). Equivalently, \( a \mid b \) and \( b \mid a \) both hold, or \((a) = (b) \). Write \( a \sim b \) if \( a \) and \( b \) are associates.

A **greatest common divisor** (gcd) of a set of elements \( S \subseteq R \) is an element \( d \in R \) such that

G1. \( d \mid a \) for all \( a \in S \), and

G2. if \( c \mid a \) for all \( a \in S \) then \( c \mid d \).

Greatest common divisors are unique up to multiplication by units. To see this, let \( d, d' \) be two gcds. Then condition G2 with \( c = d' \) and G1 with \( d = d' \) imply \( d' \mid d \). Similarly \( d \mid d' \), so \( d' = ud \) for some unit \( u \in R^\times \).
Lemma 7.2  If $R$ is a PID then gcds of any $S \subseteq R$ exist. Indeed, if $(S) = (d)$ then $d$ is a gcd of $S$ and hence can be written in the form $d = \sum_{i=1}^{r} c_i a_i$, for some $a_i \in S$, $c_i \in R$.

Proof. Since $R$ is a PID, $(S) = (d)$ for some $d$. If $a \in S$ then $a \in (S) = (d)$, so $d \mid a$. If $c \mid a$ for all $a \in S$, then $a \in (c)$ for all $a \in S$, so $(S) = (d) \subseteq (c)$. Hence $c \mid d$. Thus $d$ is a gcd of $S$. \[\square\]

Note: In an arbitrary ID, gcds may not exist, and even if they do, they may not be a linear combination of elements of $S$. For example the elements 2 and $X$ in $\mathbb{Z}[X]$ have 1 as a gcd, but 1 is not of the form $2c_1 + Xc_2$, $c_1, c_2 \in \mathbb{Z}[X]$. For an example where the gcd does not exist, consider $R = \mathbb{Z}[\sqrt{-5}]$. If $a \in R$ then $|a|^2 \in \mathbb{Z}$. Hence if $a \mid b$ in $R$ then $|a|^2 \mid |b|^2$ in $\mathbb{Z}$. Now let $x = -3(3 - \sqrt{-5}) = (1 + 2\sqrt{-5})(1 + \sqrt{-5})$ and $y = -7(1 + \sqrt{-5}) = (1 - 2\sqrt{-5})(3 - \sqrt{-5})$. Then $1 + \sqrt{-5}$ and $3 - \sqrt{-5}$ are two common factors of $x$ and $y$. If $d$ is a gcd of $x$ and $y$, then $|d|^2$ must be a factor of $|x|^2 = 2.3^2.7$ and $|y|^2 = 2.3.7^2$. On the other hand, $|d|^2$ must be a multiple of $|1 + \sqrt{-5}|^2 = 2.3$ and $|3 - \sqrt{-5}|^2 = 2.7$. Thus $|d|^2 = 2.3.7 = 42$. However, if $d = \alpha + \beta \sqrt{-5}$ then $|d|^2 = \alpha^2 + 5\beta^2$, which is never equal to 42.

The Euclidean Algorithm

We can turn Lemma 1 into an algorithm in the case when $R$ is a ED. Assume we need to find the gcd of $a_0 = a$ and $a_1 = b$. Inductively define $a_{n+1}$ for $n \geq 1$ and $a_n \neq 0$ by

$$a_{n-1} = q_na_n + a_{n+1}, \quad d(a_{n+1}) < d(a_n) \text{ or } a_{n+1} = 0$$

Since the $d(a_n)$ are a sequence of decreasing non-negative integers, eventually $a_{n+1} = 0$. However $a_{i+1} \in (a_i, a_{i-1})$ and $a_{i-1} \in (a_i, a_{i+1})$ imply the two ideals $(a_{i-1}, a_i)$ and $(a_i, a_{i+1})$ are equal. Hence $(a_0, a_1) = (a_n, a_{n+1}) = (a_n)$ and $a_n$ is a gcd of $a_0$ and $a_1$.

This algorithm is called the Euclidean Algorithm. For more than two elements, one can calculate the gcd inductively by using $\text{gcd}(c_1, c_2, \ldots, c_r) = \text{gcd}(c_1, \text{gcd}(c_2, \ldots, c_r))$.

Exercises

1. Prove that $\text{gcd}(c_1, \ldots, c_r) = \text{gcd}(c_1, \text{gcd}(c_2, \ldots, c_r))$ provided the gcds on the RHS exist. What is $\text{gcd}()$?

2. Let $R = \mathbb{Z}[\omega]$ where $\omega = \frac{1}{2}(1 + \sqrt{-3})$. Show that $R = \{a + b\omega : a, b \in \mathbb{Z}\}$ and that $R$ is a ED.

3. Use the Euclidean algorithm to find the gcd of $7 - 3i$ and $5 + 3i$ in $\mathbb{Z}[i]$.

4. Determine $((\mathbb{Z}/n\mathbb{Z})[X])^\times$. [Hint: Consider the case $n = p^r$ first.]

5. Solve the congruences

$$x \equiv i \mod 1 + i \quad x \equiv 1 \mod 2 - i$$

in $\mathbb{Z}[i]$ (use Chinese Remainder Theorem).
An element \( a \in R \) is **irreducible** if \( a \neq 0 \), \( a \notin R^\times \), and \( a = bc \) implies \( b \in R^\times \) or \( c \in R^\times \).

An element \( a \in R \) is a **prime** if \( a \neq 0 \), \( a \notin R^\times \) and \( a \mid bc \) implies \( a \mid b \) or \( a \mid c \).

**Lemma 8.1** Let \( R \) be an ID, and \( a \in R \). Then
1. \( a \) is a prime element iff \((a)\) is a non-zero prime ideal,
2. \( a \) is irreducible iff \((a)\) is maximal among proper principal ideals (i.e., \((a) \subseteq (b) \) implies \((b) = (a)\) or \((b) = R\)),
3. if \( a \) is prime then \( a \) is irreducible,
4. if \( a \) is irreducible and \( R \) is a PID then \( a \) is prime.

**Proof.**
1. If \( a \) is prime and \( bc \in (a) \) then \( a \mid bc \). Hence \( a \mid b \) or \( a \mid c \), so either \( b \in (a) \) or \( c \in (a) \). Also, \( a \neq 0 \), \( a \notin R^\times \) implies \((a) \neq (0), R\). Conversely, if \((a)\) is a prime ideal and \( a \mid bc \), then \( bc \in (a) \), so either \( b \in (a) \) or \( c \in (a) \), so either \( a \mid b \) or \( a \mid c \) and \((a) \neq (0), R\) implies \( a \neq 0 \), \( a \notin R^\times \).
2. If \( a \in R \) is irreducible and \( (a) \subseteq (b) \) then \( a = bc \), so either \( c \in R^\times \) and \((b) = (a)\) or \( b \in R^\times \) and \((b) = R\). Conversely if \((a)\) is maximal among all proper principal ideals and \( a = bc \) then \( (a) \subseteq (b) \), so either \((a) = (b)\) and \( c \) is a unit or \((b) = R\) and \( b \) is a unit.
3. If \( a \) is a prime and \( a = bc \) then \( a \mid bc \). Thus either \( a \mid b \) and \( c \in R^\times \), or \( a \mid c \) and \( b \in R^\times \).
4. By part 2, \((a)\) is a maximal ideal. Hence \((a)\) is prime and so \( a \) is prime.

A ring \( R \) is a **Unique Factorization Domain** (UFD) if \( R \) is an ID such that

**U1.** Every \( a \in R \setminus \{0\} \) can be written in the form \( a = up_1 \ldots p_r \) where \( u \in R^\times \) and the \( p_i \) are irreducible.

**U2.** Any two such factorizations are unique in the sense that if \( up_1 \ldots p_r = vq_1 \ldots q_s \) then \( r = s \) and there is a permutation \( \pi \in S_r \) such that \( p_i \sim q_{\pi(i)} \) for all \( i \).

**Lemma 8.2** \( R \) is a UFD iff \( R \) is an ID satisfying

A. there is no infinite sequence \((a_i)_{i \in \mathbb{N}} \) with \( a_{i+1} \mid a_i \) and \( a_{i+1} \not\sim a_i \), and

B. every irreducible is prime.

**Proof.**
A \( \Rightarrow \) U1. Suppose \( a_1 \in R \) has no such factorization. Then \( a_1 \) is neither a unit nor irreducible, so \( a_1 = bc \), \( b, c \notin R^\times \), and either \( b \) or \( c \) also has no factorization into irreducibles. Assume \( b \) has no factorization into irreducibles and set \( a_2 = b \). Repeating this process we get a sequence \( a_i \) with \( a_{i+1} \mid a_i \) and \( a_{i+1} \not\sim a_i \).

B \( \Rightarrow \) U2. Since \( p_1 \) is prime and \( p_1 \mid vq_1 \ldots q_s \), we must have \( p_1 \mid q_i \) for some \( i \). But \( q_i \) is irreducible, so \( p_1 \sim q_i \). Cancelling a factor of \( p_1 \) from both sides (\( R \) is an ID) and using induction on \( r \) gives the result.

U1 and U2 \( \Rightarrow \) A and B is clear. 

15
A ring is Noetherian if every sequence of ideals \( I_i \) with \( I_i \subseteq I_{i+1} \) is eventually constant, \( I_n = I_{n+1} = \ldots \), for some \( n \).

**Lemma 8.3** \( R \) is Noetherian iff every ideal is finitely generated.

**Proof.** \( \Leftarrow \): Let \( I = \bigcup I_n \). Then \( I \) is an ideal, so \( I = (d_1, \ldots, d_r) \) for some \( d_i \in R \). But then there is an \( n_i \) with \( d_i \in I_n \). Let \( n = \max n_i \), so that \( I = (d_1, \ldots, d_r) \subseteq I_n \subseteq I_{n+1} \subseteq \cdots \subseteq I \), and so \( I_n = I_{n+1} = \ldots \).

\( \Rightarrow \): Assume \( I \) is not finitely generated. Then (using Axiom of choice), pick inductively \( d_n \in I \setminus (d_1, \ldots, d_{n-1}) \). Then \( I_n = (d_1, \ldots, d_n) \) is a strictly increasing sequence of ideals. \( \square \)

**Theorem** Every PID is a UFD.

**Proof.** Every ideal in a PID is finitely generated (by one element), so PID \( \Rightarrow \) Noetherian. By considering the ideals \( (a_i) \), Noetherian rings satisfy condition A of Lemma 8.2. Lemma 8.1 part 4 implies condition B of Lemma 8.2, so PID \( \Rightarrow \) UFD. \( \square \)

GCDs and factorizations

**Lemma 8.4** If \( R \) is a UFD and \( S \subseteq R \) then a gcd of \( S \) exists.

**Proof.** The relation \( \sim \) is an equivalence relation on the set of irreducibles in \( R \). So by choosing a representative irreducible from each equivalence class we can construct a set \( P \) of pairwise non-associate irreducible elements of \( R \). We can write any element \( a \in R \) as \( u \prod_{p \in P} p^{n_p} \) and if \( b = v \prod_{p \in S} p^{m_p} \) then U2 implies \( a \mid b \) iff \( n_p \leq m_p \) for all \( p \). Write each \( a_i \in S \) as \( a_i = u_i \prod_{p \in P} p^{n_i,p} \). If we let \( d = \prod_{p \in P} p^{m_p} \) with \( m_p = \min_{i \in S} n_{i,p} \) then it is clear that \( d \) is a gcd for \( S \).

A partial converse to Lemma 8.4 is true.

**Lemma 8.5** If \( R \) is an ID in which the gcd of any pair of elements exists then every irreducible is prime.

**Proof.** First we prove that if gcds exist then \( \gcd(ab, ac) \sim a \gcd(b, c) \). Let \( e = \gcd(ab, ac) \) and \( d = \gcd(b, c) \). Then \( d \mid b, c \), so \( ad \mid ab, ac \), so \( ad \mid e \). Writing \( e = adu \) then \( e \mid ab, ac \), so \( du \mid b, c \), so \( du \mid d \). Thus \( u \in R^\times \) and \( e \sim ad \) (or \( d = 0 = e \)).

Now let \( p \) be an irreducible and assume \( p \nmid a, b \). Then \( \gcd(p, b) \sim 1 \) since the gcd must be a factor or \( p \) and \( p \nmid b \). Hence \( \gcd(p, ab) \mid \gcd(ap, ab) \sim a \). But \( \gcd(p, ab) \mid p \), so \( \gcd(p, ab) \mid \gcd(a, p) \sim 1 \). Hence \( \gcd(p, ab) \sim 1 \) and \( p \nmid ab \). Hence \( p \) is prime. \( \square \)

**Lemma 8.6** If \( R \) is an ID in which every set \( S \) has a gcd which can be written in the form \( \sum r_i a_i \) for some \( a_i \in S, r_i \in R \), then \( R \) is a PID.

**Proof.** Let \( I \) be an ideal and write \( I = (S) \) for some \( S \) (e.g., \( S = I \)). Let \( d = \sum r_i a_i \) be a gcd of \( S \). Then \( d \mid a \) for all \( a \in S \). Hence \( a \in (d) \), so \( S \subseteq (d) \). Thus \( I \subseteq (d) \). However \( d = \sum r_i a_i \in I \). Then \( (d) \subseteq I \). Hence \( I = (d) \) is principal. \( \square \)
Assume throughout this section that $R$ is a UFD.

Let $f(X) = \sum_{i=0}^{n} a_i X^i \in R[X]$. Define the content of $f(X)$ to be $c(f) = \gcd\{a_0, a_1, \ldots, a_n\}$. Note that if $f \neq 0$ then $c(f) \neq 0$. We call $f$ primitive iff $c(f) \sim 1$.

Note that monic polynomials are primitive, but not conversely, e.g. $2X + 3 \in \mathbb{Z}[X]$.

**Lemma (Gauss)** If $R$ is a UFD and $f, g \in R[X]$ are primitive, then so is $fg$.

**Proof.** Assume otherwise and let $p$ be a prime dividing $c(fg)$. Reducing the polynomials mod $p$ we get $\bar{f}, \bar{g} \in (R/(p))[X]$ with $\bar{f}, \bar{g} \neq 0$, but $\bar{f}\bar{g} = \bar{f\bar{g}} = 0$ (the map $f \mapsto \bar{f}$ $R[X] \to (R/(p))[X]$ is a special case of the evaluation homomorphism $ev_{\pi, X}$ where $X$ is sent to $X$ and $ev_{\pi, X}$ acts as the projection map $\pi : \mathbb{R} \to \mathbb{R}/(p)$ on constants). Now $p$ is prime, so $(p)$ is a prime ideal and $R/(p)$ is an ID. Hence $f, \bar{g} \neq 0$ implies $\bar{f}\bar{g} \neq 0$, a contradiction. \hfill $\square$

**Corollary 9.1** If $R$ is a UFD then $c(fg) \sim c(f)c(g)$.

**Proof.** The result clearly holds if $f$ or $g$ is zero, so assume $f, g \neq 0$ and hence $c(f) \neq 0$. Since $\gcd\{a_0\} \sim a \gcd\{a_i\}$, $c(af) \sim ac(f)$ for all $a \in R$. But $f/c(f) \in R[X]$, so $c(f)c(f/c(f)) = c(f)$ and so $f/c(f)$ is primitive. Now $fg/(c(f)c(g)) = (f/c(f))(g/c(g))$ is primitive. Hence $c(fg) \sim c(f)c(g)c(f/g/c(f)c(g)) \sim c(f)c(g)$. \hfill $\square$

**Lemma 9.2** If $\deg f > 0$ and $f$ is irreducible in $R[X]$ then $f$ is irreducible in $F[X]$, where $F = \text{Frac} R$ is the field of fractions of $R$.

**Proof.** Suppose $f = gh$ in $F[X]$. By multiplying by denominators, there exist non-zero $a, b \in R$ with $ag, bh \in R[X]$. Thus $abf = (ag)(bh) \in R[X]$ and $c(abf) \sim c(ag)c(bh)$. But $f = c(f)(f/c(f))$ is a factorization of $f$ in $R[X]$ and if $\deg f > 0$, $f/c(f) \notin (R[X])^\times = R^\times$. Thus $c(f) \in R^\times$ and so $c(abf) \sim ab$. Now $ab/c(ag)c(bh) = u \in R^\times$ and $f = (u^{-1}ag/c(ag))(bh/c(bh))$ is a factorization of $f$ in $R[X]$. Hence either $\deg g = 0$ or $\deg f = 0$ and so $g$ or $h$ is a unit in $F[X]$. \hfill $\square$

**Lemma 9.3** If $R$ is a UFD then $f \in R[X]$ is irreducible iff either

(a) $f \in R$ is an irreducible in $R$, or
(b) $f$ is primitive in $R[X]$ and irreducible in $F[X]$.

**Proof.** Assume first that $\deg f = 0$. If $f = ab$ in $R$, $f = ab$ in $R[X]$. Conversely, if $f = gh$ in $R[X]$ then $\deg g = \deg h = 0$, so $f = gh$ in $R$. Since $R^\times = (R[X])^\times$, irreducibility in $R[X]$ is equivalent to irreducibility in $R$. Assume now that $\deg f > 0$. If $f$ is irreducible in $R[X]$ then by the previous lemma, $f$ is irreducible in $F[X]$. Also, $f = c(f)(f/c(f))$, so $c(f) \in (R[X])^\times = R^\times$ and $f$ is primitive. Conversely, if $f$ is primitive and irreducible in $F[X]$ and $f = gh$ in $R[X]$, then $f = gh$ in $F[X]$, so wlog $g \in (F[X])^\times \cap R[X] = R$. But then $g \mid c(f)$ in $R$, so $g \in R^\times = (R[X])^\times$. Thus $f$ is irreducible in $R[X]$. \hfill $\square$
Theorem 9.4 If $R$ is a UFD then $R[X]$ is a UFD.

Proof. Write $f = c(f)f'$ where $f'$ is primitive. Now $c(f) = u p_1 \ldots p_r$ where $u \in R^\times = (R[X])^\times$ and $p_i$ are irreducible in $R$. If $f' = gh$ with $g, h \not\in (R[X])^\times = R^\times$ then $c(g)c(h) \sim 1$, so $g, h$ are primitive and $\deg g, \deg h > 0$ (since otherwise either $g$ or $h$ would lie in $R^\times$). By induction on the degree, $f'$ is the product of irreducible primitive polynomials $f' = \prod f_i$. Hence $f$ has a factorization into irreducibles.

Now assume $f = u p_1 \ldots p_r f_1 \ldots f_i = v q_1 \ldots q_a g_1 \ldots g_u$ where $u, v \in R^\times$, $p_i, q_j$ are irreducible in $R$ and $f_i, g_j$ are primitive and irreducible in $F[X]$. The ring $F[X]$ is a PID, so is a UFD. The elements $u p_1 \ldots p_r$ and $v q_1 \ldots q_a$ are units in $F[X]$, so $t = u$ and wlog $f_i = \gamma_i g_i$ for some $\gamma_i \in (F[X])^\times = F \setminus \{0\}$. Write $\gamma_i = a_i/b_i$ with $a_i, b_i \in R$. Now $b_i f_i = a_i g_i$, so $b_i \sim c(b_i f_i) = c(a_i g_i) \sim a_i$. Thus $\gamma_i \in R^\times$ and $f_i \sim g_i$ in $R[X]$. Now $c(f) \sim u p_1 \ldots p_r \sim v q_1 \ldots q_a$, so by unique factorization in $R$, $r = s$ and wlog $p_i \sim q_i$ in $R$ and hence in $R[X]$. Hence the factorization of $f$ is unique in $R[X]$.

Factorization methods

Evaluation method: If $g \mid f$ in $R[X]$ then $g(c) \mid f(c)$ in $R$ for all $c \in R$.

Example: If $f = x^2 - 4x + 1 \in \mathbb{Z}[x]$, then $f(\pm 2) = 1$. If $f = gh$ then we can assume wlog that $g$ is linear. But then $g(\pm 2) = \pm 1$. The only linear polynomials with this property are $\pm X/2$ which do not lie in $\mathbb{Z}[X]$. Hence $f$ is irreducible in $\mathbb{Z}[X]$ (and hence also in $\mathbb{Q}[X]$).

Reduction mod $p$: If $f = gh$ in $R[X]$ and $p$ is a prime then $\bar{f} = \bar{g}\bar{h}$ in $(R/(p))[X]$.

Example: If $f = x^4 - x^2 + 4x + 3 \in \mathbb{Z}[x]$, then if $p = 2$, $\bar{f} = x^4 + x^2 + 1 = (x^2 + x + 1)(x^2 + x + 1)$ in $(\mathbb{Z}/2\mathbb{Z})[x]$ and if $p = 3$ then $\bar{f} = x^4 - x^2 + x = x(x^3 - x + 1)$ in $(\mathbb{Z}/3\mathbb{Z})[x]$. In $\mathbb{Z}[x]$, $f$ cannot factor as a product of two quadratics (since there is no quadratic factor mod 3), nor can it have a linear factor (no linear factor mod 2), hence $f$ is irreducible in $\mathbb{Z}[X]$.

Lemma (Eisenstein’s irreducibility criterion) Assume $R$ is a UFD, $f = \sum_{i=0}^n a_n x^n \in R[X]$, is primitive, and $p$ is a prime such that $p \not\mid a_n$, $p \mid a_i$ for $i < n$ and $p^2 \not\mid a_0$. Then $f$ is irreducible in $R[X]$.

Proof. Suppose $f = gh$. Then $\bar{g} = a_n x^n$ in $(R/(p))[X]$. Thus $\bar{g} = a x^i$ and $\bar{h} = b x^j$ for some $a, b \in R/(p)$ and $i + j = n$. But $\deg g + \deg h = n$ and $i \leq \deg g$, $j \leq \deg h$. Hence $i = \deg g$ and $j = \deg h$. If $g$ and $h$ are not units in $R[X]$ and $f$ is primitive then $\deg g, \deg h > 0$. Hence $\bar{g}(0) = \bar{h}(0) = 0$, so $p \mid g(0), h(0)$. Thus $p^2 \mid g(0)h(0) = f(0) = a_0$, a contradiction. Hence $f$ is irreducible.

Exercises

1. Show that for $p$ a prime in $\mathbb{Z}$, $f(X) = 1 + X + \ldots + X^{p-1} = (X^p - 1)/(X - 1)$ is irreducible in $\mathbb{Q}[X]$ [Hint: consider $f(X+1)$ and use Eisenstein’s criterion].

2. Let $f = x^3 - x + 1$. Show that $(\mathbb{Z}/3\mathbb{Z})[X]/(f)$ is a field with 27 elements.
A polynomial \( f(X_1, \ldots, X_n) \in R[X_1, \ldots, X_n] \) is called symmetric if
\[
f(X_1, \ldots, X_n) = f(X_{\pi(1)}, \ldots, X_{\pi(n)})
\]
for any permutation \( \pi \in S_n \).

**Examples** \( X_1^2 + X_2^2 + X_3^2 \) and \( X_1X_2 + X_2X_3 + X_3X_1 \) are symmetric polynomials in the ring \( \mathbb{Z}[X_1, X_2, X_3] \), however \( X_1^2X_2 + X_2^2X_3 + X_3^2X_1 \) is not symmetric (consider the permutation \( \pi = (12) \)).

The elementary symmetric polynomials \( \sigma_r \in R[X_1, \ldots, X_n] \) are defined by \( \sigma_r = \sum_{i_1 < i_2 < \cdots < i_r} X_{i_1} \cdots X_{i_r} = \sum_{|S|=r} \prod_{i \in S} X_i \) where in the second expression the sum is over all subsets \( S \) of \( \{1, \ldots, n\} \) of size \( r \).

**Examples** For \( n = 3 \), \( \sigma_0 = 1 \), \( \sigma_1 = X_1 + X_2 + X_3 \), \( \sigma_2 = X_1X_2 + X_2X_3 + X_3X_1 \), \( \sigma_3 = X_1X_2X_3 \).

Note: \((X + X_1)(X + X_2) \cdots (X + X_n) = X^n + \sigma_1X^{n-1} + \sigma_2X^{n-2} + \cdots + \sigma_n \).

Define the degree of \( cX_1^{a_1} \cdots X_n^{a_n} \in R[X_1, \ldots, X_n], c \neq 0 \), as the \( n \)-tuple \((a_1, \ldots, a_n) \). More generally define the degree of \( f = \sum_{a_1, \ldots, a_n} c_{a_1, \ldots, a_n} X_1^{a_1} \cdots X_n^{a_n} \) as the maximum value of \((a_1, \ldots, a_n)\) over all \( c_{a_1, \ldots, a_n} \neq 0 \), where \( n \)-tuples are ordered lexicographically: \((a_1, \ldots, a_n) < (b_1, \ldots, b_n)\) iff there exists an \( i \) such that \( a_i < b_i \) and \( a_j = b_j \) for all \( j < i \).

**Example** In \( R[X_1, X_2, X_3] \), \( \deg(X_1^2X_2^3 + X_1^3X_3) = (7, 0, 1) \).

In \( R[X_1, \ldots, X_n] \), \( \deg \sigma_r = (1, \ldots, 1, 0, \ldots, 0) \), where there are \( r \) ones and \( n - r \) zeros.

**Lemma 10.1** The lexicographic ordering on \( \mathbb{N}^n \) is a well ordering: \( \mathbb{N}^n \) is totally ordered and every non-empty \( S \subseteq \mathbb{N}^n \) has a minimal element.

**Proof.** To prove every \( S \neq \emptyset \) has a minimal element, inductively construct sets \( S_i \) with \( S_0 = S \) and \( S_i \) equal to the set of elements \((a_1, \ldots, a_n)\) of \( S_{i-1} \) for which \( a_i \) is minimal. It is clear that \( S_i \neq \emptyset \) and the (unique) element of \( S_n \) is the minimal element of \( S \). \( \square \)

**Lemma 10.2** If \( f \in R[X_1, \ldots, X_n] \) is symmetric and \( \deg f = (a_1, \ldots, a_n) \) then \( a_1 \geq a_2 \geq \cdots \geq a_n \).

**Proof.** Assume otherwise and let \( a_i < a_j \) with \( i > j \). Then if \( \pi = (ij) \), \( f(X_1, \ldots, X_n) = f(X_{\pi(1)}, \ldots, X_{\pi(n)}) \) has a term with degree \((a_{\pi(1)}, \ldots, a_{\pi(n)})\) which is larger than \((a_1, \ldots, a_n)\), contradicting the definition of the degree. \( \square \)

**Lemma 10.3** If \( f, g \in R[X_1, \ldots, X_n] \) and \( f, g \) are monic (the term with degree equal to \( \deg f \) or \( \deg g \) has coefficient 1) then \( \deg fg = \deg f + \deg g \) where addition of degrees is performed componentwise: \((a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n)\).
Proof. Prove that in the lexicographical ordering, \( a < b \) and \( c \leq d \) imply \( a + c < b + d \). The rest of the proof is the same as for the one variable case. \( \square \)

Theorem 10.1 The polynomial \( f \in R[X_1, \ldots, X_n] \) is symmetric iff \( f \in R[\sigma_1, \ldots, \sigma_n] \).

Clearly \( \sigma_i \) is symmetric, and the set of symmetric polynomials forms a subring of the ring \( R[X_1, \ldots, X_n] \). Hence every element of \( R[\sigma_1, \ldots, \sigma_n] \) is symmetric. We now need to show every symmetric polynomial can be written as a polynomial in \( \sigma_1, \ldots, \sigma_n \). We use induction on \( \deg f \). Let \( f \) be a counterexample with minimal \( \deg f \) (using Lemma 1). Let \( \deg f = (a_1, \ldots, a_n) \) and let the leading term have coefficient \( c \in R \). Then \( g = c\sigma_1^{a_1} \sigma_2^{a_2} \cdots \sigma_n^{a_n} \) has \( \deg g = (a_1, \ldots, a_n) = \deg f \) (by Lemma 3) and the same leading coefficient \( c \). Thus \( \deg(f - g) < \deg f \). Now \( g \) is symmetric, so \( f - g \) is symmetric. By induction on \( \deg f \), \( f - g \in R[\sigma_1, \ldots, \sigma_n] \). But \( g \in R[\sigma_1, \ldots, \sigma_n] \). Hence \( f \in R[\sigma_1, \ldots, \sigma_n] \), contradicting the choice of \( f \).

If \( \alpha \in R' \) and \( R \) is a subring of \( R' \), we call \( \alpha \) algebraic over \( R \) if the map \( \text{ev}_\alpha : R[X] \to R' \) is not injective, i.e., there exists a non-zero \( f(X) \in R[X] \) with \( f(\alpha) = 0 \). More generally we say \( \alpha_1, \ldots, \alpha_n \) are algebraically dependent if \( \text{ev}_{\alpha_1,\ldots,\alpha_n} : R[X_1, \ldots, X_n] \to R' \) is not injective, or equivalently there exists a non-zero polynomial \( f \in R[X_1, \ldots, X_n] \) with \( f(\alpha_1, \ldots, \alpha_n) = 0 \). We say \( \alpha_1, \ldots, \alpha_n \) are algebraically independent over \( R \) if they are not algebraically dependent.

Theorem 10.2 The elements \( \sigma_1, \ldots, \sigma_n \) are algebraically independent over \( R \). The elements \( X_i \) are algebraic over \( R[\sigma_1, \ldots, \sigma_n] \).

Proof. Assume \( \sum c_{a_1,\ldots,a_n} \sigma_1^{a_1} \cdots \sigma_n^{a_n} = 0 \) in \( R[X_1, \ldots, X_n] \). Among the (finite set of) \( (b_1, \ldots, b_n) \) such that \( c_{b_1,\ldots,b_n} \neq 0 \), pick one such that \( (b_1 + \cdots + b_n, b_2 + \cdots + b_n, \ldots, b_n) \) is maximal in the lexicographical ordering. The map sending \( (a_1, \ldots, a_n) \) to \( (a_1 + \cdots + a_n, a_2 + \cdots + a_n, \ldots, a_n) \) is an injection \( \mathbb{N}^d \to \mathbb{N}^d \), so this \( (b_1, \ldots, b_n) \) is uniquely determined. Now \( \deg \sum c_{a_1,\ldots,a_n} \sigma_1^{a_1} \cdots \sigma_n^{a_n} = (b_1 + \cdots + b_n, b_2 + \cdots + b_n, \ldots, b_n) \) contradicting \( \sum c_{a_1,\ldots,a_n} \sigma_1^{a_1} \cdots \sigma_n^{a_n} = 0 \). Thus \( \sigma_1, \ldots, \sigma_n \) are algebraically independent. The elements \( X_i \) are algebraic over \( R[\sigma_1, \ldots, \sigma_n] \) since they are roots of \( X_n - \sigma_1 X^{n-1} + \cdots + \sigma_n = 0 \). \( \square \)

As a consequence of Theorem 2, any symmetric polynomial \( f \in R[X_1, \ldots, X_n] \) can be written as \( g(\sigma_1, \ldots, \sigma_n) \) with \( g \) a unique element of \( R[X_1, \ldots, X_n] \). For example, \( X_1^2 + X_2^2 + X_3^2 = \sigma_1^2 - 2\sigma_2 \).

Exercises

1. Let \( \delta = \prod_{i<j} (X_i - X_j) \in \mathbb{Z}[X_1, \ldots, X_n] \). Show that \( \delta^2 \) is symmetric and for \( n = 3 \) express \( \delta^2 \) in terms of \( \sigma_1, \sigma_2, \sigma_3 \).

2. Let \( f(X) = X^3 - 3X + 5 \) have complex roots \( \alpha_1, \alpha_2, \alpha_3 \). Find a polynomial with complex roots \( \alpha_1^2, \alpha_2^2, \alpha_3^2 \).