# Math 346 - Algebra II

March 7, 2016

# 1 Rings

## 1.1 Basic definitions

Let us first recall the definition of a group, which you learned in Algebra I.

**Definition 1.1.** A set $G$ along with a binary operation $*$ is a **group** if there is an element $e$ of $G$ such that:

1. For any $a \in G$, $a * e = e * a = a$ and there exists $b \in G$ such that $a * b = b * a = e$;
2. For any $a, b, c \in G$, $a * (b * c) = (a * b) * c$.

If furthermore $a * b = b * a$ for any $a, b \in G$, then $G$ is an **abelian group**. The element $e$ is called the **identity** of the group and the element $b$ such that $a * b = e$ is called the inverse of $a$ and written as $a^{-1}$.

Examples of groups include the set $\mathbb{Z}$ of integers with addition, the set $\mathbb{Q}$ of rational numbers with addition, the set $\mathbb{Q}^\times$ of nonzero rational numbers with multiplications, the set $\mathbb{Z}/m\mathbb{Z}$ of congruence classes modulo $m$ with addition, the set of increasing functions on $\mathbb{R}$ with composition. All of these groups except for the last one are abelian. All but the fourth one are infinite. You probably also learned that all finite abelian groups are in fact products of $\mathbb{Z}/m\mathbb{Z}$.

The sets $\mathbb{Z}, \mathbb{Q}, \mathbb{Z}/m\mathbb{Z}$ have one more operation, namely multiplication. A ring is roughly speaking a set with two binary operations, addition and multiplication, with some compatibility conditions.

**Definition 1.2.** A set $R$ with two binary operations $+, \cdot$ is a **ring** if there are elements 0 and 1 in $R$ such that:

1. $R$ along with $+$ and 0 forms an abelian group;
2. For any $a, b, c \in R$, $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ and $a \cdot 1 = 1 \cdot a = a$;
3. For any $a, b, c \in R$, $(a + b) \cdot c = a \cdot c + b \cdot c$ and $a \cdot (b + c) = a \cdot b + a \cdot c$.

If furthermore $a \cdot b = b \cdot a$ for any $a, b \in R$, then $R$ is a **commutative ring**. The additive inverse of an element $a$ in $R$ is denoted $-a$.

Notation-wise, we write $a - b$ for $a + (-b)$ and sometimes omit the $\cdot$ and write $ab$ for $a \cdot b$. It is easy to check the following basic properties about ring operations.

**Proposition 1.3.** Let $R$ be a ring. Then

1. For any $a, b \in R$, $(-a)b = -(ab) = a(-b)$;
2. For any $a, b \in R$, $(-a)(-b) = ab$;
3. For any $a \in R$, $0 \cdot a = a \cdot 0 = 0$.

Note we do not ask for every element to have a multiplicative inverse. A **unit** of a ring is an element that has a multiplicative inverse. That is, $u \in R$ is a unit if and only if there exists $v \in R$ such that $uv = vu = 1$. The set of units of a ring $R$ is denoted by $R^\times$. It forms a group along with $\cdot$ and 1. It is an abelian group if $R$ is a commutative ring.

## 1.2 Examples of rings

The sets $\mathbb{Z}, \mathbb{Q}, \mathbb{Z}/m\mathbb{Z}$ with the usual addition and multiplication form commutative rings. Similarly we also have the ring $\mathbb{R}$ of real numbers and the ring $\mathbb{C}$ of complex numbers.

The set $\mathbb{Z}[i] = \{a + bi | a, b \in \mathbb{Z}\}$ of Gaussian integers is a commutative ring with addition and multiplication defined as follows:

$$\begin{aligned} (a+bi) + (c+di) &= (a+c) + (b+d)i, \\ (a+bi) \cdot (c+di) &= (ac-bd) + (ad+bc)i. \end{aligned}$$

The set $\mathbb{Z}[x] = \{a_n x^n + \cdots + a_1 x + a_0 : a_i \in \mathbb{Z}\}$ of polynomials with integer coefficients is a commutative ring with addition and multiplication defined as follows:

$$\begin{aligned} \sum_i a_i x^i + \sum_i b_i x^i &= \sum_i (a_i + b_i) x^i, \\ \sum_i a_i x^i \cdot \sum_i b_i x^i &= \sum_i \left( \sum_{j+k=i} a_j b_k \right) x^i. \end{aligned}$$

If we allow for an infinite sum of powers of $x$, we get the ring of (formal) power series $\mathbb{Z}[[x]] = \{\sum_{i=0}^{\infty} a_i x^i | a_i \in \mathbb{Z}\}$. Addition and multiplication are defined using the same formula as above. Note in the formula for multiplication, the sum $\sum_{j+k=i} a_j b_k$ is a finite sum.

The set $M_n(\mathbb{Z})$ of $n \times n$ matrices with integer coefficients is a noncommutative ring when $n \geq 2$. It is of course the same as $\mathbb{Z}$ when $n = 1$. We denote an $n \times n$ matrix by $[a_{ij}]$. Then addition and multiplication are defined as follows:

$$\begin{aligned} [a_{ij}] + [b_{ij}] &= [a_{ij} + b_{ij}], \\ [a_{ij}] \cdot [b_{ij}] &= [\sum_k a_{ik} b_{kj}]. \end{aligned}$$

What are the groups of units in these examples? For these latter four examples, one can replace $\mathbb{Z}$ by any ring $R$ and obtain rings $R[i]$, $R[x]$, $R[[x]]$ and $M_n(R)$.

If $R$ and $S$ are two rings with operations $0_R, 1_R, +_R, \cdot_R, 0_S, 1_S, +_S, \cdot_S$, then one can put a ring structure on the set $R \times S = \{(r, s) : r \in R, s \in S\}$ by

$$\begin{aligned} (r_1, s_1) + (r_2, s_2) &= (r_1 +_R r_2, s_1 +_S s_2), \\ (r_1, s_1) \cdot (r_2, s_2) &= (r_1 \cdot_R r_2, s_1 \cdot_S s_2), \\ 0_{R \times S} &= (0_R, 0_S), \\ 1_{R \times S} &= (1_R, 1_S). \end{aligned}$$

## 1.3 Fields and integral domains

**Definition 1.4.** A **field** is a commutative ring $R$ where $1 \neq 0$ and where every nonzero element is a unit. That is, $R^\times = R - \{0\}$.

Which of the above rings are fields? Not many: $\mathbb{Q}$, $\mathbb{Z}/p\mathbb{Z} =: \mathbb{F}_p$ when $p$ is a prime number. The ring $\mathbb{Z}[i]$ is a not a field, but $\mathbb{Q}[i]$ is as we can set

$$(a + bi)^{-1} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2} i.$$

If $F_1$ and $F_2$ are two fields, is $F_1 \times F_2$ a field?

**Definition 1.5.** An nonzero element $a$ of a commutative ring $R$ is a **zero-divisor** if there is a nonzero element $b$ such that $ab = 0$. An **integral domain** is a commutative ring with no zero-divisors.

Examples of integral domains: $\mathbb{Z}$, $\mathbb{Z}[i]$, $\mathbb{Z}[x]$. How to prove it? It is obvious for $\mathbb{Z}$. For $\mathbb{Z}[i]$, we define the norm of an element by $N(a+bi) = a^2+b^2$ and so an element is 0 if and only if its norm is 0. It is then easy to check that if $\alpha, \beta \in \mathbb{Z}[i]$, then $N(\alpha\beta) = N(\alpha)N(\beta)$. Hence if $\alpha\beta = 0$, then either $\alpha$ or $\beta$ is 0.

For $\mathbb{Z}[x]$ or more generally $R[x]$ where $R$ is an integral domain, we define the degree of a nonzero polynomial $f(x) = \sum_i a_i x^i$, denoted $\deg(f)$, by the largest $n$ such that $a_n \neq 0$. The convention is to define the degree of the zero polynomial to be $-\infty$. The first nonzero term $a_n x^n$ is called the leading term with $a_n$ called the leading coefficient. If the leading coefficient is 1, we say the polynomial is monic. Then since $R$ is an integral domain, we see that $\deg(fg) = \deg(f) + \deg(g)$. The degree can be used to measure the "size" of a polynomial as we can see from the following division algorithm.

**Proposition 1.6.** Let $R$ be a ring and let $f(x) \in R[x]$ be nonzero such that the leading coefficient of $f$ is a unit. Then for any $g(x) \in R[x]$, there exists polynomials $q(x), r(x) \in R[x]$ such that

$$g(x) = f(x)q(x) + r(x)$$

with $\deg(r) < \deg(f)$.

Proof by induction on the degree of $g$. □

**Proposition 1.7.**

1. A field is an integral domain.
2. A finite integral domain is a field.

To prove the first statement, let $F$ be a field and let $a$ be an arbitrary nonzero element. Suppose $ab = 0$ for some $b \in F$. Multiplying by $a^{-1}$ gives $b = 0$ and so $a$ is not a zero-divisor.

For the second statement, let $R$ be a finite integral domain and let $a$ be an arbitrary nonzero element. We need to show there exists some $b \in R$ such that $ab = 1$. Let $\varphi_a$ denote the map $R \to R$ sending $r$ to $ar$. If $\varphi_a(r) = \varphi_a(s)$ for $r, s \in R$, then $ar = as$ and so $a(r-s) = 0$. Since $a$ is not a zero-divisor, we get $r = s$. Hence $\varphi_a$ is injective. Since $R$ is finite, we see that $\varphi_a$ is also surjective. (For if otherwise, say $R$ has $n$ elements, then $\varphi_a$ sends $n$ elements to less than $n$ images. By the pigeonhole principle, there exists at least 2 elements having the same image under $\varphi_a$ contradicting the injectivity of $\varphi_a$.) In particular, the element 1 is in the image of $\varphi_a$. Hence there exists $b$ such that $ab = \varphi_a(b) = 1$. □

## 1.4 Homework

**Exercise 1.4.1.** Let $R$ be a ring where 0 is a unit. Describe the ring $R$. That is, what are the elements, how are the binary operations defined, what are 0 and 1?

**Exercise 1.4.2.** Let $R$ be a ring. Prove that $R^\times$ along with multiplication and 1 forms a group and that when $R$ is commutative, $R^\times$ is an abelian group.

**Exercise 1.4.3.** Prove the following basic properties of a ring.

1. For any $a, b \in R$, $(-a)b = -(ab) = a(-b)$;
2. For any $a, b \in R$, $(-a)(-b) = ab)$;
3. For any $a \in R$, $0 \cdot a = a \cdot 0 = 0$.

**Exercise 1.4.4.** Show that $\mathbb{Z}[i]^\times = \{1, -1, i, -i\}$ and $R[x]^\times = R^\times$ if $R$ is an integral domain. Is the statement $R[x]^\times = R^\times$ still true when $R$ is not an integral domain? Prove it if yes, give a counterexample if no.

**Exercise 1.4.5.** Show that for any ring $R$,

$$R[[x]]^\times = \{\sum_{i=0}^{\infty} a_i x^i : a_0 \in R^\times\}.$$

# 2 Homomorphisms and Ideals

In the last proof of last section, we made use a map $\varphi_a$ between rings even though it totally disrespects the ring operations on both sides. A map that does preserve all the ring operations is called a homomorphism.

**Definition 2.1.** A map $\varphi : R \to T$ between two rings is a **homomorphism** if:

1. for any $a, b \in R$, $\varphi(a+b) = \varphi(a) + \varphi(b), \varphi(ab) = \varphi(a)\varphi(b)$,
2. $\varphi(1) = 1$.

A homomorphism is called an **isomorphism** if it is a bijection. Or equivalently, if there exists a homomorphism $\phi : T \to R$ such that $\varphi \circ \phi = \mathrm{id}_T$ and $\phi \circ \varphi = \mathrm{id}_R$. When $R = T$, an isomorphism between them is also called an **automorphism**.

Note it follows from these conditions that $\varphi(0) = 0$. One can also check that if $\varphi$ is an isomorphism then $\varphi^{-1}$ is also an isomorphism. The composite of two homomorphisms is a homomorphism.

## 2.1 Examples of homomorphisms

The map from $\mathbb{Z}$ to $\mathbb{Z}/m\mathbb{Z}$ sending an integer $a$ to its congruence class $[a]$ modulo $m$ is a ring homomorphism. In fact, for any ring $R$, there is a unique homomorphism from $\mathbb{Z}$ to $R$. Why? Because 1 has to go to 1 and 2 has to go to $1+1$ and so on. The distributive law implies that this map also preserves multiplication.

Let $\sqrt{-1}$ denote a solution to $x^2 + 1 = 0$ in $\mathbb{C}$. Then the map $\mathbb{Z}[i] \to \mathbb{C}$ sending $a + bi$ to $a + b\sqrt{-1}$ is a homomorphism.

The inclusion $R \to R[x]$ viewing an element of $R$ as a constant polynomial is a homomorphism. In general, a subset $S$ or a ring $R$ is called a **subring** of $R$ if $S$ is a ring and the natural inclusion $S \to R$ is a homomorphism. Equivalently, a subring is a subset of $R$ closed under all the ring operations, including the unary operations 0 and 1.

The map $\mathbb{Z}[x] \to \mathbb{Z}[i]$ sending a polynomial $f(x)$ to $f(i)$ is a homomorphism. More generally, given any ring homomorphism $\varphi : R \to R'$ and an element $\alpha \in R'$, there is a unique ring homomorphism $\widetilde{\varphi} : R[x] \to R'$ extending $\varphi$ and sending $x$ to $\alpha$:

$$\widetilde{\varphi}(\sum_i a_i x^i) = \sum_i \varphi(a_i)\alpha^i.$$

A more-or-less trivial application of this extension property shows that any homomorphism $\varphi : R \to R'$ of rings extend to a unique homomorphism $\varphi[x] : R[x] \to R'[x]$ sending $x$ to $x$. Indeed, apply the above property to the composite $R \to R' \to R'[x]$ and $\alpha = x \in R'[x]$.

A less trivial application is the following isomorphism

$$(R[x])[y] \cong (R[y])[x].$$

Indeed, the inclusion $R \to R[y]$ extends to $R[x] \to (R[y])[x]$ sending $x$ to $x$. Apply the above property with $\alpha = y \in (R[y])[x]$ gives the extension $(R[x])[y] \to (R[y])[x]$. Likewise one can construct a map in the other direction. The composition of these two maps is a homomorphism on, say, $(R[x])[y]$ sending $x$ to $x$ and $y$ to $y$. By uniqueness, such a map is the identity map. Because of this isomorphism, we write $R[x, y]$ instead and more generally $R[x_1, \ldots, x_n]$.

Are the rings $(R[x])[[y]]$ and $(R[[y]])[x]$ isomorphic?

The Chinese Remainder Theorem states that when $m = ab$ and $\gcd(a, b) = 1$, the map $\mathbb{Z}/m\mathbb{Z} \to \mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z}$ sending the congruence class $[n]_m$ modulo $m$ to $([n]_a, [n]_b)$ is a bijection. It is easy to check this is a ring homomorphism and so is an isomorphism.

## 2.2 Ideals

**Definition 2.2.** For any ring homomorphism $\varphi : R \to T$, its **kernel**, denoted $\ker \varphi$, is the preimage of 0. That is,

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

The kernel of a ring homomorphism tells us whether the map is injective or not.

**Proposition 2.3.** Let $\varphi : R \to T$ be a ring homomorphism. Then $\varphi$ is injective if and only if $\ker \varphi = \{0\}$.

For any $a, b \in R$, $\varphi(a) = \varphi(b)$ if and only if $\varphi(a - b) = 0$ if and only if $a - b \in \ker \varphi$. □

What properties does $\ker \varphi$ have?

1. $0 \in \ker \varphi$.

2. If $a, b \in \ker \varphi$, then $a + b \in \ker \varphi$.

3. If $a \in \ker \varphi$ and $r \in R$, then $ra, ar \in \ker \varphi$.

The first two properties are not surprising since $\ker \varphi$ is the kernel of $\varphi$ when viewed as a group homomorphism.

**Definition 2.4.** A nonempty subset $I$ of a ring $R$ is an **ideal** if:

1. For any $a, b \in I$, $a + b \in I$;

2. For any $a \in I$ and $r \in R$, $ar, ra \in I$. In particular, when $r = 0$, we get $0 \in I$.

### Examples

1. There are two trivial ideals in every ring: $\{0\}$ and $R$. A **proper** ideal is an ideal strictly contained in $R$.

2. When $R$ is commutative, for any $a \in R$, the set $aR = \{ar : r \in R\}$ is an ideal of $R$. Sometimes it is denoted $(a)$. It is also called the ideal generated by $a$ since it is the smallest ideal of $R$ containing $a$. If $I$ is an ideal and $a \in I$, then $I \supset (a)$. Likewise, we have the notations $(a_1, \ldots, a_n) = a_1R + \cdots a_nR$ and $(I, a) = I + aR$ if $I$ is an ideal. An ideal that can be expressed as $(a)$ for some $a \in R$ is called a **principal ideal**. An integral domain in which every ideal is a principal ideal is called a **principal ideal domain** or **PID**.

3. Every ideal of $\mathbb{Z}$ has the form $a\mathbb{Z}$ for some nonnegative integer $a$. Hence for any ring $R$, the kernel of the natural map $\mathbb{Z} \to R$ has the form $m\mathbb{Z}$ for some nonnegative integer $m$, called the **characteristic** of $R$. The characteristic of an integral domain is either 0 or a prime number.

4. If $F$ is a field, then every nonzero ideal of $F[x]$ has the form $(f(x))$ for some (monic) polynomial $f(x)$. Indeed, for any nonzero ideal, let $f(x)$ be an element in it of minimal degree then apply division to show every element is a multiple of $f(x)$.

5. What are the ideals of $F[[x]]$? Answer: $(x^n)$ for nonnegative integers $n$.

6. The ideal $(2, x)$ of $\mathbb{Z}[x]$ is not principal. Hence $\mathbb{Z}[x]$ is not a principal ideal domain.

7. Not every ideal is finitely generated. For example, take the ring $\mathbb{C}[x_1, x_2, \ldots]$, that is, a polynomial with infinitely many variables. Its elements are finite sums of finite products of the variables. The ideal $(x_1, x_2, \ldots)$ generated by all the variables if not finitely generated. A ring where every ideal is finitely generated is called **Noetherian**. Polynomial rings over a field of finitely many variables are Noetherian. We will come back to this later.

8. An ideal $I$ equals $R$ if and only if it contains a unit. As a consequence, a ring is a field if and only if it contains only two ideals. Furthermore, every homomorphism from a field to a ring is injective.

**Operations on ideals**

Given two ideals $I, J$ of a commutative ring $R$, we can take

$$\begin{aligned} I + J &= \{a + b : a \in I, b \in J\}, \\ I \cap J &= \{a : a \in I, a \in J\}, \\ IJ &= \{\sum_{i=1}^{n} a_i b_i : a_i \in I, b_i \in J\} \subset I \cap J. \end{aligned}$$

It is not hard to check that all three are ideals.

It is now very tempting to think of the set of all ideals equipped with addition multiplication as a ring except there are no additive inverses. We can however mimic the definition of prime numbers. We say an ideal $I$ divides an ideal $J$, and write $I \mid J$, if $I \supset J$. (A number $n$ divides $m$ if and only if $(n) \supset (m)$.)

**Definition 2.5.** Let $R$ be a commutative ring. A proper ideal $\mathfrak{p}$ of $R$ is a **prime ideal** if whenever $\mathfrak{p} \mid IJ$ for ideals $I, J$, either $\mathfrak{p} \mid I$ or $\mathfrak{p} \mid J$.

We are probably more familiar with elements than ideals. The following proposition translates the definition of a prime ideal in terms of elements.

**Proposition 2.6.** Let $R$ be a commutative ring and let $\mathfrak{p}$ be a proper ideal. Then $\mathfrak{p}$ is a prime ideal if and only if whenever $ab \in \mathfrak{p}$ for elements $a, b \in R$, either $a \in \mathfrak{p}$ or $b \in \mathfrak{p}$.

Suppose first $\mathfrak{p}$ is a prime ideal and $a, b \in R$ with $ab \in \mathfrak{p}$. Then $\mathfrak{p} \supset (ab) = (a)(b)$ and so by definition, either $\mathfrak{p} \supset (a)$ or $\mathfrak{p} \supset (b)$. Hence either $a \in \mathfrak{p}$ or $b \in \mathfrak{p}$.

Conversely, suppose $\mathfrak{p}$ is a proper ideal such that whenever $ab \in \mathfrak{p}$ for elements $a, b \in R$, either $a \in \mathfrak{p}$ or $b \in \mathfrak{p}$. Let $I, J$ be two ideals such that $\mathfrak{p} \supset IJ$. If $\mathfrak{p}$ does not contain either $I$ or $J$, then there exists elements $a \in I$ and $b \in J$ such that $a \notin \mathfrak{p}$ and $b \notin \mathfrak{p}$. Hence $ab \notin \mathfrak{p}$ but $ab \in IJ \subset \mathfrak{p}$. Contradiction. □

There is another definition of prime numbers: an integer greater than 1 with no proper divisors. The analogue definition for ideals is:

**Definition 2.7.** Let $R$ be a ring. A proper ideal $\mathfrak{m}$ of $R$ is a **maximal ideal** if whenever $I \mid \mathfrak{m}$ for some ideal $I$, either $I = R$ or $I = \mathfrak{m}$. In other words, there are no proper ideals strictly larger than $\mathfrak{m}$.

For example, the ideal $(x)$ of $\mathbb{Z}[x]$ is a prime ideal while the ideal $(2, x)$ is a maximal ideal. Next time we will see a characterization of prime and maximal ideals in terms of their quotients.

## 2.3 Homework

**Exercise 2.3.1.** (11.3.7) Determine all the automorphisms of $\mathbb{Z}[x]$.

**Exercise 2.3.2.** (11.3.9) (a) An element of a ring $R$ is called *nilpotent* if some power is zero. Prove that if $x$ is nilpotent, then $1 + x$ is a unit.
(b) Suppose that $R$ has prime characteristic $p \neq 0$. Prove that if $a$ is nilpotent, then $1 + a$ is unipotent, that is, some power of $1 + a$ is 1.

**Exercise 2.3.3.** (11.6.8) Let $I$ and $J$ be ideals of a commutative ring $R$ such that $I + J = R$.

1. Prove that $IJ = I \cap J$.

2. (Chinese Remainder Theorem) Prove that for any pair $a, b$ of elements of $R$, there is an element $x$ such that $x - a \in I$ and $x - b \in J$.

**Exercise 2.3.4.** Let $M$ be an abelian group with addition $+$. Let $\mathrm{End}(M)$ be the set of group homomorphisms from $M$ to $M$. Give $\mathrm{End}(M)$ the structure of a ring.

**Exercise 2.3.5.** Show that in a commutative ring, every maximal ideal is a prime ideal. Show that in a PID, every prime ideal is a maximal ideal.

It is not a priori clear that maximal ideals exist. To show its existence, we need Zorn's Lemma from logic theory.

**Definition 2.8.** A **partially order set** is a set $S$ along with an ordering $\leq$ such that:

1. $a \leq a$,

2. If $a \leq b$ and $b \leq c$, then $a \leq c$,

3. If $a \leq b$ and $b \leq a$, then $a = b$.

An element $a$ of $S$ is **maximal** if whenever $a \leq b$ for some $b \in S$, we have $a = b$.

For example, if $S$ is the set of proper ideals of a commutative ring $R$ and we define $I \leq J$ if $I \subset J$, then a maximal element of $S$ is the same as a maximal ideal of $R$.

**Definition 2.9.** A subset $T$ of a partially ordered set $S$ is **totally ordered** if for any $a, b \in T$, either $a \leq b$ or $b \leq a$. An **upper bound** of a subset $T$ of a partially order set $S$ is an element $a \in S$ such that $b \leq a$ for every $b \in T$.

**Exercise* 2.3.6.** Let $S$ be the set of proper ideals of a commutative ring $R$ and define $I \leq J$ if $I \subset J$. Let $T$ be a totally ordered subset of $S$. Show that $\bigcup_{I \in T} I$ is a proper ideal of $R$.

**Theorem 2.10.** (Zorn's Lemma) A nonempty partially order set has a maximal element if every totally ordered subset has an upper bound.

If $R$ is not the zero ring, the zero ideal $\{0\}$ is a proper ideal and by Exercise 2.3.5, every totally ordered subset has an upper bound. Therefore by Zorn's Lemma, every nonzero ring has a maximal ideal.

**Definition 2.11.** Let $R$ be a commutative ring. The set of all nilpotent element of $R$ is called the *nilradical* of $R$, denoted $\mathrm{Nil}(R)$.

**Exercise* 2.3.7.** Let $R$ be a commutative ring. Show that $\mathrm{Nil}(R)$ is the intersection of all the prime ideals of $R$. (In particular, it is an ideal.) Hint: The nontrivial part is proving that if $a$ is not nilpotent, then there is a prime ideal not containing it. Let $S$ be the partially ordered set of ideals $I$ that do not contain any (positive) power of $a$ and define $I \leq J$ if $I \subset J$. Show that $S$ satisfies the hypothesis of Zorn's Lemma. Let $\mathfrak{p}$ be a maximal element of $S$ and show that it is a prime ideal.

# 3 Quotients

Let $R$ be a ring and let $I$ be an ideal. Viewing $I$ as an abelian subgroup of $R$, we can form the quotient group $R/I$ consisting of cosets $a + I$. The extra ring structure of $R$ gives $R/I$ a structure of a ring with multiplication given by $(a + I)(b + I) = ab + I$ and the multiplicative identity given by $1 + I$. To check that multiplication is well-defined, we need to show that if $a' + I = a + I$ and $b' + I = b + I$, then $ab + I = a'b' + I$. This follows from the fact that ideals are closed under multiplication by elements of the ring. There is a natural map, which is a priori a group homomorphism, $\pi : R \to R/I$ sending $a$ to $a + I$. The above ring structure on $R/I$ is the unique ring structure such that $\pi$ is a ring homomorphism.

## 3.1 Universal property

Another way to characterize the quotient ring is to say that every ring homomorphism out of $R$ with kernel containing $I$ factors uniquely through $R/I$.

**Theorem 3.1.** Let $\varphi : R \to R'$ be a ring homomorphism such that $\ker\varphi \supset I$. Then there exists a unique homomorphism $\bar{\varphi} : R/I \to R'$ such that $\bar{\varphi}\pi = \varphi$.

$$\begin{tikzcd}
R \arrow[rr, "\varphi"] \arrow[dr, "\pi"'] & & R' \\
& R/I \arrow[ur, dashed, "\exists!\bar{\varphi}"'] &
\end{tikzcd}$$

Uniqueness is clear. Existence is proved by showing the map $\bar{\varphi} : R/I \to R'$ sending $a + I$ to $\varphi(a)$ is a well-defined ring homomorphism. □

**Theorem 3.2.** (Isomorphism Theorem) Let $\varphi : R \to R'$ be a surjective ring homomorphism. Then $\bar{\varphi}$ induces an isomorphism $R/\ker\varphi \cong R'$. More generally, if $\varphi : R \to R'$ is a ring homomorphism, then $R/\ker\varphi \cong \operatorname{im}(\varphi)$.

**Theorem 3.3.** (Correspondence Theorem) Let $R$ be a ring with ideal $I$. Then there is a bijection between the set of ideals of $R$ containing $I$ and the set of ideals of $R/I$ given by taking image and preimage of the natural map $\pi : R \to R/I$. Moreover, if an ideal $J$ of $R$ containing $I$ corresponds to an ideal $J/I := \pi(J)$ of $R/I$, then $R/J \cong (R/I)/(J/I)$.

The proof is a simple definition chase. Here are the necessary points to check:

1. If $J$ is an ideal of $R$ containing $I$, then $J/I$ is an ideal of $R/I$.
2. If $\mathscr{I}$ is an ideal of $R/I$, then $\pi^{-1}(\mathscr{I})$ is an ideal of $R$ containing $I$.
3. If $J$ is an ideal of $R$ containing $I$, then $\pi^{-1}(J/I) = J$. If $\mathscr{I}$ is an ideal of $R/I$, then $\pi^{-1}(\mathscr{I})/I = \mathscr{I}$.
4. If $J$ is an ideal of $R$ containing $I$, then $R/J \cong (R/I)/(J/I)$.

For the last one, take the natural map $R \to R/I \to (R/I)/(J/I)$ and show that the kernel is $J$. □

### Example: Computing quotients

Let us consider the map $\varphi : \mathbb{Z}[x] \to \mathbb{Z}[i]$ sending $f(x)$ to $f(i)$. What is the kernel of this map? Suppose instead we consider the map $\widetilde{\varphi} : \mathbb{Q}[x] \to \mathbb{Q}[i]$ defined the same way. Then we know since $\mathbb{Q}[x]$ is a PID, $\ker\widetilde{\varphi}$ is the ideal generated by a monic polynomial of minimal degree in it. By definition of $\mathbb{Q}[i]$, no linear polynomial vanishes at $i$ and the quadratic polynomial $x^2 + 1$ vanishes at $i$. Hence $\ker\widetilde{\varphi} = (x^2 + 1)$ and so $\ker\varphi = (x^2 + 1)\mathbb{Q}[x] \cap \mathbb{Z}[x]$. For any $f(x) \in \mathbb{Z}[x]$, since the leading coefficient of $x^2 + 1$ is a unit in $\mathbb{Z}$, we may apply the division algorithm to get a quotient $q(x) \in \mathbb{Z}[x]$ and a remainder $r(x) \in \mathbb{Z}[x]$ such that

$f(x) = (x^2+1)q(x) + r(x)$. Now if $f(x) \in (x^2+1)\mathbb{Q}[x]$, then $r(x) = 0$ and so $f(x) \in (x^2+1)\mathbb{Z}[x]$. Hence $\ker\varphi = (x^2+1)\mathbb{Z}[x]$. Therefore by the Isomorphism theorem,

$$\mathbb{Z}[x]/(x^2+1) \cong \mathbb{Z}[i].$$

Now $i-2 \in \mathbb{Z}[i]$ generates an ideal $(i-2)$. What is the quotient $\mathbb{Z}[i]/(i-2)$? Under the correspondence between ideals of $\mathbb{Z}[x]$ containing $(x^2+1)$ and ideals of $\mathbb{Z}[x]/(x^2+1)$, the ideal $(i-2)$ of $\mathbb{Z}[i]$ corresponds to the ideal $(x-2, x^2+1)$ of $\mathbb{Z}[x]$ and so

$$\mathbb{Z}[i]/(i-2) \cong \mathbb{Z}[x]/(x-2, x^2+1).$$

To compute $\mathbb{Z}[x]/(x-2, x^2+1)$, we may quotient out by $(x-2)$ first and get a natural map $\pi : \mathbb{Z}[x] \to \mathbb{Z}[x]/(x-2) \cong \mathbb{Z}$. Note the last isomorphism is obtained from applying the isomorphism theorem to the map $\mathbb{Z}[x] \to \mathbb{Z}$ sending $x$ to 2. The ideal $(x-2, x^2+1)$ of $\mathbb{Z}[x]$ containing $(x-2)$ corresponds to the ideal $\pi((x-2, x^2+1)) = (5)$ of $\mathbb{Z}$. Therefore,

$$\mathbb{Z}[i]/(i-2) \cong \mathbb{Z}[x]/(x-2, x^2+1) \cong \mathbb{Z}/5\mathbb{Z}.$$

**Example: Adjoining elements**

Let $\varphi : R \to R'$ be a ring homomorphism and let $\alpha$ be an element of $R'$. Let $\widetilde{\varphi} : R[x] \to R'$ denote the extension of $\varphi$ sending $x$ to $\alpha$. We denote the image of $\widetilde{\varphi}$ by $R[\alpha]$. Then $R[\alpha]$ is the smallest subring of $R'$ containing $\alpha$ and the image of $\varphi$. It consists of elements of the form $(\varphi[x])(g)(\alpha)$ for polynomials $g \in R[x]$. If $I$ is the kernel of $\widetilde{\varphi}$, then $R[\alpha] \cong R[x]/I$.

Take $\varphi$ to be the natural inclusion of $R = \mathbb{Z}$ in $R' = \mathbb{C}$ and $\alpha = \sqrt{-1}$. Then we get the ring $\mathbb{Z}[\sqrt{-1}] \cong \mathbb{Z}[x]/(x^2+1) \cong \mathbb{Z}[i]$. Since $\sqrt{-1}^2 = -1$, we see that we only need to take polynomials $g \in \mathbb{Z}[x]$ of degree at most 1 in order for $g(\sqrt{-1})$ to sweep out $\mathbb{Z}[\sqrt{-1}]$. If we take $\alpha = \sqrt[3]{2}$, then we get the ring $\mathbb{Z}[\sqrt[3]{2}] \cong \mathbb{Z}[x]/(x^3-2)$. To compute this ideal $(x^3-2)$, we work again over $\mathbb{Q}$ using the fact that $x^3-2$ has no rational root. What polynomials $g \in \mathbb{Z}[x]$ do we need to get all the elements of $\mathbb{Z}[\sqrt[3]{2}]$ via $g(\sqrt[3]{2})$?

We can also take $\varphi$ to be the natural inclusion of $R = \mathbb{Z}$ in $R' = \mathbb{Q}$ and $\alpha = \frac{1}{2}$. Then we get the ring $\mathbb{Z}[\frac{1}{2}]$ which consists of elements of the form $a/2^n$ for integers $a, n$. The kernel of the map $\mathbb{Z}[x] \to \mathbb{Z}[\frac{1}{2}]$ is the ideal $(2x-1)$. So we have $\mathbb{Z}[\frac{1}{2}] \cong \mathbb{Z}[x]/(2x-1)$. To get all the elements of $\mathbb{Z}[\frac{1}{2}]$, we have to use polynomials of arbitrary degree.

In some sense, $\mathbb{Z}[\sqrt{-1}]$ and $\mathbb{Z}[\sqrt[3]{2}]$ are finite over $\mathbb{Z}$ while $\mathbb{Z}[\frac{1}{2}]$ is infinite over $\mathbb{Z}$. We will define later in what sense.

**Definition 3.4.** Fix a ring homomorphism $\varphi : R \to R'$. An element $\alpha \in R'$ is **integral** over $R$ if there is a positive integer $N$ such that

$$R[\alpha] = \{(\varphi[x])(g)(\alpha) : g \in R[x], \deg(g) \leq N\} = \varphi(R) + \varphi(R)\alpha + \varphi(R)\alpha^2 + \cdots + \varphi(R)\alpha^N.$$

Equivalently, $\alpha \in R'$ is **integral** over $R$ if there is a monic polynomial $f \in R[x]$ such that $(\varphi[x])(g)(\alpha) = 0$.

**Definition 3.5.** If $\varphi$ is the natural inclusion of $\mathbb{Z}$ or $\mathbb{Q}$ in $\mathbb{C}$, then a complex number $\alpha$ is an **algebraic number** if it is integral over $\mathbb{Q}$; an **algebraic integer** if it is integral over $\mathbb{Z}$.

Algebraic number theory is the theory of algebraic numbers. What a surprise!

Sometimes we want to add an element $\alpha$ to a ring $R$ but we don't have a priori an ambient ring $R'$ that contains $\alpha$ or if we don't want to specify $R'$. For example, what if we want to adjoint the inverse of $[3]$ to $\mathbb{Z}/6\mathbb{Z}$? The above construction suggests that if we have a polynomial $f(x) \in R[x]$ that we want $\alpha$ to be a root of, then we can form $R[\alpha]$ by taking $R[x]/(f(x))$. In our example, we can take $\mathbb{Z}/6\mathbb{Z}[[3]^{-1}]$ to be $\mathbb{Z}/6\mathbb{Z}[x]/(3x-1)$.

What is the ring $\mathbb{Z}/6\mathbb{Z}[x]/(3x-1)$? Chinese Remainder Theorem gives $\mathbb{Z}/6\mathbb{Z} \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ and from this one obtains $\mathbb{Z}/6\mathbb{Z}[x]/(3x-1) \cong \mathbb{Z}/2\mathbb{Z}$. Hence the natural map $R \to R[\alpha]$ is not necessarily injective. The problem is that the intersection $(f(x)) \cap R$ in $R[x]$ might not be 0. It will be 0 if $f$ is monic. Generally when we adjoint elements, we would prefer the polynomial $f$ to be monic.

## 3.2 Prime ideals and Maximal ideals

**Theorem 3.6.** Let $R$ be a commutative ring and let $I$ be an ideal. Then the correspondence between ideals of $R$ containing $I$ and ideals of $R/I$ restrict to a bijection between prime ideals and between maximal ideals. As a consequence:

1. $I$ is a prime ideal if and only if $R/I$ is an integral domain;
2. $I$ is a maximal ideal if and only if $R/I$ is a field.
3. Every proper ideal $I$ is contained in a maximal ideal.

It is worth remarking that a commutative ring $R$ is an integral domain if and only if the zero ideal $(0)$ is a prime ideal. The rest of this theorem is easy.

### 3.2.1 Examples

1. The maximal ideals of $\mathbb{Z}$ are the principal ideals generated by prime numbers.
2. The maximal ideals of $F[x]$ where $F$ is a field are the principal ideals generated by monic irreducible polynomials. If $F$ is algebraically closed, that is if every nonconstant nonzero polynomial has a root in $F$, then the irreducible polynomials are the linear polynomials. So there is a bijection between maximal ideals of $F[x]$ and $F$. In particular, the maximal ideals of $\mathbb{C}[x]$ are $(x-\lambda)$ for $\lambda \in \mathbb{C}$.
3. What about $\mathbb{C}[x,y]$? (Next time.)
4. The ideal $(6, 3x-1)$ is maximal in $\mathbb{Z}[x]$ since the quotient $\mathbb{Z}/2\mathbb{Z}$ is a field.
5. What are all the maximal ideals of $\mathbb{Z}[x]$? Let $I$ be a maximal ideal of $\mathbb{Z}[x]$. Then $I \cap \mathbb{Z}$ is an ideal of $\mathbb{Z}$. In fact it is a nonzero ideal, for if $I \cap \mathbb{Z} = 0$, then for any nonzero integer $n$, the ideal $(I, n)$ is proper and contains $I$. Hence there is a (positive) integer $m$ such that $I \cap \mathbb{Z} = m\mathbb{Z}$. Since $I \supset (m)$, we see by the correspondence theorem that it is sufficient to classify maximal ideals of $\mathbb{Z}[x]/(m) = \mathbb{Z}/m\mathbb{Z}[x]$. By the Chinese Remainder Theorem and Exercise 3.3.5, it remains to classify maximal ideals of $\mathbb{Z}/p^a\mathbb{Z}[x]$ where $p$ is a prime number and $a$ is a positive integer. Let $J$ be a maximal ideal of $\mathbb{Z}/p^a\mathbb{Z}[x]$. Then as above, $\mathbb{Z}/p^a\mathbb{Z} \cap J$ is an ideal of $\mathbb{Z}/p^a\mathbb{Z}$ and there is an injection

   $$\mathbb{Z}/p^a\mathbb{Z}/(\mathbb{Z}/p^a\mathbb{Z} \cap J) \hookrightarrow \mathbb{Z}/p^a\mathbb{Z}[x]/J.$$

   Since $J$ is a maximal ideal and hence a prime ideal, the quotient has no zero-divisors and so $\mathbb{Z}/p^a\mathbb{Z}/(\mathbb{Z}/p^a\mathbb{Z} \cap J)$ also has no zero-divisors. In other words, it is an integral domain and so $\mathbb{Z}/p^a\mathbb{Z} \cap J$ is a prime ideal of $\mathbb{Z}/p^a\mathbb{Z}$. Again by the correspondence theorem, prime ideals of $\mathbb{Z}/p^a\mathbb{Z}$ are in bijection with prime ideals of $\mathbb{Z}$ containing $p^a\mathbb{Z}$. There is only one such prime ideal, namely $p\mathbb{Z}$. Hence $\mathbb{Z}/p^a\mathbb{Z} \cap J = p\mathbb{Z}/p^a\mathbb{Z}$, or by an abuse of notation, $p \in J$. Again as above, it remains to consider maximal ideals of the quotient $\mathbb{Z}/p^a\mathbb{Z}[x]/(p) = \mathbb{F}_p[x]$ which are the principal ideals generated by irreducible polynomials. The ideal $(6, 3x-1)$ corresponds to the ideal $(x-1) \times (1)$ of $\mathbb{Z}/2\mathbb{Z}[x] \times \mathbb{Z}/3\mathbb{Z}[x]$.

## 3.3 Homework

**Exercise 3.3.1.** Let $\varphi : R \to R'$ be a fixed ring homomorphism. Show that if $\alpha, \beta \in R'$ are integral over $R$, then for any $f(x,y) \in R[x,y]$, the element $f(\alpha,\beta)$ is also integral over $R$. Note by $f(\alpha,\beta)$, we mean first applying $\varphi$ to all the coefficients of $f$ to turn it into a polynomial over $R'$ and then evaluating it at $(\alpha,\beta)$. As a consequence, the subset $\bar{R}$ of $R'$ consisting of integral elements over $R$ is a subring.

**Exercise 3.3.2.** Let $\varphi : R \to R'$ be a ring homomorphism. Suppose $\alpha \in R'$ is integral over $R$ and suppose $\beta \in R'$ is integral over $R[\alpha]$. Show that $\beta$ is integral over $R$.

**Exercise 3.3.3.** (11.6.7) Show that the quotient ring $\mathbb{Z}[x]/(2x)$ is isomorphic to the subring of $\mathbb{F}_2[x] \times \mathbb{Z}$ of pairs $(f(x), n)$ such that $f(0) \equiv n \pmod 2$.

**Exercise 3.3.4.** Let $R$ be a commutative ring and let $I$ be an ideal of $R$. Define the radical of $I$, denoted $\sqrt{I}$, to be

$$\sqrt{I} := \{r \in R : r^n \in I \text{ for some } n \geq 1\}.$$

Show that $\sqrt{I}$ is the intersection of all the prime ideals containing $I$.

**Exercise 3.3.5.** Let $R_1$ and $R_2$ be two rings. Show that all the maximal ideals of $R_1 \times R_2$ are given by $\mathfrak{m}_1 \times R_2$ and $R_1 \times \mathfrak{m}_2$ for some maximal ideals $\mathfrak{m}_1, \mathfrak{m}_2$ of $R_1, R_2$.

**Exercise 3.3.6.** (11.M.3) Let $R$ denote the set of sequences $a = (a_1, a_2, a_3, \ldots)$ of real numbers that are eventually constant. Addition and multiplication are defined componentwise. Prove that $R$ is a ring and determine its maximal ideals.

# 4 Localization

There are two motivations for the localization construction:

1. We know how to study ideals of a ring $R$ containing a fixed ideal $I$ by studying the quotient ring $R/I$. What about ideals disjoint from a given set?

2. When studying the kernel of the map $\mathbb{Z}[x] \to \mathbb{Z}[i]$, we went to $\mathbb{Q}[x]$ because $\mathbb{Q}[x]$ is a PID. What is the relationship between $\mathbb{Z}$ and $\mathbb{Q}$?

## 4.1 Construction

Let $R$ be a commutative ring. Let $S$ be a subset of $R$ such that:

$$\begin{aligned}
&0 \notin S, 1 \in S \\
&S \text{ contains no zero-divisors} \\
&S \text{ is multiplicatively closed, that is, } a, b \in S \Rightarrow ab \in S
\end{aligned} \tag{1}$$

Denote an element of $R \times S$ suggestively by $r/s$ for $r \in R$, $s \in S$. Define a relation $\sim$ on $R \times S$ by

$$\frac{r_1}{s_1} \sim \frac{r_2}{s_2} \iff r_1 s_2 = r_2 s_1.$$

It is routine to check that $\sim$ is an equivalence relation, that is it is symmetric, reflexive and transitive (this requires $S$ to contain no zero-divisors). Denote the set of equivalence classes by $R[S^{-1}]$. Then $R[S^{-1}]$ is a ring with the following ring operations:

1. $0_{R[S^{-1}]} = \dfrac{0}{1}$, $1_{R[S^{-1}]} = \dfrac{1}{1}$;
2. $\dfrac{r_1}{s_1} + \dfrac{r_2}{s_2} = \dfrac{r_1 s_2 + r_2 s_1}{s_1 s_2}$;
3. $\dfrac{r_1}{s_1}\dfrac{r_2}{s_2} = \dfrac{r_1 r_2}{s_1 s_2}$.

It is called the **localization** of $R$ away from $S$.

**Example**. Suppose $R$ is an integral domain. Then the set $S = R - \{0\}$ of nonzero elements satisfies condition (1). The localization $R[S^{-1}]$ consists of (equivalence classes of) elements of form $r/s$ where $s \neq 0$. Every nonzero element is a unit. Hence $R[S^{-1}]$ is a field, called the **field of fraction** of $R$ and denoted $\mathrm{Frac}(R)$.

The field of fraction of the ring of integers is the field of rational numbers: $\mathrm{Frac}(\mathbb{Z}) = \mathbb{Q}$. If $F$ is a field, then the field of fraction of the ring of polynomials over $F$ is the field of rational functions: $\mathrm{Frac}(F[x]) = F(x)$.

## 4.2 Universal Property

**Theorem 4.1.** Let $R$ be a commutative ring and let $S$ be a subset satisfying (1). Then the natural map $\iota : R \to R[S^{-1}]$ sending $r$ to $r/1$ is a ring homomorphism it is universal among all homomorphisms of $R$ into another (commutative) ring $T$ where elements $S$ becomes units. That is, for any homomorphism $\varphi : R \to T$ such that $\varphi(s) \in T^\times$ for all $s \in S$, there is a unique homomorphism $\phi : R[S^{-1}] \to T$ such that $\phi\iota = \varphi$.

$$\begin{tikzcd}
R \arrow[rr, "\varphi"] \arrow[dr, "\iota"'] & & T \\
& R[S^{-1}] \arrow[ur, dashed, "\exists!\phi"'] &
\end{tikzcd}$$

Proof by checking definitions. Define $\phi(r/s) = \varphi(r)\varphi(s)^{-1}$. □

We remark that $R[S^{-1}]$ is the smallest ring containing $R$ in which all the elements of $S$ are invertible. In other words, we simply adjoined the inverses of all the elements of $S$, hence the notation $R[S^{-1}]$. The ideals of $R[S^{-1}]$ should then be the same as ideals of $R$ except that if it contains an element of $S$, it will contain a unit and is thus the entire ring.

**Theorem 4.2.** There is a bijection between the set of prime ideals of $R$ disjoint from $S$ and the set of prime ideals of $R[S^{-1}]$.

$$\begin{array}{rcl}\{\text{prime ideals of } R \text{ disjoint from } S\} & \leftrightarrow & \{\text{prime ideals of } R[S^{-1}]\} \\ I & \mapsto & (I) \\ J \cap R & \leftarrow\!\shortmid & J\end{array}$$

Let $J$ be an ideal of $R[S^{-1}]$. If $J$ contains $r/s$ for some $r \in R, s \in S$, then $J$ contains $r/s \cdot s = r$. Conversely, if $J$ contains some $r \in R$, then it contains $r/s = r \cdot 1/s$ for all $s \in S$. Hence

$$J = \{\frac{r}{s} : r \in J \cap R, s \in S\}.$$

From this description of ideals of $R[S^{-1}]$, it is easy to see that the composites of the above two maps are identities. □

**Corollary 4.3.** If $R$ is a PID, then so is $R[S^{-1}]$.

Let $J$ be any ideal of $R[S^{-1}]$, then $J \cap R = rR$ for some $r \in R$ and so $J = rR[S^{-1}]$. □

**Examples**

1. Let $R$ be an integral domain (or more generally a commutative ring) and let $\mathfrak{p}$ be a prime ideal. Then the set $S = R - \mathfrak{p}$ of elements of $R$ not in $\mathfrak{p}$ satisfies condition (1). Indeed, the definition of a prime ideal says that if $ab \in \mathfrak{p}$, then either $a \in \mathfrak{p}$ or $b \in \mathfrak{p}$. Hence if $a, b \notin \mathfrak{p}$, then $ab \notin \mathfrak{p}$. The localization $R[S^{-1}]$ is usually denoted $R_{\mathfrak{p}}$, called the localization of $R$ at $\mathfrak{p}$. Proper ideals of $R_{\mathfrak{p}}$ correspond to ideals of $R$ disjoint from $S$, that is ideals of $R$ contained in $\mathfrak{p}$. There is then a unique maximal ideal, namely the ideal $\mathfrak{p}R_{\mathfrak{p}}$ generated by $\mathfrak{p}$.

   **Definition 4.4.** A **local ring** is a ring $R$ that has exactly one maximal ideal $\mathfrak{m}$. The field $R/\mathfrak{m}$ is called the **residue field**.

   Therefore, the ring $R_{\mathfrak{p}}$ is a local ring and its residue field is $\text{Frac}(R/\mathfrak{p})$.

2. Let $R$ be a ring and suppose $a \in R$ is not a zero-divisor. Then the set $S = \{a^n : n \geq 0\}$ satisfies condition (1). The localization $R[S^{-1}]$ is isomorphic to $R[a^{-1}] = R[x]/(ax - 1)$.

These two constructions are extremely important in the modern schematic language of algebraic geometry.

## 4.3 Hilbert's Nullstellensatz

The main underlying theorem in classical algebraic geometry is the following result:

**Theorem 4.5.** (Hilbert's Nullstellensatz) Maximal ideals of $\mathbb{C}[x_1, \ldots, x_n]$ are precisely the ideals $(x_1 - a_1, \ldots, x_n - a_n)$ for $a_1, \ldots, a_n \in \mathbb{C}$. In other words, there is a bijection between $\mathbb{C}[x_1, \ldots, x_n]$ and $\mathbb{C}^n$.

We first show that the ideals $(x_1 - a_1, \ldots, x_n - a_n)$ are maximal. For this it suffices to consider $a_i = 0$ for all $i$. Consider the surjective homomorphism $\mathbb{C}[x_1, \ldots, x_n] \to \mathbb{C}$ obtained by evaluating a polynomial at $(0, 0, \ldots, 0)$. It is easy to see that a polynomial $f$ is in the kernel if and only if its constant term is 0 if and only if $f \in (x_1, \ldots, x_n)$. Hence $(x_1, \ldots, x_n)$ is the kernel with quotient $\mathbb{C}$. As $\mathbb{C}$ is a field, we see that $(x_1, \ldots, x_n)$ is a maximal ideal.

Conversely, let $\mathfrak{m}$ be a maximal ideal of $\mathbb{C}[x_1, \ldots, x_n]$. Then $\mathfrak{m} \cap \mathbb{C}[x_1]$ is either the entire $\mathbb{C}[x_1]$ or a prime ideal because of the injection

$$\mathbb{C}[x_1]/(\mathbb{C}[x_1] \cap \mathfrak{m}) \hookrightarrow \mathbb{C}[x_1, \ldots, x_n]/\mathfrak{m} =: F.$$

If $\mathfrak{m} \cap \mathbb{C}[x_1] = \mathbb{C}[x_1]$ then in particular $\mathfrak{m}$ contains 1 and so is not a proper ideal. Since the nonzero prime ideals of $\mathbb{C}[x_1]$ are all of the form $(x_1 - a_1)$ for some $a_1 \in \mathbb{C}$, we see that it suffices to show $\mathfrak{m} \cap \mathbb{C}[x_1] \neq (0)$ for then $\mathfrak{m}$ will contain $(x_1 - a_1, \ldots, x_n - a_n)$ for some $a_1, \ldots, a_n \in \mathbb{C}$.

We drop the subscript 1 and suppose for a contradiction that $\mathfrak{m} \cap \mathbb{C}[x] = (0)$. Then we have an injection $\mathbb{C}[x] \hookrightarrow F$. Since $F$ is a field, this injection extends to an injection of the field of fraction of $\mathbb{C}[x]$, that is we have an injection $\mathbb{C}(x) \hookrightarrow F$. We will show this is a contradiction by showing that as a vector space over $\mathbb{C}$, $F$ has countable dimension while $\mathbb{C}(x)$ has uncountable dimension.

Indeed, $F$ is spanned as a $\mathbb{C}$-vector space by the images of the monomials $x_1^{e_1} \cdots x_n^{e_n}$ for integers $e_1, \ldots, e_n$. Hence $F$ has countable dimension over $\mathbb{C}$. For $\mathbb{C}(x)$, we claim that the set

$$\{\frac{1}{x - \lambda} : \lambda \in \mathbb{C}\}$$

is linearly independent over $\mathbb{C}$. Indeed if $c_1, \ldots, c_k, \lambda_1, \ldots, \lambda_k \in \mathbb{C}$ such that

$$\sum_{i=1}^{k} \frac{c_i}{x - \lambda_i} = 0,$$

then multiplying both sides by $\prod(x - \lambda_i)$ gives a polynomial on the left-hand-side that evaluates to $c_i \prod_{j \neq i}(\lambda_i - \lambda_j)$ at $\lambda_i$. Hence in order for this to be 0, we must have $c_i = 0$ for all $i = 1, \ldots, k$. □

Classical algebraic geometry studies the simultaneous zeros of several polynomials. In other words, given polynomials $f_1, \ldots, f_k \in \mathbb{C}[x_1, \ldots, x_n]$, what are the set of $n$-tuples $(a_1, \ldots, a_n)$ such that $f_i(a_1, \ldots, a_n) = 0$ for all $i = 1, \ldots, k$. If we set $\mathfrak{m}$ to be the maximal ideal $(x_1 - a_1, \ldots, x_n - a_n)$, then to say a polynomial $f$ vanishes at $(a_1, \ldots, a_n)$ is the same as saying $f$ is in the kernel of the evaluation map, that is $f \in \mathfrak{m}$. Therefore, classical algebraic geometry can be expressed as: given an ideal $I$ of $\mathbb{C}[x_1, \ldots, x_n]$, what are the set of maximal ideals containing $I$.

**Corollary 4.6.** Suppose $f_1, \ldots, f_k \in \mathbb{C}[x_1, \ldots, x_n]$ are such that they have no simultaneous zeros. Then there exists $g_1, \ldots, g_k \in \mathbb{C}[x_1, \ldots, x_n]$ such that $f_1 g_1 + \cdots f_k g_k = 1$.

The ideal $(f_1, \ldots, f_k)$ is not contained in any maximal ideal. Hence it is the entire ring. In particular, it contains 1. □

## 4.4 Homework

**Exercise 4.4.1.** The condition that $S$ contains no zero-divisors can be removed. Let $R$ be a commutative ring and let $S$ be a multiplicatively-closed subset containing 1 and not containing 0. Define a relation $\sim$ on $R \times S$ by

$$\frac{r_1}{s_1} \sim \frac{r_2}{s_2} \iff (r_1 s_2 - r_2 s_1) s_0 = 0$$

for some $s_0 \in S$. Show that this relation is transitive and check that the operations defined in class give the set of equivalence classes the structure of a ring.

**Exercise 4.4.2.** Let $R$ be a commutative ring and let $S$ be a multiplicatively-closed subset containing 1 and not containing 0. Suppose $I$ is an ideal of $R$ disjoint from $S$. Let $\bar{S}$ denote the image of $S$ under the natural map $R \to R/I$. Show that $R/I[\bar{S}^{-1}] \cong R[S^{-1}]/(I)$.

Modern schematic algebraic geometry studies instead the set of prime ideals of a commutative ring $R$. Let Spec($R$) denote the set of all prime ideals of $R$. ("Spec" is short for Spectrum.) We now define the **Zariski** topology on Spec($R$). For every subset $E$ of $R$, let $V(E)$ denote the subset of Spec($R$) consisting of all prime ideals of $R$ containing $E$. We say a subset of Spec($R$) is **closed** if it has the form $V(E)$ for some $E$. (A subset is open if its complement is closed.)

**Exercise* 4.4.3.** Show that the above definition of closed sets satisfies the axioms of closed sets in a topological space. That is:

1. The whole space $\operatorname{Spec}(R)$ and the empty set $\emptyset$ are closed;

2. If $V_1$ and $V_2$ are closed, then $V_1 \cup V_2$ is closed;

3. If $\{V_\lambda\}_{\lambda\in\Lambda}$ is an arbitrary set of closed sets, then the intersection $\bigcap_{\lambda\in\Lambda} V_\lambda$ is closed.

**Exercise* 4.4.4.** Show that the closure of the singleton $\{\mathfrak{p}\}$ in $\operatorname{Spec}(R)$, that is the intersection of all the closed subsets containing $\mathfrak{p}$, is $V(\mathfrak{p})$.

**Exercise* 4.4.5.** Describe the topological space $\operatorname{Spec}(\mathbb{Z})$. That is, describe (explicitly) all the points and all the open sets.

You will find that in $\operatorname{Spec}(\mathbb{Z})$, every nonempty open set contains the ideal $(0)$. Hence the singleton $\{(0)\}$ is a dense subset of $\operatorname{Spec}(\mathbb{Z})$. We can also see this from Exercise 4.4.4 as every prime ideal of $\mathbb{Z}$ contains $(0)$, so the closure of $\{(0)\}$ is the entire space.

What about open sets? For every $f \in R$, let $\operatorname{Spec}(R)_f$ denote the complement of $V((f))$. Then $\operatorname{Spec}(R)_f$ consists of prime ideals of $R$ that don't contain $f$.

**Exercise* 4.4.6.** Show that the sets $\operatorname{Spec}(R)_f$ form a basis of open sets for the Zariski topology. That is, every open set contains a set of the form $\operatorname{Spec}(R)_f$ or equivalently, for every closed set $V$, there is a set of the form $\operatorname{Spec}(R)_f$ such that $V \cap \operatorname{Spec}(R)_f = \emptyset$.

**Exercise* 4.4.7.** Show that $\operatorname{Spec}(R)$ is quasi-compact. That is, every open covering of $\operatorname{Spec}(R)$ has a finite subcovering. Hint: First show that it suffices to consider coverings by sets of the form $\operatorname{Spec}(R)_f$. What does it mean to say the sets $\operatorname{Spec}(R)_{f_\lambda}$ as $\lambda$ runs through some set $\Lambda$ cover $\operatorname{Spec}(R)$?

# 5 Unique Factorization Domain (UFD)

## 5.1 Irreducible and prime elements of a ring

We used the two definitions of prime numbers to define maximal and prime ideals of a ring. We can equally use them to define irreducible and prime elements of a ring.

**Definition 5.1.** Let $R$ be an integral domain. Say an element $a$ divides an element $b$, and write $a \mid b$, if $b = aq$ for some $q \in R$. Say $a$ is a proper divisor of $b$ if $b = aq$ and neither $a$ nor $q$ is a unit. Say $a$ is an **associate** of $b$ if $b$ is a unit multiple of $a$, or equivalently $a \mid b$ and $b \mid a$.

**Definition 5.2.** Let $R$ be an integral domain. An **irreducible** element of $R$ is a nonzero nonunit element with no proper divisors. A **prime** element of $R$ is a nonzero nonunit element $p$ such that if $p \mid ab$ for some $a, b \in R$, then $p \mid a$ or $p \mid b$.

**Proposition 5.3.** Let $R$ be an integral domain. If $p$ is a prime, then $p$ is irreducible.

Suppose $p = ab$ for $a, b \in R$. We need to show either $a$ or $b$ is a unit. A priori we have $a \mid p$ and $b \mid p$. Hence it suffices to show either $p \mid a$ or $p \mid b$ for then we have either $b$ is a unit or $a$ is a unit, respectively. Now $p = ab$ implies that $p \mid ab$ and so either $p \mid a$ or $p \mid b$ because $p$ is a prime. □

If you still remember, to prove the equivalence of these two definitions of prime numbers, we need the notion of gcd which one can define through the Euclidean algorithm. In general, such an algorithm is not available. In fact, in general, there are irreducible elements that are not primes!

For example, let $F$ be a field and let $R$ be the subset of $F[x]$ consisting of polynomials such that the coefficient of $x$ is 0. It is easy to check that $R$ is a subring of $F[x]$. (Here is a funny proof: The coefficient of $x$ of a polynomial $f(x)$ is $f'(0)$. The product rule $(fg)' = f'g + fg'$ shows that $R$ is closed under multiplication.) The element $x^2$ is irreducible since $R$ has no polynomials of degree 1. However $x^2$ is not prime because $x^2 \mid x^3 \cdot x^3$ but $x^2$ does not divide $x^3$ in $R$. Similarly, $x^3$ is also irreducible but not prime. As the next proposition shows, we can blame this to the fact that the ideal $(x^2, x^3)$ is not principal.

**Proposition 5.4.** In a PID, every irreducible element is prime.

Let $R$ be a PID and let $r \in R$ be irreducible. Suppose $r \mid ab$ for some $a, b \in R$. If $r \mid a$, then we are done. Suppose $r \nmid a$. Then the ideal $(r, a)$ strictly contains $(r)$. Since $R$ is a PID, there exists some $d \in R$ such that $(r, a) = (d)$. Then as $r \in (r, a)$, we see that $d \mid r$. Since $r$ is irreducible and $(r) \subsetneq (d)$, we see that $d$ is a unit. In other words, $(r, a) = R$. Hence there exists $x, y \in R$ such that $rx + ay = 1$. Multiply both sides by $b$ gives $r \mid b$. □

## 5.2 Factorization

Being a PID is a very strong condition. As it turns out, there is a broader class of integral domains in which irreducible is equivalent to prime. Recall the Fundamental Theorem of Arithmetic: every positive integer has a factorization into a product of prime numbers, unique up to reordering.

**Proposition 5.5.** Prime factorizations are unique. That is, if $R$ is an integral domain and if $p_1, \ldots, p_t, q_1, \ldots, q_s$ are prime elements such that $p_1 \cdots p_t$ and $q_1 \cdots q_s$ are associates, then up to reordering, $t = s$ and $p_i$ is an associate of $q_i$ for every $i$.

Since $p_1$ divides $q_1 \cdots q_s$, we see $p_1$ divides $q_i$ for some $i$. Up to reordering, we may assume $p_1 \mid q_1$. By Exercise 5.3.2, $p_1$ and $q_1$ are associates. We can then cancel them from both sides to conclude that $p_2 \cdots p_t$ and $q_2 \cdots q_s$ are associates. Apply induction. □

Even though prime factorizations are unique, they may not exist. Any irreducible element that is not a prime provides one such example since they have no proper divisors. What about factorization into irreducibles? Ignoring existence problem, we know already it is not always unique. In our subring $R$ of $F[x]$ consisting of polynomials with no $x$ term, the element $x^6$ factors as $x^2 \cdot x^2 \cdot x^2$ and $x^3 \cdot x^3$ into irreducibles.

**Definition 5.6.** An integral domain $R$ is an **Unique Factorization Domain** (UFD) if every nonzero element has a factorization into irreducibles, unique up to reordering and up to multiplication by units.

Does every element a priori have a factorization into irreducibles? Start with an element $r$. If it is not irreducible, we can factor it as $r_0r_1$. Then we can do the same with $r_0$ and $r_1$. How do we know this process will terminate? Well it might not. As one example, take the ring $\mathbb{C}[x_1, x_2, \ldots]/(x_1 - x_2^2, x_2 - x_3^2, \ldots)$. Then the above process will not terminate if we start with $x_1$, or with any element. As another perhaps more explicit example, take the set $\bar{\mathbb{Z}} \subset \mathbb{C}$ of algebraic integers. By Exercise 3.3.1, $\bar{\mathbb{Z}}$ is a subring of $\mathbb{C}$. By Exercise 3.3.2, we see that if $\alpha \in \bar{\mathbb{Z}}$, then $\sqrt{\alpha}$ is also integral and likewise $\sqrt[2^n]{\alpha}$ for any positive integer $n$. These two rings in fact have no irreducibles! As another example, we take $R = \mathbb{C}[x, y/x, y/x^2, \ldots]$. That is, elements of $R$ are $\mathbb{C}$-linear combinations of monomials $x^my^n$ where $m$ can be negative if $n$ is positive. Then $x$ is irreducible but $y/x = x(y/x^2)$ is not. Hence the element $y$ has no factorization into irreducibles. The problem with these rings is that they are horribly "infinite".

**Definition 5.7.** A ring $R$ is **Noetherian** if it contains no infinite ascending chain of ideals. That is, if $I_1 \subset I_2 \subset I_3 \subset \cdots$ is an ascending chain of ideals, then there is some $N$ such that $I_n = I_N$ for all $n \geq N$.

**Proposition 5.8.** A ring $R$ is Noetherian if and only if every ideal is finitely generated.

Suppose first that $R$ is Noetherian. Let $I$ be an ideal of $R$. Take an element $a_1$ of $I$ and set $I_1 = (a_1)$. If $I_1 = I$, then we are done. Suppose not. Take an element $a_2$ of $I - I_1$ and set $I_2 = (a_1, a_2)$. Because $R$ is Noetherian, this process cannot continue forever. Hence there exists a positive integer $n$ and elements $a_1, \ldots, a_n$ of $I$ such that $I = (a_1, \ldots, a_n)$.

Conversely suppose every ideal is finitely generated. Let $I_1 \subset I_2 \subset I_3 \subset \cdots$ be an ascending chain of ideals. By Exercise 2.3.6, the union $I = \cup_{i=1}^{\infty} I_i$ is an ideal. Let $a_1, \ldots, a_n$ be elements of $I$ such that $I = (a_1, \ldots, a_n)$. For each $i = 1, \ldots, n$, there is a positive $m_i$ such that $a_i \in I_{m_i}$. Let $N$ be the maximum of $m_1, \ldots, m_n$. Then $a_i \in I_N$ for every $i$. Hence $I = I_N$ and so the chain stabilizes after $I_N$. □

**Corollary 5.9.** Every PID is Noetherian.

**Proposition 5.10.** Suppose $R$ is a Noetherian integral domain. Then every nonzero element of $R$ has a factorization into irreducibles.

If $r_n \mid \cdots \mid r_1$ is a sequence of elements each dividing the next, then we get an ascending sequence of ideals $(r_1) \subset (r_2) \subset \cdots \subset (r_n)$. The Noetherian condition implies that the process of finding proper divisors terminates after finitely many steps. This shows that every nonzero nonunit element has an irreducible divisor. Let $b_0$ be an element of $R$. Let $a_1$ be an irreducible divisor of $b_0$ and let $b_1 \in R$ be such that $b_0 = a_1b_1$. If $b_1$ is unit, then we are done. Suppose not, let $a_2$ be an irreducible divisor of $b_1$ and let $b_2 \in R$ be such that $b_1 = a_2b_2$. Repeating this process, we get a sequence $b_1, b_2, \ldots$ where $b_{i+1} \mid b_i$. Again by the Noetherian condition, this process cannot continue forever. Say it ends after defining $b_n$. Then $b_n$ is irreducible and we have $b_0 = a_1 \cdots a_nb_n$ is a factorization into irreducibles. □

**Corollary 5.11.** Every PID is a UFD.

In a PID, primes are the same as irreducibles (Proposition 5.4). Hence factorization into irreducibles is the same as factorization into primes. Uniqueness then follows from Proposition 5.5. Existence follows from Corollary 5.9 and Proposition 5.10. □

**Proposition 5.12.** Suppose $R$ is a UFD. Then every irreducible element is prime.

Let $r \in R$ be an irreducible element. Suppose $r \mid ab$ for some $a, b \in R$. By unique factorization, $r$ is associated with some irreducible factor of $a$ or $b$. Hence either $r \mid a$ or $r \mid b$. □

**Proposition 5.13.** An integral domain is an UFD if and only if every nonzero element has a unique factorization into primes, unique up to reordering and multiplication by units.

If $R$ is an UFD, then irreducibles are the same as primes. Conversely if every nonzero element has a factorization into primes, then every irreducible element is prime and so again there is no difference between primes and irreducibles. □

An UFD doesn't have to be Noetherian. For example consider the polynomial ring $F[x_1, x_2, \ldots]$ in infinitely many variables over a field $F$. We will show shortly that the polynomial ring in finitely many variables over an UFD is an UFD. Any element of $F[x_1, x_2, \ldots]$ is contained in $F[x_1, x_2, \ldots, x_N]$ for some large $N$ and so has a factorization into irreducibles in $F[x_1, x_2, \ldots, x_N]$. It is easy to see using degrees that irreducibles in $F[x_1, x_2, \ldots, x_N]$ remain irreducible in $F[x_1, x_2, \ldots]$. Hence we have existence. Any two factorization also can be viewed as taking place in $F[x_1, x_2, \ldots, x_N]$ for some large $N$ and so are the same up to reordering and units. Since the units in $F[x_1, x_2, \ldots]$ and $F[x_1, x_2, \ldots, x_N]$ are simply the nonzero constants, we obtain uniqueness.

## 5.3 Homework

**Exercise 5.3.1.** Let $R$ be an integral domain and let $p$ be a nonzero element. Show that $p$ is a prime element if and only if $(p)$ is a prime ideal.

**Exercise 5.3.2.** Let $R$ be an integral domain and let $p, q$ be two prime elements. Suppose $p \mid q$. Show that $p$ and $q$ are associates.

**Exercise 5.3.3.** Show that the ring $\mathbb{C}[x_1, x_2, \ldots]/I$, where $I$ is the ideal generated by $x_1 - x_2^2, x_2 - x_3^2, \ldots$, is an integral domain. Hint: An element of this ring is a coset $f + I$ where $f$ is a polynomial that involves finitely many variables.

**Exercise 5.3.4.** Show that an UFD is integrally closed in its field of fraction. That is, let $R$ be an UFD and let $F = \mathrm{Frac}(R)$ be its field of fractions. Let $\varphi : R \to F$ denote the natural inclusion. Show that if $\alpha \in F$ is integral over $R$, then $\alpha \in R$.

## 5.4 Polynomial rings over UFD

The goal of this section is to prove the following theorem.

**Theorem 5.14.** Suppose $R$ is an UFD. Then $R[x]$ is also an UFD.

**Corollary 5.15.** Suppose $R$ is an UFD and $n$ is a positive integer. Then $R[x_1, \ldots, x_n]$ is an UFD.

We will abuse notation and write $=$ for equality up to multiplication by units. Fix an UFD $R$. Every element of $R$ has a unique factorization of the form $p_1^{a_1} \cdots p_n^{a_n}$ where $p_i$ is irreducible (prime). We define the greatest common divisor and the least common multiple of two elements $\alpha, \beta$ by:

$$\begin{aligned}
\alpha &= p_1^{a_1} \cdots p_n^{a_n}, \\
\beta &= p_1^{b_1} \cdots p_n^{b_n}, \\
\gcd(\alpha, \beta) &:= p_1^{\min(a_1, b_1)} \cdots p_n^{\min(a_n, b_n)}, \\
\mathrm{lcm}(\alpha, \beta) &:= p_1^{\max(a_1, b_1)} \cdots p_n^{\max(a_n, b_n)}.
\end{aligned}$$

Let $F = \mathrm{Frac}(R)$ be the field of fraction of $R$. Since $F[x]$ is a PID, everything in $F[x]$ behaves as we expect. For any $f(x) = r_n x^n + \cdots r_1 x + r_0 \in R[x] - \{0\}$, define its **content**, denoted $C(f)$, to be $\gcd(r_n, \ldots, r_0)$. If $C(f) = 1$, we say $f$ is **primitive**.

**Lemma 5.16.** For every nonzero polynomial $f \in F[x]$, there exists some $\alpha \in F$ (unique up to $R^\times$) such that $\alpha f \in R[x]$ is primitive. We write $\widetilde{f}$ for the element $\alpha f$.

For existence, multiply $f$ by the product (or lcm) of all the denominators to get a polynomial in $R[x]$, then divide it by its content to get a primitive polynomial in $R[x]$. For uniqueness, suppose $f_1$ and $f_2$ are two primitive polynomials in $R[x]$ that differ by a constant in $F$. Apply unique factorization to show that this constant must be a unit in $R$. □

The next result allows us to use divisibility results in $F[x]$ to obtain divisibility results in $R[x]$.

**Proposition 5.17.** Suppose $g \in R[x]$ and $f \in F[x] - \{0\}$ and $f \mid g$ in $F[x]$. Then $\widetilde{f} \mid g$ in $R[x]$.

**Proposition 5.18.** (Gauss's Lemma) The product of primitive polynomials is primitive.

Let $f, g \in R[x]$ be two primitive polynomials. For any prime element $p \in R$, we have the natural quotient map $R \to R/pR$ which extends to a map $R[x] \to R/(p)[x]$ of polynomials. For any polynomial $h \in R[x]$, let $\bar{h}_p$ denote the image of $h$ under this map. Then $p$ divides the content $C(h)$ if and only if $\bar{h}_p = 0$. Hence $h$ is primitive if and only if $\bar{h}_p \neq 0$ for every prime $p \in R$. By Exercise 5.3.1, the ideal $(p)$ is a prime ideal and so $R/(p)$ and $R/(p)[x]$ are integral domains. For any prime $p$ of $R$, since $f, g$ are primitive, the polynomials $\bar{f}_p, \bar{f}_p$ are nonzero and so the polynomials $\overline{fg}_p$ is also nonzero. Hence $fg$ is primitive. □

We now prove Proposition 5.17. The statement is trivial if $g = 0$. We assume $g \neq 0$ from now on. Since $f \mid g$ in $F[x]$, we have $\widetilde{f} \mid g$ in $F[x]$. Let $h \in F[x]$ be a polynomial such that $g = \widetilde{f}h$. Let $r \in F^\times$ be an element such that $\widetilde{h} = rh$. Then there exists elements $a, b \in R$ such that $r = a/b$ and $\gcd(a, b) = 1$. We have now

$$ag = b\widetilde{f}\widetilde{h}.$$

Since $\widetilde{f}$ and $\widetilde{h}$ are primitive, so is their product. Taking content on both sides gives $aC(g) = b$. Hence $a \mid b$ and so $r = 1/s$ for some $s \in R$. Then $g = \widetilde{f}\widetilde{h}s$ is divisible by $\widetilde{f}$. □

**Corollary 5.19.** If $g \in R[x]$ has a proper divisor in $F[x]$, then it has a proper divisor in $R[x]$.

**Corollary 5.20.** The irreducible elements of $R[x]$ are the irreducible elements of $R$ along with nonconstant primitive polynomials that are irreducible over $F$.

It is easy to see irreducible elements of $R$ and nonconstant primitive polynomials that are irreducible over $F$ are irreducible elements in $R[x]$. Conversely let $g \in R[x]$ be an irreducible element. If $g$ is a constant then the proper divisors of $g$ in $R[x]$ are the same as proper divisors of $g$ in $R$ for degree reasons. Suppose now $g$ is a nonconstant polynomial. Then $C(g) \mid g$ and so $C(g)$ is a unit because $g/C(g)$ is a nonconstant polynomial and hence is not a unit. Hence $g$ is primitive. Corollary 5.19 then requires $g$ to have no proper divisor in $F[x]$. Hence $g$ is an irreducible polynomial over $F$. □

We now prove Theorem 5.14. Uniqueness of factorization follows from the uniqueness over $F[x]$ and that two primitive polynomials differing by an element of $F^\times$ in fact differ by an element of $R^\times$. For existence, let $g$ be any nonzero element of $R[x]$. First factor its content $C(g)$ into irreducibles in $R$, which by Corollary 5.20 are also irreducibles in $R[x]$. Next factor $\widetilde{g} = g/C(g)$ in $F[x]$ as $\widetilde{g} = f_1^{a_1} \cdots f_n^{a_n}$ where $f_i \in F[x]$ is irreducible in $F[x]$ for each $i$. By Gauss's Lemma, $\widetilde{f_1}^{a_1} \cdots \widetilde{f_n}^{a_n}$ is a primitive polynomial in $R[x]$ that is a scalar multiple of $f_1^{a_1} \cdots f_n^{a_n}$. Hence by Proposition 5.17, $\widetilde{f_1}^{a_1} \cdots \widetilde{f_n}^{a_n} \mid \widetilde{g}$. Comparing degree and content shows that $\widetilde{f_1}^{a_1} \cdots \widetilde{f_n}^{a_n} = \widetilde{g}$. Each $\widetilde{f_i}$ is an irreducible element of $R[x]$ by Corollary 5.20. Therefore, multiplying $\widetilde{f_1}^{a_1} \cdots \widetilde{f_n}^{a_n}$ with the irreducible factorization of $C(g)$ gives an irreducible factorization of $g$. □

The situation is a lot more complicated for the ring $R[[x]]$ of power series rings over an UFD $R$. The main problem is Lemma 5.16. Since a power series can have infinitely many terms, the denominators of the coefficients of a power series over $F$ can get arbitrarily big. Another, albeit sillier, problem is that a power series over $R$ with nonzero constant term becomes a unit in $F[[x]]$. It is known that the localization $R$ of $\mathbb{C}[x, y, z]/(x^2 + y^3 + z^8)$ at the prime ideal $(x, y, z)$ is an UFD but $R[[t]]$ is not an UFD. It is unknown whether the ring $\mathbb{C}[x_1, x_2, \ldots][[t]]$ is an UFD or not! The following result can be a final project.

**Theorem 5.21.** Suppose $R$ is a PID. Then $R[[x]]$ is an UFD.

## 5.5 Euclidean domains

The reason why $\mathbb{Z}$ and $F[x]$ are PIDs can be attributed to their division algorithms: for any $a, b$ with $a$ nonzero, there exists a quotient $q$ and a remainder $r$ such that $b = aq + r$ and such that $r$ is smaller than $a$. To generalize this idea, we need a notion of size. For $\mathbb{Z}$, this is the absolute value. For $F[x]$, this is the degree.

**Definition 5.22.** An integral domain $R$ is an **Euclidean domain** if there is a map $\sigma : R - \{0\} \to \mathbb{Z}_{\geq 0}$ such that for any $a, b \in R$ with $a$ nonzero, there exists elements $q, r \in R$ such that $b = aq + r$ and such that either $r = 0$ or $\sigma(r) < \sigma(a)$.

**Proposition 5.23.** An Euclidean domain is a PID.

Let $R$ be an Euclidean domain with size function $\sigma$. Let $I$ be a nonzero ideal. Then $\sigma(I - \{0\})$ is a subset of the nonnegative integers and hence has a minimal element. Let $a \in I - \{0\}$ be an element such that $\sigma(a)$ is minimal. For any $b \in I$, divide $b$ by $a$ to get a quotient $q$ and a remainder $r$. Then $r = b - aq \in I$. By the minimality assumption on $\sigma(a)$, we see that $r = 0$ and so $a \mid I$. Therefore, $I = (a)$. □

**Proposition 5.24.** The ring $\mathbb{Z}[i]$ is an Euclidean domain with $\sigma(a + bi) = a^2 + b^2$.

We extend this size function to the field $\mathbb{Q}[i]$. It is easy to check that $\sigma$ is multiplicative, i.e., $\sigma(\alpha\beta) = \sigma(\alpha)\sigma(\beta)$ for any $\alpha, \beta \in \mathbb{Q}[i]$. Moreover, $\sigma(\alpha) = 0$ if and only if $\alpha = 0$. Given $\alpha, \beta \in \mathbb{Z}[i]$ with $\alpha \neq 0$, write the quotient $\beta/\alpha$ as $u + vi$ for $u, v \in \mathbb{Q}$. Let $s, t$ be integers closest to $u, v$. Then $|s - u| \leq \frac{1}{2}$ and $|t - v| \leq \frac{1}{2}$. Set $\gamma = s + ti \in \mathbb{Z}[i]$ and set $\delta = \beta - \alpha\gamma = \alpha((u - s) + (v - t)i)$. Then

$$\sigma(\delta) = \sigma(\alpha)(|u - s|^2 + |v - t|^2) \leq \frac{1}{2}\sigma(\alpha) < \sigma(\alpha). \qquad \square$$

## 5.6 Homework

**Exercise 5.6.1.** Show that a Noetherian integral domain is an UFD if and only if the intersection of two principal ideals is principal.

**Exercise 5.6.2.** Let $R$ be a commutative ring and let $S$ be a multiplicatively closed subset containing 1 and not containing 0. Show that if $R$ is a Noetherian, then so is $R[S^{-1}]$. Show that if $R$ is an UFD, then so is $R[S^{-1}]$.

**Exercise 5.6.3.** Let $R$ be a Noetherian integral domain and let $p$ be a prime element of $R$. Show that if $R[1/p]$ is an UFD, then $R$ is an UFD. Hint: Show that the intersection $\cap_{i=0}^{\infty}(p^i) = (0)$ and define for any $r \in R - \{0\}$, its $p$-adic valuation $\mu_p(r)$ by the largest nonnegative integer $n$ such that $r \in (p^n)$.

**Exercise 5.6.4.** Let $R$ be a Noetherian integral domain and let $I$ be a subset of $R$ such that every element of $I$ is prime. Let $S(I)$ be the multiplicatively closed subset of $R$ generated by $I$, that is, elements of $S$ have the form $p_1^{a_1} \cdots p_n^{a_n}$ for $p_i \in I$ and nonnegative integers $a_i$. Show that if $R[S(I)^{-1}]$ is an UFD, then so is $R$. Hint: Apply Exercise 5.6.3 and induction to deal with the case when $I$ is finite. Then notice that any element of $R[S(I)^{-1}]$ belongs to $R[S(I_1)^{-1}]$ for some finite subset $I_1$ of $I$ and similarly for any factorization.

**Exercise 5.6.5.** Let $R$ be an UFD and let $I$ be the set of all primes in $R$. Show (without using results proved in today's class) that every element of $I$ is also prime in $R[x]$ and that $R[x][S(I)^{-1}] \cong R[S(I)^{-1}][x] = F[x]$ where $F = \mathrm{Frac}(R)$. This gives another proof that $R[x]$ is an UFD whenever $R$ is an UFD. Why doesn't this work for $R[[x]]$?

**Exercise 5.6.6.** Show that the ring $\mathbb{Z}[\sqrt{-2}]$ with $\sigma(a + b\sqrt{-2}) = a^2 + 2b^2$ is an Euclidean domain.

# 6 Modules

When proving Hilbert's Nullstellensatz, we used the theory of vector spaces over $\mathbb{C}$. In general, one can do linear algebra over an arbitrary ring $R$.

**Definition 6.1.** Let $R$ be a commutative ring. An **$R$-module** is an abelian group $M$ along with a ring homomorphism $\varphi : R \to \mathrm{End}(M)$. For any $r \in R$ and any $m \in M$, we write $rm$ for the image of $m$ under $\varphi(r)$. More explicitly, we have for any $r, s \in R$ and any $m, n \in M$:

1. $1m = m$,
2. $(rs)m = r(sm)$,
3. $(r + s)m = rm + sm$,
4. $r(m + n) = rm + rn$.

A **submodule** is an abelian subgroup $N$ closed under the ring action, that is $rn \in N$ for any $r \in R$ and any $n \in N$.

## 6.1 Examples

1. When $R = F$ is a field, then an $R$-module is the same as an $F$-vector space. Submodules are vector subspaces.
2. When $R = \mathbb{Z}$, then the ring homomorphism $\varphi$ is automatic since every ring admits a unique map from $\mathbb{Z}$. A $\mathbb{Z}$-module is then the same as an abelian group and submodules are subgroups.
3. For any commutative ring, the ring $R$ itself is an $R$-module, called the free $R$-module of rank 1. For any $r \in R$, multiplication by $r$ from $R$ to $R$ is a group homomorphism and so is an element of $\mathrm{End}(R)$. The map $\varphi : R \to \mathrm{End}(R)$ sends $r$ to the multiplication by $r$ map. A submodule is an abelian subgroup of $R$ closed under multiplication by $R$. In other words, a submodule of $R$ is an ideal.
4. Let $\alpha : R \to R'$ be a ring homomorphism, then any $R'$-module $M$ can also be viewed as an $R$-module. In particular, $R'$ itself can be viewed as an $R$-module.
5. If $R$ is an integral domain and $F = \mathrm{Frac}(R)$ is its field of fractions, then the natural inclusion $R \to F$ gives $F$ the structure of an $R$-module. Submodules of $F$ are called **fractional ideals** of $R$.
6. Given any family $\{M_\lambda\}_{\lambda\in\Lambda}$ of $R$-modules, we define an $R$-module and a submodule, the direct product and the direct sum:
$$\begin{aligned}\prod_{\lambda\in\Lambda} M_\lambda &= \{f : \Lambda \to \bigcup_{\lambda\in\Lambda} M_\lambda : f(\lambda) \in M_\lambda, \forall\lambda \in \Lambda\},\\ \bigoplus_{\lambda\in\Lambda} M_\lambda &= \{f \in \prod_{\lambda\in\Lambda} M_\lambda : f(\lambda) = 0 \text{ for all but finitely many } \lambda \in \Lambda\}.\end{aligned}$$
Given $f, g \in \prod_{\lambda\in\Lambda} M_\lambda$, define $f + g$ by $(f + g)(\lambda) = f(\lambda) + g(\lambda)$ and for any $r \in R$, define $rf$ by $(rf)(\lambda) = r(f(\lambda))$.
7. Let $M$ be an $R$-module and let $N$ be a submodule. Let $M/N$ denote the quotient as abelian groups. Define an action of $R$ on $M/N$ by $r(m + N) = rm + N$ for any $r \in R$ and any $m \in M$. Since $N$ is a submodule, this action is well-defined.
8. Let $M$ be an $R$-module and let $S$ be multiplicatively closed subset of $R$ containing 1 and not containing 0. Denote an element of $M \times S$ suggestively by $m/s$. Define a relation $\sim$ on $M \times S$ by
$$\frac{m_1}{s_1} \sim \frac{m_2}{s_2} \iff s_0(s_2m_1 - s_1m_2) = 0 \text{ for some } s_0 \in S.$$

Then $\sim$ is an equivalence relation. The set $M[S^{-1}]$ of equivalence classes has the structure of an abelian group with addition being

$$\frac{m_1}{s_1} + \frac{m_2}{s_2} = \frac{m_1 s_2 + m_2 s_1}{s_1 s_2}.$$

Define an action of $R[S^{-1}]$ on $M[S^{-1}]$ by

$$\frac{r}{s}\frac{m}{s'} = \frac{rm}{ss'}.$$

Then $M[S^{-1}]$ is an $R[S^{-1}]$-module. It is also an $R$-module via the natural embedding $R \to R[S^{-1}]$.

9. Let $M$ be an $R$-module and let $\varphi : R \to \mathrm{End}(M)$ be the ring action map. Let $\alpha$ be an arbitrary element of $\mathrm{End}(M)$. Then $\varphi$ extends to a ring homomorphism $R[x] \to \mathrm{End}(M)$ sending $x$ to $\alpha$ and this gives $M$ the structure of an $R[x]$-module.

## 6.2 Homomorphism

**Definition 6.2.** Let $M, N$ be two $R$-modules for a commutative ring $R$. An $R$-module homomorphism $\varphi : M \to N$ is a group homomorphism respecting the $R$-actions. That is, for any $r \in R$ and $m \in M$,

$$\varphi(rm) = r\varphi(m).$$

Similar definitions of isomorphisms, automorphisms, kernels and images.

Denote by $\mathrm{Hom}_R(M, N)$ the set of $R$-module homomorphisms between two $R$-modules $M$ and $N$. When $M = N$, we write $\mathrm{End}_R(M)$ instead.

**Proposition 6.3.** $\mathrm{Hom}_R(M, N)$ is an $R$-module while $\mathrm{End}_R(M)$ is a subring of $\mathrm{End}(M)$.

We write down the action of $R$ on $\mathrm{Hom}_R(M, N)$. The rest is easy. For any $r \in R$ and $\varphi \in \mathrm{Hom}_R(M, N)$, define $r\varphi$ by

$$(r\varphi)(m) := r(\varphi(m)) = \varphi(rm). \qquad \square$$

Everything about quotient ideals generalizes to quotient modules.

**Theorem 6.4.** Let $R$ be a commutative ring. Let $M$ be an $R$-module and let $N$ be a submodule. The natural quotient map $M \to M/N$ is an $R$-module homomorphism and every $R$-module homomorphism out of $M$ with kernel containing $N$ factors uniquely through this map. It also induces a bijection between submodules of $M$ containing $N$ and submodules of $M/N$. If $N'$ is a submodule of $M$ containing $N$, then $(M/N)/(N'/N) \cong M/N'$.

**Theorem 6.5.** Let $\varphi : M \to N$ be an $R$-module homomorphism. Then $M/\ker\varphi \cong \mathrm{im}(\varphi)$.

**Definition 6.6.** Let $R$ be a commutative ring. An $R$-module $M$ is **finitely generated** if there is an integer $n$ and a surjection $R^n \to M$, or equivalently there exists elements $m_1, \ldots, m_n$ of $M$ such that $M = Rm_1 + \cdots Rm_n$. A finitely generated $R$-module is **free** if there is an integer $n$ and an isomorphism $R^n \to M$.

Suppose $R$ is an integral domain and set $S = R - \{0\}$. If $M$ is a finitely generated $R$-module, then $M[S^{-1}]$ is a finitely generated $F = \mathrm{Frac}(R)$ module, that is, a finite dimensional $F$-vector space. Hence if $n$ and $m$ are two positive integers such that $R^n \cong R^m$, then $n = m$. Therefore, if $M$ is a finitely generated free $R$-module, then the integer $n$ such that there is an isomorphism $R^n \cong M$ is unique and is called the **rank** of $M$. A finitely generated free submodule of $R$ is a principal ideal.

**Corollary 6.7.** A Noetherian integral domain $R$ is a PID if and only if every submodule of $R$ is free.

**Definition 6.8.** An $R$-module $M$ is **Noetherian** if there are no infinite ascending chain of submodules.

**Proposition 6.9.** An $R$-module $M$ is Noetherian if and only if every submodule is finitely generated.

## 6.3 Torsion

The analogue of a zero-divisor for modules is torsion.

**Definition 6.10.** Let $R$ be a commutative ring and let $M$ be an $R$-module. An element $m \in M$ is **torsion** if there exists some $r \in R - \{0\}$ such that $rm = 0$. Let $M^{\text{tor}}$ be the subset of torsion elements of $M$. An $R$-module $M$ is called **torsion-free** if $M^{\text{tor}} = \{0\}$. It is called **torsion** if $M^{\text{tor}} = M$.

**Proposition 6.11.** Let $R$ be an integral domain and let $M$ be an $R$-module. Then $M^{\text{tor}}$ is a submodule of $M$ and $M/M^{\text{tor}}$ is torsion free.

To show $M^{\text{tor}}$ is a submodule of $M$, we need to show it is closed under addition and the ring action. Let $m$ be a torsion element and let $r$ be an element of $M$. Let $r_0$ be a nonzero element of $R$ such that $r_0 m = 0$. Then $r_0(rm) = r(r_0 m) = 0$. Hence $rm \in M^{\text{tor}}$. Let $m_1, m_2$ be two torsion elements. Let $r_1, r_2$ be two nonzero elements of $R$ such that $r_1 m_1 = 0 = r_2 m_2$. Since $R$ is an integral domain, $r_1 r_2$ is nonzero and $r_1 r_2(m_1 + m_2) = 0$. Hence $m_1 + m_2 \in M^{\text{tor}}$.

Suppose $m + M^{\text{tor}}$ is a torsion coset in $M/M^{\text{tor}}$. Let $r \in R - \{0\}$ be an element such that $r(m + M^{\text{tor}}) = 0$. That is, $rm \in M^{\text{tor}}$. Let $s \in R - \{0\}$ be an element such that $s(rm) = 0$. Then $(rs)m = 0$ and $rs$ is nonzero since $R$ is an integral domain. Hence $m$ is torsion and so $m + M^{\text{tor}} = 0$. □

**Proposition 6.12.** Let $R$ be a PID. Then every finitely generated torsion free $R$-module is free.

Let $M$ be a finitely generated torsion free $R$-module. Let $S = R - \{0\}$ be the set of nonzero elements of $S$. The condition that $M$ is torsion free translates into the injectivity of the map $M \to M[S^{-1}]$. Now $M[S^{-1}]$ is a finitely generated module over a field $F = \text{Frac}R$ and so is an $F$-vector space. We are now reduced to showing that every sub-$R$-module of $F^n$ is free over $R$.

When $n = 1$, you will show in Exercise 6.4.3 that every submodule of $F$ is of the form $Ra$ for some $a \in F$. If $a = 0$, then $M = 0$. Otherwise, $Ra \cong R$ as $R$-modules.

Suppose now $n > 1$ and $M$ is a submodule of $F^n$. Let $\text{pr}_n$ denote the projection map onto the last coordinate. Then $\text{pr}_n(M)$ is a submodule of $F$ and so is of the form $Ra$ for some $a \in F$. If $a = 0$, then $M \subset F^{n-1}$ and by induction, $M$ is free. Suppose from now on that $a \neq 0$. Let $m \in M$ be an element such that $\text{pr}_n(b) = a$. We claim that $M \cong (\ker \text{pr}_n \cap M) \times Rb$ for then $\ker \text{pr}_n \cap M$ is a submodule of $F^{n-1}$ and so is free by induction.

Define a map $\varphi : (\ker \text{pr}_n \cap M) \times Rb \to M$ by $\varphi(m, rb) = m + rb$. It is easy to check that $\varphi$ is an $R$-module homomorphism. For any $m' \in M$, there is some $r \in R$ such that $\text{pr}_n(m') = ra$. Then $m' - rb \in \ker \text{pr}_n \cap M$. Hence $\varphi$ is surjective. On the other hand, if $m + rb = 0$ for some $m \in \ker \text{pr}_n \cap M$, then applying $\text{pr}_n$ gives $ra = 0$ and so $r = 0$ since $a \neq 0$ and $R$ is a domain. Then $m = 0$ and so $\varphi$ is also injective. □

**Corollary 6.13.** Let $R$ be a PID. Let $M$ be a finitely generated $R$-module generated by $n$ elements. Then every submodule of $M$ can be generated by at most $n$ elements.

Let $\varphi : R^n \to M$ be a surjective $R$-module homomorphism. Let $N$ be a submodule of $M$. It suffices to show that $\varphi^{-1}(N)$ can be generated by at most $n$ elements. Since $R$ is Noetherian, so is $R^n$ and so $\varphi^{-1}(N)$, being a submodule, is finitely generated. Since $R$ is an integral domain, $R^n$ is torsion free and so is its submodule $\varphi^{-1}(N)$. Hence by Proposition 6.12, $\varphi^{-1}(N)$ is a free submodule of $R^n$. Localize away from 0 shows that the rank of $\varphi^{-1}(N)$ is at most the rank of $R^n$, which is $n$. □

## 6.4 Homework

**Exercise 6.4.1.** (Universal properties of direct product and direct sum) Let $R$ be a commutative ring and let $\{M_\lambda\}_{\lambda \in \Lambda}$ be a family of $R$-modules. Prove the following statements.

1. There are $R$-module homomorphisms $\text{pr}_\lambda : \prod_{\lambda \in \Lambda} M_\lambda \to M_\lambda$ such that if $N$ is an $R$-module with $R$-module homomorphism $\varphi_\lambda : N \to M_\lambda$, there is a unique $R$-homomorphism $\varphi : N \to \prod_{\lambda \in \Lambda} M_\lambda$ such that $\text{pr}_\lambda \circ \varphi = \varphi_\lambda$ for all $\lambda \in \Lambda$.

2. There are $R$-module homomorphisms $\iota_\lambda : M_\lambda \to \oplus_{\lambda \in \Lambda} M_\lambda$ such that if $N$ is an $R$-module with $R$-module homomorphism $\varphi_\lambda : M_\lambda \to N$, there is a unique $R$-homomorphism $\varphi : \oplus_{\lambda \in \Lambda} M_\lambda \to N$ such that $\varphi \circ \iota_\lambda = \varphi_\lambda$ for all $\lambda \in \Lambda$.

When $\Lambda$ is finite, these two modules are the same and we sometimes use the notation $\times$ instead.

**Exercise 6.4.2.** Let $R$ be a commutative ring and let $S$ be a multiplicatively closed subset containing 1 and not containing 0. Let $M$ be a Noetherian $R$-module. Show that $M[S^{-1}]$ is a Noetherian $R[S^{-1}]$-module.

**Exercise 6.4.3.** Let $R$ be a PID with field of fraction $F = \mathrm{Frac}(R)$. Show that the finitely generated submodules of $F$ are all of the form $Ra$ for some $a \in F$. Hint: Let $M$ be a submodule and consider the set $I = \{r \in R : rM \subset R\}$. Show that $I$ is an ideal of $R$.

**Exercise 6.4.4.** Let $R$ be a commutative ring. Let $M$ be an $R$-module with submodule $N$. Show that $M$ is Noetherian if and only if both $N$ and $M/N$ are Noetherian. Show as a consequence that if $M$ and $M'$ are Noetherian $R$-modules, then so is $M \times M'$.

# 7 Structure Theory of Modules over PIDs

The goal of this section is to classify all finitely generated modules over a PID $R$.

**Theorem 7.1.** Let $R$ be a PID and let $M$ be a finitely generated module over $R$. Then there exists prime elements $p_1, \ldots, p_k$ of $R$, integers $r$, $t_i, n_{i,j}$ for $i = 1, \ldots, k$ and $j = 1, \ldots, t_i$ such that

$$M \cong R^r \times \prod_{i=1}^{k} \prod_{j=1}^{t_i} R/(p_i^{n_{i,j}}).$$

When $R = \mathbb{Z}$, modules become abelians and we have the following familiar classification theorem for finitely generated abelian groups.

**Corollary 7.2.** Finitely generated abelian groups are direct products of $\mathbb{Z}$ and $\mathbb{Z}/m\mathbb{Z}$.

## 7.1 Torsion Decomposition

Let $M$ be a finite generated module over $R$. Recall we have a submodule $M^{\text{tor}}$ of torsion element and a torsion free quotient $M/M^{\text{tor}}$. We can express this as a short exact sequence

$$0 \to M^{\text{tor}} \to M \to M/M^{\text{tor}} \to 0.$$

In general, a sequence of homomomorhisms $M_1 \to M_2 \to \cdots \to M_n$ is called an **exact sequence** if the image of one map is the kernel of the next map. A short exact sequence is an exact sequence of the form $0 \to A \to B \to C \to 0$. This is equivalent to saying that $A$ is isomorphic to a sub-object of $B$ with quotient isomorphic to $C$. Exact sequences are foundational objects in homological algebra.

Since $M$ is finitely generated, so is $M/M^{\text{tor}}$. Since $R$ is a PID, we have by Proposition 6.12 that $M/M^{\text{tor}}$ is free. One of the nice thing about free modules is that homomorphisms out of them can be very easily defined.

**Proposition 7.3.** Let $\varphi : N \to R^n$ be a surjective $R$-module homomorphism. Then $\varphi$ admits a splitting. That is, there exists an $R$-module homomorphism $\phi : R^n \to N$ such that $\varphi \circ \phi = \text{id}$.

Write $e_i$ for the element of $R^n$ that is 0 at all the coordinate except for the $i$-th coordinate and it is 1 at the $i$-th coordinate. Then to define an $R$-module homomorphism out of $R^n$, it suffices to define the images of $e_i$ and there are no restrictions on these images (hence the meaning of "free"). Define $\phi(e_i)$ to be any of the preimages $\varphi^{-1}(e_i)$. □

For any commutative ring $R$, an $R$-module is called **projective** if every surjection to it admits a splitting.

**Proposition 7.4.** Let $M$ be a finitely generated module of a PID $R$. Then $M \cong M^{\text{tor}} \times M/M^{\text{tor}}$.

The canonical surjection $\varphi : M \to M/M^{\text{tor}}$ admits a splitting $\phi : M/M^{\text{tor}} \to M$. Let $\alpha$ denote the map $M^{\text{tor}} \times M/M^{\text{tor}} \to M$ sending $(m_1, m_2)$ to $m_1 + \phi(m_2)$. It is easy to see that $\alpha$ is an $R$-module homomorphism. For any $m \in M$, since $\phi$ is a splitting, we see that $m - \phi(\varphi(m)) \in M^{\text{tor}}$. Hence $m = \alpha(m - \phi(\varphi(m)), \varphi(m))$ and so $\alpha$ is surjective. On the other hand, if $\alpha(m_1, m_2) = 0$, then applying $\varphi$ to $m_1 + \phi(m_2)$ gives $m_2 = 0$ and then $m_1 = 0$. Hence $\alpha$ is injective. □

Since $M/M^{\text{tor}}$ is a finitely generated free module over $R$, there exists a (unique) integer $r$ such that $M/M^{\text{tor}} \cong R^r$. It remains now to classify torsion modules over $R$.

## 7.2 Primary Decomposition

For any prime $p$ of $R$, we say a module $M$ is $p$-**primary** if for every $m \in M$, there is some $n$ such that $p^n m = 0$. For any module $M$, we let $M[p^\infty]$ denote its $p$-primary part:

$$M[p^\infty] = \{m \in M : p^n m = 0, \text{ for some positive integer } n\}.$$

**Proposition 7.5.** Let $R$ be a PID. Then every torsion $R$-module is a direct sum of its primary submodules. If the $R$-module is finitely generated, then only a finite number of primary submodules are needed.

We say two prime elements are equivalent if they are associates. This is an equivalence relation and let $\Lambda$ denote the set of equivalence classes. We fix a prime $p_\lambda$ in each equivalence class $\lambda$. (Note this requires the Axiom of Choice, which is equivalent to Zorn's Lemma. The statement for finitely generated modules does not require this.)

Let $\sum_{\lambda\in\Lambda} M[p_\lambda^\infty]$ denote the submodule of $M$ consisting of elements of the form $m_{\lambda_1} + \cdots + m_{\lambda_k}$ where $m_{\lambda_i} \in M[p_{\lambda_i}^\infty]$. We claim that

$$M = \sum_{\lambda\in\Lambda} M[p_\lambda^\infty].$$

Indeed, let $m$ be any nonzero element of $m$. Then there exists a nonzero element $r_0 \in R$ such that $r_0 m = 0$ as $M$ is torsion. We factor $r_0$ in $R$ into $r_0 = u p_{\lambda_1}^{a_{\lambda_1}} \cdots p_{\lambda_k}^{a_{\lambda_k}}$ for some unit $u$ and positive integers $a_{\lambda_i}$. For each $i$, we set $q_i = r_0/p_{\lambda_i}^{a_{\lambda_i}}$. Then $\gcd(q_1, \dots, q_k) = 1$ or equivalently, the ideal $(q_1, \dots, q_k)$ is the entire ring. Let $r_1, \dots, r_k$ be elements of $R$ such that $q_1 r_1 + \cdots + q_k r_k = 1$. We set $m_{\lambda_i} = q_i r_i m$. Then $m = \sum_{i=1}^k m_{\lambda_i}$ and $p_{\lambda_i}^{a_{\lambda_i}} m_{\lambda_i} = r_i r_0 m = 0$ implying that $m_{\lambda_i} \in M[p_{\lambda_i}^\infty]$.

We have just proved that the natural map $\oplus_{\lambda\in\Lambda} M[p_\lambda^\infty] \to M$ is surjective. To show it is injective, it suffices to show that

$$M[p_{\lambda_0}^\infty] \cap \sum_{i=1}^k M[p_{\lambda_i}^\infty] = 0$$

for any $\lambda_0, \dots, \lambda_k \in \Lambda$. Any element $m$ of the intersection is annihilated by $p_{\lambda_0}^{a_{\lambda_0}}$ and by $\prod_{i=1}^k p_{\lambda_k}^{a_{\lambda_k}}$ for integers $a_{\lambda_i}$. These two elements generate the entire ring and so some $R$-linear combination of them gives 1 and it annihilates $m$. Hence $m = 0$.

If $M$ is finitely generated, then there is an element $r \in R$ such that $rm = 0$ for every $m \in M$. Indeed, take a finite generating set $\{m_1, \dots, m_k\}$ for $M$ and for each $i$, take some nonzero $r_i$ in $R$ that annihilates $m_i$. Take $r$ to be the lcm of $r_1, \dots, r_k$. The primes needed for the primary decomposition are the primes dividing $r$. Therefore, only finitely many summands are needed. □

## 7.3 Cyclic Decomposition

We are now left to classify all finitely generated $p$-primary $R$-modules. Note we know they are finitely generated because $R$ is Noetherian and so are $R^n$ and its quotient $M$.

**Proposition 7.6.** Let $R$ be a PID and let $p$ be a prime element of $R$. Let $M$ be a finitely generated $p$-primary $R$-module. Let $t$ be the smallest integer such that $p^t$ annihilates every element of $M$. Let $a$ be an element of $M$ annihilated by $p^t$ but not by $p^{t-1}$. Then there is a submodule $M'$ of $M$ such that the natural map $M' \oplus Ra \to M$ is an isomorphism.

In order for such a map to be an isomorphism, we need $M' \cap Ra = 0$ and $M' + Ra = M$. Let $M'$ be a maximal submodule of $M$ satisfying $M' \cap Ra = 0$. (How do we know $M'$ exist?) Suppose for a contradiction that $M' + Ra \subsetneq M$. Our plan is to exhibit a larger submodule $M''$ of $M$ such that having nontrivial intersection with $Ra$ is not obvious. Let $b$ be an element of $M$ not in $M' + Ra$. Some power of $p$ times $b$ is in $M' + Ra$, so we may assume without loss of generality that $pb \in M' + Ra$. We could take $M'' = M' + Rb$ but it will intersect $Ra$ for obvious reasons and so will not give us valuable information.

Say $pb = m + ra$ for some $m \in M'$ and some $r \in R$. Multiplying by $p^{t-1}$ gives $0 = p^{t-1}m + p^{t-1}ra$. Hence $p^{t-1}ra$ is in $M' \cap Ra$ and so is 0. Hence $p \mid r$. Write $r = pr'$ for some $r' \in R$. Then we have $p(b - r'a) = m$. Let $c = b - r'a$ and take $M'' = M' + Rc$. Since $b \notin M' + Ra$, we see that $c \notin M'$. Hence $M''$ strictly contains $M'$. By maximality of $M'$, we have a nontrivial intersection $M'' \cap Ra \neq 0$. Let $m'$ be an element of $M'$ and $r_1, r_2$ be elements of $R$ such that $m' + r_1 c = r_2 a \neq 0$. Since $pc \in M'$ and $M' \cap Ra = 0$, we must have $p \nmid r_1$. Let $x, y \in R$ be elements with $px + r_1 y = 1$. Multiplying by $c$, we have $pc \in M'$ and $r_1 c \in M' + Ra$ and so $c \in M' + Ra$. But then $b = c + r'a$ is also in $M' + Ra$. Contradiction. □

Rename the submodules $M'$ and $Ra$ by $M_1$ and $Ra_1$ and suppose $a_1$ is annihilated by $p^{t_1}$. Then we have

$$M \cong M_1 \times Ra_1 \cong M_1 \times R/(p^{t_1}).$$

Since $M$ is Neotherian and $p$-primary, its submodule $M_1$ is finitely generated and $p$-primary. We may then decompose $M_1$ as $M_2 \oplus Ra_2 \cong M_2 \times R/(p^{t_2})$. This process cannot continue forever for then we would have an infinite ascending chain of submodules:

$$Ra_1 \subsetneq Ra_1 + Ra_2 \subsetneq Ra_1 + Ra_2 + Ra_3 \subsetneq \cdots .$$

**Corollary 7.7.** Let $R$ be a PID and let $p$ be a prime element of $R$. Let $M$ be a finitely generated $p$-primary $R$-module. Then there are integers $t_1, \ldots, t_k$ such that

$$M \cong R/(p^{t_1}) \times \cdots \times R/(p^{t_k}).$$

We have now completed the proof of Theorem 7.1.

## 7.4 Homework

The goal of this homework is show that for finitely generated modules over a local ring, projective and free are the same. Recall that over a commutative ring $R$, a module $M$ is projective if every surjection $\varphi : N \to M$ admits a splitting. That is, there exists a module homomorphism $\phi : M \to N$ such that $\varphi \circ \phi = \mathrm{id}$.

**Exercise 7.4.1.** Let $R$ be a commutative ring and let $M$ be an $R$-module generated by $n$ elements. Let $I$ be an ideal of $R$ and write $IM$ for the submodule of $M$ consisting of elements of the form $a_1 m_1 + \cdots + a_k m_k$ for $a_i \in I$ and $m_i \in M$. Suppose $\varphi \in \mathrm{End}_R(M)$ such that $\varphi(M) \subset IM$. Show that there exists $b_1, \ldots, b_n \in I$ such that

$$\varphi^n + b_1 \varphi^{n-1} + \cdots b_n = 0$$

in the ring $\mathrm{End}_R(M)$. Hint: the determinant trick. (When $R$ is a field, this is the Cayley-Hamilton theorem.)

**Exercise 7.4.2.** Let $R$ be a commutative ring and let $M$ be a finitely generated $R$-module. Let $I$ be an ideal of $R$. If $M = IM$, then there exists some $r \in R$ such that $r \in 1 + I$ and $rM = 0$.

**Definition 7.8.** Let $R$ be a commutative ring. The intersection of all maximal ideals of $R$ is called the **Jacobson radical** of $R$, denoted $\mathrm{rad}(R)$.

**Exercise* 7.4.3.** Let $R$ be a commutative ring. Show that $r \in \mathrm{rad}(R)$ if and only if $1 + rs$ is a unit for every $s \in R$.

As a corollary, if the ideal $I$ in Exercise 7.4.2 is contained in $\mathrm{rad}(R)$, then the module $M$ is zero. This is the Nakayama Lemma.

**Exercise* 7.4.4.** Let $R$ be a local ring with maximal ideal $\mathfrak{m}$ and residue field $k = R/\mathfrak{m}$. Let $M$ be a finitely generated $R$-module. Then $M/\mathfrak{m}M$ is a finite dimensional $k$-vector space. Let $\bar{m}_1, \ldots, \bar{m}_k$ be a basis for $M/\mathfrak{m}M$ as a $k$-vector space and let $m_1, \ldots, m_k$ be elements of $M$ mapping to $\bar{m}_1, \ldots, \bar{m}_k$ under the natural map $M \to M/\mathfrak{m}M$. Show that $M$ is generated by $m_1, \ldots, m_k$. This is called a minimal basis for $M$.

**Exercise* 7.4.5.** Let $R$ be a local ring with maximal ideal $\mathfrak{m}$ and residue field $k = R/\mathfrak{m}$. Let $M$ be a finitely generated $R$-module with minimal basis $m_1, \ldots, m_k$. Let $\varphi : R^k \to M$ be the $R$-module homomorphism sending $(r_1, \ldots, r_k)$ to $r_1 m_1 + \cdots r_k m_k$. Suppose $\varphi$ admits a splitting. Show that $\varphi$ is an isomorphism.

Since such splittings always exist for projective modules, we have shown that a finitely generated projective module over a local ring is free. As it turns out, the assumption on finite generation is not necessary.

# 8 Fields

We now begin the study of fields. Recall fields are commutative rings in which every nonzero element is invertible. Examples include $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{F}_p$, $F(t)$.

## 8.1 Field extensions

If $E$ is a field containing another field $F$, then $E$ is a **field extension** of $F$, denoted by $E/F$. The inclusion $F \to E$ gives $E$ the structure of an $F$-module, that is an $F$-vector space. The dimension $\dim_F(E)$ (or the rank of $E$ as a free $F$-module) is the **degree** of the extension, denoted $[E : F]$. We $E/F$ is finite if $[E : F] < \infty$, otherwise we say it is infinite.

For example $[\mathbb{C} : \mathbb{R}] = 2$ as $\mathbb{C} = \mathbb{R} + \mathbb{R}\sqrt{-1}$ while $\mathbb{R}/\mathbb{Q}$ is infinite. The extension $F(t)/F$ is also infinite.

**Proposition 8.1.** If $E/K$ and $K/F$ are finite field extensions, then $E/F$ is also finite and

$$[E : F] = [E : K][K : F].$$

Let $\{u_i\}_{i=1}^n$ be a basis of $E$ over $K$ and let $\{v_j\}_{j=1}^m$ be a basis of $K$ over $F$. Then $\{u_iv_j\}_{i=1}^n{}_{j=1}^m$ is a basis of $E$ over $F$. □

Given a field extension $E/F$ and an element $\alpha \in E$, write $F[\alpha]$ for the smallest subring of $E$ containing $F$ and $\alpha$ and write $F(\alpha)$ for the smallest subfield of $E$ containing $F$ and $\alpha$. If $E = F(\alpha)$ for some $\alpha \in E$, then $E$ is a **simple** extension of $F$. As we will prove later, every finite extension of $\mathbb{Q}$ is simple!

**Proposition 8.2.** If $E/F$ is a finite field extension, then there are elements $\alpha_1, \ldots, \alpha_n \in E$ such that

$$F \subsetneq F(\alpha_1) \subsetneq \cdots \subsetneq F(\alpha_1, \ldots, \alpha_n) = E.$$

We induct on the degree $[E : F]$. If $[E : F] = 1$, then there is nothing to prove. In general, take any $\alpha_1 \in E - F$. Then $[F(\alpha_1) : F] > 1$ and so $[E : F(\alpha_1)] < [E : F]$. Apply induction to $E/F(\alpha_1)$. □

Since $F(\alpha)$ contains $F[\alpha]$ and every nonzero element of $F[\alpha]$ becomes invertible, we see that $F(\alpha)$ contains the field of fraction of $F[\alpha]$. Conversely, the field of fraction of $F[\alpha]$ is a subfield of $E$ containing $F$ and $\alpha$ and so contains $F(\alpha)$ by minimality. Hence $F(\alpha) = \mathrm{Frac}(F[\alpha])$. Sometimes the ring $F[\alpha]$ is already a field. Recall an element $\alpha$ is algebraic over $F$ if there is a nonzero monic polynomial $f(x) \in F[x]$ such that $f(\alpha) = 0$. Otherwise, we say $\alpha$ is transcendental.

Equivalently, consider the natural surjection $\varphi : F[x] \to F[\alpha]$. Its kernel $I$ is an ideal of $F[x]$ and so has the form $(p(x))$ for some $p(x) \in F[x]$. Then $\alpha$ is transcendental if and only if $I = 0$. In this case, $F[\alpha] \cong F[x]$ and $F(\alpha) \cong F(x)$. If $I \neq 0$, then $p(x)$ is a nonconstant polynomial. Without loss of generality, assume $p$ is monic. Since $F[\alpha]$ is an integral domain, we see that $I$ is a prime ideal and equivalently a maximal ideal. Hence $F[\alpha] \cong F[x]/(p(x))$ is a field and $p$ is irreducible. This polynomial $p$ is the **minimal polynomial** of $\alpha$. It is the monic polynomial of minimal degree that vanishes at $\alpha$. Suppose $\deg(p) = d$, then $\{1, \alpha, \alpha^2, \ldots, \alpha^{d-1}\}$ is a basis of $F[\alpha]$ over $F$. That is, $[F[\alpha] : F] = \deg(p)$.

Conversely, for any irreducible polynomial $p(x)$, there is a finite field extension of $F$ in which $p$ has a root. Namely the field $F[x]/(p)$ in which the coset $x + (p)$ is a root of $p$. This is known as Kronecker's Theorem.

**Corollary 8.3.** Let $E/F$ be a field extension. Then an element $\alpha \in E$ is algebraic over $F$ if and only if $F(\alpha)/F$ is finite.

**Corollary 8.4.** If $E/F$ is a finite field extension, then every element of $E$ is algebraic over $F$.

Given an extension $E/F$, the set of elements of $E$ algebraic over $F$ forms a subfield containing $F$, called the **algebraic closure** of $F$ in $E$. You proved this earlier in a homework. We now give a similar proof. Suppose $\alpha, \beta$ are algebraic over $F$. Then $F(\alpha)$ and $F(\beta)$ are finite over $F$. Any basis of $F(\alpha)$ over $F$ gives a spanning set of $F(\alpha, \beta)$ over $F(\beta)$. Hence $F(\alpha, \beta)$ is finite over $F(\beta)$ and so is finite over $F$. Then all elements of $F(\alpha, \beta)$, in particular $\alpha+\beta$, $\alpha\beta$, $\alpha/\beta$, are all algebraic over $F$. If every element of $E$ is algebraic over $F$, then we say the extension $E/F$ is algebraic. Corollary 8.4 implies that every finite extension is algebraic. The converse is not true.

Consider the extension $\mathbb{C}/\mathbb{Q}$. Denote by $\bar{\mathbb{Q}}$ the algebraic closure of $\mathbb{Q}$ in $\mathbb{C}$. Then $\bar{\mathbb{Q}}$ is algebraic over $\mathbb{Q}$, but not finite.

## 8.2 Examples

1. What is the degree of the extension $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ over $\mathbb{Q}$? It strictly contains $\mathbb{Q}(\sqrt{2})$ and so has degree divisible by and larger than 2. Check that $\mathbb{Q}(\sqrt{2}, \sqrt{3}) = \mathbb{Q}(\sqrt{2}+\sqrt{3})$. Find a degree 4 polynomial vanishing at $\sqrt{2}+\sqrt{3}$. This shows that the degree is 4. There are three intermediate fields $\mathbb{Q}(\sqrt{2})$, $\mathbb{Q}(\sqrt{3})$ and $\mathbb{Q}(\sqrt{6})$. As we will see with Galois theory later, these three fields correspond to the 3 proper subgroups of $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

2. Suppose $E/F$ is a field extension and suppose $\alpha, \beta \in E$ are algebraic over $F$. Write $[F(\alpha) : F] = m$ and $[F(\beta) : F] = n$. Suppose $\gcd(m, n) = 1$, then $[F(\alpha, \beta) : F] = mn$. Indeed, it is divisible by $m$ and $n$ and so by $mn$ since they are coprime. On the other hand, it is at most $mn$. As an example, the extension $\mathbb{Q}(\sqrt[3]{2}, i)/\mathbb{Q}$ has degree 6.

3. Let $p$ be a prime and let $\zeta_p = \cos(2\pi/p) + i\sin(2\pi/p)$ be a primitive $p$-th root of unity in $\mathbb{C}$. That is, $\zeta_p \neq 1$ is a root of $x^p - 1$. The extension $\mathbb{Q}(\zeta_p)$ is called the $p$-th cyclotomic extension of $\mathbb{Q}$. What is its degree? The $p$-th cyclotomic polynomial is

$$\Phi_p(x) = \frac{x^p-1}{x-1} = x^{p-1} + x^{p-2} + \cdots + x + 1.$$

We show it is irreducible, which will then imply that $[\mathbb{Q}(\zeta_p) : \mathbb{Q}] = p-1$. It suffices to consider

$$\Phi_p(y+1) = \frac{(y+1)^p - 1}{y} = y^{p-1} + \binom{p}{1} y^{p-2} + \cdots + \binom{p}{p-1}.$$

Note that the leading coefficient is not divisible by $p$ while all the other coefficients are divisible by $p$ and the constant term is not divisible by $p^2$. We now prove a general lemma, called Eisenstein's Criterion, that shows that such polynomials are irreducible.

**Lemma 8.5.** Let $R$ be an UFD with field of fraction $F$ and let $p$ be a prime of $R$. Let $f(x) = a_n x^n + \cdots + a_0 \in R[x]$ be a polynomial such that $p \nmid a_n$, $p \mid a_i$ for all $i = 0, \ldots, n-1$ and $p^2 \nmid a_0$. Then $f(x)$ is irreducible in $F[x]$.

It suffices to show that $f(x)$ does not factor properly in $R[x]$. Suppose we have

$$f(x) = (b_k x^k + \cdots + b_0)(c_m x^m + \cdots + c_0)$$

with $k$ and $m$ both nonzero. Then $p \mid b_0 c_0$ but $p^2 \nmid b_0 c_0$. Without loss of generality, assume $p \mid b_0$ and $p \nmid c_0$. Let $l$ be the smallest positive integer such that $p \nmid b_l$. Since $p \nmid a_n$, we know this $l$ exists and $l \leq k < n$. Then $a_l = b_l c_0 + b_{l-1} c_1 + \cdots b_0 c_l$ is not divisible by $p$. Contradiction. □

## 8.3 Algebraically closed fields

The field $\mathbb{C}$ is special in that the fundamental theorem of algebra says that every nonzero nonconstant polynomial has a root in $\mathbb{C}$. In other words, the irreducible polynomials are the linear ones but $\mathbb{C}[x]/(x-a) \cong \mathbb{C}$ for any $a \in \mathbb{C}$. Hence $\mathbb{C}$ has only one algebraic extension, namely $\mathbb{C}$ itself.

**Definition 8.6.** A field is **algebraically closed** if every nonzero nonconstant polynomial has a root. Equivalently, it has no proper algebraic extensions.

**Theorem 8.7.** Every field $F$ is contained in an algebraically closed field.

The trick is to start from any field $F$, and construct a field $F_1$ such that every nonzero nonconstant polynomial with coefficients in $F$ has a root in $F_1$. Then construct a field $F_2$ from $F_1$ similarly. Their union $F_\infty = \cup_{i=1}^{\infty} F_i$ will do the job. Indeed any polynomial in $F_\infty$ has finitely many coefficients and hence belong to some $F_N[x]$ and so has a root in $F_{N+1} \subset F_\infty$.

Let $S$ be the set of monic irreducible polynomials in $F[x]$. For each $p \in S$, let $x_p$ denote an indeterminant indexed by $p$. Let $R$ be the polynomial ring $F[x_p]_{p \in S}$ in a lot of variables indexed by $S$. Let $I$ be the ideal of $R$ generated by $p(x_p)$ for every $p \in S$. Then every irreducible polynomial in $F$ has a root in $R/I$. It then remains to show that $I$ is a proper ideal, for then there is a maximal ideal $\mathfrak{m}$ containing it and we take $F_1$ to be $R/\mathfrak{m}$.

Suppose for a contradiction that $1 = r_1 p_1(x_{p_1}) + \cdots + r_m p_m(x_{p_m})$ for some $p_1, \ldots, p_m \in S$ and $r_1, \ldots, r_m \in R$. Setting all the indeterminants $x_p$ for $p \neq p_1, \ldots, p_m$ to 0 gives an equality in $F[x_{p_1}, \ldots, x_{p_m}]$. By Kronecker's Theorem, there is a finite extension $L$ of $F$ where $p_1, \ldots, p_m$ all have a root, namely $\alpha_1, \ldots, \alpha_m$ respectively. There is a ring homomorphism $L[x_{p_1}, \ldots, x_{p_m}] \to L$ sending $x_{p_i}$ to $\alpha_i$ that sends $1 = r_1 p_1(x_{p_1}) + \cdots + r_m p_m(x_{p_m})$ to 0. Contradiction. □

**Corollary 8.8.** Every field $F$ is contained in an algebraically closed field $E$ such that $E/F$ is algebraic. The field $E$ is called an algebraic closure of $F$.

Let $F_\infty$ be an algebraically closed field containing $F$. Let $E$ be the subfield of elements algebraic over $F$. Any element $\alpha$ of $F_\infty$ algebraic over $E$ is the root of a polynomial with coefficient in some finite subextension $K$ over $F$. Then $K(\alpha)$ is finite over $F$ and so $\alpha$ is algebraic over $F$ and so lies in $E$. Hence $E$ is algebraically closed. □

## 8.4 Homework

**Exercise 8.4.1.** Show that for any positive integer $n$, $\cos(2\pi/n)$ and $\sin(2\pi/n)$ are algebraic numbers. Are they algebraic integers?

**Exercise 8.4.2.** Show that the polynomial $x^2 + y^3 + z^5$ is irreducible in $\mathbb{C}[x, y, z]$. Hint: You may use the following statement without proof (we will prove it in class soon): Let $F$ be a field and let $f(x) \in F[x]$. Then $f$ has no repeated factors over the algebraic closure $\bar{F}$ if and only if $f$ and its (formal) derivative $f'$ are coprime.

**Exercise 8.4.3.** Show that algebraic closures are isomorphic. That is, let $F$ be a field and let $E_1$ and $E_2$ be two algebraically closed field containing $F$. Suppose both $E_1$ and $E_2$ are algebraic over $F$. Show that there is an isomorphism between them that acts as the identity on $F$ (such an isomorphism is called an $F$-isomorphism). Hint: Let $\Sigma$ be the set of pairs $(L_1, L_2)$ where $L_i$ is a subfield of $E_i$ containing $F$ and where $L_1$ and $L_2$ are isomorphic by an $F$-isomorphism. Define a partial order on $\Sigma$ by $(L_1, L_2) \leq (K_1, K_2)$ if $L_1 \subset K_1$ and $L_2 \subset K_2$. Apply Zorn's Lemma to obtain a maximal element in $\Sigma$ and show that that element is $(E_1, E_2)$.

# 9 Splitting Fields

As we saw last time, a natural way to obtain a field extension $E/F$ is to start with an irreducible polynomial $f(x) \in F[x]$ and take $E = F[x]/(f(x))$. This field $E$ contains a root of $f(x)$, namely the coset of $x$. What if we want all the roots of $f(x)$? That is, what if we want $f(x)$ to factor as a product of linear terms?

**Definition 9.1.** Let $E/F$ be a field extension and let $f(x) \in F[x]$. We say that $f(x)$ **splits** in $E$ if it is a product of linear polynomials in $E[x]$. We say that $E$ is a **splitting field** of $F$ is $f(x)$ splits in $E$ and there is no proper subfield of $E$ over which $f(x)$ splits.

For example, a splitting field of $x^2 - 2$ over $\mathbb{Q}$ is $\mathbb{Q}(\sqrt{2})$. For $x^3 - 2$, we need $\mathbb{Q}(\sqrt[3]{2}, \zeta_3)$. For $x^2 - t$ over $F(t)$, we can take $F(\sqrt{t})$.

## 9.1 Existence

**Proposition 9.2.** Let $F$ be a field and let $f(x)$ be a polynomial in $F[x]$. Then there exists a finite extension $E/F$ over which $f(x)$ splits. Let $\alpha_1, \dots, \alpha_n$ be all the roots of $f(x)$ in $E$, then $F(\alpha_1, \dots, \alpha_n)$ is a splitting field of $f(x)$.

We induct on the degree of $f(x)$. When $\deg(f) \leq 1$, we take $E = F$. More generally, suppose $\deg(f) > 1$. Then there exists a finite extension $L/F$ over which $f(x)$ has a root, say $\alpha$. Then there exists a polynomial $g(x) \in L[x]$ such that $f(x) = (x - \alpha)g(x)$. (Divide $f(x)$ by $x - \alpha$ to get a quotient $g(x)$ and a remainder $r(x)$ that has to be a constant, then plug in $x = \alpha$.) Now $\deg(g) < \deg(f)$ so by the induction hypothesis, there exists a finite extension $E/L$ over which $g(x)$ splits. Then $f(x)$ also splits over $E$. The second statement is obvious. □

**Proposition 9.3.** Let $F$ be a field and let $f(x)$ be a polynomial in $F[x]$ of degree $n$. Suppose $E/F$ is a splitting field of $f(x)$. Then $[E : F] \mid n!$.

Without loss of generality, assume $f(x)$ is monic. We again induct on the degree $n$ of $f(x)$. The statement is obvious when $n \leq 1$. Suppose now $n > 1$. Let $\alpha \in E - F$ be a root of $f(x)$ and let $g(x)$ be a polynomial of $F(\alpha)[x]$ with $f(x) = (x - \alpha)g(x)$. Then $g(x)$ splits over $E$ and not over any subfield of $E$ for then $f(x)$ would also be split over it. Hence $E$ is a splitting field of $g(x)$. By induction hypothesis as $\deg(g) = n - 1 < n$, we have $[E : F(\alpha)] \mid (n-1)!$.

If $f(x)$ is irreducible, then it is the minimal polynomial of $\alpha$ over $F$. Hence $[F(\alpha) : F] = n$ and so $[E : F] \mid n!$. If $f(x)$ is not irreducible, then there are polynomials $h_1(x), h_2(x) \in F[x]$ such that $f = h_1 h_2$ and $\deg(h_i) < \deg(f)$. Let $K \subset E$ be a splitting field of $h_1(x)$ over $F$. Then $E/K$ is a splitting field of $h_2(x)$. By induction hypothesis, we have $[K : F] \mid (\deg(h_1))!$ and $[E : K] \mid (\deg(h_2))!$. Their product then divides $(\deg(h_1) + \deg(h_2))! = n!$. □

## 9.2 Uniqueness

We have been careful and using the phrase "a splitting field" instead of "the splitting field" because from the definitions, it is not a priori clear that all splitting fields are isomorphic. The goal of this section is show that in fact they are isomorphic. The following slightly generalized version will help us in the induction.

**Proposition 9.4.** Let $\varphi : F_1 \to F_2$ be an isomorphism of fields and let $\varphi[x] : F_1[x] \to F_2[x]$ denotes its natural extension. Let $f_1(x) \in F_1[x]$ and let $f_2(x)$ be its image under $\varphi[x]$. Let $E_1$ be a splitting field of $f_1$ over $F_1$ and let $E_2$ be a splitting field of $f_2$ over $F_2$. Then there exists a field isomorphism $\widetilde{\varphi} : E_1 \to E_2$ extending $\varphi$.

As before, we induct on $\deg(f_1)$. When $\deg(f_1) \leq 1$, the statement is obvious. Suppose now $\deg(f_1) > 1$. Let $p_1(x)$ be an irreducible factor of $f_1(x)$ (which may be $f_1(x)$ itself). Set $p_2 = \varphi[x](p_1)$. Then we have an isomorphism $\varphi'$ of fields $F_1[x]/(p_1(x)) \cong F_2[x]/(p_2(x))$. Let $\alpha_1 \in E_1$ be a root of $p_1$ and let $\alpha_2 \in E_2$ be a root of $p_2$. Then we have

$$\varphi' : F_1(\alpha_1) \cong F_1[x]/(p_1(x)) \cong F_2[x]/(p_2(x)) \cong F_2(\alpha_2).$$

Let $h_1(x) \in F_1(\alpha_1)[x]$ be the polynomial such that $f_1(x) = (x - \alpha_1)h_1(x)$. Set $h_2(x) = \varphi'[x](h_1)$. Then $f_2(x) = (x-\alpha_2)(h_2(x))$. Moreover, $E_1/F_1(\alpha_1)$ is a splitting field of $h_1(x)$ while $E_2/F_2(\alpha_2)$ is a splitting field of $h_2(x)$. Since $\deg(h_1) < \deg(f_1)$, by induction hypothesis, the isomorphism $\varphi'$ extends to an isomorphism $\widetilde{\varphi} : E_1 \to E_2$. □

**Corollary 9.5.** Let $E_1, E_2$ be two splitting fields of $f(x)$ over $F$. Then there exists an isomorphism $E_1 \to E_2$ fixing $F$.

**Corollary 9.6.** Let $E/F$ be the splitting field of a polynomial $f(x) \in F[x]$. Let $g(x)$ be an irreducible polynomial in $F[x]$ with a root in $E$. Then $g(x)$ also splits over $E$.

Let $\alpha \in E$ be a root of $g(x)$. Let $K/E$ be a splitting field over $g(x)$. Then $g(x)$ factors as $c(x - \alpha_1)\cdots(x-\alpha_n)$ with $c \in F$, $\alpha_i \in K$ and $\alpha_1 = \alpha$. We need to show all the $\alpha_i$'s lie in $E$. Fix some $i = 2, \dots, n$. Since $g(x)$ is irreducible, there is an isomorphism $\varphi : F(\alpha) \to F[x]/(g(x)) \to F(\alpha_i)$. Since $\varphi$ is the identity on $F$, the extension $\varphi[x]$ fixes $g(x)$ and $f(x)$. Since $K$ is the splitting field of $g(x)f(x)$ over $F(\alpha)$ and over $F(\alpha_i)$. By Proposition 9.4, there is an isomorphism $\widetilde{\varphi} : K \to K$ extending $\varphi$. Since $\widetilde{\varphi}$ is the identity on $F$, it sends roots of $f(x)$ to roots of $f(x)$. Since $E$ is generated by the roots of $f(x)$, $\widetilde{\varphi}$ preserves the subfield $E$. Hence $\alpha_i = \widetilde{\varphi}(\alpha) \in E$. □

## 9.3 Normal extensions

**Definition 9.7.** An algebraic extension $E/F$ is **normal** if every irreducible polynomial $f(x) \in F[x]$ either has no root or splits in $E$.

Corollary 9.6 gives many examples of normal extensions, namely just take the splitting field of some polynomial. Our next result shows that all finite normal extensions arise this way.

**Proposition 9.8.** Let $E/F$ be a finite normal extension. Then $E$ is the splitting field of some polynomial $f(x) \in F[x]$ over $F$.

Since $E/F$ is finite, there are elements $\alpha_1, \dots, \alpha_n \in E$ such that $E = F(\alpha_1, \dots, \alpha_n)$. For each $i = 1, \dots, n$, let $p_1(x)$ denote the minimal polynomial of $\alpha_i$ in $F[x]$. Let $f(x) = p_1(x)\cdots p_n(x)$. Since $E/F$ is normal and each $p_i(x)$ has a root in $E$, we see that $p_i(x)$ and hence $f(x)$ split over $E$. Since $E$ is generated by the roots of $f(x)$, we see that $E$ is the splitting field of $f(x)$. □

**Corollary 9.9.** Suppose $E/F$ is a quadratic extension, that is an extension of degree 2. Then $E/F$ is normal.

Let $\alpha \in E - F$. Then $E = F(\alpha)$. Let $p(x)$ be the minimal polynomial of $\alpha$. Then $p$ has degree 2 and so has the form $x^2 + ax + b$ for some $a, b \in F$. Since $\alpha$ is a root of $b$, we see that the other root of $p$ is $-a - \alpha \in E$. Hence $E$ is the splitting field of $p(x)$. □

For example, the extensions $\mathbb{Q}(\sqrt{2})/\mathbb{Q}$ and $\mathbb{Q}(\sqrt[4]{2})/\mathbb{Q}(\sqrt{2})$ are quadratic and so are normal. The total extension $\mathbb{Q}(\sqrt[4]{2})/\mathbb{Q}$ is not normal since the polynomial $x^4 - 2$ is irreducible (by Eisenstein's criterion with $p = 2$) and it has two complex roots not contained in $\mathbb{Q}(\sqrt[4]{2})$. Hence in general, if $E/K$ and $K/F$ are normal, then $E/F$ is not necessarily normal.

**Proposition 9.10.** Suppose $E/F$ is a normal extension and $K$ is an intermediate field. Then $E/K$ is normal.

Let $f(x)$ be an irreducible polynomial in $K[x]$ with a root $\alpha$ in $E$. Since $E/F$ is algebraic, $\alpha$ is algebraic over $F$. Let $p(x)$ be the minimal polynomial of $\alpha$ over $F$. Since $p(x)$ is a polynomial in $K[x]$ vanishing at $\alpha$, we see that $f(x) \mid p(x)$. Since $E/F$ is normal and it contains a root of $p(x)$, it also contains all the roots of $p(x)$ and so all the roots of $f(x)$. □

The extension $K/F$ is generally not normal. For example $E = \mathbb{Q}(\sqrt[3]{2}, \zeta_3)$, $K = \mathbb{Q}(\sqrt[3]{2})$ and $F = \mathbb{Q}$. Then $E/F$ is normal but $K/F$ is not.

Consider the extension $\mathbb{F}_p(\sqrt[p]{t})/\mathbb{F}_p(t)$. We claim this is normal. Let $f(x) \in \mathbb{F}_p(t)[x]$ be the (irreducible) polynomial $x^p - t$. Then $f(x)$ splits in $\mathbb{F}_p(\sqrt[p]{t})$ as $f(x) = (x - \sqrt[p]{t})^p$. Hence $\mathbb{F}_p(\sqrt[p]{t})$ is the splitting field of $x^p - t$ over $\mathbb{F}_p(t)$ and so the extension is normal. Any automorphism $\varphi$ of $\mathbb{F}_p(\sqrt[p]{t})$ fixing $\mathbb{F}_p(t)$ has to send $\sqrt[p]{t}$ to itself since it is the only root of $f(x)$. Hence $\varphi$ is the identity map.

# 10 Separable extensions

The last example from the previous section involves an irreducible polynomial $f(x)$ that has a lot of repeated roots over a field extension.

**Definition 10.1.** Let $F$ be a field. An irreducible polynomial $f(x) \in F[x]$ is **separable** if it has no repeated roots in any extension $E/F$. In general, a polynomial is **separable** if all of its irreducible factors are separable.

## 10.1 Formal derivatives

The arithmetic of polynomials over a field is very similar to the arithmetic of integers. There is one advantage the polynomials have over the integers, namely formal derivatives.

**Definition 10.2.** For any $f(x) = a_n x^n + \cdots + a_0 \in F[x]$, its **formal derivative** is the polynomial $f'(x) = na_n x^{n-1} + \cdots + 2a_2 x + a_1$.

This formal derivative satisfies all the usual properties of the derivatives you learned in Calculus. It is an $F$-module homomorphism $F[x] \to F[x]$ and it satisfies the Leibniz rule: $(fg)' = f'g + fg'$.

**Lemma 10.3.** Let $f(x) \in F[x]$ and let $\alpha \in F$. Then

1. $x - \alpha \mid f(x)$ in $F[x]$ if and only if $f(\alpha) = 0$;
2. $(x-\alpha)^2 \mid f(x)$ in $F[x]$ if and only if $f(\alpha) = f'(\alpha) = 0$.

In particular, $f(x)$ has a repeated factor over some extension of $F$ if and only if $f(x)$ and $f'(x)$ have a common factor over some extension of $F$, i.e. if and only if $\gcd(f, f') \neq 1$.

For the first statement, divide $f(x)$ by $x - \alpha$. There exists polynomials $q(x)$ and $r(x)$ such that $f(x) = (x-\alpha)q(x) + r(x)$ where $\deg(r) < 1$. That is, $r(x) = r$ is a constant. Setting $x = \alpha$ then gives $f(\alpha) = r$. Hence $x - \alpha \mid f(x)$ if and only if $r = 0$ if and only if $f(\alpha) = 0$.

For the second statement, we may suppose $x - \alpha \mid f(x)$. Write $f(x) = (x-\alpha)g(x)$ for some $g(x) \in F[x]$. Taking derivative gives $f'(x) = g(x) + (x-\alpha)g'(x)$. Hence $(x-\alpha)^2 \mid f$ if and only if $x - \alpha \mid g$ if and only if $x - \alpha \mid f'$. □

**Corollary 10.4.** Suppose $f(x) \in F[x]$ is a polynomial of degree $d$. Then $f$ has at most $d$ roots in $F$.

Induct on $d$. When $d \leq 1$, the result is obvious. In general, if $f$ has no roots, then we are done. Otherwise, let $\alpha$ be a root of $f$. Then $x - \alpha \mid f$ and so there exists $g(x)$ such that $f(x) = (x-\alpha)g(x)$. Now $g$ has degree $d-1$ and so has at most $d-1$ roots. Any root of $f$ is either $\alpha$ or a root of $g$. Hence $f$ has at most $d$ roots. □

**Corollary 10.5.** Suppose $f(x) \in F[x]$ is an irreducible polynomial. Then $f(x)$ is separable if and only if $f'(x) \neq 0$.

Since $f(x)$ is irreducible, $\gcd(f, f') \neq 1$ if and only if $\gcd(f, f') = f$. Since $\deg(f') < \deg(f)$, we have $f \mid f'$ if and only if $f' = 0$. □

What polynomials have zero derivatives? Obviously the constant polynomials do. Are there any more? The answer to this question depends on the characteristic of the field $F$. Recall the characteristic of a ring $R$ is the nonnegative integer $m$ such that the kernel of the canonical map $\mathbb{Z} \to R$ is $m\mathbb{Z}$. If $R$ is an integral domain, then its characteristic is either 0 or a prime number $p$. If $\mathrm{char}(F) = 0$, then $\mathbb{Z}$ injects into $F$ and so $F$ contains $\mathbb{Q} = \mathrm{Frac}(\mathbb{Z})$. If $\mathrm{char}(F) = p$, then $F$ contains $\mathbb{F}_p$. They are called the prime field of $F$.

**Proposition 10.6.** Let $f(x)$ be a polynomial in $F[x]$. Then:

1. If $\mathrm{char}(F) = 0$, then $f' = 0$ if and only if $f = c$ is a constant;
2. If $\mathrm{char}(F) = p$, then $f' = 0$ if and only if $f(x) = g(x^p)$ for some polynomial $g \in F[x]$.

Suppose $f$ is not a constant. Then $f' = 0$ if and only if $na_n = 0$ whenever $a_n \neq 0$ if and only if $n = 0$ whenever $a_n \neq 0$ if and only if the only nonzero monomials that appear are $x^n$ where $n$ is a multiple of $\mathrm{char}(F)$. □

**Definition 10.7.** A field $F$ is **perfect** if every polynomial over $F$ is separable.

**Corollary 10.8.** Let $F$ be a field. Then:

1. If $\mathrm{char}(F) = 0$, then $F$ is perfect;

2. If $\mathrm{char}(F) = p$ and $F$ is perfect if and only if $F^p = F$ (every element of $F$ is a $p$-th power).

If $\mathrm{char}(F) = 0$, then the only irreducible polynomials with zero derivatives are the constants and they obviously have no repeated roots.

Suppose now $\mathrm{char}(F) = p$. Suppose first $F^p = F$. Let $f(x)$ be an irreducible polynomial with zero derivative. Then there exists a polynomial $g(x) \in F[x]$ such that $f(x) = g(x^p)$. Write $g(x) = b_n x^n + \cdots + b_0$. Since $F^p = F$, there exists $c_0, \dots, c_n$ such that $c_i^p = b_i$ for each $i = 0, \dots, n$. Then

$$g(x^p) = c_n^p x^{np} + \cdots + c_0^p = (c_n x^n + \cdots c_0)^p.$$

Since $f(x)$ is irreducible, we must have $n = 0$ and so $f$ is a constant.

Suppose now $F^p \neq F$. Take any $\alpha \in F - F^p$ and take $f(x) = x^p - \alpha$. It is easy to see $f(x)$ has repeated roots over $E = F(\sqrt[p]{\alpha})$. It remains to show that $f(x)$ is irreducible. Suppose $f(x) = g(x)h(x)$ for $g, h \in F[x]$. Without loss of generality, assume $g, h$ are monic. Write $\beta = \sqrt[p]{\alpha} \notin F$. Then $f(x) = (x - \beta)^p$ and so $g(x) = (x - \beta)^r$ and $h(x) = (x - \beta)^s$ for some nonnegative integers $r, s$ with $r + s = p$. Expanding $g$ gives $g(x) = x^r - rx^{r-1}\beta + \cdots$ . Hence $r\beta \in F$. Since $\beta \notin F$, we have $r = 0$ or $r = p$. Hence either $g$ or $h$ is a unit. Therefore $f(x)$ is irreducible. (As $f(x)$ has only one root, we say it is purely inseparable.) □

## 10.2 Classification of finite fields

Let $\bar{\mathbb{F}}_p$ be the algebraic closure of $\mathbb{F}_p$. For any $q = p^m$, let $\mathbb{F}_q$ denote the subset of $\bar{\mathbb{F}}_p$ consisting of roots of $x^q - x$. Then it is easy to check that $\mathbb{F}_q$ is a subfield using the equality $(a + b)^{p^n} = a^{p^n} + b^{p^n}$ for $a, b$ in a field of characteristic $p$.

**Theorem 10.9.** Every finite field is isomorphic to $\mathbb{F}_q$ for some $q = p^n$.

Let $F$ be a finite field. Then it doesn't contain $\mathbb{Q}$ and so has characteristic $p$ for some prime $p$. Again by finiteness, $F$ is a finite extension of $\mathbb{F}_p$, say of degree $n$. Then $F \cong \mathbb{F}_p^n$ as $\mathbb{F}_p$-modules and so has $q = p^n$ elements. The group $F^\times$ has $q - 1$ elements and so by the Euler-phi theorem, every element of $F^\times$ satisfies $x^{q-1} = 1$. Hence every element of $F$ is a root of $x^q - x$. Since $x^q - x$ has nonzero derivative, it has no repeated roots. Hence $F$ is the splitting field of $x^q - x$. By the uniqueness of splitting fields, we have $F \cong \mathbb{F}_q$. □

**Corollary 10.10.** Every finite field is perfect.

Consider $\mathbb{F}_q$ for $q = p^n$. Then for any $a \in \mathbb{F}_q$, $a = a^{p^n} = (a^{p^{n-1}})^p$ is a $p$-th power. □

## 10.3 Separable extensions

**Definition 10.11.** Let $E/F$ be an algebraic extension. An element $\alpha \in E$ is **separable** if its minimal polynomial is separable. If every element of $E$ is separable, we say $E/F$ is a **separable extensions**.

**Corollary 10.12.** Algebraic extensions of perfect fields are separable. (In particular, separable extensions are not necessarily normal.)

**Proposition 10.13.** Suppose $E/F$ is the splitting field of a separable polynomial $f(x)$, then $E/F$ is separable.

(An incomplete proof) Take any $\alpha \in E$ with minimal polynomial $p(x)$. Let $\alpha_1, \dots, \alpha_n$ be the roots of $p(x)$ in $E$. Let $g(x) = (x - \alpha_1) \cdots (x - \alpha_n)$. It suffices to show that $g = p$. Since $g \mid p$, it remains to show that $g(x) \in F[x]$. The coefficients of $g$ are the elementary symmetric polynomials in $\alpha_1, \dots, \alpha_n$. Any automorphism of $E$ fixing $F$ will permute these roots and hence fix the coefficients of $g$. By "separable descent", which we will discuss next week, this means that the coefficients of $g$ are in $F$. □

**Corollary 10.14.** Suppose $E = F(\alpha_1, \dots, \alpha_n)$ is a finite extension of $F$ where every $\alpha_i$ is separable. Then $E/F$ is separable.

For each $i$, let $f_i(x)$ be the minimal polynomial of $\alpha_i$. Let $K/F$ be the splitting field of $f(x) = f_1(x) \cdots f_n(x)$. Since each $\alpha_i$ is separable, $f(x)$ is separable and so $K/F$ is separable. Thus, $E$, being a subfield of $K$, is also separable. □

**Corollary 10.15.** Let $E/F$ be an algebraic extension. Let $L$ be the set of the all elements of $E$ that are separable over $F$. Then $L$ is a subfield of $E$, called the separable closure of $F$ in $E$.

Let $\alpha, \beta$ be any two elements over $E$ separable over $F$. By the above Corollary, $F(\alpha, \beta)$ is separable over $F$. Hence $\alpha + \beta, \alpha\beta, \alpha/\beta$ are all separable. □

# 11 Primitive element theorem

The goal of this section is to prove the following theorem.

**Theorem 11.1.** Let $E/F$ be a finite separable extension. Then there exists an element $\alpha \in E$ such that $E = F(\alpha)$.

Such an element $\alpha$ is called a primitive element. The extension $E/F$ is said to be simple.

For example, if $p$ and $q$ are two distinct primes, then $\mathbb{Q}(\sqrt{p}, \sqrt{q}) = \mathbb{Q}(\sqrt{p}+\sqrt{q})$. Here is an extension that doesn't have any primitive element. Let $p$ be a prime and let $F$ be a field of characteristic $p$. Consider the extension $F(\sqrt[p]{s}, \sqrt[p]{t})/F(s,t)$. There is an intermediate field $F(\sqrt[p]{s}, t)$. The minimial polynomial of $\sqrt[p]{s}$ over $F(s,t)$ is $x^p - s$. Hence $[F(\sqrt[p]{s}, t) : F(s,t)] = p$ and likewise $[F(\sqrt[p]{s}, \sqrt[p]{t}) : F(\sqrt[p]{s}, t)] = p$. So our extension has degree $p^2$. Now for any $u \in F(\sqrt[p]{s}, \sqrt[p]{t})$, there are polynomials $f(x,y)$ and $g(x,y)$ in $F[x,y]$ such that $u = f(\sqrt[p]{s}, \sqrt[p]{t})/g(\sqrt[p]{s}, \sqrt[p]{t})$. Then

$$u^p = \frac{f(\sqrt[p]{s}, \sqrt[p]{t})^p}{g(\sqrt[p]{s}, \sqrt[p]{t})^p} = \frac{\widetilde{f}(s,t)}{\widetilde{g}(s,t)} \in F(s,t).$$

This shows that every element has degree at most $p$ and so cannot generate an extension of degree $p^2$. There are infinitely many irreducible elements $f(s,t)$ in the UFD $F[s,t]$ The polynomial $x^p - f(s,t)$ is then irreducible by Eisenstein's criterion. Suppose for now $F = F^p$. Then there exists a polynomial $g(x,y) \in F[x,y]$ such that $g(\sqrt[p]{s}, \sqrt[p]{t})^p = f(s,t)$. The field $F(s,t)(g(\sqrt[p]{s}, \sqrt[p]{t}))$ is an intermediate extension of degree $p$. As the next result shows, the main problem of this extension being not simple is that there are infinitely many intermediate extensions.

**Theorem 11.2.** A finite extension $E/F$ is simple if and only if it has finitely many intermediate fields.

We prove this result separately depending on whether $F$ is finite or infinite. First suppose $F$ is finite. Then clearly there are only finitely many intermediate fields since there are only finitely many subsets! It remains to show that every finite extension of finite fields is simple. For this, it suffices to show that for any prime $p$ and any positive integer $n$, there exists $\alpha \in \mathbb{F}_{p^n}$ such that $\mathbb{F}_{p^n} = \mathbb{F}_p(\alpha)$. Write $q = p^n$. Every nonzero element of $\mathbb{F}_q$ satisfies $x^{q-1} = 1$. As shown in the above example, we need to show there exists some $\alpha \in \mathbb{F}_q^\times$ of order $q-1$.

Any element of $\mathbb{F}_q^\times$ has order dividing $q-1$. If $\beta$ is an element of order $d \mid q-1$, then $1, \beta, \beta^2, \ldots, \beta^{d-1}$ are distinct roots of $x^d - 1$. Since $x^d - 1$ has degree $d$, it has at most $d$ roots in $\mathbb{F}_q$ by Corollary 10.4 and we have just written them all down. Which of them have order $d$? The order of $\beta^m$ is $d/\gcd(m,d)$. Hence the number of order $d$ elements, assuming it is nonzero, is the Euler-phi function $\phi(d)$. Without the assumption, we have that the number of order $d$ elements is at most $\phi(d)$. The total number of elements of order less than $q-1$ is then $\sum_{d|q-1, 1\le d<q-1} \phi(d)$. We now show that

$$\sum_{d|n, 1\le d\le n} \phi(d) = n \tag{2}$$

for any positive integer $n$. Applying this to $n = q-1$ then shows that elements of order $q-1$ exist.

To prove (2), we simply run the above argument for counting order $d$ elements again but for the cyclic group $\mathbb{Z}/n\mathbb{Z}$. For each $d \mid n$, order $d$ elements exist and so there are $\phi(d)$ elements of order $d$. The left hand side then counts the total number of elements of $\mathbb{Z}/n\mathbb{Z}$ which we know is $n$.

We now consider the case $F$ is infinite. Suppose first $E = F(\alpha)$ is simple. Let $K$ be any intermediate extension. Then $E = K(\alpha)$. Let $g(x) = x^m + c_{m-1}x^{m-1} + \cdots c_0$ be the minimal polynomial of $\alpha$ over $K$. Then $[E : K] = m$. Let $L = F(c_0, \ldots, c_{m-1}) \subset K$. Then $g(x) \in L[x]$ and is irreducible. Hence $g$ is also the minimal polynomial of $\alpha$ over $L$. Hence $[E : L] = m$; but as $L \subset K$, this implies that $K = L = F(c_0, \ldots, c_{m-1})$. In other words, the intermediate extensions are determined by the minimal polynomial of $\alpha$ over it. Let $f(x)$ be the minimal polynomial of $\alpha$ over $F$. Then $g \mid f$ over $E$. There are only finitely many possible factors of $f$ over any field. Hence there are finitely many intermediate fields.

Suppose conversely now that there are only finitely many intermediate fields. Since $E/F$ is finite, there are elements $\alpha_1, \ldots, \alpha_n \in E$ such that $E = F(\alpha_1, \ldots, \alpha_n)$. Without loss of generality, $E = F(\alpha, \beta)$ as the general case follows by induction on $n$. Consider the intermediate fields $F(\alpha + \lambda\beta)$ for every $\lambda \in F$. Since $F$ is infinite, there are infinitely many $\lambda$. By assumption on finiteness, there exists distinct $\lambda_1, \lambda_2$ such that $F(\alpha + \lambda_1\beta) = F(\alpha + \lambda_2\beta) =: K$. Then $\beta = (\lambda_1 - \lambda_2)^{-1}((\alpha + \lambda_1\beta) - (\alpha + \lambda_2\beta)) \in K$ and $\alpha = (\alpha + \lambda_1\beta) - \lambda_1\beta \in K$. Hence $K = E$. □

Theorem 11.2 tells us that in order to prove Theorem 11.1, we need to show that a finite separable extension has finitely many intermediate fields. This is what Galois theory will allow us to do. We will begin the study of Galois theory next time. Galois theory deals with extensions that are separable and normal. Our extension is only separable. How to make it normal?

**Definition 11.3.** A **normal closure** of a finite extension $E/F$ is a finite normal extension $N/F$ such that $E \subset N$ and no proper intermediate extension between $N$ and $E$ is normal over $F$.

For example a normal closure of $\mathbb{Q}(\sqrt[3]{2})$ is $\mathbb{Q}(\sqrt[3]{2}, \zeta_3)$.

**Proposition 11.4.** Every finite extension $E/F$ has a normal closure, unique up to isomorphisms fixing $E$. Moreover, if $E/F$ is separable, its normal closure is also separable.

Since $E/F$ is finite, there are elements $\alpha_1, \ldots, \alpha_n \in E$ such that $E = F(\alpha_1, \ldots, \alpha_n)$. For each $i = 1, \ldots, n$, let $f_i(x)$ denote the minimal polynomial of $\alpha_i$. Let $f(x) = f_1(x) \cdots f_n(x)$ and let $N/F$ be the splitting field of $f(x)$. Then $N/F$ is normal. Any proper subfield of $N$ must miss some root of $f_i(x)$ for some $i$, but if it contains $E$, then it contains $\alpha_i$ and so cannot be normal. We have proved the existence. Note if $E/F$ is separable, then $f(x)$ is separable and so its splitting field is also separable.

If $N'/F$ is another normal closure of $E/F$, then as $f_i(x)$ has roots in $E$ for each $i$, we see by normality that $f(x)$ also splits in $N'$. By minimality, $N'$ is then the splitting field of $f(x)$. Hence there is an isomorphism $N \to N'$ fixing $E$. □

## 11.1 Homework

The goal of this homework is to learn about tensor products.

**Exercise 11.1.1.** Let $R, R_1, R_2$ be commutative rings. Show that to give ring homomorphisms $\varphi_1 : R \to R_1$ and $\varphi_2 : R \to R_2$ is the same as giving a ring homomorphism $\varphi : R \to R_1 \times R_2$.

The tensor product $R_1 \otimes R_2$ will be a product where all the arrows in the above Exercise are reversed. We start with the tensor product of modules. Let $R$ be a commutative ring and let $M$ and $N$ be two $R$-modules. Then $M \times N$ is an $R$-module. Let

$$\mathbb{Z}^{M \times N} = \bigoplus_{(m,n) \in M \times N} \mathbb{Z}(m, n)$$

be the free abelian group with basis $M \times N$, where each $\mathbb{Z}(m, n)$ means a copy of $\mathbb{Z}$ indexed by the element $(m, n)$ of $M \times N$. Let $I$ be the subgroup generated by elements of the form

$$(m_1 + m_2, n) - (m_1, n) - (m_2, n), \quad (m, n_1 + n_2) - (m, n_1) - (m, n_2), \quad (rm, n) - (m, rn).$$

Let $M \otimes_R N$ denote the quotient $M \times N/I$. We denote the image of each $(m, n)$ by $m \otimes n$ and call these *elemental tensors*. Then elements of $M \otimes_R N$ are finite sums of elemental tensors. The above quotient procedure translates into the following identities:

$$\begin{aligned}(m_1 + m_2) \otimes n &= m_1 \otimes n + m_2 \otimes n \\ m \otimes (n_1 + n_2) &= m \otimes n_1 + m \otimes n_2 \\ (rm) \otimes n &= m \otimes (rn).\end{aligned}$$

We may then give $M \otimes_R N$ an $R$-module structure by setting $r.(m \otimes n) = (rm) \otimes n$.

**Exercise 11.1.2.** Compute $\mathbb{Z}/2\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/2\mathbb{Z}$ and $\mathbb{Z}/2\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/3\mathbb{Z}$.

Suppose $M, N, S$ are three $R$-modules. Then a map $f : M \times N \to S$ is $R$-**bilinear** if

$$\begin{aligned} f(m_1 + m_2, n) &= f(m_1, n) + f(m_2, n) \\ f(m, n_1 + n_2) &= f(m, n_1) + f(m, n_2) \\ f(rm, n) &= f(m, rn) = rf(m, n). \end{aligned}$$

Note in particular $f$ is a not a map of $R$-modules.

**Exercise 11.1.3.** (Universal property of tensor product) Let $R$ be a commutative ring and let $M, N, S$ be $R$-modules. Suppose $f : M \times N \to S$ is $R$-bilinear. Let $\alpha : M \times N \to M \otimes_R N$ be the natural map sending $(m, n)$ to $m \otimes n$. Show that there exists a unique $R$-module homomorphism $\beta : M \otimes_R N \to S$ such that $\beta\alpha = f$.

**Exercise 11.1.4.** Let $R$ be a commutative ring with ideal $I$ and a multiplicatively closed subset $S$ containing 1 but not containing 0. Let $M$ be an $R$-module. Show that $M \otimes_R R/I \cong M/IM$ and $M \otimes_R R[S^{-1}] \cong M[S^{-1}]$.

If $M$ and $N$ are themselves rings, then $M \otimes_R N$ also has a ring structure defined by $(m_1 \otimes n_1)(m_2 \otimes n_2) = (m_1 m_2) \otimes (n_1 n_2)$ and extended by the distributive law. Check in private that this multiplication is well-defined.

**Exercise 11.1.5.** Let $R$ and $T$ be two rings and let $\alpha : R \to T$ be a ring homomorphism so we may view $T$ as an $R$-module. Show that $T \otimes_R R[x] \cong T[x]$ as rings.

**Exercise 11.1.6.** Let $\mathbb{R} \to \mathbb{C}$ be the natural inclusion. Show that as rings $\mathbb{C} \otimes_R \mathbb{C} \cong \mathbb{C} \times \mathbb{C}$.

Since every ring has a unique $\mathbb{Z}$-module structure, we may write $R_1 \otimes R_2$ to mean $R_1 \otimes_{\mathbb{Z}} R_2$.

**Exercise 11.1.7.** Let $F_1$ and $F_2$ be two fields. Show that $F_1 \otimes F_2 = 0$ if and only if $F_1$ and $F_2$ have different characteristic.

**Exercise 11.1.8.** Let $L/F$ be a finite separable extension and let $K/F$ be any field extension. Give necessary and sufficient condition for $L \otimes_F K$ to be a field. (Hint: The primitive element theorem.)

Let $R, R_1, R_2$ be three commutative rings and let $\alpha_1 : R \to R_1$ and $\alpha_2 : R \to R_2$ be ring homomorphisms. Let $T$ be another commutative ring and let $\varphi_1 : R_1 \to T$ and $\varphi_2 : R_2 \to T$ be two ring homomorphisms.

**Exercise 11.1.9.** Suppose $\varphi_1\alpha_1 = \varphi_2\alpha_2$. Show that the map $f : R_1 \times R_2 \to T$ defined by $f(r_1, r_2) = \varphi_1(r_1)\varphi_2(r_2)$ is $R$-bilinear. Once we have this, the universal property for tensor products then gives a $R$-module homomorphism $R_1 \otimes R_2 \to T$. Show that this map is also a ring homomorphism.

# 12 Galois extensions

## 12.1 Automorphisms of a field extension

**Definition 12.1.** Let $E/F$ be a field extension. The automorphism group of $E/F$, denoted $\mathrm{Aut}(E/F)$, is the set of all field automorphisms of $E$ that are identity on $F$. It obviously forms a group under composition.

For example, $\mathrm{Aut}(\mathbb{Q}(i)/\mathbb{Q}) = \mathbb{Z}/2\mathbb{Z}$. Any automorphism of $\mathbb{Q}(i)$ fixing $\mathbb{Q}$ has to send $i$ to $\pm i$ and a field homomorphism from $\mathbb{Q}(i)$ to $\mathbb{Q}(i)$ is determined by its value on $i$ and on $\mathbb{Q}$. The automorphism group $\mathrm{Aut}(\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}) = 1$ since $\sqrt[3]{2}$ is the only root of $x^3 - 2$ in $\mathbb{Q}(\sqrt[3]{2})$.

What about $\mathrm{Aut}(\mathbb{Q}(\sqrt[3]{2}, \zeta_3)/\mathbb{Q})$? Any such automorphism sends $\sqrt[3]{2}$ to $\sqrt[3]{2}\zeta_3^i$ for $i = 0, 1, 2$ and sends $\zeta_3$ to $\zeta_3^j$ for $j = 1, 2$. Assuming that every possible combination of $i$ and $j$ is possible, then by computing some compositions, it is easy to see that $\mathrm{Aut}(\mathbb{Q}(\sqrt[3]{2}, \zeta_3)/\mathbb{Q}) \cong S_3$. In these examples, we have been using the following lemmas repeatedly.

**Lemma 12.2.** Let $f(x) \in F[x]$ and $\varphi \in \mathrm{Aut}(E/F)$. Let $\alpha \in E$ be a root of $f(x)$, then $\varphi(\alpha)$ is also a root of $f(x)$.

**Lemma 12.3.** Let $E = F(\alpha_1, \dots, \alpha_n)$ be a field extension of $F$. Suppose $\varphi_1, \varphi_2 \in \mathrm{Aut}(E/F)$ are such that $\varphi_1(\alpha_i) = \varphi_2(\alpha_i)$ for every $i = 1, \dots, n$, then $\varphi_1 = \varphi_2$.

**Corollary 12.4.** Suppose $E/F$ is a finite extension. Then $\mathrm{Aut}(E/F)$ is a finite group.

There exists $\alpha_1, \dots, \alpha_n \in E$ such that $E = F(\alpha_1, \dots, \alpha_n)$. Since $E/F$ is finite and so algebraic, every $\alpha_i$ has a minimal polynomial $p_i(x)$ of degree $d_i$. Any element of $\mathrm{Aut}(E/F)$ is determined by its action on $\alpha_1, \dots, \alpha_n$ and each $\alpha_i$ has to be sent to a root of $p_i$ in $E$ of which there are at most $d_i$. □

**Proposition 12.5.** Suppose $E/F$ is the splitting field of $f(x) \in F[x]$. Then $\#\mathrm{Aut}(E/F) \leq [E : F]$ with equality if and only if $f(x)$ is separable.

As always, we induct on the degree of $f(x)$. If $\deg(f) \leq 1$, then the result is obvious. Suppose $\deg(f) > 1$. Let $\alpha \in E$ be a root of a $f(x)$. Let $\alpha_1 = \alpha, \dots, \alpha_d$ be all the roots of the minimal polynomial $p(x)$ of $\alpha$. There are isomorphisms $\varphi_i : F(\alpha) \to F(\alpha_i)$ fixing $F$ and sending $\alpha$ to $\alpha_i$. Then as $E$ is the splitting field of $f/(x - \alpha)$ and $f/(x - \alpha_i)$ over $F(\alpha)$ and $F(\alpha_i)$ respectively, there are isomorphisms $\widetilde{\varphi}_i : E \to E$ extending $\varphi_i$. Then as sets, $\mathrm{Aut}(E/F) \cong \{\varphi_1, \dots, \varphi_d\} \times \mathrm{Aut}(E/F(\alpha))$. Indeed, given any $\phi : \mathrm{Aut}(E/F)$, it sends $\alpha$ to some $\alpha_i$ and so $\phi\widetilde{\varphi}_i^{-1} \in \mathrm{Aut}(E/F(\alpha))$. Now $E/F(\alpha)$ is the splitting field of $f(x)/(x - \alpha)$ and so by induction, $\#\mathrm{Aut}(E/F(\alpha)) \leq [E : F(\alpha)]$. Moreover, $d \leq \deg(p) = [F(\alpha) : F]$. Hence we have the desired inequality.

Equality happens if and only if $d = \deg(p)$ and $\#\mathrm{Aut}(E/F(\alpha)) = [E : F(\alpha)]$. The former happens if and only if $p$ has no repeated factors and the latter happens if and only if $f(x)/(x-\alpha)$ is separable over $F(\alpha)$. These two combined are equivalent to $f(x)$ being separable (i.e. $\alpha$ is not a repeated root and $f(x)/(x - \alpha)$ has no repeated roots). □

Since $\mathbb{Q}(\sqrt[3]{2}, \zeta_3)/\mathbb{Q}$ is the splitting field of $x^3 - 2$ which is separable, we have that the automorphism group has size 6. The next proposition says that we didn't even need to compute some composition to know that it is $S_3$.

**Proposition 12.6.** Suppose $E/F$ is the splitting field of $f(x) \in F[x]$. Suppose $f(x)$ has $n$ distinct roots. Then $\mathrm{Aut}(E/F) \leq S_n$. If $f(x)$ is irreducible, then $\mathrm{Aut}(E/F)$ is a transitive subgroup of $S_n$. If furthermore $f(x)$ is separable, then $\mathrm{Aut}(E/F)$ has order divisible by $n$.

The splitting field $E$ is generated by the $n$ roots of $f(x)$. An element of $\mathrm{Aut}(E/F)$ is determined then by how it permutes the roots of $f(x)$. Hence we have the first statement.

We say a subgroup $G$ of $S_n$ is **transitive** if for every $i \neq j = 1, \dots, n$, there exists $g \in G$ such that $g(i) = j$. If $f(x)$ is irreducible, then as we saw in the above proof, there exists an automorphism of $E$ fixing $F$ sending any root of $f$ to any other root. Hence $\mathrm{Aut}(E/F)$ is a transitive subgroup of $S_n$. Moreover for any root $\alpha$ of $f$ in $E$, the intermediate extension $[F(\alpha) : F]$ has degree $n$. Hence if $f$ is separable, then $\#\mathrm{Aut}(E/F) = [E : F]$ is divisible by $n$. □

**Corollary 12.7.** Let $f(x)$ be an irreducible polynomial over $\mathbb{Q}$ of degree $p$ for some prime $p$. Suppose $f$ has exactly 2 nonreal roots in $\mathbb{C}$. Then $\mathrm{Gal}(f) \cong S_p$.

As you learned in group theory, $S_n$ is generated by a transposition and an $n$-cycle. In our case, the transposition comes from complex conjugation as it fixes all the real roots and swaps the two nonreal roots. Since $\mathrm{Gal}(f)$ has order divisible by $p$, it contains an element of order $p$ and it must be a $p$-cycle in $S_p$. □

## 12.2 Fixed fields

Let $E/F$ be a field extension. For any $\varphi \in \mathrm{Aut}(E/F)$, write $E^\varphi = \{a \in E : \varphi(a) = a\}$. For any $G \leq \mathrm{Aut}(E/F)$, write $E^G = \cap_{\varphi \in G} E^\varphi$. It is easy to see $E^G$ is an intermediate field between $E$ and $F$.

**Proposition 12.8.** (Separable descent) Let $f(x) \in F[x]$ be a separable polynomial and let $E/F$ be its splitting field. Let $G = \mathrm{Aut}(E/F)$. Then $E^G = F$.

Set $L = E^G$. Since $F \subset L$, we have $\mathrm{Aut}(E/L) \subset \mathrm{Aut}(E/F)$. On the other hand, by definition of $L$, any element of $\mathrm{Aut}(E/F)$ also fixes $L$. Hence $\mathrm{Aut}(E/L) = \mathrm{Aut}(E/F)$. Since $E/F$ is the splitting field of $f(x)$, by viewing $f(x) \in L[x]$ we see that $E/L$ is the splitting field of $f(x)$. Since $f(x)$ is separable (over $F$ and hence over any extension of $F$), we have

$$[E : L] = \#\mathrm{Aut}(E/L) = \#\mathrm{Aut}(E/F) = [E : F].$$

Therefore $L = F$. □

**Proposition 12.9.** Let $E$ be a field and let $G$ be a finite subgroup of the group $\mathrm{Aut}(E)$ of automorphisms of $E$. Then $E/E^G$ is normal and separable.

Write $F = E^G$. Let $\alpha$ be any element of $E$. We need to show that its minimal polynomial over $F$ is separable and splits in $E$. Consider the orbit of $\alpha$ under $G$. Write $\{\varphi(\alpha) : \varphi \in G\} = \{\alpha_1, \ldots, \alpha_m\}$ for some integer $m \leq n$. Set $f(x) = (x - \alpha_1) \cdots (x - \alpha_m)$. Since the elements of $G$ permutes the factors of $f(x)$, we see that $f(x) \in F[x]$. Clearly $f(x)$ is separable and splits in $E$. It remains to show that it is irreducible. Let $g(x)$ be a factor of $f(x)$ in $F[x]$. Without loss of generality, $g(x) = (x - \alpha_1) \cdots (x - \alpha_l)$. If $1 \leq l < m$, then there exists $\varphi \in G$ such that $\{\alpha_1, \ldots, \alpha_l\} \neq \{\varphi(\alpha_1), \ldots, \varphi(\alpha_l)\}$. Then $\varphi(g(x)) \neq g(x)$ contradicting $g(x) \in F[x]$. □

## 12.3 Galois extensions

**Definition 12.10.** An algebraic extension $E/F$ is **Galois** if it is normal and separable. Ihis case, we write $\mathrm{Gal}(E/F)$ instead of $\mathrm{Aut}(E/F)$, called the **Galois group** of the extension.

By what we proved about normal and separable extensions, a finite Galois extension is the splitting field of some separable polynomial $f(x)$. If the degree of $f$ is $n$, then the Galois group is a subgroup of $S_n$.

For example, let $E/\mathbb{Q}$ be the splitting field of $(x^2 - 2)(x^2 - 3)(x^2 - 5)$. Then $E = \mathbb{Q}(\sqrt{2}, \sqrt{3}, \sqrt{5})$ is a degree 8 extension of $\mathbb{Q}$. An element of $\mathrm{Gal}(E/\mathbb{Q})$ is determined by the signs $\sqrt{2} \mapsto \pm\sqrt{2}$, $\sqrt{3} \mapsto \pm\sqrt{3}$, $\sqrt{5} \mapsto \pm\sqrt{5}$. Hence $\mathrm{Gal}(E/\mathbb{Q}) \cong (\mathbb{Z}/2\mathbb{Z})^3$.

**Proposition 12.11.** Let $E$ be a field and let $G$ be a finite subgroup of the group $\mathrm{Aut}(E)$ of automorphisms of $E$. Then $E/E^G$ is a finite Galois extension with Galois group $G$.

Again write $F = E^G$. Since clearly $G \leq \mathrm{Aut}(E/F)$, it remains to show that $[E : F] \leq n = \#G$. Suppose for a contradiction that $[E : F] > n$. Let $\alpha_1, \ldots, \alpha_{n+1}$ be elements of $E$ linearly independent over $F$. Consider the following linear system in the variables $x_1, \ldots, x_{n+1}$:

$$\varphi(\alpha_1)x_1 + \cdots + \varphi(\alpha_{n+1})x_{n+1} = 0, \qquad \forall \varphi \in G. \tag{3}$$

This is a system of $n$ equations in $n + 1$ unknowns and so has nontrivial solutions. Let $(c_1, \ldots, c_{n+1})$ be a solution in $E^{n+1}$ with the least positive number of nonzero entries. Without loss of generality, $c_1, \ldots, c_r \neq 0$

and $c_{r+1} = \cdots = c_{n+1} = 0$ for some $r \leq n$. Dividing everything by $c_r$, we may assume $c_r = 1$. So we have for every $\varphi \in G$,

$$\varphi(\alpha_1)c_1 + \cdots + \varphi(\alpha_{r-1})c_{r-1} + \varphi(\alpha_r)c_r = 0. \tag{4}$$

Our goal is to construct another solution with less zeros. If all of $c_1, \ldots, c_{r-1}$ are in $F$, then $c_i = \varphi(c_i)$ and so $\varphi(\alpha_1 c_1 + \cdots + \alpha_r c_r) = 0$. However, this means that $\alpha_1 c_1 + \cdots + \alpha_r c_r = 0$ contradicting the $F$-linear independence of $\alpha_1, \ldots, \alpha_r$. Hence without loss of generality, we assume $c_1 \notin F$. Then there exists $\phi \in G$ such that $\phi(c_1) \neq c_1$. Applying $\phi$ to (4) and using the fact that $\phi\varphi$, as $\varphi$ runs through elements of $G$, also runs through elements of $G$, we have (remember $c_r = 1$)

$$\varphi(\alpha_1)\phi(c_1) + \cdots + \varphi(\alpha_{r-1})\phi(c_{r-1}) + \varphi(\alpha_r)c_r = 0. \tag{5}$$

Subtracting (5) from (4) shows that $(c_1 - \phi(c_1), \ldots, c_{r-1} - \phi(c_{r-1}), 0, \ldots, 0)$ also gives a nontrivial solution to (3) but it has at most $r - 1$ but at least 1 nonzero entries. Contradiction. □

# 13 Galois correspondence

Galois theory describes all the intermediate fields over a finite Galois extension via the subgroups of Galois group.

## 13.1 Fundamental theorem of Galois theory

**Theorem 13.1.** Let $E/F$ be a finite Galois extension with Galois group $G = \mathrm{Gal}(E/F)$. Then there is an order reversing bijection between the set of intermediate fields of $E/F$ and subgroups of $G$: an intermediate field $L$ is mapped to $\mathrm{Gal}(E/L)$; a subgroup $H$ of $G$ is mapped to $E^H$. Moreover, if $L_1 \subset L_2$ are two intermediate fields and $H_1 \leq H_2$ are two subgroups of $G$, then

$$[\mathrm{Gal}(E/L_1) : \mathrm{Gal}(E/L_2)] = [L_2 : L_1], \qquad [E^{H_1} : E^{H_2}] = [H_2 : H_1].$$

The fact that these two maps are inverses of each other follows from Propositions 12.8 and 12.11:

$$E^{\mathrm{Gal}(E/L)} = L, \qquad \mathrm{Gal}(E/E^H) = H.$$

The statement on the indices follows from Proposition 12.5: $[E : E^H] = \#H$. □

Note if $E/F$ is Galois, then so is $E/L$ for any intermediate field $L$. Recall this follows because if $p(x) \in L[x]$ has a root $\alpha$ in $E$, then it divides the minimal polynomial of $\alpha$ over $F$, which by normality of $E/F$ splits in $E$ and by separability has no repeated roots. The extension $L/F$ is still separable but as we saw may not be normal.

**Proposition 13.2.** Let $E/F$ be a finite Galois extension. Let $L$ be an intermediate field. Then for any $\varphi \in \mathrm{Gal}(E/F)$, $\varphi(L)$ is another intermediate field and $\mathrm{Gal}(E/\varphi(L)) = \varphi \mathrm{Gal}(E/L)\varphi^{-1}$.

It is easy to check that every element of $\varphi\mathrm{Gal}(E/L)\varphi^{-1}$ fixes $\varphi(L)$ and so lies in $\mathrm{Gal}(E/\varphi(L))$. Conversely we also have $\varphi^{-1}\mathrm{Gal}(E/\varphi(L))\varphi \subset \mathrm{Gal}(E/L)$ applying the above with $L$ replaced by $\varphi(L)$ and $\varphi$ replaced by $\varphi^{-1}$. □

**Corollary 13.3.** Let $E/F$ be a finite Galois extension. Let $L$ be an intermediate field. Then $L/F$ is Galois if and only if $\mathrm{Gal}(E/L)$ is a normal subgroup of $\mathrm{Gal}(E/F)$. In this case, $\mathrm{Gal}(L/F) = \mathrm{Gal}(E/F)/\mathrm{Gal}(E/L)$.

By the above proposition, $\mathrm{Gal}(E/L)$ is a normal subgroup of $\mathrm{Gal}(E/F)$ if and only if $\varphi(L)$ and $L$ have the same Galois group if and only if (by the Galois correspondence) $\varphi(L) = L$. Now suppose $f(x) \in F[x]$ is an irreducible polynomial with a root $\alpha$ in $L$. Since $E/F$ is normal, $E$ contains all the roots of $f(x)$. For any root $\beta$ of $f(x)$, there is an isomorphism $\delta : F(\alpha) \to F(\beta)$ fixing $F$. Since $E/F(\alpha)$ and $E/F(\beta)$ are splitting fields of $f(x)$, we see that there is an automorphism of $E$ extending $\delta$. That is, there exists some $\phi \in \mathrm{Gal}(E/F)$ sending $\alpha$ to $\beta$. If $L$ is preserved by every element of $\mathrm{Gal}(E/F)$, then $L$ contains all the roots of $f(x)$. Conversely if $L$ is not preserved by some element $\varphi \in \mathrm{Gal}(E/F)$, let $\alpha \in L$ be an element such that $\varphi(\alpha) \notin L$. Let $f(x)$ be the minimal polynomial of $\alpha$ over $F$. Then $\varphi(\alpha)$ is also a root of $f(x)$. Hence $L/F$ is not normal.

Suppose now $L/F$ is Galois. Then we have a well-defined restriction map $\mathrm{Gal}(E/F) \to \mathrm{Gal}(L/F)$. The kernel of this map is by definition $\mathrm{Gal}(E/L)$. □

**Proof of the Primitive element theorem**: Let $E/F$ be a finite separable extension. We need to show it has finitely many intermediate fields. Let $N/F$ be its normal closure. Then $N/F$ is a finite Galois extension. Any intermediate field of $E/F$ is also one for $N/F$. The intermediate fields of $N/F$ correspond to the subgroups of $\mathrm{Gal}(N/F)$. Since $\mathrm{Gal}(N/F)$ is finite, it has only finitely many subgroups. □

## 13.2 Examples

1. Let $p$ be a prime number and let $q = p^n$. Then $\mathbb{F}_q/\mathbb{F}_p$ is a Galois extension of degree $n$. Let $\sigma$ denote the Frobenius automorphism of $\mathbb{F}_q$ sending $a$ to $a^p$. Then $\sigma \in \mathrm{Gal}(\mathbb{F}_q/\mathbb{F}_p)$. For any positive integer $m$,

$\sigma^m$ sends $a$ to $a^{p^m}$. Since every element of $\mathbb{F}^q$ is fixed by raising to the $p^n$-th power, we see that $\sigma$ has order $n$. Therefore $\mathrm{Gal}(\mathbb{F}_q/\mathbb{F}_p) = \langle\sigma\rangle \cong \mathbb{Z}/n\mathbb{Z}$.

Subgroups of $\mathbb{Z}/n\mathbb{Z}$ are generated by $\sigma^d$ for $d \mid n$. Subfields of $\mathbb{F}_q$ containing $\mathbb{F}_p$ are given by $\mathbb{F}_{p^d}$ for $d \mid n$. The correspondence is evident.

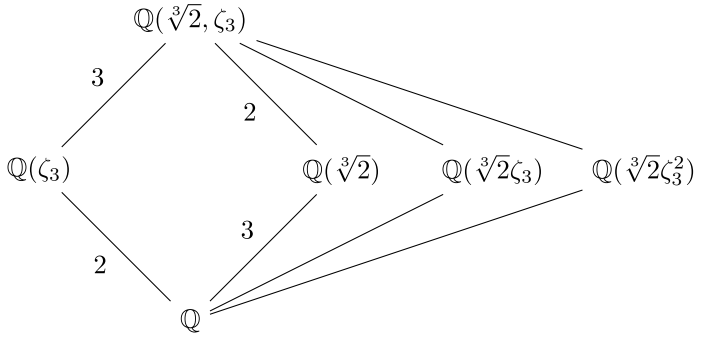
2. Consider the extension $\mathbb{Q}(\sqrt[3]{2}, \zeta_3)/\mathbb{Q}$. This is the splitting field of $x^3 - 2$. The Galois group is $S_3$ generated by $\sigma, \tau$ where

$$\sigma : \begin{array}{ccc} \sqrt[3]{2} & \mapsto & \sqrt[3]{2}\zeta_3 \\ \zeta_3 & \mapsto & \zeta_3 \end{array} \quad \tau : \begin{array}{ccc} \sqrt[3]{2} & \mapsto & \sqrt[3]{2} \\ \zeta_3 & \mapsto & \zeta_3^2. \end{array}$$

In terms of presentations, we have

$$\mathrm{Gal}(\mathbb{Q}(\sqrt[3]{2}, \zeta_3)/\mathbb{Q}) = \langle \sigma, \tau | \sigma^3 = \tau^2 = 1, \tau\sigma\tau = \sigma^2 \rangle.$$

It has one subgroup $\langle\sigma\rangle$ of order 3. The corresponding field is $\mathbb{Q}(\zeta_3)$. This subgroup is normal and so is the field extension $\mathbb{Q}(\zeta_3)/\mathbb{Q}$. There are three non-normal subgroups $\langle\tau\rangle$, $\langle\tau\sigma\rangle$, $\langle\tau\sigma^2\rangle$ of order 2. Their corresponding fields are $\mathbb{Q}(\sqrt[3]{2})$, $\mathbb{Q}(\sqrt[3]{2}\zeta_3)$, $\mathbb{Q}(\sqrt[3]{2}\zeta_3^2)$. They are not normal.


$\mathbb{Q}(\sqrt[3]{2}, \zeta_3)$
3
2
$\mathbb{Q}(\zeta_3)$
$\mathbb{Q}(\sqrt[3]{2})$
$\mathbb{Q}(\sqrt[3]{2}\zeta_3)$
$\mathbb{Q}(\sqrt[3]{2}\zeta_3^2)$
3
2
$\mathbb{Q}$

3. Consider the extension $\mathbb{Q}(\sqrt[5]{7}, \zeta_5)/\mathbb{Q}$. This is the splitting field of $x^5 - 7$. Write $\alpha = \sqrt[5]{7}$ and $E = \mathbb{Q}(\sqrt[5]{7}, \zeta_5)$. Since $\alpha$ has degree 5 and $\zeta_5$ has degree 4 over $\mathbb{Q}$, we see that $[E : \mathbb{Q}] = 20$ and so $G = \mathrm{Gal}(E/\mathbb{Q})$ is a group of order 20. The elements of $G$ are then uniquely determined by two integers $i = 0, \ldots, 4$ and $j = 1, \ldots, 4$: write $\varphi_{i,j}$ for the element

$$\varphi_{i,j} : \alpha \mapsto \alpha\zeta_5^i, \zeta_5 \mapsto \zeta_5^j.$$

Define $\sigma, \tau \in G$ by

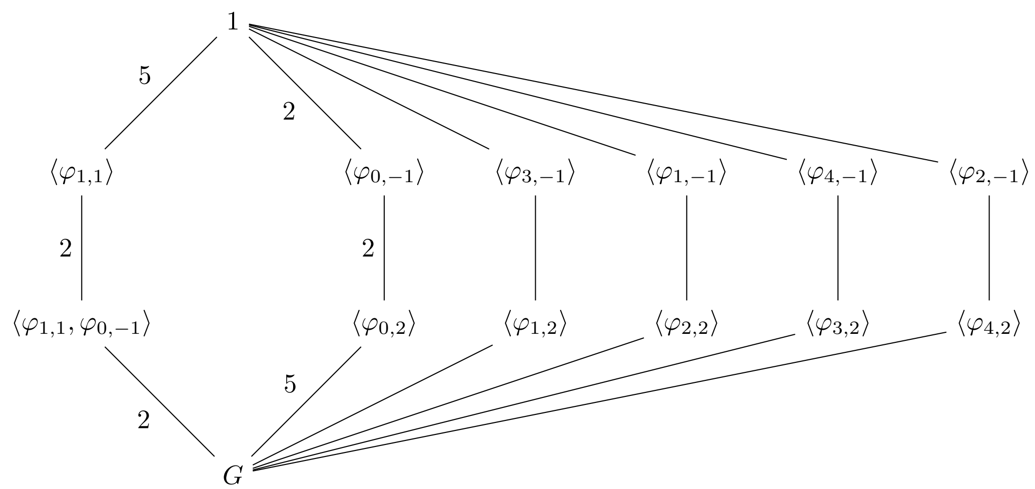
$$\sigma = \varphi_{1,1} : \begin{array}{ccc} \alpha & \mapsto & \alpha\zeta_5 \\ \zeta_5 & \mapsto & \zeta_5 \end{array} \quad \tau = \varphi_{0,2} : \begin{array}{ccc} \alpha & \mapsto & \alpha \\ \zeta_5 & \mapsto & \zeta_5^2. \end{array}$$

Then it is easy to check that $\tau\sigma = \sigma^2\tau$. Hence $G$ has the presentation

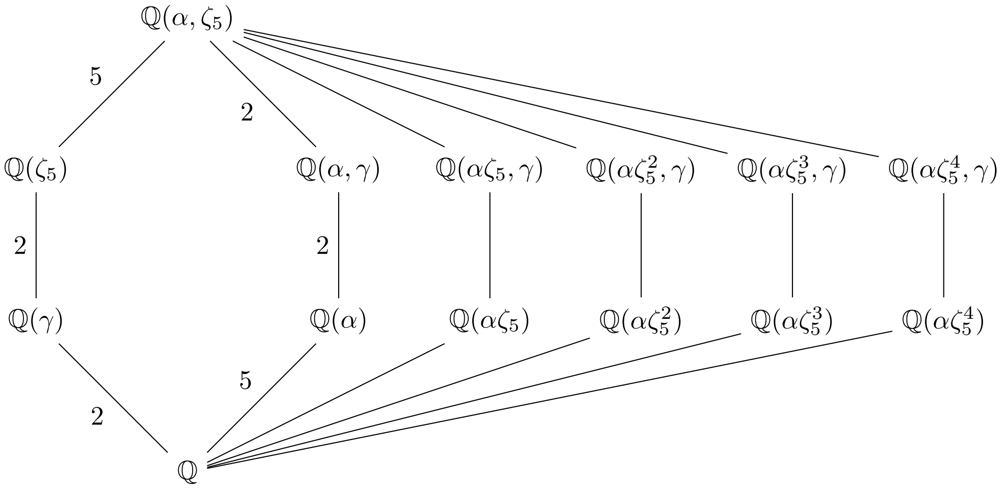
$$G = \langle \sigma, \tau | \sigma^5 = \tau^4 = 1, \tau\sigma = \sigma^2\tau \rangle.$$

What are the subgroups of $G$? They have order $1, 2, 4, 5, 10, 20$. The subgroups of order 5 are Sylow 5-subgroups of $G$. The number of Sylow 5-subgroups divides $\#G = 20$ and is congruent to 1 modulo 5. Hence there is a unique subgroup $P_5 = \langle\sigma\rangle$ of order 5 and it is normal. Any subgroup $H$ of order 10 has to contain $P_5$. An element $\sigma^m\tau^n$ is in $H$ if and only if $\tau^n$ is in $H$. Since $H$ contains no elements of order 4, we see that only $1, \tau^2 \in H$. In other words, there is a unique subgroup of order 10 and it is $H = \langle\sigma, \tau^2\rangle$.

Subgroups of order 4 are the Sylow 2-subgroups. The number of them divides 20 and is congruent to 1 modulo 2. That is, there are either 5 or 1 of them. We can write down one of them, namely $\langle\tau\rangle$. By computing the conjugates $\sigma^m\tau\sigma^{-m}$ for $m = 1, 2, 3, 4$, we find four more: $\langle\varphi_{1,2}\rangle$, $\langle\varphi_{2,2}\rangle$, $\langle\varphi_{3,2}\rangle$, $\langle\varphi_{4,2}\rangle$. Subgroups of order 2 are contained in the Sylow 2-subgroups. Each of the above 5 Sylow 2-subgroup contains a unique subgroup of order 2. We now have the following subgroup lattice:


1
5
2
$\langle\varphi_{1,1}\rangle$
$\langle\varphi_{0,-1}\rangle$
$\langle\varphi_{3,-1}\rangle$
$\langle\varphi_{1,-1}\rangle$
$\langle\varphi_{4,-1}\rangle$
$\langle\varphi_{2,-1}\rangle$
2
2
$\langle\varphi_{1,1},\varphi_{0,-1}\rangle$
$\langle\varphi_{0,2}\rangle$
$\langle\varphi_{1,2}\rangle$
$\langle\varphi_{2,2}\rangle$
$\langle\varphi_{3,2}\rangle$
$\langle\varphi_{4,2}\rangle$
5
2
$G$

To get the subfield lattice, we compute the fixed fields for each of these subgroups making use of the indices. For example, $\zeta_5$ is fixed by $\varphi_{1,1}$ and $\mathbb{Q}(\zeta_5)/\mathbb{Q}$ is a degree 4 extension. So $\mathbb{Q}(\zeta_5) = E^{\varphi_{1,1}}$. Let $\gamma = \zeta_5 + \zeta_5^{-1}$. Then $\gamma$ is fixed by $\varphi_{0,-1}$. Note $\gamma$ is quadratic over $\mathbb{Q}$ as $\gamma^2+\gamma-1 = 0$. Hence $\mathbb{Q}(\gamma) = E^H$. Similarly working through the other subgroups gives the following subfield lattice:


$\mathbb{Q}(\alpha,\zeta_5)$
5
2
$\mathbb{Q}(\zeta_5)$
$\mathbb{Q}(\alpha,\gamma)$
$\mathbb{Q}(\alpha\zeta_5,\gamma)$
$\mathbb{Q}(\alpha\zeta_5^2,\gamma)$
$\mathbb{Q}(\alpha\zeta_5^3,\gamma)$
$\mathbb{Q}(\alpha\zeta_5^4,\gamma)$
2
2
$\mathbb{Q}(\gamma)$
$\mathbb{Q}(\alpha)$
$\mathbb{Q}(\alpha\zeta_5)$
$\mathbb{Q}(\alpha\zeta_5^2)$
$\mathbb{Q}(\alpha\zeta_5^3)$
$\mathbb{Q}(\alpha\zeta_5^4)$
5
2
$\mathbb{Q}$

## 13.3 Homework

The inverse Galois problem asks for the existence of a Galois extension over $\mathbb{Q}$ with a prescribed Galois group. This is still very open! In the first three problems, we show the existence of a Galois extension $E/F$ where $F$ is some extension of $\mathbb{Q}$ (not necessarily $\mathbb{Q}$) with a prescribed Galois group.

**Exercise 13.3.1.** Let $G$ be a finite group. Show that $G$ can be viewed as a subgroup of some symmetric group $S_n$. That is, show that there exists an integer $n$ and an injective group homomomorphism $G \to S_n$.

**Exercise 13.3.2.** Let $n$ be any positive integer at least 2. Show that there exists an irreducible polynomial over $\mathbb{Q}$ that has precisely 2 complex roots and $n-2$ real roots. (Hint: I think this is tricky.)

**Exercise 13.3.3.** Let $G$ be a finite group. Show that there exists a Galois extension $E/F$ with $\mathrm{Gal}(E/F) \cong G$. (Hint: We showed in class that if $n$ in the above proposition is prime, then the Galois group of the splitting field is $S_p$.)

Now some problems on Galois extensions.

**Exercise 13.3.4.** Determine the subgroup and the subfield diagram for the splitting field of $f(x) = (x^2 - 2)(x^2 - 3)(x^2 - 5)$.

**Exercise 13.3.5.** Suppose $K = \mathbb{Q}(\sqrt{2 + \sqrt{2}})$. Show that $K/\mathbb{Q}$ is Galois and determine is Galois group.

**Exercise 13.3.6.** Let $f(x)$ be an irreducible quartic (degree 4) polynomial over $\mathbb{Q}$ whose splitting field $E/\mathbb{Q}$ has Galois group $S_4$. Let $\alpha$ be a root of $f(x)$. Show that $\mathbb{Q}(\alpha)/\mathbb{Q}$ has no proper intermediate fields. (Hint: What is $\mathrm{Gal}(E/\mathbb{Q}(\alpha))$?)

**Exercise 13.3.7.** Consider the field $E = \mathbb{C}(t)$ of rational functions in one variable over $\mathbb{C}$. Let $\zeta_3 = e^{2\pi i/3}$. Define two $\mathbb{C}$-automoprhisms $\sigma, \tau$ of $E$ by $\sigma(t) = \zeta_3 t$ and $\tau(t) = 1/t$. Show that the subgroup $G$ of $\mathrm{Aut}(\mathbb{C}(t)/\mathbb{C})$ generated by $\sigma$ and $\tau$ is isomorphic to $S_6$. Show that $E^G = \mathbb{C}(t^3 + t^{-3})$.

**Exercise 13.3.8.** (Extra fun/credit) Prove that the Galois group for the splitting field of $x^p - 2$ over $\mathbb{Q}$, where $p$ is a prime, is isomorphic to the group

$$G = \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} : a \in \mathbb{F}_p^\times, b \in \mathbb{F}_p \right\}$$

**Exercise 13.3.9.** (More extra fun/credit) Let $K$ be the splitting field of $x^8 - 2$ over $\mathbb{Q}$. Describe the Galois group $\mathrm{Gal}(K/\mathbb{Q})$. Draw the subgroup and subfield diagram.

# 14 Galois groups of polynomials

Let $f(x)$ be a separable polynomial over a field $F$ with splitting field $E$. Then $E/F$ is a Galois extension and we define the Galois group $\mathrm{Gal}(f)$ of $f$ to be $\mathrm{Gal}(E/F)$. The goal of this section is to give an algorithm for computing the Galois group of irreducible polynomials of degrees $2, 3, 4$.

## 14.1 Quadratic polynomial and discriminants

Let $f(x) = x^2 + bx + c$ be a quadratic polynomial over a field $F$. Then its Galois group is a subgroup of $S_2$. It is the trivial group if $f$ is not irreducible and it is $S_2$ if $f$ is irreducible. When the characteristic of $F$ is not 2, then the discriminant $D = b^2 - 4c$ tells us whether $f$ is irreducible or not: $f$ is irreducible if and only if $D$ is not a square.

When the characteristic of $F$ is 2, then $D$ is always a square and it doesn't say anything about the irreducibility of $f$. The polynomial $x^2 + x + 1 \in \mathbb{F}_2[x]$ is (the unique) irreducible of degree 2.

One can define the discriminant in general for any separable polynomial $f(x) = a_n x^n + \cdots + a_0 \in F[x]$. Let $E$ be its splitting field and let $\alpha_1, \ldots, \alpha_n$ denote the roots of $f$ in $E$ (not necessarily distinct). Then the discriminant is defined

$$\Delta(f) = a_n^{2n-2} \prod_{i<j} (\alpha_i - \alpha_j)^2.$$

It is easy to check that the discriminant of $f(x) = ax^2 + bx + c$ is $b^2 - 4ac$. Moreover $\Delta(f) = 0$ if and only if $f$ has repeated roots. Any $\varphi \in \mathrm{Gal}(E/F)$ permutes the roots of $f$ and so leaves $\Delta(f)$ fixed. Hence $\Delta(f) \in F$. In fact since it is symmetric in the roots of $f$, it can be written as a polynomial in the elementary symmetric functions of the roots of $f$ which are themselves simply $\pm a_i/a_n$. The extra factor of $a_n^{2n-2}$ makes sure that $\Delta(f)$ is a polynomial in the coefficients of $f$.

Note if $f$ has repeated factors, then the product $g(x)$ of $x - \alpha_i$ for distinct $\alpha_i$ is a polynomial over $F$ with the same splitting field as $f$. So for the purpose of computing Galois groups, we may assume $f$ is square-free.

**Proposition 14.1.** *Let $F$ be a field of characteristic not 2 and let $f(x)$ be a square-free polynomial over $F$ with degree $n$. THen $\Delta(f)$ is a square in $F$ if and only if $\mathrm{Gal}(f) \leq A_n$, the subgroup of $S_n$ consisting of even permutations.*

We know already that $\mathrm{Gal}(f)$ is a subgroup of $S_n$. Let $\varphi$ be any element of $\mathrm{Gal}(f)$ viewed also as an element of $S_n$ via its permutation action on the roots of $f$. Suppose $\Delta(f) = d^2$ for some $d \in F$. Then $d = \pm a_n^{n-1} \prod_{i<j} (\alpha_i - \alpha_j)$ and without loss of generality we take the positive sign. Since $d \in F$, $\varphi$ fixes it and so $\varphi$ can only changes the order of an even number of pairs $i, j$. That is the number of pairs $(i, j)$ with $i < j$ and $\varphi(i) > \varphi(j)$ is even. Hence $\varphi$ is an even permutation.

Conversely we set $d = a_n^{n-1} \prod_{i<j} (\alpha_i - \alpha_j) \in E$. If every element of $\mathrm{Gal}(f)$ is an even permutation, then it fixes $d$ and so $d \in F$. Hence $\Delta(f)$ is a square. □

## 14.2 Cubic polynomials

Let $f(x) = x^3 + bx^2 + cx + d$ be a cubic polynomial over $F$. Then it is reducible if and only if it has a root in $F$ in which case we are reduced to the quadratic case. We suppose that $f$ is irreducible. Its discriminant is

$$\Delta(f) = b^2c^2 - 4c^3 - 4b^3d - 27d^2 + 18bcd.$$

Of course no one is going to remember this! When the characteristic of $F$ is not 3, then by making a change of variable $x \mapsto x - b/3$ we may assume $f(x)$ has the form $x^3 + Ax + B$ in which case the discriminant is simply $-4A^3 - 27B^2$. As we have seen, $\mathrm{Gal}(f)$ is a transitive subgroup of $S_3$ of order divisible by 3. There are only two such subgroups: $A_3 \cong \mathbb{Z}/3\mathbb{Z}$ and $S_3$ and we can distinguish them by the square-ness of $\Delta(f)$ if the characteristic of $F$ is not 2.

**Proposition 14.2.** *Let $F$ be a field of characteristic not 2 or 3. Let $f(x) = x^3 + Ax + B$ be an irreducible polynomial over $F$. Then $\mathrm{Gal}(f)$ is $S_3$ if $-4A^3 - 27B^2$ is not a square; $\mathbb{Z}/3\mathbb{Z}$ otherwise.*

## 14.3 Quartic polynomials

Let $F$ be a field of characteristic not 2 and let $f(x) = x^4+bx^3+cx^2+dx+e$ be an irreducible polynomial over $F$. By making a change of variable $x \mapsto x-b/4$, we may assume $f(x)$ has the form $x^4+cx^2+dx+e$. Then $\mathrm{Gal}(f)$ is a transitive subgroup of $S_4$ of order divisible by 4. There are 5 of them: $S_4$, $A_4$, $D_4$, $V$, $\mathbb{Z}/4\mathbb{Z}$; where $D_4$ is the dihedral group of order 8 and it contains the Klein group $V = \{1, (12)(34), (13)(24), (14)(23)\} \cong \mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$ with index 2.

Let $\alpha_1, \ldots, \alpha_4$ denote the roots of $f(x)$ in its splitting field $E$. Consider

$$\begin{aligned} u &= \alpha_1\alpha_2 + \alpha_3\alpha_4, \\ v &= \alpha_1\alpha_3 + \alpha_2\alpha_4, \\ w &= \alpha_1\alpha_4 + \alpha_2\alpha_3. \end{aligned}$$

It is easy to see that $S_4$ permutes the set $\{u, v, w\}$ and the subgroup of $S_4$ that fixes $u, v, w$ is $V$. Define

$$g(x) = (x-u)(x-v)(x-w) = x^3 - cx^2 - 4ex + 4ce - d^2$$

. Then the coefficients of $g$ are fixed by every element of $S_4$ and hence by every element of $\mathrm{Gal}(f)$. Hence $g \in F[x]$. Moreover, notice that

$$u - v = (\alpha_1 - \alpha_4)(\alpha_2 - \alpha_3)$$

and similarly for the other two differences. Therefore, $\Delta(g) = \Delta(f)$. The polynomial $g$ is called the **cubic resolvent** of $f$. What is the relation between the Galois groups of $f$ and $g$? Let $L = F(u, v, w)$ be the splitting field of $g$. Then $\mathrm{Gal}(E/L)$ is the subgroup of $\mathrm{Gal}(E/F)$ fixing $u, v, w$. Hence $\mathrm{Gal}(E/L) = \mathrm{Gal}(f)\cap V$ and so

$$\mathrm{Gal}(g) \cong \mathrm{Gal}(f)/(\mathrm{Gal}(f)\cap V).$$

We have the following table:

| $G$ | $S_4$ | $A_4$ | $D_4$ | $V$ | $\mathbb{Z}/4\mathbb{Z}$ |
|---|---|---|---|---|---|
| $G\cap V$ | $V$ | $V$ | $V$ | $V$ | $\mathbb{Z}/2\mathbb{Z}$ |
| $G/G\cap V$ | $S_3$ | $\mathbb{Z}/3\mathbb{Z}$ | $\mathbb{Z}/2\mathbb{Z}$ | 1 | $\mathbb{Z}/2\mathbb{Z}$ |

From this table, we see that unless $g$ factors into a linear term times a quadratic term, the Galois group of $g$ tells us exactly what the Galois group of $f$ is.

Suppose now $g$ factors as a linear polynomial times an irreducible quadratic polynomial. We will use the fact that the Galois groups of $E$ over $L$ are different to distinguish them. Suppose without loss of generality that $u \in F$ is a root of $g$ over $F$. Then we have

$$\begin{aligned} x^2 + (c-u) &= x^2 + (\alpha_1+\alpha_2)(\alpha_3+\alpha_4) = (x-(\alpha_1+\alpha_2))(x-(\alpha_3+\alpha_4)) \\ x^2 - ux + e &= x^2 - (\alpha_1\alpha_2+\alpha_3\alpha_4)x + e = (x-\alpha_1\alpha_2)(x-\alpha_3\alpha_4). \end{aligned}$$

When $\mathrm{Gal}(E/F) \cong \mathbb{Z}/4\mathbb{Z}$, its intersection with $V$ is one of the three order 2 subgroups of $V$. In order $u$ to be fixed by the entire Galois group, we must have $\mathrm{Gal}(f) = \langle(1324)\rangle$ and so $\mathrm{Gal}(f)\cap V = \{1, (12)(34)\}$. As a consequence $\alpha_1\alpha_2, \alpha_1+\alpha_2, \alpha_3\alpha_4, \alpha_3+\alpha_4$ lie inside $L$. That is, the polynomials $x^2+(c-u)$ and $x^2-ux+e$ split over $L$.

Conversely, if $x^2+(c-u)$ and $x^2-ux+e$ split over $L$, then we claim that $\mathrm{Gal}(f)\cap V \neq V$ for if otherwise, we must have $\alpha_1\alpha_2 = \alpha_3\alpha_4 = \delta$ and $\alpha_1+\alpha_2 = \alpha_3+\alpha_4 = \gamma$ which would imply that $\alpha_1, \ldots, \alpha_4$ are all roots of $x^2 - \gamma x + \delta$. Contradiction. Therefore, we must have $\mathrm{Gal}(f)\cap V \cong \mathbb{Z}/2\mathbb{Z}$ and so $\mathrm{Gal}(f) \cong \mathbb{Z}/4\mathbb{Z}$. We summarize our result in the following proposition.

**Proposition 14.3.** Let $f(x) = x^4+cx^2+dx+e$ be an irreducible polynomial over a field $F$ of characteristic not 2. Let $g(x) = x^3 - cx^2 - 4ex + 4ce - d^2$ be its cubic resolvent. Then

$$\mathrm{Gal}(f) \cong \begin{cases} S_4 & \text{if } \mathrm{Gal}(g) \cong S_3, \\ A_4 & \text{if } \mathrm{Gal}(g) \cong \mathbb{Z}/3\mathbb{Z}, \\ D_4 \text{ or } \mathbb{Z}/4\mathbb{Z} & \text{if } \mathrm{Gal}(g) \cong \mathbb{Z}/2\mathbb{Z}, \\ V & \text{if } \mathrm{Gal}(g) = 1. \end{cases}$$

When $\mathrm{Gal}(g) \cong \mathbb{Z}/2\mathbb{Z}$, let $u$ denote the root of $g$ in $F$. Then $\mathrm{Gal}(f) \cong \mathbb{Z}/4\mathbb{Z}$ if and only if both $x^2 + (c - u)$ and $x^2 - ux + e$ split over the splitting field of $g$.

**Examples:**

1. Consider $f(x) = x^4 - 2x - 2 \in \mathbb{Q}[x]$. Then it is irreducible by Eisenstein's criterion. Its cubic resolvent is $g(x) = x^3 + 8x - 4$ which one can easily check has no rational roots. (If $p/q$ is a root with $p, q$ coprime integers, then $p \mid 4$ and $q \mid 1$.) The discriminant of $g$ is $-155.4^2$ which is not a square. Hence $\mathrm{Gal}(g) \cong S_3$ and $\mathrm{Gal}(f) \cong S_4$.

2. Consider $f(x) = x^4 - 10x^2 + 1 \in \mathbb{Q}[x]$. Then via the quadratic formula, we can find all the roots $\pm\sqrt{2} \pm \sqrt{3}$. From this one can already see that $\mathrm{Gal}(f) \cong V$. If we were to follow our algorithm, we would compute the resolvent cubic $g(x) = x^3 + 10x^2 - 4x - 40 = (x + 10)(x - 2)(x + 2)$. Hence $\mathrm{Gal}(g) = 1$ and $\mathrm{Gal}(f) \cong V$.

3. Consider $f(x) = x^4 + 5x + 5 \in \mathbb{Q}[x]$. Then it is irreducible by Eisenstein's criterion. Its cubic resolvent is $g(x) = x^3 - 20x - 25 = (x - 5)(x^2 + 5x + 5)$. So $\mathrm{Gal}(g) \cong \mathbb{Z}/2\mathbb{Z}$. The rational root of $g$ is $u = 5$ and so we look at the polynomials $x^2 - 5$ and $x^2 - 5x + 5$ both of which split over $L = \mathbb{Q}(\sqrt{5})$. Hence $\mathrm{Gal}(f) \cong \mathbb{Z}/4\mathbb{Z}$.

# 15 Cyclotomic extensions and abelian extensions

We have seen the cyclotomic extension $\mathbb{Q}(\zeta_p)/\mathbb{Q}$ before, as the splitting field of the $p$-th cyclotomic polynomial $\Phi_p(x) = x^{p-1} + \cdots + 1$. In this section, we consider the extension $\mathbb{Q}(\zeta_n)/\mathbb{Q}$ where $n$ is an arbitrary positive integer.

## 15.1 The $n$-th cyclotomic polynomial

When $n$ is not a prime, the polynomial $x^{n-1} + \cdots + 1$ is no longer irreducible. For example when $n = 4$, we have

$$x^3 + x^2 + x + 1 = (x^2 + 1)(x + 1).$$

In this case $\zeta_4 = \sqrt{-1}$ with minimal polynomial $x^2 + 1$. How do we get rid of the extra factor of $x + 1$? In the multiplicative group $\langle \zeta_n \rangle$ generated by $\zeta_n$, the order of $\zeta_n^d$ is $n/\gcd(n, d)$. Hence the primitive $n$-th roots of unities correspond to $d$ with $\gcd(n, d) = 1$. This suggests we define

$$\Phi_n(x) = \prod_{\substack{1\le d\le n \\ \gcd(n,d)=1}} (x - \zeta_n^d) = \prod_{\substack{\alpha\in\langle\zeta_n\rangle \\ o(\alpha)=n}} (x - \alpha).$$

From this, it is immediate that

$$x^n - 1 = \prod_{d|n} \Phi_d(x).$$

A priori, $\Phi_n(x)$ is a polynomial with coefficient in $\mathbb{Q}(\zeta_n)$. Any $\varphi \in \mathrm{Aut}(\mathbb{Q}(\zeta_n)/\mathbb{Q})$ preserves the group $\langle \zeta_n \rangle$ and the order of the elements in it. Since $\mathbb{Q}(\zeta_n)/\mathbb{Q}$ is separable, this shows that $\Phi_n(x) \in \mathbb{Q}[x]$ by separable descent. It is also not hard to see that in fact $\Phi_n(x) \in \mathbb{Z}[x]$. Indeed let $c$ be an integer such that $c\phi_n(x)$ is a primitive integral polynomial. Since $\Phi_n(x) \mid x^n - 1$ in $\mathbb{Q}[x]$, we have $c\Phi_n(x) \mid x^n - 1$ in $\mathbb{Z}[x]$ (Gauss Lemma, see the section on UFDs). Looking at the leading coefficient gives $c = \pm 1$ and so $\Phi_n(x) \in \mathbb{Z}[x]$. What is nontrivial is that $\Phi_n(x)$ is irreducible.

**Proposition 15.1.** For any positive integer $n$, $\Phi_n(x)$ is irreducible.

Let $f(x) \in \mathbb{Q}[x]$ be the minimal polynomial of $\zeta_n$, Write $x^n - 1 = f(x)g(x)$ for some $g \in \mathbb{Q}[x]$. Then similar to above, both $f$ and $g$ are in $\mathbb{Z}[x]$. Let $S$ denote the subset of $\langle \zeta_n \rangle$ that are roots of $f$.

**Lemma 15.2.** If $\alpha \in S$, then for any prime number $p$ with $\gcd(n, p) = 1$, we have $\alpha^p \in S$.

Suppose for a contradiction that for some prime $p$ with $\gcd(n, p) = 1$, $\alpha^p \notin S$. Then $g(\alpha^p) = 0$. Consider the reduction $\mathbb{Z}[x] \to \mathbb{F}_p[x]$. Let $\bar{f}, \bar{g}$ denote the reductions of $f$ and $g$ respectively. Then $\bar{f}\bar{g} = x^n - 1$ which has no repeated roots since $\gcd(nx^{n-1}, x^n - 1) = 1$ in $\mathbb{F}_p[x]$. Hence $\gcd(\bar{f}, \bar{g}) = 1$. On the other hand, since $g(\alpha^p) = 0$, $g$ is divisible by the minimal polynomial of $\alpha$, which we know to be $f$. Hence there exists a polynomial $h(x) \in \mathbb{Z}[x]$ such that $g(x^p) = f(x)h(x)$. (A priori $h$ is in $\mathbb{Q}[x]$ but since both $f$ and $g$ are monic integral polynomial, it follows that $h$ is also integral.) Taking reduction mod $p$ gives $(\bar{g}(x))^p = \bar{g}(x^p) = \bar{f}(x)\bar{h}(x)$ which forces $\bar{f}$ and $\bar{g}$ to have common factors. Contradiction. □

Now for any positive integer $k$ coprime to $n$, we can write $k = p_1 p_2 \cdots p_m$ as a product of primes (not necessarily distinct) all coprime to $n$. Then $\zeta_n^{p_1} \in S$ and so $\zeta_n^{p_1 p_2} \in S$ and so on by the above Lemma. Hence $\zeta_p^k$ is a root of $f(x)$ which implies that $\Phi_n(x) \mid f(x)$ in $\mathbb{Q}[x]$. Conversely it is obvious that $f(x) \mid \Phi_n(x)$. Therefore $\Phi_n(x) = f(x)$ is irreducible. □

**Corollary 15.3.** For any positive integer $n$, $\mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \cong (\mathbb{Z}/n\mathbb{Z})^\times$.

Any automorphism of $\mathbb{Q}(\zeta_n)$ over $\mathbb{Q}$ has to send $\zeta_n$ to $\zeta_n^k$ for some $k$ coprime to $n$. Conversely as $\Phi_n(x)$ is irreducible, such maps exist for any $k$ coprime to $n$. □

## 15.2 Quadratic and abelian extensions of $\mathbb{Q}$

We have seen that quadratic extensions are splitting fields of quadratic polynomials. Suppose $E$ is the splitting field of $f(x) = ax^2 + bx + c$ with $a, b, c \in \mathbb{Z}$. Then it is easy to see that $E = \mathbb{Q}(\sqrt{b^2 - 4ac})$. Hence quadratic extensions are all of the form $\mathbb{Q}(\sqrt{D})$ where $D$ is square-free.

**Proposition 15.4.** Let $E$ be any quadratic extension over $\mathbb{Q}$. Then there exists a cyclotomic extension $\mathbb{Q}(\zeta_n)$ containing $E$.

Suppose $E = \mathbb{Q}(\sqrt{D})$ where $D$ is square-free. Factor $D$ as $p_1 p_2 \cdots p_m$ into distinct primes. It suffices to prove the proposition for $D = \pm p$ for if $\sqrt{p_1} \in \mathbb{Q}(\zeta_{n_1})$ and $\sqrt{p_2} \in \mathbb{Q}(\zeta_{n_2})$, then $\sqrt{p_1 p_2} \in \mathbb{Q}(\zeta_{n_1}, \zeta_{n_2}) \subset \mathbb{Q}(\zeta_{n_1 n_2})$. For $p = 2$, we have $(1+i)^2 = 2i$ and so $\sqrt{2i} \in \mathbb{Q}(\zeta_4)$. Since $\sqrt{i} \in \mathbb{Q}(\zeta_8)$, we have $\sqrt{\pm 2} \in \mathbb{Q}(\zeta_8)$.

Suppose now $p$ is odd. Then the discriminant of the $p$-th cyclotomic polynomial is

$$\Delta(\Phi_p(x)) = \prod_{1 \le i < j \le p-1} (\zeta_p^i - \zeta_p^j)^2 = (-1)^{\frac{p-1}{2}} p^{p-2}.$$

This shows that $\sqrt{\pm p} \in \mathbb{Q}(\zeta_{4p})$. □

**Theorem 15.5.** (Kronecker-Weber) Let $E/\mathbb{Q}$ be a finite Galois extension such that its Galois group is abelian. Then there exists a cyclotomic extension $\mathbb{Q}(\zeta_n)$ containing $E$.

You will learn this result in your algebraic number theory class. For this case, we prove a somewhat weaker result.

**Theorem 15.6.** Let $A$ be a finite abelian group. Then there exists a cyclotomic extension $\mathbb{Q}(\zeta_n)$ and a subextension $E$ such that $\mathrm{Gal}(E/\mathbb{Q}) \cong A$.

Suppose $A \cong \mathbb{Z}/k_1\mathbb{Z} \times \cdots \times \mathbb{Z}/k_l\mathbb{Z}$. Suppose we can find primes $p_1 < p_2 < \cdots < p_l$ such that $p_i \equiv 1 \pmod{k_i}$ for $i = 1. \ldots, l$. Then setting $n = p_1 \cdot p_l$, we have

$$G = \mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \cong (\mathbb{Z}/n\mathbb{Z})^\times \cong (\mathbb{Z}/p_1\mathbb{Z})^\times \times \cdots \times (\mathbb{Z}/p_l\mathbb{Z})^\times \cong (\mathbb{Z}/(p_1-1)\mathbb{Z}) \times \cdots \times (\mathbb{Z}/(p_l-1)\mathbb{Z}).$$

Let $H$ be a subgroup of $G$ such that $G/H \cong A$. Then $E = \mathbb{Q}(\zeta_n)^H$ does the job. □

It remains now to show the existence of such a sequence of prime numbers.

## 15.3 Primes in arithmetic progression

It is a theorem of Dirichlet using analytic tools that for any pair of coprime integers $a, m$, there exists infinitely many primes congruent to $a$ modulo $m$. We need a weaker version of this where $a = 1$.

**Proposition 15.7.** For any positive integer $m$, there exists infinitely many primes $p$ congruent to 1 modulo $m$.

The condition $m \mid p-1$ is equivalent to the existence of an integer $\alpha$ coprime to $p$ such that its congruence class mod $p$ has order $m$ in $\mathbb{F}_p^\times$. Heuristically speaking, since the roots of $\Phi_m(x)$ are the primitive $m$-roots of unity, this is equivalent to asking for $\Phi_m(\alpha) \equiv 0 \pmod{p}$. To make this precise, we prove the following lemma.

**Lemma 15.8.** Suppose $p$ is a prime and $m$ is a positive integer coprime to $p$. Suppose $\alpha$ is an integer with $\Phi_m(\alpha) \equiv 0 \pmod{p}$. Then $m \mid p-1$.

Since $x^m - 1$ is a product of $\Phi_d(x)$ for $d \mid m$, we see that $\alpha^m \equiv 1$ (mod ()$p$). We claim that the order of $\alpha$ in $\mathbb{F}_p^\times$ is $m$ for then $m \mid |\mathbb{F}_p^\times|$. Suppose for a contradiction that $o(\alpha) = d < m$. Then $d \mid m$ and so $\alpha$ is a root of $x^d - 1$ modulo $p$. Then $\alpha$ is a root of some $\Phi_{d'}(x)$ for some $d' \mid d$. However, this means that $\alpha$ is a repeated root of $x^m - 1$ modulo $p$. Since $m$ and $p$ are coprime, $x^m - 1$ has no repeated factors modulo $p$. Contradiction. □

It now remains to show that the sequence $\Phi_m(1), \Phi_m(2), \ldots$ has infinitely many prime divisors. This statement is true for any nonconstant polynomial. □

**Lemma 15.9.** Let $f(x) \in \mathbb{Z}[x]$ be a nonconstant polynomial. Then the sequence $f(1), f(2), \ldots$ has infinitely many prime divisors.

Suppose for a contradiction that there are only finitely many prime divisors, namely $p_1, \ldots, p_k$. Let $s$ be a (nonnegative) integer such that $f(s) = l \neq 0$. Let $g(x) = f(s + lp_1 \cdots p_k x)/l$. Then $g(0) = 1$ and the coefficients of $x^t$ is divisible by $l^{t-1}$. Hence $g(x) \in \mathbb{Z}[x]$. Moreover for any $n \in \mathbb{Z}$, $g(n) \equiv 1 \pmod{p_1 \cdots p_k}$ and so none of $p_1, \ldots, p_k$ divides $g(n)$. That is, $g(n)$ has prime factors other than $p_1, \ldots, p_k$. Hence $f(s + lp_1 \cdots p_k n)$ being a multiple of $g(n)$ also has prime factors other than $p_1, \ldots, p_k$. Contradiction. □

We remark that for $m = 2$, $\Phi_2(x) = x + 1$ and we take $s = 0$ which gives $g(x) = p_1 \cdots p_k x + 1$. This is the usual proof of the infinitude of primes. When $m = 4$, we have $\Phi_4(x) = x^2 + 1$ and we take $s = 0$ which gives $g(x) = (p_1 \cdots p_k x)^2 + 1$. This is the usual proof of the infinitude of primes congruent to 1 modulo 4.

# 16 Solvable extensions and solvability by radicals

The goal of this section is to establish the classic result that there is in general no algebraic formula for the roots of a polynomial of degree at least 5. We start by giving the formulae for cubic and quartic polynomials.

## 16.1 The cubic and the quartic formulae

Let $F$ be a field of characteristic not 2 or 3. Any cubic polynomial then can be written in the form $x^3+bx+c$ for $b, c \in F$ after a change of variable. We look for solutions of the form $u+v$ for some $u, v$ to be determined later. The condition we need is

$$u^3+v^3+(3uv+b)(u+v)+c=0.$$

Suppose we further require that $3uv+b=0$. Then the above equation becomes $u^3+v^3=-c$. Write $\alpha=u^3$ and $\beta=v^3$. Then

$$\begin{aligned} \alpha+\beta &= -c, \\ \alpha\beta &= -(b/3)^3. \end{aligned}$$

Hence $\alpha, \beta$ are roots of $x^2+cx-(b/3)^3=0$. We can then use the quadratic formula to compute $\alpha$ and $\beta$. Since $\alpha=u^3$ and $\beta=v^3$, we get three solutions each for $u$ and $v$. Imposing the condition $uv=-b/3$ then gets us down to 3 pairs.

**Proposition 16.1.** The solutions of $x^3+bx+c=0$ are of the form

$$\begin{aligned} x_1 &= \left(-\frac{c}{2}+\sqrt{\frac{c^2}{4}+\frac{b^3}{27}}\right)^{\frac{1}{3}}+\left(-\frac{c}{2}-\sqrt{\frac{c^2}{4}+\frac{b^3}{27}}\right)^{\frac{1}{3}}, \\ x_2 &= \zeta_3\left(-\frac{c}{2}+\sqrt{\frac{c^2}{4}+\frac{b^3}{27}}\right)^{\frac{1}{3}}+\zeta_3^2\left(-\frac{c}{2}-\sqrt{\frac{c^2}{4}+\frac{b^3}{27}}\right)^{\frac{1}{3}}, \\ x_3 &= \zeta_3^2\left(-\frac{c}{2}+\sqrt{\frac{c^2}{4}+\frac{b^3}{27}}\right)^{\frac{1}{3}}+\zeta_3\left(-\frac{c}{2}-\sqrt{\frac{c^2}{4}+\frac{b^3}{27}}\right)^{\frac{1}{3}}, \end{aligned}$$

where the cube roots were chosen so that

$$\left(-\frac{c}{2}+\sqrt{\frac{c^2}{4}+\frac{b^3}{27}}\right)^{\frac{1}{3}}\left(-\frac{c}{2}-\sqrt{\frac{c^2}{4}+\frac{b^3}{27}}\right)^{\frac{1}{3}}=-\frac{b}{3}.$$

We now consider quartic polynomials. After a change of variable, we may assume our polynomial has the form $x^4+cx^2+dx+e$. Suppose that $\alpha_1, \ldots, \alpha_4$ are the four roots. Then $\alpha_1+\alpha_2+\alpha_3+\alpha_4=0$. The resolvent cubic is $g(x)=x^3-cx^2-4ex-d^2+4ce$ whose roots are

$$\begin{aligned} u &= \alpha_1\alpha_2+\alpha_3\alpha_4, \\ v &= \alpha_1\alpha_3+\alpha_2\alpha_4, \\ w &= \alpha_1\alpha_4+\alpha_2\alpha_3. \end{aligned}$$

We can use our cubic formula to compute these $u, v, w$. We now express the roots in terms of $u, v, w$. Note that

$$u+v=(\alpha_1+\alpha_4)(\alpha_2+\alpha_3)=-(\alpha_1+\alpha_4)^2$$

and similarly for the other two pairs. Hence we have

$$\begin{aligned} \alpha_1+\alpha_4 &= \pm\sqrt{-u-v}, \\ \alpha_1+\alpha_3 &= \pm\sqrt{-u-w}, \\ \alpha_1+\alpha_2 &= \pm\sqrt{-v-w}. \end{aligned}$$

Then

$$\begin{aligned}2\alpha_1 &= (\alpha_1+\alpha_4)+(\alpha_1+\alpha_3)+(\alpha_1+\alpha_2)\\ &= (-u-v)^{\frac{1}{2}}+(-u-w)^{\frac{1}{2}}+(-v-w)^{\frac{1}{2}}.\end{aligned}$$

It appears that there are 8 choices for the signs. However, the condition

$$(\alpha_1+\alpha_4)(\alpha_1+\alpha_3)(\alpha_1+\alpha_2) = -d$$

cuts it down to 4.

**Proposition 16.2.** The solutions of $x^4+cx^2+dx+e=0$ are of the form

$$\begin{aligned}x_1 &= \frac{1}{2}((-u-v)^{\frac{1}{2}}+(-u-w)^{\frac{1}{2}}+(-v-w)^{\frac{1}{2}}),\\ x_2 &= \frac{1}{2}(-(-u-v)^{\frac{1}{2}}-(-u-w)^{\frac{1}{2}}+(-v-w)^{\frac{1}{2}}),\\ x_3 &= \frac{1}{2}(-(-u-v)^{\frac{1}{2}}+(-u-w)^{\frac{1}{2}}-(-v-w)^{\frac{1}{2}}),\\ x_4 &= \frac{1}{2}((-u-v)^{\frac{1}{2}}-(-u-w)^{\frac{1}{2}}-(-v-w)^{\frac{1}{2}}).\end{aligned}$$

where the square roots are chosen so that

$$(-u-v)^{\frac{1}{2}}(-u-w)^{\frac{1}{2}}(-v-w)^{\frac{1}{2}}) = -d.$$

Of course, no one is going to remember these formulae! The key thing to note that these formulae involve only taking roots (besides the usual addition, subtraction, multiplication and division). Equivalently, the polynomials split in some extension $L/F$ formed out of only adjoining elements of the form $\sqrt{\alpha}$. These are called radical extensions.

**Definition 16.3.** A finite extension $E/F$ is a **radical extension** if there is a chain of subfields

$$F = F_0 \subset F_1 \subset F_2 \subset \cdots \subset F_m = E$$

such that $F_i = F_{i-1}(\sqrt[d_i]{\alpha_i})$ for some positive integer $d_i$ and element $\alpha_i \in F_{i-1}$ for all $i$. We say a polynomial $f(x) \in F[x]$ is **solvable by radicals** if it splits in some radical extension.

To study radical extensions, we start with their building blocks, namely cyclic extensions.

## 16.2 Cyclic extensions

**Definition 16.4.** A Galois extension $E/F$ is called **cyclic** if its Galois group is cyclic.

**Proposition 16.5.** (Dirichlet's Lemma) Let $K$ and $L$ be fields and let $\varphi_1,\ldots,\varphi_n$ be distinct homomorphisms from $L$ to $K$. Then they are linearly independent. That is, if $c_1,\ldots,c_n \in K$ such that $c_1\varphi_1(\alpha)+\cdots c_n\varphi_n(\alpha) = 0$ for all $\alpha \in L$, then $c_1 = \cdots = c_n = 0$.

Suppose for a contradiction that they are linearly dependent. Consider a minimal (in terms of cardinality) subset of them that is linearly dependent. Without loss of generality, there exists $c_1,\ldots,c_m \in K$ not all 0 such that

$$c_1\varphi_1(\alpha)+\cdots+c_n\varphi_n(\alpha) = 0, \quad \forall\alpha \in L.$$

By minimality, none of the $c_i$ is 0. Let $\beta \in L$ be such that $0 \neq \varphi_1(\beta) \neq \varphi_2(\beta)$. Then we have for every $\alpha \in L$,

$$\begin{aligned}0 &= c_1\varphi_1(\alpha\beta)+\cdots+c_n\varphi_n(\alpha\beta),\\ 0 &= c_1\varphi_1(\alpha)\varphi_1(\beta)+\cdots+c_n\varphi_n(\alpha)\varphi_1(\beta).\end{aligned}$$

Subtracting gives

$$c_2(\varphi_2(\beta)-\varphi_1(\beta))\varphi_2+\cdots+c_n(\varphi_n(\beta)-\varphi_1(\beta))\varphi_n = 0.$$

Since $c_2(\varphi_2(\beta)-\varphi_1(\beta)) \neq 0$, we have a contradiction to the minimality assumption. □

**Theorem 16.6.** (Kummer extensions) Let $F$ be a field and let $n$ be a positive integer such that $\mathrm{char}(F) \nmid n$. Assume that $F$ contains all the $n$-th roots of unity. Then:

1. If $E/F$ is a cyclic extension of degree $n$, then $E = F(\sqrt[n]{\alpha})$ for some $\alpha \in F$.

2. Conversely if $E = F(\sqrt[n]{\alpha})$ for some $\alpha \in F$, then $E/F$ is cyclic of degree $d$ dividing $n$ and $\sqrt[n]{\alpha}^d \in F$.

We prove the second statement first since it will give us an idea on how to find $\alpha$. Write $\beta \in E$ such that $\beta^n = \alpha$. Let $f(x)$ be the minimal polynomial of $\beta$ over $F$ of degree $d$. We claim that $\beta^d \in F$ and so $f(x) = x^d - \beta^d$ for degree reasons. Indeed $f(x)$ divides $x^n - \alpha$. All of its roots are then of the form $\beta\zeta_n^i$ for some $i$. Multiplying all of them gives $f(0) = \beta^d\zeta_n^k$ for some $k$. Since $\zeta_n \in F$ and $f(x) \in F$, we get that $\beta^d \in F$.

Next we prove that $d \mid n$. Divide $n$ by $d$ to get a quotient $q$ and remainder $r$. Then $\beta^r = \beta^n(\beta^d)^{-q} \in F$. If $r \neq 0$, then $\beta$ is also a root of $x^r - \beta^r \in F[x]$ which forces $f(x)$ to divide $x^r - \beta^r$. Since $d > r$, this is a contradiction. Hence $d \mid n$. Write $n = md$. Then the roots of $f(x) = x^d - \beta^d$ are $\beta, \beta\zeta_n^m, \ldots, \beta\zeta_n^{m(d-1)}$. Hence $E = F(\beta)$ is splitting field of $f(x)$. Since $\mathrm{char}(F) \nmid n$, we see that $f(x)$ is separable. Hence $E/F$ is Galois. For any $[l] \in \mathbb{Z}/d\mathbb{Z}$, the map $\varphi_l : E \to E$ sending $\beta$ to $\beta\zeta_n^{ml}$ is an $F$-automorphism. Hence $\mathrm{Gal}(E/F) \cong \mathbb{Z}/d\mathbb{Z}$. Note the generator of the Galois group sending $\beta$ to $\beta\zeta_d$.

We now consider the first statement. Let $\varphi$ be a generator of $\mathrm{Gal}(E/F)$. Then $\varphi^n = \mathrm{id}$. Our proof of the second statement suggest to look for $\beta \in E$ such that $\varphi(\beta) = \beta\zeta_n$. A moment of smartness suggests taking

$$\beta = u + \zeta_n^{-1}\varphi(u) + \cdots + \zeta_n^{-(n-1)}\varphi^{n-1}(u)$$

for some $u \in E$. Then it is easy to check that $\varphi(\beta) = \beta\zeta_n$. Dirichlet's Lemma on the linear independence of $\{1, \varphi, \varphi^2, \cdots, \varphi^{n-1}\}$ implies that there exists some $u$ such that $\beta \neq 0$. We now fix this $u$ and then $\beta$. Then $\beta, \beta\zeta_n, \cdots, \beta\zeta_n^{n-1}$ are conjugates of each other since there are elements in the Galois group sending them to each other. Hence they have the same minimal polynomial $f(x)$. As a result, $[F(\beta) : F] \geq n = [E : F]$ and so $E = F(\beta)$. It remains now to show that $\beta^n \in F$. For this we use separable descent: for any $\varphi^i \in \mathrm{Gal}(E/F)$, we have $\varphi^i(\beta^n) = (\beta\zeta_n^i)^n = \beta^n$. □

For completeness, we describe what happens when $\mathrm{char}(F) \mid n$.

**Theorem 16.7.** (Artin-Schreier extensions) Let $F$ be a field of characteristic $p$. Then cyclic extensions of degree $p$ are splitting fields of irreducible polynomials of the form $x^p - x - \alpha$ for some $\alpha \in F$.

Since every group of size $p$ is cyclic, it suffices to classify degree $p$ Galois extensions of $F$. Suppose first $f(x) = x^p - x - \alpha$ is irreducible. Let $\beta$ be a root of $f$ in some splitting field $E$. It is easy to check that the roots of $f(x)$ are in fact $\beta, \beta+1, \beta+2, \ldots, \beta+(p-1)$. Hence $E = F(\beta)$ has degree $p$ over $F$. Since $f(x)$ has no repeated roots, $E/F$ is Galois.

For the converse, suppose $E/F$ is Galois of degree $p$. Let $\varphi$ be a generator of $\mathrm{Gal}(E/F)$. Then as before, we seek an element $\beta \in E$ such that $\varphi(\beta) = \beta + 1$. Applying Dirichlet's Lemma on the linear independence of the elements of $\mathrm{Gal}(E/F)$ gives a $v \in E$ such that $\gamma = v + \varphi(v) + \cdots + \varphi^{p-1}(v) \neq 0$. Then observe that $\varphi(\gamma) = \gamma$ and so $\varphi^i(\gamma) = \gamma$ for all $i$. Hence $\gamma \in F$. Let $u = v/\gamma$. Then $u + \varphi(u) + \cdots + \varphi^{p-1}(u) = 1$. Set

$$\beta = -\varphi(u) - 2\varphi^2(u) - \cdots - (p-1)\varphi^{p-1}(u).$$

Then it is easy to check that $\varphi(\beta) = \beta + 1$. Arguing as before, we see that $\beta, \beta+1, \ldots, \beta+(p-1)$ are conjugates of each other and so $E = F(\beta)$ is the splitting field of the minimal polynomial $f(x)$ of $\beta$. Apply separable descent to show that $\beta^p - \beta =: \alpha \in F$. Hence $\beta$ is a root of $x^p - x - \alpha$ and so for degree reasons $x^p - x - \alpha = f(x)$. □

Since in general a cyclic extension can be decomposed into a chain of cyclic extensions, we see that if $E/F$ is cyclic of degree $n$, $p \mid n$ and if $F$ contains all the $n$-th roots of unity, then we have a chain

$$F = F_0 \subset F_1 \subset \cdots \subset F_m \subset E$$

where each $F_i/F_{i-1}$ is an Artin-Schreier extension of degree $p$ while $E/F_m$ is a Kummer extension of degree coprime to $p$.

## 16.3 Radical extensions

The goal of this section is to prove the following theorem.

**Theorem 16.8.** *Let $F$ be a field of characteristic 0. A nonzero polynomial $f(x)$ is solvable by radicals if and only if its Galois group $\mathrm{Gal}(f)$ is solvable.*

Recall a group is solvable if and only if it has a chain of subgroups with successive quotients being cyclic groups. Last time we saw that cyclic extensions are exactly obtained by adjoining the roots of some elements. This gives a chain of radical extensions such that $f$ splits over the final extension. We just need to be a little careful since our classification of cyclic extensions require the base field to contain all the roots of unities.

Let $E/F$ be the splitting field of $f$ with solvable Galois group $G$ of size $n$. Let $L/E$ be the splitting field of $x^n - 1$. Then $L = E(\zeta_n)$. Let $K = F(\zeta_n)$. Then $K/F$ is radical and $f$ splits in $L$. So it suffices to show that $L/K$ is radical. Since $L/F$ is the splitting field of $f(x)(x^n - 1)$, we have that $L/F$ is Galois and so is $L/K$. Since $E/F$ is normal, any $K$-automorphism of $L$ restricts to an $F$-automorphism of $E$. Hence we have a group homomorphism $\Phi : \mathrm{Gal}(L/K) \to \mathrm{Gal}(E/F)$. Any element in the kernel of $\Phi$ is an automorphism of $L$ fixing $E$ and $\zeta_n$ and so must be the identity map. Hence $\Phi$ is an injection. Since $\mathrm{Gal}(E/F)$ is solvable, so is $\mathrm{Gal}(L/K)$. Let

$$\mathrm{Gal}(L/K) = H_0 \geq H_1 \geq H_2 \geq \cdots \geq H_m = \{1\}$$

be a sequence of subgroups with $H_{i+1}$ normal inside $H_i$ and $H_i/H_{i+1}$ cyclic of size $d_i$. For each $i$, we let $L_i = L^{H_i}$. We then get a chain of cyclic extensions

$$K = L_0 \subset L_1 \subset L_2 \subset \cdots \subset L_m = L.$$

Each $L_i/L_{i-1}$ is a Kummer extension since $K$ contains all the $n$-th roots of unity and so all the $d_i$-th roots of unities. Hence $L/K$ is radical.

Suppose conversely that $f(x)$ is solvable by radicals. Then there exists a field extension $E/F$ and chain of subfields

$$F = F_0 \subset F_1 \subset F_2 \subset \cdots \subset F_m = E$$

such that each $F_i = F_{i-1}(\sqrt[d_i]{\alpha_i})$ for some $\alpha_i \in F_{i-1}$. Let $N/F$ denote the normal closure of $E/F$. We claim that $N/F$ is also radical. Indeed, by the Primitive Element Theorem, there exists $\beta \in E$ such that $E = F(\beta)$. Let $p(x)$ be the minimal polynomial of $\beta$ over $F$ and let $\beta_1 = \beta, \beta_2, \ldots, \beta_n$ denote the roots of $p(x)$ in $N$. Then $N = E(\beta_2, \ldots, \beta_n)$. For each $j = 2, \ldots, n$, let $\sigma_j \in \mathrm{Aut}(N/F)$ be an $F$-automorphism of $N$ sending $\beta$ to $\beta_j$. If $K_1, K_2$ are two subfields of $N$, we write $K_1K_2$ for the smallest subfield of $N$ containing $K_1$ and $K_2$. Consider the following chain of subfields:

$$\begin{aligned}
F &= F_0 \subset F_1 \subset \cdots \subset F_m = E = F(\beta_1)\\
&= F(\beta_1)\sigma_2(F_0) \subset F(\beta_1)\sigma_2(F_1) \subset \cdots \subset F(\beta_1)(\sigma_2(F_m)) = F(\beta_1, \beta_2)\\
&= F(\beta_1, \beta_2)\sigma_3(F_0) \subset F(\beta_1, \beta_2)\sigma_3(F_1) \subset \cdots \subset F(\beta_1, \beta_2)(\sigma_3(F_m)) = F(\beta_1, \beta_2, \beta_3)\\
&\subset \cdots \subset F(\beta_1, \beta_2, \ldots, \beta_n)\\
&= N.
\end{aligned}$$

It is easy to see that each subextension is a Kummer extension. Hence we may assume without loss of generality that $E/F$ is Galois.

Let $n = [E : F]$ and let $L = E(\zeta_n)$ be the splitting field of $x^n - 1$ and let $K = F(\zeta_n)$. Then $K/F$ is Galois with $\mathrm{Gal}(K/F) \leq \mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q})$. Hence it is abelian and thereby solvable. Now $\mathrm{Gal}(E/F)$ is a quotient of $\mathrm{Gal}(L/F)$ while $\mathrm{Gal}(L/K)$ is a normal subgroup with quotient $\mathrm{Gal}(K/F)$. Hence it remains to show that $\mathrm{Gal}(L/K)$ is solvable.

For each $i = 0, \ldots, m$, we let $K_i = F_i(\zeta_n)$. Then we have a chain of cyclic extensions

$$K = K_0 \subset K_1 \subset \cdots \subset K_m = L$$

which corresponds to a decreasing chain

$$\mathrm{Gal}(L/K) \geq \mathrm{Gal}(L/K_1) \geq \cdots \geq \mathrm{Gal}(L/K_m) = \{1\}$$

with cyclic quotients. Therefore, $\mathrm{Gal}(L/K)$ is solvable. □

## 16.4 Insolvability of degree 5 or more polynomials via radicals

Degree 4 or less polynomials are solvable by radicals because their Galois groups are subgroups of $S_4$ and $S_4$ is solvable. In general, a degree $n$ polynomial will have Galois group $S_n$. For example as we have seen, an irreducible polynomial over $\mathbb{Q}$ of prime degree $p$ that has exactly two real roots has Galois group $S_p$.

**Theorem 16.9.** When $n \geq 5$, $S_n$ is not solvable. Therefore, in general a degree 5 or more polynomial is not solvable by radicals.

Since subgroups of solvable groups are solvable, it suffices to show that $A_n$ is not solvable. Recall that $A_n$ is generated by 3-cycles as

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

Hence it suffices show that any normal subgroup of $A_n$ contains a 3-cycle and therefore all 3-cycle which would imply that it is the entire $A_n$. You can look this up anywhere, for example here.

Let $N$ be nontrivial normal subgroup of $A_n$ with $n \geq 5$. Suppose first that $N$ contains an element that contains a cycle of length at least 4. Say $\alpha = (1234\cdots m)\cdots \in N$ with $m \geq 4$. Conjugating by $(123)$ gives $\beta = (2314\cdots m) \in N$. Now $\beta\alpha^{-1} = (241) \in N$.

Suppose now $N$ contains no such element. Suppose next that $N$ contains an element $\alpha$ that contains a cycle of length 3. Then $\alpha^2$ has cycle structure $[3,\ldots,3,1,\ldots,1]$. If $\alpha^2$ is a 3-cycle, then we are done. Otherwise, $\alpha^2$ has the form $(123)(456)\cdots$. Conjugating by $(124)$ gives $\beta = (243)(156)\cdots \in N$. Then $\beta(\alpha^2)^{-1} = (12534) \in N$. Contradiction.

Suppose now $N$ contains no elements that contain a cycle of length at least 3. Take any $\alpha \in N$. Then either $\alpha = (12)(34)$ or $\alpha = (12)(34)(56)\cdots$. Conjugating by $(13)(25)$ gives $\beta = (35)(14)$ or $(35)(14)(26)$ in $N$. Then $\beta\alpha = (12453)$ or $(163)(245)$. Contradiction. □

# 17 Ruler-Compass construction and Gauss's Theorem

Everybody has played around with rulers and compasses to see what shapes can be constructed. For example, it is possible to draw equilateral triangles, squares (this requires starting with a line $L_1$ and a point $P$ on $L_1$, draw a line $L_2$ perpendicular to $L_1$ and intersect $L_1$ at $P$. It is then a natural question to ask: which regular $n$-gons can be made? There were papers of enormous length proving that regular 17-gon and 257-gon can be made. We will learn a theorem of Gauss using Galois theory that completely answers this question.

## 17.1 Constructibility of points in $\mathbb{R}^2$

We start with the two points $(0,0)$ and $(1,0)$ on $\mathbb{R}^2$ and ask what other points can be constructed using rulers and compasses. We first make a general definition. Let $S$ be a set of points in $\mathbb{R}^2$. An $S$-line is a line $L$ that is spanned by 2 points on $S$, or equivalently if $\#(S \cap L) \geq 2$. An $S$-circle is a circle $C$ whose center is in $S$ and whose radius belongs to the set $\{|p-q| : p, q \in S\}$. Then $S$-lines and $S$-circles are exactly what can be constructed "directly" from $S$ using rulers and compasses. We denote by $S'$ the set of points that either lie on 2 $S$-lines, 1 $S$-line and 1 $S$-circle, or 2 $S$-circles. (Note an $S$-circle can have radius 0 in which case it is just a point. So $S'$ also contains $S$.)

As an example, it is easy to see that

$$\{(0,0),(1,0)\}' = \{(0,0),(1,0),(-1,0),(2,0),(\frac{1}{2},\frac{\sqrt{3}}{2}),(\frac{1}{2},-\frac{\sqrt{3}}{2})\}.$$

We say a point $P \in \mathbb{R}^2$ is **contructible** if there exists a sequence of points $p_1, \ldots, p_{m-1}, p_m = P$ such that $p_i \in \{(0,0),(1,0),p_1,\ldots,p_{i-1}\}'$. We say a real number $\alpha \in \mathbb{R}$ is contructible if $(\alpha, 0)$ is constructible.

It is easy to see from the above example that all integers are constructible. Then by using similar triangles, one can show that all rational numbers are constructible. We can generalize this method by saying that if $\alpha, \beta \in \mathbb{R}$ are constructible, then $\alpha/\beta$ is constructible. It is obvious that $\alpha \pm \beta$ are constructible. Hence we have shown that the set of real constructible numbers is a subfield of $\mathbb{R}$.

For any subset $S$ of $\mathbb{R}^2$, we write $\mathbb{Q}(S)$ for the subfield of $\mathbb{R}$ generated by the coordinates of points of $S$. Then $S$-lines have the form $ax+by=c$ while $S$-circles have the form $(x-d)^2+(y-e)^2 = (f_1-f_2)^2+(g_1-g_2)^2$ where $a, b, c, d, e, f_1, f_2, g_1, g_2$ are all in $\mathbb{Q}(S)$. The intersection of 2 $S$-line is a point both of whose coordinates lie in $\mathbb{Q}(S)$. The intersection of an $S$-line with an $S$-circle amounts to solving a quadratic formula and so we get 2 points defined over a quadratic extension of $\mathbb{Q}(S)$. The difference of 2 $S$-circle is a linear equation with coefficients in $\mathbb{Q}(S)$ and so the intersection of two $S$-circles is also defined over a quadratic extension of $\mathbb{Q}(S)$. We have now proved the forward direction of the following theorem.

**Theorem 17.1.** Let $P = (\alpha, \beta) \in \mathbb{R}^2$. Then $P$ is constructible if and only if there exists a chain of quadratic extensions $\mathbb{Q} = F_0 \subset F_1 \subset F_2 \subset \cdots \subset F_n \subset \mathbb{R}$ such that $\alpha, \beta \in F_n$.

For the backwards direction, it suffices to show that if every element of $F$ is constructible and if $E/F$ is a quadratic extension, then every element of $E$ is constructible. For this, we know that $E = F(\sqrt{\gamma})$ for some $\gamma \in F$. Since the set of constructible numbers forms a field, it suffices to show that $\sqrt{\gamma}$ is constructible. Draw a circle that passes through $(-1,0)$ and $(\gamma, 0)$ with center on the $x$-axis. Then the intersection of this circle with the $y$-axis gives $(0, \pm\sqrt{\gamma})$. □

**Corollary 17.2.** If $\alpha \in \mathbb{R}$ is constructible, then it is algebraic and the degree of its minimal polynomial is a power of 2.

By the above theorem, $\mathbb{Q}(\alpha)$ is a subfield of some field extension $F_n/\mathbb{Q}$ of degree a power of 2. □

As some quick example, we see that the cube cannot be duplicated: it is impossible to draw a cube with twice the volume of a given cube. This amounts to the constructibility of $\sqrt[3]{2}$ which has degree 3 over $\mathbb{Q}$ and so is not constructibility. Similarly, the circle cannot be squared: it is impossible to draw a square that has the same area as the unit circle. This amounts to the constructibility of $\sqrt{\pi}$ which is not even algebraic.

**Theorem 17.3.** Let $\alpha \in \mathbb{R}$ be algebraic with minimal polynomial $f$. Then $\alpha$ is constructible if and only if $\mathrm{Gal}(f)$ is a 2-group.

Suppose first that $\mathrm{Gal}(f)$ is a 2-group. Then by the Sylow theorem, there exists a chain $1 \leq G_1 \leq G_2 \leq \cdots \leq G_m = \mathrm{Gal}(f)$ where $[G_{i+1} : G_i] = 2$ for all $i$. The corresponding chain of subfields of the splitting field of $f$ is then a chain of quadratic extensions. Hence $\alpha$ is constructible.

Conversely, suppose $\alpha$ is constructible. Then let $\mathbb{Q} = F_0 \subset F_1 \subset F_2 \subset \cdots \subset F_n \subset \mathbb{R}$ be a chain of quadratic extensions such that $\alpha \in F_n$. The splitting field of $f$ is the normal closure of $\mathbb{Q}(\alpha)$ which is contained in the normal closure of $F_n$. It then suffices to show that the normal closure of $F_n$ has degree a power of 2 over $\mathbb{Q}$. By the primitive element theorem, there exists $\beta \in F_n$ such that $F_n = \mathbb{Q}(\beta)$. Let $\beta_1 = \beta, \beta_2, \ldots, \beta_m$ be the conjugates of $\beta$ in the normal closure $E$ of $F_n$. Then there exists $\mathbb{Q}$-automorphisms $\sigma_j$ of $E$ sending $\beta$ to $\beta_j$ for $j = 1, \ldots, m$. We now have the following chain of (at most) quadratic extensions;

$$\begin{aligned}
\mathbb{Q} &= F_0 \subset F_1 \subset \cdots \subset F_n = \mathbb{Q}(\beta_1) \\
&= \mathbb{Q}(\beta_1)\sigma_2(F_0) \subset \mathbb{Q}(\beta_1)\sigma_2(F_1) \subset \cdots \subset \mathbb{Q}(\beta_1)(\sigma_2(F_n)) = \mathbb{Q}(\beta_1, \beta_2) \\
&= \mathbb{Q}(\beta_1, \beta_2)\sigma_3(F_0) \subset \mathbb{Q}(\beta_1, \beta_2)\sigma_3(F_1) \subset \cdots \subset \mathbb{Q}(\beta_1, \beta_2)(\sigma_3(F_n)) = \mathbb{Q}(\beta_1, \beta_2, \beta_3) \\
&\subset \cdots \subset \mathbb{Q}(\beta_1, \beta_2, \ldots, \beta_m) \\
&= E.
\end{aligned}$$

This shows that $[E : \mathbb{Q}]$ is a power of 2. □

As an example, if one take a quartic polynomial that has Galois group $S_4$, then its roots will not be constructible even though their minimal polynomial has degree 4.

## 17.2 Constructibility of regular $n$-gon

For convenience, we say a complex number $\gamma = \alpha + i\beta$ is constructible if the point $(\alpha, \beta) \in \mathbb{R}^2$ is constructible. Then a regular $n$-gon is constructible if and only if $\zeta_n$ is constructible. The cyclotomic extension $\mathbb{Q}(\zeta_n)$ contains a real subextension of index 2, namely $\mathbb{Q}(\zeta_n + \zeta_n^{-1})$. To see $[\mathbb{Q}(\zeta_n) : \mathbb{Q}(\zeta_n + \zeta_n^{-1})] = 2$, note that $\zeta_n^{-1}$ is the complex conjugate of $\zeta_n$ and so $(\zeta_n + \zeta_n^{-1})/2$ is the real part of $\zeta_n$. To get the imaginary part of $\zeta_n$, we only need to take a square root since $|\zeta_n| = 1$. Suppose $\zeta_n$ satisfies the quadratic polynomial $ax^2 + bx + c = 0$ with $a, b, c \in \mathbb{Q}(\zeta_n + \zeta_n^{-1})$. Then we see that $\zeta_n$ is constructible if and only if $\zeta_n + \zeta_n^{-1})$ is constructible if and only if the Galois group of its minimal polynomial is a 2-group if and only if $[\mathbb{Q}(\zeta_n) : \mathbb{Q}]$ is a power of 2 if and only if $\phi(n)$ is a power of 2. Using the formula for $\phi(n)$, we now have the following theorem of Gauss

**Theorem 17.4.** A regular $n$-gon is constructible if and only if $n$ has the form $2^\alpha p_1 p_2 \cdots p_k$ where $p_i$ are distinct primes of the form $2^{b_i} + 1$. These are called Fermat primes.

Note in order for $2^{b_i} + 1$ to be a prime, $b_i$ cannot have any odd factions. So they are themselves of the form $2^{c_i}$. This is why they are called Fermat primes. There are only 5 known Fermat primes!

# 18 Introduction to Category Theory

## 18.1 Basic definitions

**Definition 18.1.** A **category** $\mathcal{C}$ consists of:

1. a set Obj whose elements are called the objects of $\mathcal{C}$,
2. for each $X, Y \in \mathrm{Obj}$, a set $\mathrm{Hom}(X, Y)$ whose elements are called the morphisms from $X$ to $Y$,
3. for any $X, Y, Z \in \mathrm{Obj}$, a map, called composition,

$$\mathrm{Hom}(X, Y) \times \mathrm{Hom}(Y, Z) \to \mathrm{Hom}(X, Z)$$

   and denoted $(f, g) \mapsto g \circ f$,

such that:

1. $\circ$ is associative,
2. for each $X \in \mathrm{Obj}$, there exists $\mathrm{id}_X \in \mathrm{Hom}(X, X)$ such that for any $f \in \mathrm{Hom}(X, Y)$, $f \circ \mathrm{id}_X = f = \mathrm{id}_Y \circ f$.

   Note that $\mathrm{id}_X$ is unique.

**Examples:**

1. A set $S$ can be viewed as a category where $\mathrm{Obj} = S$ and all the morphisms are identities. Such a category is called **discrete**.
2. The category **Set** of sets where the objects are sets and the morphisms are simply maps.
3. The category **Grp** of groups where the objects are groups and the morphisms are group homomorphisms. Similarly, the cateogry **Rng** of rings, **Fld** of fields. For any commutative ring $R$, we have the cateogry $\mathbf{Mod}_R$ of $R$-modules. Note when thinking categorically, even though the elements of Obj can themselves be sets, it doesn't make sense to talk about the elements of these sets.
4. The cateogory **Top** of topological spaces and continuous maps.
5. A directed graph can be viewed as a category where the objects are the vertices and morphisms are (roughly) the edges. Roughly in the sense that for each vertex $X$, we add $\mathrm{id}_X$ to $\mathrm{Hom}(X, X)$ and for each $X, Y$, we add 0 to $\mathrm{Hom}(X, Y)$ and we define the compositions to be 0 unless one of the factors is id.
6. Every ordered set is a category where the objects are the elements of the set and there is a unique morphism from $X$ to $Y$ if and only if $X \leq Y$.
7. Given any category $\mathcal{C}$, one can reverse all the arrows and obtain the opposite category $\mathcal{C}^{\mathrm{op}}$.

A morphism $f \in \mathrm{Hom}(X, Y)$ is an isomorphism if there exists an inverse $g \in \mathrm{Hom}(Y, X)$. A morphism $f \in \mathrm{Hom}(X, Y)$ is a monomorphism (think injective) if for any $g_1, g_2 \in \mathrm{Hom}(Z, X)$, $f \circ g_1 = f \circ g_2$ if and only if $g_1 = g_2$. It is an epimorphism (think surjective) if $g_1, g_2 \in \mathrm{Hom}(Y, Z)$, $g_1 \circ f = g_2 \circ f$ if and only if $g_1 = g_2$. A category is called a **groupoid** if all the morphisms are isomorphisms. This name is chosen so that one can view a group as a groupoid with one object!

An object $P$ of a category $\mathcal{C}$ is an **initial** object if for all objects $X$, there is a unique morphism from $P$ to $X$. It is a **terminal** object if for all objects $X$, there is a unique morphism from $X$ to $P$. It is a **zero** objects if it is both initial and terminal. In **Rng**, $\mathbb{Z}$ is an initial object and the zero ring is the terminal object. Initial/terminal objects are unique up to unique isomorphisms.

A category $\mathcal{C}'$ is a **subcategory** of $\mathcal{C}$ if the obvious holds: objects form a subset, morphisms form a subset. It is a full subcateogry if the morphisms include everything.

**Definition 18.2.** Let $\mathcal{C}$ and $\mathcal{C}'$ be two categories. A **functor** $F : \mathcal{C} \to \mathcal{C}'$ consists of a map between their sets of objects and map $F : \mathrm{Hom}(X, Y) \to \mathrm{Hom}(FX, FY)$ between their sets of morphisms such that

$$F(\mathrm{id}_X) = \mathrm{id}_{FX} \qquad F(f \circ g) = F(f) \circ F(g).$$

A functor is sometimes called a covariant functor. A contravariant functor is a functor from $\mathcal{C}^{\mathrm{op}}$ to $\mathcal{C}'$.

**Examples:**

1. The identity map is a functor from a category to itself.
2. The operation of the taking the group of units of a ring is a functor from **Rng** to **Grp**.
3. The operation of taking a ring $R$ and turning it into $R[x]$ is a functor from **Rng** to itself.
4. If one forgets about all the algebraic structures, one gets the forgetful functors from **Rng**, **Grp**, **Fld** to **Set**.
5. For any category $\mathcal{C}$ and an object $X$, the functor $\mathrm{Hom}(X, \cdot)$ is a functor from $\mathcal{C}$ to **Set**: it sends an object $Y$ to the set $\mathrm{Hom}(X, Y)$ of morphisms between $X$ and $Y$ and it sends a morphism via composition. Such a functor from $\mathcal{C}$ to **Set** is called **representable**. Yoneda's lemma says that if a functor is representable, then it is represented by a unique object up to a unique isomorphism.

A functor is full (resp. faithful, resp. fully faithful) if its map on morphisms is surjective (resp. injective, resp. bijective). A functor is essentially surjective if it is surjective on objects. A functor $F : \mathcal{C} \to \mathcal{C}'$ is an **equivalence of categories** if it is fully faithful and essentially surjective.

Now one may think that an equivalence of categories should mean that there exists an inverse $G$ such that $F \circ G$ and $G \circ F$ are the identities. However this is not the most flexible definition as we can add more "isomorphic objects" to a category and would still like to the say the two categories are equivalent - after all giving isomorphic groups different names do not change the mathematics of group theory.

**Definition 18.3.** A **natural transformation** $\alpha : F \to G$ between two functors $F, G : \mathcal{C} \to \mathcal{C}'$ is a map $\alpha_X : FX \to GX$ for each object $X$ of $\mathcal{C}$ such that for any $f : X \to Y$ in $\mathcal{C}$, the following diagram commute:

$$\begin{tikzcd}
FX \arrow[r, "\alpha_X"] \arrow[d, "F(f)"'] & GX \arrow[d, "G(f)"] \\
FY \arrow[r, "\alpha_Y"] & GY
\end{tikzcd}$$

If each $\alpha_X$ is an isomorphism in $\mathcal{C}'$, then we say $F$ and $G$ are naturally isomorphic. Or we say $FX \cong GX$ functorial in $X$.

**Theorem 18.4.** A functor $F : \mathcal{C} \to \mathcal{C}'$ is an equivalence of categories if and only if there exists a functor $G : \mathcal{C}' \to \mathcal{C}$ such that $F \circ G$ and $G \circ F$ are naturally isomorphic to the identities.

## 18.2 Yoneda's lemma and the adjoint functor

**Proposition 18.5.** (Yoneda's lemma) Let $\mathcal{C}$ be a category and let $X$ and $X'$ be two objects. Then to give a natural transformation from $\mathrm{Hom}(X, \cdot)$ to $\mathrm{Hom}(X', \cdot)$ is the same as giving a morphism $f : X' \to X$.

It is easy to see that given $f : X' \to X$, one can define a natural transformation from $\mathrm{Hom}(X, \cdot)$ to $\mathrm{Hom}(X', \cdot)$ by pre-composing with $f$. Conversely, let $\alpha$ be a natural transformation. Then $\alpha_X$ is a map $\alpha_X : \mathrm{Hom}(X, X) \to \mathrm{Hom}(X', X)$ and we set $f$ to be the image of $\mathrm{id}_X$. Let $Y$ be an arbitrary object of $\mathcal{C}$

and let $\beta : X \to Y$ be a morphism. We claim that $\alpha_Y(\beta) = \beta \circ f$. The morphism $\beta$ gives the following commutative diagram:

$$\begin{tikzcd}
\mathrm{Hom}(X, X) \arrow[r, "\alpha_X"] \arrow[d] & \mathrm{Hom}(X', X) \arrow[d] \\
\mathrm{Hom}(X, Y) \arrow[r, "\alpha_Y"] & \mathrm{Hom}(X', Y)
\end{tikzcd}$$

where the vertical maps are post-composition with $\beta$. Chasing where $\mathrm{id}_X$ goes in this diagram gives the desired equality. □

There is also the contravariant version of this where $\mathrm{Hom}(X, \cdot)$ is replaced by $\mathrm{Hom}(\cdot, X)$.

**Corollary 18.6.** Let $F : \mathcal{C} \to \mathbf{Set}$ be a functor such that there exists objects $X$ and $X'$ of $\mathcal{C}$ and natural isomorphisms $F \cong \mathrm{Hom}(X, \cdot)$ and $F \cong \mathrm{Hom}(X', \cdot)$. Then there is a unique isomorphism $\theta : X' \xrightarrow{\sim} X$ making the following diagram commute:

$$\begin{tikzcd}
& F \arrow[dl, "\sim"'] \arrow[dr, "\sim"] & \\
\mathrm{Hom}(X, \cdot) \arrow[rr, "\theta", "\sim"'] & & \mathrm{Hom}(X', \cdot)
\end{tikzcd}$$

**Definition 18.7.** Let $F : \mathcal{C} \to \mathcal{C}'$ and $G : \mathcal{C}' \to \mathcal{C}$ be two functors. Then we say $F$ is a left adjoint of $G$ and $G$ is a right adjoint of $F$ if

$$\mathrm{Hom}(FX, Y) \cong \mathrm{Hom}(X, GY)$$

functorial in both $X$ and $Y$.

For example, if $G$ is the forgetful functor from **Grp** to **Set**, then its left adjoint is the operation of turning a set into the free group generated by it. What if $G$ is the forgetful functor from **Rng** to **Set**?