Groups, Rings and Fields
MATH20310

Contents

Chapter 0

Notation: Permutations

This very short chapter is about a convenient notation that somehow managed to escape being introduced in first year.

Definition 0.1. *If X is a non-empty set, we denote by S_X the set of all bijections from X to X. A bijection from X to X is also called a permutation of X.*

We will be mostly interested in the set of all permutations of $\{1,\ldots,n\}$, which we denote by S_n for short (instead of $S_{\{1,\ldots,n\}}$). So

$$S_n = \{f : \{1,\ldots,n\} \to \{1,\ldots,n\} : f \textit{ bijective}\}.$$

Two observations:

- If $\alpha \in S_n$, it is given by its values on the elements $1,\ldots,n$, for instance:

  $$\alpha(1) = a_1,\ \alpha(2) = a_2,\ \ldots,\ \alpha(n) = a_n.$$

  Since α is surjective, $\{a_1,\ldots,a_n\} = \{1,\ldots,n\}$, and since α is injective, we have $a_i \neq a_j$ if $i \neq j$. So the sequence $a_1,\ldots a_n$ is just the sequence $1,\ldots,n$ written in a possibly different order.

- A convenient way to describe α is to write it in "array form", i.e.,

  $$\alpha = \begin{pmatrix} 1 & 2 & \cdots & n \\ a_1 & a_2 & \cdots & a_n \end{pmatrix},$$

  where we write the value $a_i = \alpha(i)$ below the element i.

Example 0.2. *Define a permutation $\alpha \in S_4$ by*

$$\alpha(1) = 2,\ \alpha(2) = 3,\ \alpha(3) = 1,\ \alpha(4) = 4.$$

In "array form", it is given by:

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix}.$$

If $\alpha, \beta \in S_n$, the composition $\alpha \circ \beta$ is also in S_n (it is also a bijection from $\{1, \ldots, n\}$ to $\{1, \ldots, n\}$). We will drop the composition sign and write it as a product, so $\alpha\beta$ for $\alpha \circ \beta$.

Question: How do we multiply (i.e., compose) permutations in S_n?

Answer: From right to left. Recall that the operation is the composition of functions, and σ and τ in S_n are functions. Therefore $\sigma\tau = \sigma \circ \tau$, and $(\sigma\tau)(x) = \sigma \circ \tau(x) = \sigma(\tau(x))$ by definition of composition. So we first compute $\tau(x)$, then feed the result into σ.

Example 0.3. *Let* $\beta = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}$. *We compute* $\alpha\beta$*:*

$\alpha\beta(1) = \alpha(\beta(1)) = \alpha(4) = 4$, $\alpha\beta(2) = \alpha(\beta(2)) = \alpha(3) = 1$, *and so on (finish it yourself). The result is:*

$$\alpha\beta = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix}.$$

Observe that

$$\beta\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix},$$

so that $\alpha\beta \neq \beta\alpha$.

Do a handul of examples of such computations, to make sure that you have no problem with this.

Finally, if $\alpha \in S_n$, then α is bijective, so α has an inverse α^{-1}, which is a bijection from $\{1, \ldots, n\}$ to $\{1, \ldots, n\}$, so $\alpha^{-1} \in S_n$.

How do we determine α^{-1}? Suppose for instance that α is given by

$$\alpha = \begin{pmatrix} 1 & 2 & \cdots & n \\ a_1 & a_2 & \cdots & a_n \end{pmatrix}.$$

Then $\alpha^{-1}(a_1) = 1, \ldots, \alpha^{-1}(a_n) = n$. In other words, we read the array from the bottom to the top.

Example 0.4. *Let*

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix}.$$

Then, reading the table from bottom to top:

$$\alpha^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}.$$

We finish with a remark about three properties of the product (=composition of maps) in S_n:

1. It is associative: $(\alpha\beta)\gamma = \alpha(\beta\gamma)$ for all $\alpha, \beta, \gamma \in S_n$.

   It is actually always true in general for composition of functions. How do you prove this? This equality claims that two functions are equal. So you can check it by applying both of them to an element x, then computing the results. You will get the same for both.

2. The identity map $\mathrm{id} : \{1, \ldots, n\} \to \{1, \ldots, n\}$ is an element of S_n and has the property that, for all $\alpha \in S_n$,

   $$\alpha\,\mathrm{id} = \alpha \text{ and } \mathrm{id}\,\alpha = \alpha.$$

3. Every element of S_n has an inverse:

   $$\forall \alpha \in S_n \quad \exists \alpha^{-1} \in S_n \text{ such that } \alpha\alpha^{-1} = \mathrm{id} \text{ and } \alpha^{-1}\alpha = \mathrm{id}.$$

We will come back soon to these three properties.

Chapter 1

Group Theory

Example 1.1. *Consider the set of remainders upon division by 4 under the (binary) operation "*$+ \bmod 4$*", i.e., the set* $\{0,1,2,3\}$ *and the operation*

$$\{0,1,2,3\} \times \{0,1,2,3\} \to \{0,1,2,3\}, \quad (a,b) \mapsto (a+b) \bmod 4.$$

For instance

$$\begin{aligned} (0,1) &\mapsto 1 \textit{ since } 0+1=1 \equiv 1 \pmod{4} \\ (2,2) &\mapsto 0 \textit{ since } 2+2=4 \equiv 0 \pmod{4} \\ (2,3) &\mapsto 1 \textit{ since } 2+3=5 \equiv 1 \pmod{4} \end{aligned}$$

(where $\equiv$ *means "is congruent to"). We can record the operation in a table, called the Cayley table of this operation.*

+	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

where "$+$*" means "*$+ \bmod 4$*". We call* 0 *the identity element (of this operation), because adding* 0 *to any other element does not do anything.*

This set with the operation $+$ *is denoted* $(\mathbb{Z}_4,+)$ *(we say "*$\mathbb{Z}$ *mod 4" for* $\mathbb{Z}_4$*).*

Example 1.2. *Consider the set of non-zero reminders upon division by 5, under the (binary) operation* $\cdot \bmod 5$*, i.e.,* $\{1,2,3,4\}$ *with operation*

$$\{1,2,3,4\} \times \{1,2,3,4\} \to \{1,2,3,4\}, \quad (a,b) \mapsto (a \cdot b) \bmod 5.$$

For instance

$$\begin{aligned}(1,1) &\mapsto 1 \textit{ since } 1\cdot 1 = 1 \equiv 1 \ (\mathrm{mod} 5)\\ (3,3) &\mapsto 4 \textit{ since } 3\cdot 3 = 9 \equiv 4 \ (\mathrm{mod} 5)\\ (2,4) &\mapsto 3 \textit{ since } 2\cdot 4 = 8 \equiv 3 \ (\mathrm{mod} 5)\end{aligned}$$

The Cayley table of this operation is

$\cdot$	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

The element 1 *is the identity element (of this operation), since multiplying* 1 *with any other element does not do anything.*

This set with the operation $\cdot$ *is denoted* $(\mathbb{Z}_5^\times, \cdot)$, *where* $\mathbb{Z}_5^\times := \mathbb{Z}_5 \setminus \{0\}$.

In general...

Definition 1.3. *A group is a non-empty set* G *equipped with a (binary) operation (that we will usually denote as a product)*

$$G \times G \to G, \ (x, y) \mapsto x \cdot y$$

such that

1. $x \cdot (y \cdot z) = (x \cdot y) \cdot z$ *for all* $x, y, z \in G$ *(we say that* $\cdot$ *is associative);*

2. *There exists and element* $e \in G$ *(the identity element) such that*

   $$x \cdot e = e \cdot x = x$$

   for all $x \in G$*;*

3. *For each* $x \in G$ *there exists* $y \in G$ *(called the inverse of* x*) such that* $x \cdot y = e$ *and* $y \cdot x = e$.

Remark 1.4. *1. We will use the notation* $(G, \cdot)$ *when we want to specify at once the set and the symbol that we use for the operation (it does not need to be* $\cdot$*, it can be anything).*

2. We write xy *for* $x \cdot y$*, and* x^{-1} *for the inverse of* x.

3. *If the operation is denoted by $\cdot$ (as in the definition), we often denote the identity by* 1 *instead of* e.

   If the operation is denoted by the symbol $+$, *we will often denote the identity element by* 0 *instead of* e; *in this case, we will also denote the inverse of* x *by* $-x$.

Example 1.5. 1. $(\mathbb{Z}_4, +)$ *is a group. The identity is the element* 0.

In general $(\mathbb{Z}_n, +)$ *is a group (where* $n \in \mathbb{N}$*), the group of integers modulo* n.

2. $(\mathbb{Z}_5^\times, \cdot)$ *is a group, its identity element is* 1. *Observe that* 5 *is prime.*

   In general $(\mathbb{Z}_p^\times, \cdot)$ *is a group (more later) when* p *is prime.*

3. *In view of the properties listed at the end of Chapter 0,* $(S_n, \cdot)$ *is a group, with identity element the identity map id, and with operation the composition of maps.*

4. $\mathbb{R}^\times := \mathbb{R} \setminus \{0\}$ *under* $\cdot$ *(the usual product) is a group.*

5. $\mathrm{GL}_n(\mathbb{R})$, *the set of* $n \times n$ *invertible matrices with real entries, with operation matrix multiplication, is a group (called the general linear group).*

6. $\mathrm{SL}_n(\mathbb{R})$, *the set of* $n \times n$ *matrices with real entries and with determinant 1, with operation matrix operation, is a group (called the special linear group).*

7. *In the two previous examples,* $\mathrm{GL}_n(\mathbb{R})$ *and* $\mathrm{SL}_n(\mathbb{R})$, *we can replace* $\mathbb{R}$ *by* $\mathbb{Q}$, $\mathbb{C}$, $\mathbb{Z}_p$ *(*p *prime). We will get groups in each case.*

A particular example of group can have a operation that is not denoted by $\cdot$ (we saw some examples of this). But in the general considerations about groups we will always denote the operation by $\cdot$.

Definition 1.6. *A group* G *with the property* $xy = yx$ *for all* $x, y \in G$ *is called Abelian.*

Example 1.7. $(\mathbb{Z}_n, +)$ *is Abelian, but* $GL_n(\mathbb{R})$ *is non-abelian.*

Definition 1.8. *The number of elements of a group is called its order. It is an element of* $\mathbb{N} \cup \{\infty\}$. *We write* $|G|$ *to denote the order of* G.

Proposition 1.9. *Let* G *be a group. Then*

1. *The identity element e is unique, i.e., if e' is another identity element, then $e = e'$.*

2. *If $x \in G$, then the inverse of x is unique, i.e., if y' is another inverse of x, then $y' = x^{-1}$.*

Proof. 1. By hypothesis we have $x \cdot e' = e' \cdot x = x$ for every $x \in G$. Taking $x = e$ we get $e \cdot e' = e$. But by the definition of e we also have $e \cdot e' = e'$. So $e = e'$.

2. If y' is an inverse of x, then $xy' = e$. But we also have $xx^{-1} = e$. So $xy' = xx^{-1}$. Multiplying both sides on the left by x^{-1}, we get $x^{-1}xy' = x^{-1}xx^{-1}$, so $ey' = ex^{-1}$ and $y' = x^{-1}$. □

Definition 1.10. *Let G be a group. The order of an element $g \in G$ is the smallest positive integer n such that $g^n = e$. If no such n exists, we say that g has infinite order. The order of g is denoted by $|g|$.*

Note: If the operation on G is denoted by "+", we will write $ng = 0$ instead of $g^n = e$ (because the identity is denoted by 0 in this case, and $g^n = g \cdot g \cdots g$ is written $g + g + \cdots + g$, which is more naturally denoted by ng).

1.1 Disgression: Permutations

Proposition 1.11. *The order of S_n is $n!$ (see exercise sheet 1).*

Definition 1.12. *A permutation $\alpha \in S_n$ is called a* **cycle of length** *k if there are elements $a_1, \dots, a_k \in \{1, \dots, n\}$, all different, such that*

$$\alpha(a_1) = a_2,\ \alpha(a_2) = a_3, \dots,\ \alpha(a_{k-1}) = a_k,\ \alpha(a_k) = a_1,$$

and $\alpha(x) = x$ for all the other elements of $\{1, \dots, n\}$ (α does not move the other elements).

We write $(a_1\ a_2\ \cdots\ a_k)$ to denote the cycle α.

Remark 1.13. 1. *Observe that $(a_1\, a_2 \cdots a_k) = (a_i\, a_{i+1} \cdots a_k\, a_1\, a_2 \cdots a_{i-1})$ for every i in $\{1, \dots, k\}$ (so we can start writing the cycle where we want, as long as we "cycle").*

2. *A cycle of length one: (a), is the identity map, because it sends a to a, and it does not move the other elements.*

3. *Cycles are permutations, so they can be multiplied together or with other permutations (recall that the multiplication here is the composition of functions).*

Example 1.14. *Suppose we have* $\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 4 & 6 & 5 & 3 \end{pmatrix}$. *We look at what successive applications of* α *do:*

$$\begin{tikzcd} 1 \arrow[r, bend left, "\alpha"] & 2 \arrow[l, bend left, "\alpha"] & 3 \arrow[r, bend left, "\alpha"] & 4 \arrow[r, bend left, "\alpha"] & 6 \arrow[ll, bend left, "\alpha"] & 5 \arrow[loop right, "\alpha"] \end{tikzcd}$$

Using this, we can diretly check that

$$\alpha = (1\ 2)(3\ 4\ 6)(5),$$

where the right hand side denotes the composition of the three cycles (just compute it and write it in array form, and you will see that the result is α*).*

One final remark: As observed above, (5) *is the identity map, so actually*

$$\alpha = (1\ 2)(3\ 4\ 6),$$

i.e., there is no need to keep the cycles of length 1 in the expression.

Example 1.15. *Let*

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix}, \ \beta = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}.$$

Then, following the example above, we have

$$\begin{tikzcd} 1 \arrow[r, bend left, "\alpha"] & 2 \arrow[r, bend left, "\alpha"] & 3 \arrow[ll, bend left, "\alpha"] & 4 \arrow[loop right, "\alpha"] & \text{and } 1 \arrow[r, bend left, "\beta"] & 4 \arrow[l, bend left, "\beta"] & 2 \arrow[r, bend left, "\beta"] & 3 \arrow[l, bend left, "\beta"] \end{tikzcd}$$

so that

$$\alpha = (1\ 2\ 3)(4) = (1\ 2\ 3) \text{ and } \beta = (1\ 4)(2\ 3).$$

We compute $\alpha\beta$ *as usual for the composition of maps:*

$$\alpha\beta(1) = \alpha(\beta(1)) = \alpha(4) = 4,$$

$$\alpha\beta(2) = \alpha(\beta(2)) = \alpha(3) = 1,$$

and so on, and we obtain, in array form

$$\alpha\beta = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix} = (1\ 4\ 2)(3) = (1\ 4\ 2).$$

Definition 1.16. *Two cycles in S_n, $(a_1\ a_2\ \cdots\ a_k)$ and $(b_1\ b_2\ \cdots\ b_\ell)$, are called disjoint if $a_i \neq b_j$ for all i and j.*

Example 1.17. *The cycles* $(1\ 2\ 6)$ *and* $(3\ 5)$ *are disjoint, the cycles* $(1\ 3\ 5)$ *and* $(3\ 6)$ *are not. Computing their products, we observe:*

$$(1\ 2\ 6)(3\ 5) = (3\ 5)(1\ 2\ 6),$$

but

$$(1\ 3\ 5)(3\ 6) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 2 & 6 & 4 & 1 & 5 \end{pmatrix} = (1\ 3\ 6\ 5) \neq$$

$$(3\ 6)(1\ 3\ 5) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 2 & 5 & 4 & 1 & 3 \end{pmatrix} = (1\ 6\ 3\ 5).$$

Proposition 1.18. *Let σ and γ be two disjoint cycles in S_n. Then $\sigma\gamma = \gamma\sigma$.*

Proof. Let $\sigma = (a_1\ a_2\ \cdots a_k)$ and $\gamma = (b_1\ b_2\ \cdots b_\ell)$. We check that $\sigma\gamma = \gamma\sigma$ by checking that $\sigma\gamma(x) = \gamma\sigma(x)$, i.e., we check that they always return the same values, which mean that they are equal (they are functions). We distinguish 3 cases:

(1) $x = a_i$. Then $\sigma(x) = a_{i+1}$ (with the convention that $a_{i+1} = a_1$ is $i = k$) and $\gamma(x) = x$ (since x is not one of $b_1, \ldots, b_\ell$). Then $\sigma\gamma(x) = a_{i+1} = \gamma\sigma(x)$.

(2) $x = b_j$. Similar.

(3) $x \notin \{a_1, \ldots, a_k, b_1, \ldots, b_\ell\}$. Also similar. □

Theorem 1.19. *Every permutation in S_n can be written as a cycle or as a product of disjoint cycles.*

Proof. Let $\alpha \in S_n$. We follow the idea of Example 1.14. Pick $a_1 \in \{1, \ldots, n\}$ and consider the sequence $\alpha(a_1), \alpha^2(a_1), \alpha^3(a_1), \ldots$.

We claim that we must have $\alpha^k(a_1) = a_1$ for some k: Since the values of the $\alpha^i(a_1)$ are in $\{1, \ldots, n\}$ (which is finite), there must be repetitions: there are $i, k \in \mathbb{N}$ such that $\alpha^i(a_1) = \alpha^{i+k}(a_1)$. Applying the inverse of α^i to both sides, we get $a_1 = \alpha^k(a_1)$. We take for k the smallest integer for which $a_1 = \alpha^k(a_1)$.

Therefore the first cycle that we obtain is $(a_1\ a_2\ \cdots a_{k-1})$, where $a_i = \alpha^{i-1}(a_1)$.

Pick now $b \in \{1, \ldots, n\} \setminus \{a_1, \ldots, a_{k-1}\}$. Similarly, we have $\alpha^\ell(b) = b$ for some ℓ (we take the smallest possible) and the second cycle that we obtain is $(b_1\ b_2\ \cdots b_{\ell-1})$ where $b_i = \alpha^{i-1}(b_1)$.

We claim that the cycles $(a_1\ a_2\ \cdots a_{k-1})$ and $(b_1\ b_2\ \cdots b_{\ell-1})$ are disjoint: If not we have $\alpha^i(a_1) = \alpha^j(b_1)$ for some i, j. For instance $j < i$ (the other case is similar). Applying α^{-j} to both sides, we get $\alpha^{i-j}(a_1) = b_1$, which is impossible since $b_1 \notin \{a_1, \dots, a_{k-1}\}$.

We then continue this process until all elements of $\{1, \dots, n\}$ are used, and we get

$$\alpha = (a_1\ a_2\ \cdots a_{k-1})(b_1\ b_2\ \cdots b_{\ell-1}) \cdots (c_1\ c_2\ \cdots c_{t-1}). \qquad \square$$

In the special case of permutations in the group S_n, how do we compute the order of α, i.e., the smallest positive integer n such that $\alpha^n = \text{id}$? We only state (and use):

Theorem 1.20. *Let $\alpha \in S_n$ be written as a product of disjoint cycles. Then $|\alpha|$ is the least common multiple of the lengths of the cycles.*

(Recall that the least common multiple of some integers $k_1, \dots k_t$ is the smallest positive integer d such that k_i divides d, for $i = 1, \dots, t$.)

Example 1.21. *Given $\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 5 & 4 & 6 & 3 \end{pmatrix}$, find $|\alpha|$.*

We have $\alpha = (1\ 2)(3\ 5\ 6)$, where $(1\ 2)$ has length 2 and $(3\ 5\ 6)$ has length 3. Therefore $|\alpha| = lcm(2, 3) = 6$

Definition 1.22. *A cycle of length 2 is called a transposition.*

Note that a cycle of length 2 has order 2, so is its own invers: $(a\ b)(a\ b) = \text{id}$.

Theorem 1.23. *Every permutation in S_n ($n > 1$) can be written as a product of transpositions.*

Proof. We know that every permutation can be written as a product of (disjoint) cycles. Thus, it suffices to show that every cycle can be written as a product of transpositions. Consider $(a_1\ a_2\ \cdots a_k)$.

If $k = 1$, (a_1) is the identity, and if we pick any $x \in \{1, \dots, n\} \setminus \{a_1\}$ we have $(a_1) = (a_1\ x)(a_1\ x)$.

If $k > 1$, we have

$$(a_1\ a_2\ \cdots a_k) = (a_1\ a_k)(a_1\ a_{k-1}) \cdots (a_1\ a_2)$$

(to check it you can simply check that these two functions are equal: Apply each of them to x and check that you get the same result on both sides, considering the case $x \in \{a_1, \dots, a_k\}$ and the case $x \notin \{a_1, a_2, \dots, a_k\}$). $\square$

Example 1.24. *Decompose $(1\ 6\ 3\ 2)(4\ 5\ 7)$ as a product of transpositions:*

$$(1\ 6\ 3\ 2)(4\ 5\ 7) = (1\ 2)(1\ 3)(1\ 6)(4\ 7)(4\ 5).$$

Observe that the deomposition as product of transpositions is not unique. For instance:

$$\begin{aligned}(1\ 2\ 3\ 4\ 5) &= (1\ 5)(1\ 4)(1\ 3)(1\ 2) \text{ (4 transpositions)}\\ &= (5\ 4)(5\ 3)(5\ 2)(5\ 1) \text{ (4 transpositions)}\\ &= (5\ 4)(5\ 2)(2\ 1)(2\ 5)(2\ 3)(1\ 3) \text{ (6 transpositions).}\end{aligned}$$

In general, we state (for now, the proof will be given later).

Theorem 1.25. *If $\alpha = \beta_1 \cdots \beta_r = \gamma_1 \cdots \gamma_s$ where the β_i's and the γ_i's are transpositions, then either r and s are both even or r and s are both odd.*

Thus, we can make the following definition.

Definition 1.26. *A permutation is called even if it can be written as a product of an even number of transpositions. Otherwise it is called odd.*

We denote by A_n the set of all even permutations of S_n.

Exercise 1.27. *Prove that A_n is a group. It is called the alternating group of degree n.*

Theorem 1.28. *For $n > 1$, $|A_n| = \dfrac{n!}{2}$.*

Proof. The map $\alpha \mapsto (1\ 2)\alpha$ is a bijection from the set of odd permutations to the set of even permutations (it is injective since $(1\ 2)\alpha = (1\ 2)\beta$ implies $\alpha = \beta$ (multiply both sides on the left by $(1\ 2)$ and use that $(1\ 2)(1\ 2) = \mathrm{id}$); it is surjective since $\beta = (1\ 2)(1\ 2)\beta$, using again that $(1\ 2)(1\ 2) = \mathrm{id}$).

Therefore there are an equal number of even and odd permutations, and since every permutation is either odd or ever, we must have $|A_n| = |S_n|/2$. □

1.2 Cyclic groups

Definition 1.29. *A group G is called cyclic is there is an element $a \in G$ such that*

$$G = \langle a \rangle \overset{def}{=} \{a^n : n \in \mathbb{Z}\} = \{\ldots, a^{-2}, a^{-1}, 1, a, a^2, \ldots\}.$$

We say that a is a generator of G.

(We are assuming that the operation on G is $\cdot$; see next Remark.)

Remark 1.30. *1. If $n < 0$ we define*

$$a^n = \underbrace{a^{-1} \cdots a^{-1}}_{|n|\ times} = (a^{-1})^{|n|}.$$

2. If the operation is denoted by + we would use the notation:

$$G = \langle a \rangle \overset{def}{=} \{na : n \in \mathbb{Z}\} = \{\ldots, -2a, -1a, 0, a, 2a, \ldots\}.$$

Here, if $n < 0$, then $na = \underbrace{(-a) + \cdots + (-a)}_{|n| \text{ times}}$.

Example 1.31. *1. We consider the group $(\mathbb{Z}, +)$.*

$$\mathbb{Z} = \langle 1 \rangle = \langle -1 \rangle.$$

2. We consider the group $(\mathbb{Z}_n, +)$.

$$\mathbb{Z}_n = \{0, 1, 2, \ldots, n-1\} = \langle 1 \rangle.$$

Observe that $\mathbb{Z}_n$ may have many generators. For instance (exercise, check it):

$$\mathbb{Z}_8 = \langle 1 \rangle = \langle 3 \rangle = \langle 5 \rangle = \langle 7 \rangle.$$

Remark 1.32. *Consider an element $a \in G$ of finite order, say $|a| = n$. Then $a^n = e$, and therefore (multiplying by a):*

$$a^{n+1} = a, \ a^{n+2} = a^2, \ldots$$

$$a^{-1} = a^{n-1}, \ a^{-2} = a^{n-2}, \ldots$$

so that

$$\langle a \rangle = \{e, a, a^2, \ldots, a^{n-1}\}.$$

The last line of this remark gives us the following result:

Proposition 1.33. *Let $a \in G$ be of finite order. Then $|a| = |\langle a \rangle|$.*

In particular, if G is cyclic and generated by a, then $|G|$ is equal to the order of a.

Theorem 1.34. *Let $G = \langle a \rangle$ be a cyclic group of order n. Then*

$$G = \langle a^k \rangle \Leftrightarrow \gcd(k, n) = 1.$$

Proof. ($\Leftarrow$) Since $\gcd(k, n) = 1$ there are $u, v \in \mathbb{Z}$ such that $ku + nv = 1$. Then

$$a^1 = a^{ku+nv} = a^{ku}a^{nv} = a^{ku}(\underbrace{a^n}_{=e})^v = a^{ku} \in \langle a^k \rangle.$$

Thus $a \in \langle a^k \rangle$, and so all powers of a are in $\langle a^k \rangle$. This implies $\langle a \rangle \subseteq \langle a^k \rangle$. The other inclusion is obvious, so $G = \langle a^k \rangle$.

($\Rightarrow$) Assume that the conclusion does not hold, i.e., $\gcd(k,n) = d > 1$. Write $k = rd$ and $n = sd$. Then

$$(a^k)^s = (a^{rd})^s = (a^{rd})^r = (\underbrace{a^n}_{=e})^r = e^r = e.$$

So $|a^k| \leq s < n$. Thus, a^k is not a generator of G (cf. Proposition 1.33), a contradiction. □

Corollary 1.35. *An integer k is a generator of $\mathbb{Z}_n$ if and only if* $\gcd(k,n) = 1$.

We actually saw a special case of this in Example 1.31.

1.3 Subgroups

Definition 1.36. *Let G be a group. A subset $H \subseteq G$ is a subgroup of G if it is a group under the operation of G.*

We write $H \leq G$ to indicate that H is a subgroup of G.

Lemma 1.37. *If G is a group and H is a subgroup of G, then H has an identity element e_H (because H is a group). We have $e_H = e$.*

Proof. Since e_H is the identity of H, we have $e_H e_H = e_H$. Multiplying both sides by the inverse of the element e_H of G, we get $e_H^{-1} e_H e_H = e_H^{-1} e_H$, so $e_H = e$. □

Theorem 1.38 (Subgroup Test). *Let G be a group and let $H \subseteq G$ be a non-empty subset. Then H is a subgroup of G if and only if the following two conditions hold*

1. *If $x, y \in H$, then $xy \in H$ (we say that H is closed under the operation).*
2. *If $x \in H$, then $x^{-1} \in H$ (inverses exist in H).*

Proof. ($\Rightarrow$) If H is a subgroup, then H is closed under the operation, because in a group, the result of a product of two elements is again in the group. Also, every element $x \in H$ has an inverse y in H. We just need to check that this element y is equal to the inverse x^{-1} of x in G:

We have $xy = e_H = e$ (using Lemma 1.37). Multiplying both sides on the left by x^{-1} we get $x^{-1}xy = x^{-1}e$, so $y = x^{-1}$.

($\Leftarrow$) Suppose $H \subseteq G$ is a non-empty subset satisfying conditions 1 and 2. Since 1 is true, $\cdot$ is a binary operation on H. Also, the associtivity follows from that of G. H has inverses by 2. Finally, we check that e belongs to H: Let $x \in H$. By 2, we have $x^{-1} \in H$, and by 1: $x^{-1}x = e \in H$. □

Example 1.39. *Let G be a group and let $a \in G$. Then $\langle a \rangle = \{a^n : n \in \mathbb{Z}\}$ is a subgroup of G.*

Check: It is of course non-empty. Let $a^i, a^j \in \langle a \rangle$. Then $a^i a^j = a^{i+j} \in \langle a \rangle$. Also, $(a^i)^{-1} = a^{-i} \in \langle a \rangle$.

Example 1.40. *Recall that a^n is written na when G is additive (i.e., when we use the symbol $+$ for the operation on G).*

$\langle 2 \rangle = \{0, 2, 4, 6, 8\}$ is a cyclic subgroup of $(\mathbb{Z}_{10}, +)$.

Exercise 1.41. *Let $\sigma = (1\ 2)$ and $\rho = (1\ 2\ 3)$ in S_3. It is easy to check that $S_3 = \{id, \rho, \rho^2, \sigma, \sigma\rho, \sigma\rho^2\}$ (for instance: check that they are all different, and we know that S_3 has 6 elements). The following are subgroups of S_3:*

$$\{id, \rho, \rho^2\}, \quad \{id, \sigma\}, \quad \{id, \sigma\rho\}, \quad \{id, \sigma\rho^2\}.$$

In fact these are all the proper subgroups of S_3 (proper means different from the whole group).

Example 1.42. *Let $G = \mathbb{R}^\times$ under $\cdot$. Consider the set*

$$K = \{x \in G : x \geq 1\}.$$

If K a subgroup of G?

No, condition 2 does not hold: $2 \in K$ but $2^{-1} = \dfrac{1}{2} \notin K$.

Suppose we have

$$H = \{x \in G : x = 1 \text{ or } x \text{ is irrational}\}.$$

Is H a subgroup of G?

Again no: Condition 1 does not hold. Consider $\sqrt{2} \in H$. Then $\sqrt{2}\sqrt{2} = 2 \notin H$.

Exercise 1.43 (Finite subgroup test)**.** *Let H be a finite non-empty subset of a group G. Then $H \leq G$ if and only if $xy \in H$ for every $x, y \in H$.*

In other words, the condition that H is closed under the operation implies that if $x \in H$, then $x^{-1} \in H$. To prove this, consider the set of all powers of x, and use that since H is finite there must be repetitions.

Another important example:

Definition 1.44. *The center $Z(G)$ of a group G is the set of elements in G that commute with all the elements of G, i.e.,*

$$Z(G) := \{a \in G : ax = xa\ \forall x \in G\}.$$

Proposition 1.45. *Let G be a group. Then $Z(G) \leq G$.*

Proof. Exercise. □

1.4 Cosets

Definition 1.46. *Let G be a group and let $H \leq G$. For any $a \in G$, the set*

$$aH := \{ah : h \in H\}$$

is called the left coset of H in G containing a. Also

$$Ha := \{ha : h \in H\}$$

is called the right coset of H in G containing a. The element a is called a representative of the coset aH (or Ha).

We use the notation $|aH|$ to denote the number of elements of the set aH.

Example 1.47. *Let $G = S_3$ and consider the subgroup $H = \{id, \rho\sigma\}$ of G, where $\rho = (1\ 2\ 3)$ and $\sigma = (1\ 2)$. Then*

- *$idH = H$;*
- *$\sigma H = \{\sigma id, \sigma\rho\sigma\} = \{\sigma, \rho^2\}$ (direct computation) $= \rho^2 H$;*
- *$\rho\sigma H = \{\rho\sigma, id\} = H$ (check it);*
- *$\sigma\rho H = \{\sigma\rho, \rho\} = \rho H$ (check it).*

Observe that there are only 3 distinct cosets of H, namely H, σH, ρH, and $\dfrac{|G|}{|H|} = \dfrac{6}{2} = 3$.

Example 1.48. *Let $G = \mathbb{Z}_9$ (with operation $+$) and consider the subgroup $H = \langle 3 \rangle = \{0, 3, 6\}$ of G.*

As the operation if $+$, we write $a + H$ instead of aH. Then

- *$0 + H = \{0, 3, 6\} = 3 + H = 6 + H = H$;*
- *$1 + H = \{1, 4, 7\} = 4 + H = 7 + H$;*
- *$2 + H = \{2, 5, 8\} = 5 + H = 8 + H$.*

Observe that there are only 3 distinct cosets, and $\dfrac{|G|}{|H|} = \dfrac{9}{3} = 3$.

We now look at some properties of cosets.

Lemma 1.49. *Let $H \leq G$ and $a, b \in G$. Then*

1. *$a \in aH$;*

2. *$aH = H \Leftrightarrow a \in H$;*

3. *$aH = bH$ or $aH \cap bH = \emptyset$ (identical or disjoint);*

4. *$aH = bH \Leftrightarrow a^{-1}b \in H$;*

5. *$|aH| = |H|$;*

6. *$aH = Ha \Leftrightarrow H = aHa^{-1}$;*

7. *aH is a subgroup of G if and only if $a \in H$ (in which case $aH = H$).*

Proof. 1. $a = a \cdot e \in aH$.

2. ($\Rightarrow$) If $aH = H$, then $a = a \cdot e \in aH = H$.

   ($\Leftarrow$) If $a \in H$, then $aH \subseteq H$ as H a subgroup. Also, $h = a\underbrace{(a^{-1}h)}_{\in H} \in aH$. So $H \subseteq aH$. Thus $aH = H$.

3. Suppose $aH \cap bH \neq \emptyset$. Let $x \in aH \cap bH$. Then there exist $h_1, h_2 \in H$ such that $x = ah_1 = bh_2$. Thus $a = xh_1^{-1} = bh_2h_1^{-1}$ and

$$aH = b(\underbrace{h_2h_1^{-1}H}_{=H \text{ by } 2}) = bH.$$

4. Note that $aH = bH \Leftrightarrow H = a^{-1}bH$ (check it!) $\Leftrightarrow a^{-1}b \in H$ by 2.

5. Consider the function $f : H \to aH$, $f(h) = ah$. We claim that f is bijective (and the result will follow).

   Injective: Suppose $f(h_1) = f(h_2)$, i.e., $ah_1 = ah_2$. multiplying both sides on the left by a^{-1} gives $h_1 = h_2$.

   Surjective: By definition of aH.

6. Note that $aH = Ha \Leftrightarrow (aH)a^{-1} = (Ha)a^{-1} \Leftrightarrow aHa^{-1} = H$.

7. ($\Rightarrow$) If $aH \leq G$, then $e \in aH$. Thus $aH \cap eH \neq \emptyset$ and by 3 we get $aH = eH = H$, thus by 2, $a \in H$.

   ($\Rightarrow$) If $a \in H$, then by 2, $aH = H$ is a subgroup of G. □

We can now prove

Theorem 1.50 (Lagrange's Theorem)**.** *If G is a finite group and H is a subgroup of G, then $|H|$ divides $|G|$. Moreover*

$$\frac{|G|}{|H|} = \text{ the number of distinct (left) cosets of } H \text{ in } G.$$

Proof. Let $a_1H, a_2H, \ldots, a_rH$ denote the distinct left cosets of H in G (there are finitely many since G is finite). Then for each $a \in G$, we have $aH = a_iH$ for some i. Therefore $a \in aH = a_iH$. Thus every element of G belongs to one of the a_iH:

$$G = a_1H \cup a_2H \cup \cdots \cup a_rH.$$

By 3 of Lemma 1.49, this union is disjoint, thus

$$|G| = |a_1H| + |a_2H| + \cdots + |a_rH|.$$

Since $|a_iH| = |H|$ by Lemma 1.49 (statement 5), we have $|G| = r|H|$ where r is the number of distinct left cosets of H in G. □

Observe that the same result can obtained (in the same way) for right cosets. It only requires checking the statement corresponding to Lemma 1.49 for right cosets of H.

Definition 1.51. *We define $[G : H] := \dfrac{|G|}{|H|}$ to be the index of H in G. It is equal to the number of distinct left cosets of H in G (or the number of distinct right cosets of H in G).*

The converse of Lagrange's Theorem is false: If G is a group of order n and k divides n, then G may not have a subgroup of order k.

Corollary 1.52. *Let G be a finite group and let $a \in G$. Then $|a|$ divides $|G|$. In particular $a^{|G|} = e$.*

Proof. By Proposition 1.33, the order of a is equal to the order of $\langle a \rangle$, which is a subgroup of G (Example 1.39). The result follows by Lagrange's theorem. □

We record some applications of Lagrange's Theorem to number theory.

Lemma 1.53. *Let p be a prime number. Then $(\mathbb{Z}_p^\times, \cdot)$ is a group.*

Proof. We need to check that the product of 2 elements of $\mathbb{Z}_p^\times$ is still in $\mathbb{Z}_p^\times$, i.e., if $x, y \in \{1, \ldots, p-1\}$, then $xy \neq 0$ in $\mathbb{Z}_p$, i.e., xy is not divisible by p. This is because p is prime.

The product modulo p is associative (it comes from the product in $\mathbb{Z}$).

The element 1 is clearly an identity for the operation $\cdot$. We now show that for all $x \in \mathbb{Z}_p^\times$ there exists $y \in \mathbb{Z}_p^\times$ such that

$$xy \equiv 1 \pmod{p}. \tag{1.1}$$

This means showing that there is a solution y mod p to equation (1.1). Recall (from MATH10040): In general $ay \equiv b \pmod{n}$ has a solution if and only if $d = \gcd(a, n)$ divides b and there are d distinct solutions.

Here: $a = x$, $b = 1$, $n = p$ (note that $x \not\equiv 0 \pmod{p}$), so $d = \gcd(x, p) = 1$. Thus (1.1) has exactly one solution y modulo p. In fact $y \not\equiv 0 \pmod{p}$ since if $y \equiv 0 \pmod{p}$ then $xy \equiv 0 \pmod{p}$, a contradiction (since $xy \equiv 1 \pmod{p}$). □

Exercise 1.54 (Fermat's Little Theorem)**.** *Let p be a prime abd let $a \in \mathbb{Z}$ with $\gcd(a, p) = 1$. Then $a^{p-1} \equiv 1 (\mathrm{mod} p)$.*

Hint: Use Lemma 1.53 and Corollary 1.52.

We mentioned earlier that the converse of Lagrange's Theorem is false, i.e. that if k divides $|G|$, then G may not have a subgroup of order n. Here is an example.

Proposition 1.55. *The group A_4 has order 12 but has no subgroup of order 6.*

Proof. We assume that it does have a subgroup of order 6 and reach a contradiction.

In problem 6, Exercise sheet 2, you listed all 12 elements of A_4. Check that A_4 has exactly 8 elements of order 3 (i.e. such that $\alpha \neq \mathrm{id}$, $\alpha^2 \neq \mathrm{id}$ but $\alpha^3 = \mathrm{id}$). Suppose that H is a subgroup of A_4 of order 6. Let α be any of the 8 elements of order 3 of A_4. As

$$[A_4 : H] = \frac{12}{6} = 2, \tag{1.2}$$

we have two distinct left cosets of H in A_4.

We consider the three cosets $H, \alpha H, \alpha^2 H$ (note: since $\alpha^3 = 1$, $\alpha^3 H = H$). By (1.2), at most two of these three cosets are distinct. Therefore, two of the cosets are equal, which leads in all cases to $\alpha H = H$ (check it), so that $\alpha \in H$. Thus H contains all eight elements of order 3, a contradiction. □

1.5 Normal subgroups and Quotient groups

If G is a group and $H \leq G$, we introduced, for any $a \in G$, the left coset and right coset:

$$aH := \{ah : h \in H\}, \text{ and } Ha := \{ha : h \in H\}.$$

Can these two sets be equal?

Example 1.56. *1. Let $G = S_3$ and $H = \{id, (1\ 3)\}$. Let $a = (1\ 2) \in S_3$. Then (do the computations):*

$$(1\ 2)H = \{(1\ 2), (1\ 3\ 2)\}, \quad H(1\ 2) = \{(1\ 2), (1\ 2\ 3)\}.$$

Thus $(1\ 2)H \neq H(1\ 2)$.

2. Let $G = S_3$ and $H = \{id, (1\ 2\ 3), (1\ 3\ 2)\}$. Let $a = (1\ 2) \in S_3$. Then

$$(1\ 2)H = \{(1\ 2), (2\ 3), (1\ 3)\} = H(1\ 2).$$

Thus $(1\ 2)H = H(1\ 2)$. Acutally, for every $a \in S_3$, $aH = Ha$ (check this if you find the time).

We want to study when $aH = Ha$ for every $a \in G$

Definition 1.57. *A subgroup H of a group G is called normal if $aH = Ha$ for every $a \in G$. We denote this by $H \trianglelefteq G$.*

Example 1.58. $H = \{id, (1\ 2\ 3), (1\ 3\ 2)\} \trianglelefteq S_3$.

Theorem 1.59 (Normality test)**.** *A subgroup H of G is normal if and only if $xHx^{-1} \subseteq H$ for all $x \in G$.*

Note:

$$xHx^{-1} := \{xhx^{-1} : h \in H\}.$$

Proof. ($\Rightarrow$) If $H \trianglelefteq G$ then for any $x \in G$ and $h \in H$, there exists $h' \in H$ such that $xh = h'x$ (be careful, $xH = Hx$ does not imply $xh = hx$). Thus $xhx^{-1} = h'$ and so $xHx^{-1} \subseteq H$.

($\Leftarrow$) If $xHx^{-1} \subseteq H$ for all $x \in G$, then letting $x = a$ yields $aHa^{-1} \subseteq H$, which is equivalent to (check it) $aH \subseteq Ha$.

Letting $x = a^{-1}$ yields $a^{-1}H(a^{-1})^{-1} \subseteq H$, i.e., $a^{-1}Ha \subseteq H$, which is equivalent to (check it) $Ha \subseteq aH$.

Thus $aH = Ha$ and so $H \trianglelefteq G$. □

Note: To prove that a set H is a normal subgroup, we must prove that it is a subgroup and that it is normal.

Example 1.60. *Every subgroup H of an Abelian group G is normal. Why? Since $ab = ba$ for all $a, b \in G$, we have $ah = ha$ for $a \in G$ and $h \in H$ and thus $aH = Ha$.*

Example 1.61. $\mathrm{SL}_n(\mathbb{R}) \trianglelefteq \mathrm{GL}_n(\mathbb{R})$. *Why?*

We use the criterion above: Let $A \in \mathrm{GL}_n(\mathbb{R})$ and $B \in \mathrm{SL}_n(\mathbb{R})$. We want to show that $ABA^{-1} \in \mathrm{SL}_n(\mathbb{R})$, i.e., $\det(ABA^{-1}) = 1$.

But $\det(ABA^{-1}) = \det(A)\underbrace{\det(B)}_{=1}\det(A)^{-1} = 1.$

Exercise 1.62.

1. *Prove that $Z(G) \trianglelefteq G$.*
2. *Consider the subgroup $K = \{id, (1\ 2\ 3), (1\ 3\ 2)\}$ of A_4. Prove that K is not normal in A_4.*

Are there other situations in which $H \trianglelefteq G$? Yes!

Exercise 1.63. *Prove that if H is a subgroup of G with index $[G : H] = 2$, then $H \trianglelefteq G$.*

Hint: By Langrange's Theorem, there are two left cosets of H: H and aH, and two right cosets of H: H and Hb. Show that $aH = Hb$ (Lemma1.49, statement 3 give $G \setminus H = aH$, and the same result holds for right cosets) and that $Hb = Ha$.

Warning: $[G : H] = 2$ implies $H \trianglelefteq G$, but $H \trianglelefteq G$ does not imply $[G : H] = 2$.

Exercise 1.64. *Let $G = S_3$. Find a non-trivial (i.e. not $\{id\}$ and not G) normal subgroup of S_3.*

1.6 Quotient Groups

Let G be a group and let H be a subgroup of G.

Idea: We want to discuss a general structure whose elements will be the left cosets of H (where H is a subgroup).

Definition 1.65. *We define*

$$G/H := \{aH : a \in G\},$$

and call it G mod H, the set of distinct left cosets of H in G.

Example 1.66. *Let $G = S_3$ and $H = \{id, \rho\sigma\}$, where $\rho = (1\ 2\ 3)$ and $\sigma = (1\ 2)$. Here*

$$G/H = \{H, \sigma H, \sigma\rho H\}.$$

Theorem 1.67 (Hölder, 1889)**.** *Let H be a normal subgroup of G. Then G/H, equipped with the operation*

$$(aH)(bH) := (ab)H$$

is a group.

Before the proof, let us look again at the previous example.

Example 1.68.

$$(\sigma H)(\sigma\rho H) \overset{def}{=} (\sigma\sigma\rho)H = \sigma^2\rho H = \rho H \ (since\ \sigma^2 = id) = \sigma\rho H \ (check\ it).$$

Proof. The operation is

$$G/H \times G/H \to G/H,\ (aH, bH) \mapsto (ab)H.$$

We check the definition of group.
Associative: Let $xH, yH, zH \in G/H$. Then

$$\begin{aligned} xH(yH \cdot zH) &= xH \cdot (yz)H \\ &= x(yz)H \\ &= (xy)zH \text{ since the operation on } G \text{ is associative} \\ &= (xy)H \cdot zH \\ &= (xH \cdot yH)zH. \end{aligned}$$

Identity: We show that $eH = H$ is the identity element. For all $xH \in G/H$:

$$xH \cdot H = xHeH = (xe)H = xH, \text{ and}$$

$$H \cdot xH = eH \cdot xH = (ex)H = xH.$$

Thus, H is the identity in G/H.

Inverses: Check that the inverse of xH is $x^{-1}H$, i.e., that $(xH)(x^{-1}H) = H$ and $(x^{-1}H)(xH) = H$:

$$(xH)(x^{-1}H) = (xx^{-1})H = eH = H.$$

The other one is similar (do it).

Are we done? No! (We have not even used that H is a normal subgroup yet). We need to check that the operation if well-defined. What is this?

The underlying problem is that we can have $aH = bH$ with $a \neq b$ (cf. Lemma 1.49). So we could have $aH = a'H$ and $bH = b'H$ with $a \neq a'$ and $b \neq b'$. But since $aH = a'H$ and $bH = b'H$ we must have $(aH)(bH) = (a'H)(b'H)$, i.e., $abH = a'b'H$. In other words: We must check that the result does not depend on the particular way that we use to write the cosets (as aH or $a'H$, or as bH or $b'H$):

Since $aH = a'H$ we have $a' \in a'H = aH$, so $a' = ah_1$ for some $h_1 \in H$. Similarly $b' = bh_2$ for some $h_2 \in H$. Thus

$$\begin{aligned} a'b'H &= ah_1bh_2H \\ &= ah_1bH \text{ (since } h_2H = H) \\ &= ah_1Hb \text{ (since } bH = Hb) \\ &= aHb \text{ (since } h_1H = H) \\ &= abH \text{ (since } bH = Hb). \end{aligned}$$

□

Example 1.69. *Consider $G = (\mathbb{Z}, +)$ and $H = 4\mathbb{Z} = \{\ldots, -8, -4, 0, 4, 8, \ldots\}$. Note that $H \trianglelefteq G$ (since G is Abelian, and thus every subgroup is normal). We construct $G/H = \mathbb{Z}/4\mathbb{Z}$.*

The elements of $\mathbb{Z}/4\mathbb{Z}$ are the left cosets of H, so are of the form $a + H$ for $a \in \mathbb{Z}$ (recall that the operation is the sum, hence $a + H$ instead of aH). So

$$\begin{aligned} G/H = \mathbb{Z}/4\mathbb{Z} &= \{a + 4\mathbb{Z} : a \in \mathbb{Z}\} \\ &= \{0 + 4\mathbb{Z}, 1 + 4\mathbb{Z}, 2 + 4\mathbb{Z}, 3 + 4\mathbb{Z}\} \end{aligned}$$

Why is it so, why are there no other (distinct) cosets?

If $a \in \mathbb{Z}$ then by the division algorithm, we have $a = 4t + r$ where $0 \leq r < 4$. Thus:

$$\begin{aligned} a + 4\mathbb{Z} = 4t + r + 4\mathbb{Z} &= r + 4t + 4\mathbb{Z} \\ &= r + 4\mathbb{Z} \end{aligned}$$

using that $4t + 4\mathbb{Z} = 4\mathbb{Z}$ since $4t \in 4\mathbb{Z}$ (cf. Lemma 1.49). The result follows since $r = 0, 1, 2$ or 3.

1.7 Homomorphisms and Isomorphisms

Definition 1.70. *Let G, H be groups. A (group) homomorphism $\Phi : G \to H$ is a map satisfying*

$$\Phi(ab) = \Phi(a) \cdot \Phi(b) \text{ for all } a, b \in G. \tag{1.3}$$

Remark 1.71.

1. *Observe that the map* $\Phi : G \to H$ *given by* $\Phi(g) = e_H$ *(the identity of the group* H*) for every* $g \in G$*, is always a homomorphism (the trivial homomorphism).*

   We check it: $\Phi(ab) = e_H = e_H \cdot e_H = \Phi(a) \cdot \Phi(b)$.

2. *It is important to keep track of the operation in* G *and* H*: The left hand side of* (1.3) *uses the operation of* G*, and the right hand side uses the operation of* H*, e.g,*

$$\begin{aligned} (G, +) \text{ and } (H, \cdot) &\text{ gives } \Phi(a+b) = \Phi(a) \cdot \Phi(b) \\ (G, \cdot) \text{ and } (H, +) &\text{ gives } \Phi(a \cdot b) = \Phi(a) + \Phi(b) \\ (G, +) \text{ and } (H, +) &\text{ gives } \Phi(a+b) = \Phi(a) + \Phi(b) \end{aligned}$$

Example 1.72.

1. *Consider* $\Phi : \mathbb{Z} \to \mathbb{Z}_n$, $\Phi(m) = m \bmod n$. *Then*

$$\Phi(a+b) = (a+b) \bmod n = (a \bmod n) + (b \bmod n) = \Phi(a) + \Phi(b),$$

   so Φ *is a homomorphism.*

2. $\exp : (\mathbb{R}, +) \to (\mathbb{R}^\times, \cdot)$, $\exp(x) = e^x$. *Then*

$$\exp(x+y) = e^{x+y} = e^x e^y = \exp(x)\exp(y),$$

   so exp *is a homomorphism.*

3. $\Phi : \mathrm{GL}_n(\mathbb{R}) \to (\mathbb{R}^\times, \cdot)$, $\Phi(A) = \det A$. *We have*

$$\Phi(AB) = \det(AB) = \det(A) \cdot \det(B) = \Phi(A) \cdot \Phi(B),$$

   so Φ *is a homomorphism.*

4. $\Phi : (\mathbb{R}, +) \to (\mathbb{R}, +)$, $\Phi(x) = x^2$. *We have*

$$\Phi(x+y) = (x+y)^2 = x^2 + 2xy + y^2,$$

$$\Phi(x) + \Phi(y) = x^2 + y^2.$$

   They are not equal for every $x, y \in \mathbb{R}$*, so* Φ *is not a homomorphism.*

Lemma 1.73. *Let* $\Phi : G \to H$ *be a group homomorphism. Then for all* $x, y \in G$ *we have*

1. $\Phi(xy^{-1}) = \Phi(x)\Phi(y)^{-1}$.

2. $\Phi(e_G) = e_H$.

3. $\Phi(y^{-1}) = \Phi(y)^{-1}$.

Proof. 1. $\Phi(xy^{-1})\Phi(y) = \Phi(xy^{-1}y) = \Phi(x)$. Multiplying both sides on the right by $\Phi(y)^{-1}$ we get $\Phi(xy^{-1})\Phi(y)\Phi(y)^{-1} = \Phi(x)\Phi(y)^{-1}$, so $\Phi(xy^{-1}) = \Phi(x)\Phi(y)^{-1}$.

2. Take $x = y$ in 1 to get $\Phi(e_G) = \Phi(x)\Phi(x)^{-1} = e_H$.

3. Take $x = e_G$ in 1 to get $\Phi(y^{-1}) = \Phi(e_G \cdot y^{-1}) = \Phi(e_G)\Phi(y^{-1}) = e_H\Phi(y)^{-1} = \Phi(y)^{-1}$. □

Definition 1.74. *Let $\Phi : G \to H$ be a group homomorphism. We define*

$$\operatorname{Ker}\Phi := \{g \in G : \Phi(g) = e_H\},$$

the kernel of Φ, and

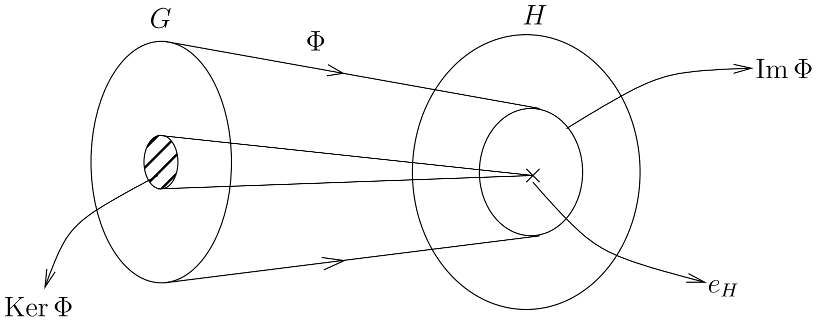
$$\begin{aligned}\operatorname{Im}\Phi &:= \{h \in H : h = \Phi(g) \textit{ for some } g \in G\}\\ &= \{\Phi(g) : g \in G\},\end{aligned}$$

the image of Φ.

Remark 1.75. *1. As sets,* $\operatorname{Ker}\Phi \subseteq G$ *and* $\operatorname{Im}\Phi \subseteq H$.

2. Φ *is surjective* $\Leftrightarrow \operatorname{Im}\Phi = H$.

3. A general picture:



Proposition 1.76. *Let $\Phi : G \to H$ be a group homomorphism. Then*

1. $\operatorname{Ker}\Phi \trianglelefteq G$*;*

2. $\operatorname{Im}\Phi \leq H$.

Proof.

1. We first check that $\operatorname{Ker}\Phi \leq G$, using the subgroup test.

   Since $\Phi(e_G) = e_H$ we have $e_G \in \operatorname{Ker}\Phi$, which is then non-empty. Let $g_1, g_2 \in \operatorname{Ker}\Phi$. Then $\Phi(g_1 g_2) = \Phi(g_1)\Phi(g_2) = e_H e_H = e_H$, so $g_1 g_2 \in \operatorname{Ker}\Phi$. Furthermore, $\Phi(g_1^{-1}) = \Phi(g_1)^{-1} = e_H^{-1} = e_H$, so $g_1^{-1} \in \operatorname{Ker}\Phi$. So $\operatorname{Ker}\Phi \leq G$.

   We use the normality test: Let $x \in G$ and $g \in \operatorname{Ker}\Phi$. We want to show that $xgx^{-1} \in \operatorname{Ker}\Phi$. We compute $\Phi(xgx^{-1}) = \Phi(x)\Phi(g)\Phi(x^{-1}) = \Phi(x)e_H\Phi(x)^{-1} = \Phi(x)\Phi(x)^{-1} = e_H$. So $x(\operatorname{Ker}\Phi)x^{-1} \subseteq \operatorname{Ker}\Phi$ for every $x \in G$, thus $\operatorname{Ker}\Phi \trianglelefteq G$.

2. We use the subgroup test. $\operatorname{Im}\Phi$ is non-empty since G is non-empty. Let $\Phi(g_1), \Phi(g_2) \in \operatorname{Im}\Phi$. Then $\Phi(g_1)\Phi(g_2) = \Phi(g_1 g_2) \in \operatorname{Im}\Phi$, and $\Phi(g_1)^{-1} = \Phi(g_1^{-1}) \in \operatorname{Im}\Phi$. □

Exercise 1.77. *Check the following.*

1. *Consider* $\Phi : (\mathbb{Z}, +) \to (\mathbb{Z}_n, +)$, $\Phi(m) = m \bmod n$. *Then*

$$\operatorname{Ker}\Phi = \langle n \rangle = \{k \cdot n : k \in \mathbb{Z}\}.$$

$$\operatorname{Im}\Phi = \mathbb{Z}_n \ \textit{(Φ is surjective).}$$

2. $\exp : (\mathbb{R}, +) \to (\mathbb{R}^\times, \cdot)$, $\exp(x) = e^x$. *Then*

$$\operatorname{Ker}\exp = \{0\}, \ \textit{and}$$

$$\operatorname{Im}\exp = \mathbb{R}_{>0} \ \textit{(the positive real numbers).}$$

3. $\Phi : \mathrm{GL}_n(\mathbb{R}) \to (\mathbb{R}^\times, \cdot)$, $\Phi(A) = \det A$. *Then*

$$\operatorname{Ker}\Phi = \mathrm{SL}_n(\mathbb{R}), \ \textit{and} \ \operatorname{Im}\Phi = \mathbb{R}^\times.$$

4. *On exercise set 3, problem 5:*

$$\varepsilon : S_n \to (\{-1, 1\}, \cdot)$$

given by

$$\varepsilon(\sigma) = \begin{cases} 1 & \textit{if } \sigma \textit{ is even} \\ -1 & \textit{if } \sigma \textit{ is odd.} \end{cases}$$

We have

$$\operatorname{Ker}\varepsilon = A_n \ \textit{and} \ \operatorname{Im}\varepsilon = \{-1, 1\}.$$

Definition 1.78. *A bijective group homomorphism is called an isomorphism. Also, two groups G and H are isomorphic if there exists an isomorphism $\Phi : G \to H$. In this case we write $G \cong H$.*

Remark 1.79. 1. *Intuitively, two groups are isomorphic if they are the "same" group. The only difference is that the elements may have different names and the symbol used for the operation may be a different one. The map Φ gives a correspondence between the elements of G and the elements of H (since it is bijective), and if we use this correspondence to "change" the names of elements of G into elements of H, the homomorphism property tells us that the operation behaves in the same way.*

2. *(Exercise; see Problem Sheets) If $\Phi : G \to H$ is an isomorphism, then so is $\Phi^{-1} : H \to G$.*

Example 1.80. 1. *Let G be the group:*

$$G = \{\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} : n \in \mathbb{Z}\}$$

with operation $\cdot$ (check this; hint: The shortest way is to show that it is a subgroup of $\mathrm{GL}_2(\mathbb{R})$).

We claim that $\mathbb{Z} \cong G$. Why?

Let $\Phi : (\mathbb{Z}, +) \to G$, $\Phi(n) = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$. We show that Φ is an isomorphism.

(1) Φ is clearly bijective.

(2) Φ is a homomorphism:

$$\Phi(m+n) = \begin{pmatrix} 1 & m+n \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} = \Phi(m)\Phi(n).$$

2. *Let $G = (\mathbb{Q}, +)$ and $H = (\mathbb{Q}^\times, \cdot)$. We claim that there is not isomorphism between G and H.*

We use contradition: Suppose $\Phi : \mathbb{Q} \to \mathbb{Q}^\times$ is an isomorphism. Consider $-1 \in \mathbb{Q}^\times$. Since Φ is surjective there is $a \in \mathbb{Q}$ such that $\Phi(a) = -1$. But then

$$-1 = \Phi(a) = \Phi(\frac{a}{2} + \frac{a}{2}) = \Phi(\frac{a}{2})\Phi(\frac{a}{2}) = [\Phi(\frac{a}{2})]^2,$$

a contradiction since no rational number squared is equal to -1.

Proposition 1.81. *A group homomorphism $\Phi : G \to H$ is injective if and only if* $\operatorname{Ker}\Phi = \{e_G\}$.

Proof. ($\Rightarrow$) Since $\Phi(e_G) = e_H$ we have $e_G \in \operatorname{Ker}\Phi$. Let $x \in \operatorname{Ker}\Phi$, i.e., $\Phi(x) = e_H = \Phi(e_G)$. Since Φ is injective we deduce $x = e_G$.

($\Leftarrow$) Let $a, b \in G$ be such that $\Phi(a) = \Phi(b)$. Then $\Phi(a)\Phi(b)^{-1} = e_H$, so $\Phi(ab^{-1}) = e_H$, i.e. $ab^{-1} \in \operatorname{Ker}\Phi = \{e_G\}$. Therefore $ab^{-1} = e_G$ and thus $a = b$. □

Corollary 1.82. *A group homomorphism $\Phi : G \to H$ is an isomorphism if and only if* $\operatorname{Ker}\Phi = \{e_G\}$ *and* $\operatorname{Im}\Phi = H$.

Given a group homomorphism $\Phi : G \to H$, we have seen the normal subgroup $\operatorname{Ker}\Phi$ of G and the subgroup $\operatorname{Im}\Phi$ of H. The next result shows that there is a link between these.

Theorem 1.83 (First Isomorphism Theorem)**.** *Let $\Phi : G \to H$ be a group homomorphism. Then the map*

$$\Psi : G/\operatorname{Ker}\Phi \to \operatorname{Im}\Phi, \quad \Psi(x \cdot \operatorname{Ker}\Phi) := \Phi(x)$$

is an isomorphism of groups.

Proof. Let $N = \operatorname{Ker}\Phi$. We first check that Ψ is well-defined: Suppose $xN = yN$. We want to show that $\Psi(xN) = \Psi(yN)$. i.e., $\Phi(x) = \Phi(y)$. Since $xN = yN$ there is $n \in N$ such that $x = yn$. Therefore $\Phi(x) = \Phi(yn) = \Phi(y)\Phi(n) = \Phi(y)e_H = \Phi(y)$.

We check that Ψ is a homomorphism: Let $xN, yN \in G/N$. Then

$$\begin{aligned}\Psi(xN \cdot yN) &= \Psi(xyN) \text{ (by definition of the product in } G/N)\\ &= \Phi(xy) \text{ (by definition of } \Psi)\\ &= \Phi(x)\Phi(y)\\ &= \Psi(xN)\Psi(yN).\end{aligned}$$

We check that Ψ is injective. Suppose $\Psi(xN) = \Psi(yN)$, i.e., $\Phi(x) = \Phi(y)$. Then $e_H = \Phi(x)^{-1}\Phi(y) = \Phi(x^{-1}y)$. So $x^{-1}y = z$ for some $z \in N$. It follows that $y = xz$ and thus $yN = xN$.

We check that Φ is surjcetive: Let $h \in \operatorname{Im}\Phi$, i.e., $h = \Phi(g)$ for some $g \in G$. Then by definition $\Phi(g) = \Psi(gN)$. □

Example 1.84. *Consider $\Phi : (\mathbb{Z}, +) \to (\mathbb{Z}_n, +)$, $\Phi(m) = m \bmod n$. We have seen*

1. *Φ is a homomorphism;*

2. $\operatorname{Ker}\Phi = \langle n\rangle = \{kn : k \in \mathbb{Z}\} = n\mathbb{Z}$*;*

3. Φ *is surjective:* $\operatorname{Im}\Phi = \mathbb{Z}_n$*.*

By the first isomorphism theorem, we have

$$\mathbb{Z}/\operatorname{Ker}\Phi \cong \operatorname{Im}\Phi = \mathbb{Z}_n,$$

so $\mathbb{Z}/\langle n\rangle \cong \mathbb{Z}_n$ *(i.e.,* $\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}_n$*).*
We already saw this! Take $n = 4$*. Then*

$$\begin{aligned}\mathbb{Z}/\langle 4\rangle = \mathbb{Z}/4\mathbb{Z} &= \{0+4\mathbb{Z}, 1+4\mathbb{Z}, 2+4\mathbb{Z}, 3+4\mathbb{Z}\}\\ &\cong \{0,1,2,3\} = \mathbb{Z}_4.\end{aligned}$$

Exercise 1.85. *Prove that* $\operatorname{GL}_n(\mathbb{R})/\operatorname{SL}_n(\mathbb{R}) \cong \mathbb{R}^\times$*.*

Answer 1. *Solution: We use the first isomorphism theorem. For this we need to find a homomorphism* $\Phi : \operatorname{GL}_n(\mathbb{R}) \to \mathbb{R}^\times$ *such that* $\ker\Phi = \operatorname{SL}_n(\mathbb{R})$ *and* $\operatorname{Im}\Phi = \mathbb{R}^\times$*.*
Let $\Phi : \operatorname{GL}_n(\mathbb{R}) \to \mathbb{R}^\times$*,* $\Phi(A) = \det A$*. We saw that* Φ *is a homomorphism,* $\operatorname{Ker}\Phi = \operatorname{SL}_n(\mathbb{R})$ *and* Φ *is surjective. Therefore (by the first isomorphism theorem)*

$$\operatorname{GL}_n(\mathbb{R})/\operatorname{Ker}\Phi \cong \operatorname{Im}\Phi, \text{ i.e.,}$$

$$GL_n(\mathbb{R})/\operatorname{SL}_n(\mathbb{R}) \cong \mathbb{R}^\times.$$

Chapter 2

Ring Theory

Idea: We have seen the notion of group, where we had only one operation. Now we will add more structure to a group in order to understand: polynomial rings, ideals, quotient rings.

Definition 2.1. *A ring is a non-empty set R equipped with two binary operations $+$ (addition) and $\cdot$ (multiplication) satisfying*

1. *R with $+$ is an Abelian group (its identity is denoted 0 and the inverse for $+$ of an element x is denoted $-x$).*
2. *$\cdot$ is associative, i.e., $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ for all $a, b, c \in R$.*
3. $$a \cdot (b + c) = a \cdot b + a \cdot c, \text{ and}$$ $$(b + c) \cdot a = b \cdot a + c \cdot a,$$ *for all $a, b, c \in R$.*

Example 2.2. *The following are rings:*

- *$\mathbb{Z}$, $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{C}$, with their usual sum and product;*
- *$\mathbb{Z}_n$ with sum and multiplication modulo n;*
- *$M_n(\mathbb{R})$, the set of all $n \times n$ matrices, with matrix addition and multiplication.*

Note: The ring depends on how one defines $+$ and $\cdot$.

Example 2.3. *Define "addition" on $\mathbb{Z}$ by $a \oplus b := a + b - 1$ and "multiplication" on $\mathbb{Z}$ by $a \odot b := a + b - ab$.*

Check that $\mathbb{Z}$ with $\oplus$ and $\odot$ is a ring.

Exercise 2.4. 1. *Prove that for $a \in R$ (a ring), the equation $a + x = 0$ has a unique solution.*

We write $-a$ for this unique solution. Thus $b - a$ means $b + (-a)$.

2. *Let S be the set of odd integers (with usual sum and product of integers). Show that S is not a ring.*

Definition 2.5. 1. *A ring with identity (or "unity") is a ring R that contains an element denoted 1 such that $a \cdot 1 = a = 1 \cdot a$ for all $a \in R$. Such a ring is also called a unitary ring.*

2. *A commutative ring is a ring R such that $a \cdot b = b \cdot a$ for all $a, b \in R$.*

Example 2.6. 1. *$\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ are all commutative rings with identity (1 in all cases).*

2. *Let $\mathbb{R}_{>0}$ be the set of all positive real numbers. Define "addition" by $a \oplus b = ab$ and "multiplication" by $a \odot b = a^{\log b}$.*

Check that this yields a commutative ring with identity.

Definition 2.7. *An integral domain is a commutative ring R with identity such that if $a, b \in R$ and $ab = 0$ then $a = 0$ or $b = 0$.*

Example 2.8. *$\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ are all integral domains.*

$\mathbb{Z}_6$ is not an integral domain (here $2 \cdot 3 = 6 = 0 \pmod 6$, but $2 \neq 0 \pmod 6$ and $3 \neq 0 \pmod 6$).

Exercise 2.9. *Show that $\mathbb{Z}_n$ is an integral domain if and only if n is prime.*

Definition 2.10. *A field is a commutative ring R with identity such that for each $a \in R \setminus \{0\}$, there is $x \in R$ such that $ax = 1$. This element x is called the (multiplicative) inverse of a and is denoted by a^{-1}.*

Example 2.11. 1. *If R is a field, then R is an integral domain. Check it.*

2. *$\mathbb{Q}$, $\mathbb{R}$, $\mathbb{Z}_p$ for p prime, $\mathbb{C}$ are all fields.*

$\mathbb{Z}$ is not a field.

Theorem 2.12. 1. *Every field F is an integral domain.*

2. *If R is an integral domain and if $a \in R \setminus \{0\}$, then $ab = ac$ in R implies $b = c$.*

3. *Every finite integral domain is a field.*

Proof. 1. As F is a field, it is a commutative ring with identity. Suppose $ab = 0$ If $a = 0$ we are done. If $a \neq 0$, then a is a unit and so $a^{-1}(ab) = a^{-1}0$, i.e., $(a^{-1}a)b = 0$, so $b = 0$.

2. If $ab = ac$ then $ab - ac = 0$, so $a(b - c) = 0$. As $a \neq 0$, then $b - c = 0$, i.e., $b = c$.

3. Recall the Pigeonhole principle: If n objects are put into n boxes in such a way that no box receives more than one object, then each box receives exactly one object.

   Reformulated: If S is a finite set (with n elements) and $f : S \to S$ is an injective map, then f is surjective.

   Let R be a finite integral domain. To show that R is a field, we have to show that if $a \in R \setminus \{0\}$, then a has an inverse, i.e., there is an element $x \in R$ such that $ax = 1$.

   Consider the map $f : R \to R$, $f(x) = ax$. By the previous item, f is injective. By the observation, f is thus surjective. In particular $1 \in \operatorname{Im} f$, i.e., there is $x \in R$ such that $f(x) = 1$, i.e., $ax = 1$. □

Definition/Proposition 2.13 (Product of rings)**.** *Let R and S be rings. Then the set*

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

with operations defined "coordinate by coordinate", i.e.,

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

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

is also a ring.

Its identity element if $0_{R\times S} = (0_R, 0_S)$.

The same statement is also true for a product of more than just 2 rings, or even an infinite number of rings, with addition and multiplication are defined coordinate by coordinate.

Definition 2.14. *If a subset S of a ring R is itself a ring under $+$ and $\cdot$ (from R), then S is a subring of R.*

Example 2.15. *1. $\mathbb{Z}$ is a subring of $\mathbb{Q}$ and of $\mathbb{C}$, $\mathbb{Q}$ is a subring of $\mathbb{C}$.*

2. $\mathbb{Z}[i] := \{a + bi : a, b \in \mathbb{Z}\}$ is a subring of $\mathbb{C}$ (check this).

3. *$J := \{0, 2\}$ is a subring of $\mathbb{Z}_4$, $K := \{0, 3, 6, 9\}$ is a subring of $\mathbb{Z}_{12}$.*

   Also, $J \times K$ is a subring of $\mathbb{Z}_4 \times \mathbb{Z}_{12}$ (recall that the operations are defined coordinate by coordinate for a product of rings).

4. *$\{1, -1, i, -i\}$ is not a subring of $\mathbb{C}$ (why?).*

Definition 2.16. *If a subset S of a field R is itself a field and contains 1_R, then S is called a subfield of R.*

Example 2.17. 1. *$\mathbb{Q}$ is a subfield of $\mathbb{R}$.*

2. *Let $d \in \mathbb{Z}$ be such that d is not a square. Then*

$$\mathbb{Q}(\sqrt{d}) := \{a + b\sqrt{d} : a, b \in \mathbb{Q}\}$$

   is a subfield of $\mathbb{C}$.

3. *Let $\{E_i\}_{i \in I}$ be a family of subfields of a field K. Then $\bigcap_{i \in I} E_i$ is a subfield of K (check this).*

Exercise 2.18 (Subring Test)**.** *Let S be a non-empty subset of a ring R such that*

1. *$a - b \in S, \quad \forall a, b \in S$;*

2. *$ab \in S, \quad \forall a, b \in S$.*

Prove that S is a subring of R.

Definition 2.19. *An element a in a ring R with identity is called a unit (or invertible) if there exists $u \in R$ such that $au = 1 = ua$. The element u is the multiplicative inverse of a and is denoted a^{-1}.*

Note the following:

1. The multiplicative inverse of a, when it exists, is unique (check it). Hence the terminology *the* multiplicative inverse.

2. The set of units in R forms a group under multiplication and is denoted by $R^\times$ (which is not necessarilly $R \setminus \{0\}$, be careful…).

3. Every non-zero element in a field F is a unit (by definition of field), i.e., $F^\times = F \setminus \{0\}$.

Example 2.20.

- *$\mathbb{Z}_p$ is a field when p is prime.*

- *$\mathbb{Z}_3[i] := \{a + bi : a, b \in \mathbb{Z}_3\}$ is a field.*

  Note that $\mathbb{Z}_3 = \{0, 1, 2\}$, so that $\mathbb{Z}_3[i]$ has 9 elements.

2.1 Homomorphisms, Isomorphisms

Definition 2.21. *Let R and S be rings. A function $f : R \to S$ is a ring homomorphism if, for all $a, b \in R$:*

1. $f(a+b) = f(a) + f(b)$;
2. $f(ab) = f(a)f(b)$.

Example 2.22. *Consider $f : \mathbb{R} \to M_2(\mathbb{R})$, $f(r) = \begin{pmatrix} 0 & 0 \\ -r & r \end{pmatrix}$. Then f is a ring homomorphism. Check:*

1. $f(r+s) = \begin{pmatrix} 0 & 0 \\ -(r+s) & r+s \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ -r & r \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ -s & s \end{pmatrix} = f(r) + f(s)$;
2. $f(rs) = \begin{pmatrix} 0 & 0 \\ -rs & rs \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ -r & r \end{pmatrix} \begin{pmatrix} 0 & 0 \\ -s & s \end{pmatrix} = f(r)f(s)$.

Example 2.23. *Consider $f : \mathbb{Z} \to \mathbb{Z}$, $f(x) = -x$. Is f a ring homomorphism?*

Check: $f(-1)f(-1) = 1 \cdot 1 = 1$, but $f(-1 \cdot -1) = f(1) = -1$. So no, f is not a ring homomorphism.

Definition 2.24. *Let R and S be rings. A function $f : R \to S$ is a ring isomorphism if f is a bijective ring homomorphism.*

If such an f exists, we say that R is isomorphic to S and write $R \cong S$.

Exactly as in the case of groups, if $f : R \to S$ is a ring isomorphism, then so is $f^{-1} : S \to R$.

Example 2.25.

1. *Let $\mathbb{Z}[\sqrt{2}] := \{a + b\sqrt{2} : a, b \in \mathbb{Z}\}$ and $H := \{\begin{pmatrix} a & 2b \\ b & a \end{pmatrix} : a, b \in \mathbb{Z}\}$. Then $\mathbb{Z}[\sqrt{2}] \cong H$ as rings.*

   Check that $f(a + b\sqrt{2}) = \begin{pmatrix} a & 2b \\ b & a \end{pmatrix}$ is an isomorphism of rings.

2. *Consider $f : \mathbb{R} \to M_2(\mathbb{R})$, $f(r) = \begin{pmatrix} 0 & 0 \\ -r & r \end{pmatrix}$. Is f an isomorphism?*

   Injective? Yes.

   Surjective? No: There is no $r \in \mathbb{R}$ such that $f(r) = I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Exercise 2.26. *Let $f : R \to S$ be a ring homomorphism. Prove that*

1. *$f(0_R) = 0_S$. Hint: Consider $f(0_R + 0_R)$.*

2. *$f(-a) = -f(a)$ for all $a \in R$. Hint: Recall that to check that $y = -x$ it suffices to check that $x + y = 0$.*

3. *$f(a - b) = f(b) - f(b)$ for all $a, b \in R$.*

If R is a ring with identity 1_R and f is surjcetive, prove that

4. *S has an identity 1_S and $f(1_R) = 1_S$.*

5. *If u is a unit in R, then $f(u)$ is a unit in S and $f(u)^{-1} = f(u^{-1}$.*

Akin to group theory, we now define.

Definition 2.27. *If $f : R \to S$ is a ring homomorphism, then*

$$\begin{aligned} \operatorname{Im} f &:= \{s \in S : s = f(r) \text{ for some } r \in R\} \\ &= \{f(r) : r \in R\}. \end{aligned}$$

the image of f, and

$$\operatorname{Ker} f := \{r \in R : f(r) = 0_S\}$$

is the kernel of f.

Note that if f is surjective, then $\operatorname{Im} f = S$.

Proposition 2.28. *If $f : R \to R$ is a ring homomorphism, then* $\operatorname{Im} f$ *is a subring of S.*

Proof. Since R is non-empty, $\operatorname{Im} f$ is non-empty. We use the subring test:

If $f(a), f(b) \in \operatorname{Im} f$, then $f(a)f(b) = f(ab) \in \operatorname{Im} f$ and $f(a) - f(b) = f(a - b) \in \operatorname{Im} f$. □

Exercise 2.29. *Prove that* $\operatorname{Ker} f$ *is a subring of R.*

2.2 Ideals and Quotient Rings

Definition 2.30. *A subring I of a ring R is an ideal if for every $r \in R$ and $a \in I$, then $ra \in I$ and $ar \in I$.*

Note: If R is commutative, we just need to check say $ra \in I$.

Example 2.31. *1. For any ring R, $\{0\}$ and R are ideals of R ("trivial ideals").*

2. *Let $\mathbb{R}[X]$ be the ring of polynomials in X with coefficients in $\mathbb{R}$ (more later about it):*

$$f \in \mathbb{R}[X] \Rightarrow f(X) = a_n X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0.$$

Consider $S = \{f(x) : f \in \mathbb{R}[X],\ a_0 = 0\}$. Check that S is an ideal of $\mathbb{R}[X]$.

3. *If A, B are ideals of a ring R, then the sum of A and B:*

$$A + B := \{a + b : a \in A,\ b \in B\}$$

is an ideal of R (check it).

4. *The set $J = \{\begin{pmatrix} 0 & 0 \\ 0 & r \end{pmatrix} : r \in R\}$ is not an ideal of $M_2(\mathbb{R})$. Why not?*

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{R}),\ \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \in J,\ \textit{but}\ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \notin J.$$

5. *Check that the subring*

$$S := \{a + bi : a, b \in \mathbb{Z},\ b \text{ is even}\}$$

of $\mathbb{Z}[i]$ is not an ideal of $\mathbb{Z}[i]$.

Theorem 2.32 (Ideal Test)**.** *A non-empty subset I of a ring R is an ideal if and only the following two conditions are satisfied*

1. *if $a, b \in I$ then $a - b \in I$;*

2. *if $r \in R$ and $a \in I$, then $ra \in I$ and $ar \in I$.*

Proof. ($\Rightarrow$) By definition of an ideal.

($\Leftarrow$) In view of the second conditions, we only need to check that I is a subring of R. For this, the only property that remains to be checked is that if $a, b \in I$, then $ab \in I$. But it is a special case of the second condition. □

Corollary 2.33. *Let R be a commutative ring, $c \in R$ and*

$$I := \{rc : r \in R\}.$$

Then I is an ideal of R.

Proof. We use the ideal test. If $r_1, r_2 \in R$ and $r_1c, r_2c \in I$, then $r_1c - r_2c = (r_1 - r_2)c \in R$ and $r(r_1c) = (rr_1)c \in I$ (since $r_1 - r_2$ and rr_1 are in R). As R is commutative, $(r_1c)r = (r_1r)c \in I$. □

Remark 2.34.

1. *The ideal $I := \{rc : r \in R\}$ is called the (principal) ideal generated by c and is denoted by $\langle c\rangle$.*
2. *In $\mathbb{Z}$, the ideal $\langle c\rangle$ is the set of all multiples of c.*
3. *In any commutative ring R with identity, $\langle 1\rangle = R$, as $r = r \cdot 1 \in \langle 1\rangle$ for every $r \in R$.*
4. *Not every ideal in a ring R is principal (= generated by one element). For instance*
$$J := \{f(X) : f(X) \in \mathbb{Z}[X],\ 3|a_0\}$$
is an ideal of $\mathbb{Z}[X]$ but is not principal (check this).

Exercise 2.35 (Important, see Exercise Sets)**.**

1. *If I is an ideal of R and I contains a unit, then $I = R$*
2. *Let R be a with identity, and let I be an ideal of R. Then*
$$I = R \Leftrightarrow 1 \in I.$$

Proposition 2.36. *Let $f : R \to S$ be a ring homomorphism. Then* $\operatorname{Ker} f$ *is an ideal of R.*

Proof. We use the ideal test. Suppose that $a, b \in \operatorname{Ker} f$, i.e., $f(a) = 0_S$ and $f(b) = 0_S$. Then $f(a - b) = f(a) - f(b) = 0_S - 0_S = 0_S$, so $a - b \in \operatorname{Ker} f$.

Now, we must show that $ra \in \operatorname{Ker} f$ for any $r \in R$ and $a \in \operatorname{Ker} f$. Well, $f(ra) = f(r)f(a) = f(r) \cdot 0_S = 0_S$, so $ra \in \operatorname{Ker} f$. Similarly $ar \in \operatorname{Ker} f$. □

Exercise 2.37. *Let $f : R \to S$ be a ring homomorphism. Prove that* $\ker f = \{0_R\}$ *if and only if f is injective.*

Definition 2.38. *Let R be a ring and $I \subseteq R$ an ideal. The quotient ring R/I is defined as follows:*

1. *$R/I := \{a+I : a \in R\}$ the set of left cosets of I in R. It is the quotient Abelian group of the group $(R, +)$ by the normal subgroup I (under $+$). In particular, addition in R/I is defined by*
$$(a + I) + (b + I) = (a + b) + I).$$

2. *Multiplication is defined by*

$$(a+I)\cdot(b+I) := (ab)+I.$$

We check that this multiplication is well-defined.

Suppose $a+I = a'+I$ and $b+I = b'+I$, so $a = a'+x$ and $b = b'+y$ with $x, y \in I$. Then

$$\begin{aligned} ab+I &= (a'+x)(b'+y)+I \\ &= a'b' + \underbrace{a'y}_{\in I} + \underbrace{xb'}_{\in I} + \underbrace{xy}_{\in I} + I \\ &\quad (\text{so } a'y + ab' + xy \in I) \\ &= a'b' + I \end{aligned}$$

Example 2.39.

$$\mathbb{Z}/4\mathbb{Z} = \{0+\langle 4\rangle, 1+\langle 4\rangle, 2+\langle 4\rangle, 3+\langle 4\rangle\}.$$

Addition: $2+\langle 4\rangle + 3+\langle 4\rangle = 5+\langle 4\rangle = 1+\langle 4\rangle$.
Multiplication: $(2+\langle 4\rangle)(3+\langle 4\rangle) = 6+\langle 4\rangle = 2+\langle 4\rangle$.

Remark 2.40. 1. *If R is commutative, so is R/I. Why?*

$$(a+I)(b+I) = ab+I = ba+I = (b+I)(a+I).$$

2. *If R has an identity 1_R, then R/I has an identity, namely 1_R+I (check this).*

Theorem 2.41 (First Isomorphism Theorem for Rings)**.** *Let $f : R \to S$ be a ring homomorphism. Then the map*

$$\phi : R/\operatorname{Ker} f \to \operatorname{Im} f,\ \phi(r+\operatorname{Ker} f) := f(r)$$

is an isomorphism of rings.

Proof. We have seen in the section on groups that ϕ is an isomorphism of abelian groups from the group $(R/\operatorname{Ker} f, +)$ to the group $(\operatorname{Im} f, +)$. So we only have to check the multiplicative property of ring homomorphism:

$$\begin{aligned} \phi((a+\operatorname{Ker} f)(b+\operatorname{Ker} f)) &= \phi(ab+\operatorname{Ker} f) \\ &= f(ab) \\ &= f(a)f(b) \\ &= \phi(a+\ker f)\cdot\phi(b+\operatorname{Ker} f). \end{aligned}$$ □

Example 2.42. *Let* $S = \{\begin{pmatrix} a & b \\ 0 & c \end{pmatrix} : a, b, c \in \mathbb{R}\}$ *and* $I = \{\begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix} : b \in \mathbb{R}\}$*. Then* S *is a ring (with the usual sum and product of matrices), and* I *is an ideal of* S *(check this).*

Prove that $S/I \cong \mathbb{R} \times \mathbb{R}$*.*

Answer 2. *Define* $f : S \to \mathbb{R} \times \mathbb{R}$ *by* $f(\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}) = (a, c)$*. Check that* f *is a surjective ring homomorphism with* $\operatorname{Ker} f = I$*. Apply the first isomorphism theorem.*

Definition 2.43. *An ideal* P *in a commutative ring* R *with identity is called prime if* $P \neq R$ *and* $bc \in P$ *implies* $b \in P$ *or* $c \in P$ *(for all* $b, c \in R$*).*

This terminology comes from the following (which is one of the main examples):

Exercise 2.44. *In* $\mathbb{Z}$*, the ideal* $\langle n \rangle$ *is prime if and only if* n *is prime.*

Definition 2.45. *An ideal* M *in a ring* R *is called maximal if* $M \neq R$ *and, whenever* J *is an ideal such that* $M \subseteq J \subseteq R$*, then* $M = J$ *or* $J = R$*.*

(I.e., the only ideal strictly larger than M *is* R*, i.e., there is no proper ideal strictly larger than* M*.)*

Exercise 2.46.

1. *Show that if* p *is prime, then* $\langle p \rangle$ *is maximal in* $\mathbb{Z}$*.*

   Hint: Show that if J *is an ideal such that* $\langle p \rangle \subsetneqq J$*, then* $1 \in J$ *(use that* p *is prime and Bezout's identity. Deduce that* $J = R$*. Conclude.*

2. *Show that* $\langle 2 \rangle$ *and* $\langle 3 \rangle$ *are maximal ideals in* $\mathbb{Z}_{36}$*.*

Theorem 2.47. *Let* R *be a commutative ring with identity and let* P *be an ideal of* R*. Then*

$$R/P \text{ is an integral domain} \Leftrightarrow P \text{ is a prime ideal.}$$

The key to this proof is the observation (cf. Lemma 1.49 2)

$$r + P = 0 (= 0 + P) \text{ in } R/P \Leftrightarrow r \in P.$$

Proof. ($\Rightarrow$) Let $a, b \in R$ be such that $ab \in P$. Then $(a+P)(b+P) = ab+P = 0 + P = 0_{R/P}$. Since R/P is an integral domain, we get $a + P = 0 + P$ or $b + P = 0 + P$, i.e., $a \in P$ or $b \in P$.

($\Leftarrow$) Note that R/P is a commutative ring with identity. Suppose that $(a+P)(b+P) = 0+P$, i.e., $ab+P = 0+P$. Then $ab \in P$. Since P is prime, $a \in P$ or $b \in P$, and thus $a + P = 0 + P$ or $b + P = 0 + P$. □

Theorem 2.48. *Let R be a commutative ring with identity, and let M be an ideal of R. Then*

$$R/M \text{ is a field} \Leftrightarrow M \text{ is a maximal ideal.}$$

Proof. ($\Rightarrow$) We show that if B is an ideal of R such that $M \subsetneqq B \subseteq R$, then $B = R$.

As observed above in Exercise 2.35, it suffices to show that $1 \in B$. Let $b \in B \setminus M$. Therefore $b + M \neq 0 + M = 0_{R/M}$. Since R/M is a field, there is $a \in R$ such that $(b + M)(a + M) = 1 + M$, i.e., $ba + M = 1 + M$, i.e., $1 = ba + m$ for some $m \in M$. But $ba \in B$ (since $b \in B$ and B is an ideal) and $m \in M \subseteq B$. Therefore $1 \in B$.

($\Leftarrow$) It suffices to show that if $b + M \in R/M$ is different from $0_{R/M} = 0 + M$, then $b + M$ has an inverse. Observe first that $b + M \neq 0 + M$ means that $b \notin M$. Consider the set

$$B := \{m + br : m \in M,\ r \in R\}.$$

This is an ideal (check it), and it propertly contains M (since it contains b). Since M is maximal, we must have $B = R$. Thus $1 \in B$, say $1 = m' + bc$ with $m' \in M$, $c \in R$. Then

$$1 + M = bc + m' + M = bc + M = (b + M) \cdot (c + M). \qquad \square$$

Remark 2.49. 1. *Since a field is always an integral domain, it follows that every maximal ideal is a prime ideal.*

2. *But not every prime ideal is maximal. For instance:*

Consider $\mathbb{Z}[X]$ the ring of polynomials with one indeterminate and coefficients in $\mathbb{Z}$. The ideal $\langle X \rangle$ is a prime ideal in $\mathbb{Z}[X]$. Why?

Define $\Phi : \mathbb{Z}[X] \to \mathbb{Z}$, $\Phi(a_n X^n + \cdots + a_1 X + a_0) = a_0$. Then Φ is a ring hommomorphism, is surjective, and $\operatorname{Ker} \Phi = \langle X \rangle$ *(check this). Thus, by the first isomorphism theorem*

$$\mathbb{Z}[X]/\langle X \rangle \cong \mathbb{Z},$$

and so is $\mathbb{Z}[X]/\langle X \rangle$ is an integral domain, since $\mathbb{Z}$ is one. Therefore, by theorem 2.47, $\langle X \rangle$ is a prime ideal.

However, $\mathbb{Z}[X]/\langle X \rangle \cong \mathbb{Z}$, so is not a field. which means that $\langle X \rangle$ is not a maximal ideal. We can also be more explicit, you can check that

$$\langle X \rangle \subsetneqq \langle X, 2 \rangle \subsetneqq \mathbb{Z}[X],$$

where $\langle X, 2 \rangle := \{r_1 X + r_2 2 : r_1, r_1 \in \mathbb{Z}[X]\}$ is an ideal in $\mathbb{Z}[X]$.

Chapter 3

Polynomial Rings

Question: Let R be a ring. How can we think of the expression

$$a_0 + a_1X + \cdots + a_nX^n,$$

where $n \geq 0$, $a_i \in R$? What is X?

Idea: The expression makes sense if the a_i's and X are elements of a larger ring.

Theorem 3.1. *If R is a ring, there exists a ring P containing an element X not in R such that*

1. *R is a subring of P;*
2. *$Xa = aX$ for every $a \in R$;*
3. *every element of P is of the form*

   $$a_0 + a_1X + \cdots + a_nX^n,$$

   for some $n \geq 0$ and $a_0, \ldots, a_n \in R$;
4. *if $n \leq m$ and $a_0 + a_1X + \cdots + a_nX^n = b_0 + b_1X + \cdots + b_mX^m$, then $a_i = b_i$ for $i \leq n$ and $b_i = 0_R$ for $i > n$;*
5. *$a_0 + a_1X + \cdots + a_nX^n = 0$ if and only if $a_i = 0_R$ for evrey i.*

This ring P is denoted by $R[X]$.

Proof. Idea of the proof: We represent polynomials by the tuples of their coefficients:

$5+6X-2X^3$ becomes $(5, 6, 0, -2)$ or, more precisely $(5, 6, 0, -2, 0, 0, \ldots)$.

$-3 + 5X - X^2$ becomes $(-3, 5, -1, 0, 0, \ldots)$.

In general a polynomial $a_0 + a_1X + \cdots + a_nX^n$ is represented by the infinite tuple

$$(a_0, a_1, \ldots, a_n, 0, 0, \ldots),$$

such that each $a_i \in R$, and only finitely many of the coefficients of the tuple are non-zero. The ring P is then:

$$\{(a_0, a_2, a_3, \ldots)|a_i \in R, \text{ for } i \geq 0, \text{ only finitely many } a_i\text{'s are non-zero}\}.$$

And we define addition and multiplication to reflect what we want: That they correspond to our "intuitive" knowledge of polynomials. We then need to check that the properties of ring hold (they do, it is just long and not very interesting to check).

The element X is represented by the tuple $(0, 1, 0, 0, \ldots)$. □

This formal construction of $R[X]$ is just there to justify once and for all that there is indeed a ring $R[X]$. We will mostly ignore it and will manipulate polynomials as usual.

We also record the following, whose formal proof would depend (we skip it) on the above construction of $R[X]$.

Proposition 3.2. *1. If R is commutative, then $R[X]$ is commutative.*

2. If R has an identity 1_R, then $R[X]$ has an identity $1_{R[X]} = 1_R$.

Definition 3.3. *Let $f(X) = a_nX^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0 \in R[X]$, with $a_n \neq 0$. Then*

1. *a_n is called the leading coefficient of $f(X)$, and*
2. *n is called the degree of $f(X)$ (it is the largest exponent of X that appears with a non-zero coefficient). We write $\deg f(X) = n$.*

Observe that $\deg f(X) = 0 \Leftrightarrow f(X) = a_0$ for $a \in R \setminus \{0_R\}$. So the polynomials of degree 0 are the constant non-zero polynomials.

The degree of the polynomial 0_R is not defined.

Theorem 3.4. *If R is an integral domain and $f(X), g(X)$ are non-zero polynomials in $R[X]$, then*

$$\deg f(X)g(X) = \deg f(X) + \deg g(X).$$

Proof. Suppose $f(X) = a_0 + \cdots + a_nX^n$ and $g(X) = b_0 + \cdots + b_mX^m$ with $a_n \neq 0_R$ and $b_m \neq 0_R$. Thus $\deg f(X) = n$ and $\deg g(X) = m$. Then

$$f(X)g(X) = a_0b_0 + (a_0b_1 + a_1b_0)X + \cdots + a_nb_mX^{n+m}.$$

Since R is an integral domain, we have $a_nb_m \neq 0$. Thus $f(X)g(X)$ is non-zero and its degree is $n + m = \deg f(X) + \deg g(X)$. □

Corollary 3.5. *If R is an integral domain, then so is $R[X]$.*

Proof. As R is a commutative ring with 1, so is $R[X]$. By the (proof of the) previous theorem, the product of nonzero polynomials in $R[X]$ is nonzero. □

Note: For any ring R (not necessarilly an integral domain) we have

$$deg\, f(X)g(X) \leq \deg f(X) + \deg g(X).$$

Example 3.6. *We work in $\mathbb{Z}_6[X]$, with $f(X) = 2X^4$ ($\deg f = 4$), and $g(X) = 1 + 3X^2$ ($\deg g = 2$). Then*

$$f(X)g(X) = (2X^4)(1 + 3X^2) = 2X^4 + 6X^6 = 2X^4,$$

so $\deg f(X)g(X) = 4 < \deg f + \deg g = 6$.

3.1 The Division Algorithm

You have seen the (Euclidean) division algorithm in $\mathbb{Z}$:

For any integers a and b with $b > 0$, there exist unique integers q and r such that

$$a = bq + r, \qquad \text{with } 0 \leq r < b.$$

q is the quotient and r the remainder in the division of a by b.

We would like to have something similar in the ring $F[X]$, where F is a field.

Theorem 3.7. *Let F be a field and let $f(X), g(X) \in F[X]$ with $g(X) \neq 0_F$. Then there exist unique polynomials $q(X)$ and $r(X)$ such that*

$$f(X) = g(X)q(X) + r(X),$$

and either $r(X) = 0_F$ or $\deg r(X) < \deg g(X)$.

Proof. We first prove the existence of $q(X)$ and $r(X)$, and consider two cases.

Case 1: If $f(X) = 0_F$ or if $\deg f(X) < \deg g(X)$. We take $q(X) = 0$ and $r(X) = f(X)$. Then $f(X) = g(X) \cdot 0_F + f(X)$.

Case 2: If $f(X) \neq 0_F$ and $\deg g(X) \leq \deg f(X)$. We proceed by induction on $\deg f(X)$. If $\deg f(X) = 0$, then $\deg g(X) = 0$, hence $f(X) = a$ and $g(X) = b$, for some $a, b \in F \setminus \{0\}$. Since F is a field, b has an inverse, and we can take $q(X) = b^{-1}a$ and $r(X) = 0$.

Assume that the result is true for $\deg f(X) < n$ (i.e., for all polynomials $f(X)$ of degree less than n). We want to show that it is true if $\deg f(X) = n$.

Write $f(X) = \sum_{i=0}^{n} a_i X^i$ with $a_n \neq 0$. Then $g(X)$ must be of the form $g(X) = \sum_{j=0}^{m} b_j X^j$ with $b_m \neq 0$ and $m \leq n$.

Since F is a field and $b_m \neq 0$, b_m is a unit. Multiply $g(X)$ by $a_n b_m^{-1} X^{n-m}$ to get

$$\begin{aligned} a_n b_m^{-1} X^{n-m} g(X) &= a_n b_m^{-1} X^{n-m}(b_m X^m + \cdots + b_0) \\ &= a_n X^n + a_n b_m^{-1} b_{m-1} X^{n-1} + \cdots + a_n b_m^{-1} b_0 X^{n-m}. \end{aligned}$$

As $\deg f(X) = n$ and $\deg a_n b_M^{-1} X^{n-m} g(X) = n$, and since these two polynomials have the same leading coefficient, we have

$$\deg(\underbrace{f(X) - a_n b_m^{-1} X^{n-m} g(X)}_{(\star)}) < n.$$

We can apply the induction hypothesis to $(\star)$ and $g(X)$. Thus, there exist $q_1(X)$ and $r(X)$ such that

$$f(X) - a_n b_m^{-1} X^{n-m} g(X) = g(X) q_1(X) + r(X),$$

with $r(X) = 0_F$ or $\deg r(X) < \deg g(X)$. Thus

$$f(X) = q(X)(\underbrace{a_n b_m^{-1} X^{n-m} + q_1(X)}_{(\star\star)}) + r(X),$$

with $r(X) = 0$ or $\deg r(X) < \deg g(X)$. Take $q(X) = (\star\star)$ and we are done.

Uniqueness of $q(X)$ and $r(X)$: Exercise. You have to show that if $f(X) = g(X)t(X) + s(X)$ with either $s(X) = 0$ or $\deg s(X) < \deg g(X)$, then $s(X) = r(X)$ and $t(X) = q(X)$. □

Remark 3.8. *Theorem 3.7 does not necessarilly hold if we work in $R[X]$ with R ring. For instance, in $\mathbb{Z}[X]$ with $f(X) = 1$ and $g(X) = 2$. Assume there are $q(X), r(X)$ in $\mathbb{Z}[X]$ with*

$$1 = 2q(X) + r(X)$$

with $r(X) = 0$ or $\deg r(X) < \deg 2$. Since $\deg 2 = 0$, we must have $r(X) = 0$. So $1 = 2q(X)$, which is not possible with $q(X) \in \mathbb{Z}[X]$.

Example 3.9. *Given $f(X) = 3X^4 + X^3 + 2X^2 + 1$ and $g(X) = X^2 + 4X + 2$ in $\mathbb{Z}_5[X]$, find $q(X), r(X) \in \mathbb{Z}_5[X]$ such that $f(X) = g(X)q(X) + r(X)$ with $r(X) = 0$ or $\deg r(X) < \deg g(X)$.*

We simply write down the usual division, following the same algorithm:

$$\begin{array}{rrrrrr}
 & & & 3X^2 & +4X & \\ \cline{2-6}
X^2+4X+2 \mid & 3X^4 & +X^3 & +2X^2 & & +1 \\
 & 3X^4 & +2X^3 & +X^2 & & \\ \cline{2-6}
 & & 4X^3 & +X^2 & & +1 \\
 & & 4X^3 & +X^2 & 3X & \\ \cline{3-6}
 & & & & 2X & +1
\end{array}$$

So:

$$3X^4 + X^3 + 2X^2 + 1 = (X^2 + 4X + 2)(\underbrace{3X^2 + 4X}_{q(X)}) + \underbrace{2X + 1}_{r(X)}.$$

Now: What happens if $r(X) = 0$ in the Division Algorithm?

3.2 Divisibility in $F[X]$

Definition 3.10. *Let F be a field and $f(X), g(X) \in F[X]$ with $g(X)$ nonzero. We say that $g(X)$ divides $f(X)$ (or that $g(X)$ is a factor of $f(X)$, or that $f(X)$ is a multiple of $g(X)$) and write $g(X)|f(X)$ if $f(X) = g(X) \cdot h(X)$ for some $h(X) \in F[X]$.*

Example 3.11. 1. *$X + 1|X^2 - 2X - 3$ in $\mathbb{Q}[X]$ since $X^2 - 2X - 3 = (X + 1)(X - 3)$.*

2. *Also, $6(X + 1)|X^2 - 2X - 3$ in $\mathbb{Q}[X]$ since $X^2 - 2X - 3 = 6(X + 1)[\frac{1}{6}(X - 3)]$.*

Remark 3.12. 1. *If $g(X)|f(X)$ then $cg(X)|f(X)$ for each $c \in F \setminus \{0\}$.*

2. *Suppose $f(X) \neq 0$. Then every divisor of $f(X)$ has degree at most $\deg f(X)$.*

   Why? Suppose $g(X)|f(X)$ and so $f(X) = g(X)h(X)$. Thus, $\deg f(X) = \deg g(X) + \deg h(X)$, which gives $0 \leq \deg g(X) \leq \deg f(X)$.

Definition 3.13. *A polynomial in $F[X]$ is monic if its leading coefficient is 1.*

Example 3.14. *$X^2 - 2X - 3$ is monic in $\mathbb{Q}[X]$ while $\frac{1}{2}X + 2$ is not monic in $\mathbb{Q}[X]$.*

Definition 3.15. *Let F be a field and $f(X), g(X) \in F[X]$ not both zero. The greatest common divisor (gcd) of $f(X)$ and $g(X)$ is the monic polynomial $d(X)$ of highest degree that divides both $f(X)$ and $g(X)$.*

Question: Why do we say “the” gcd?

Theorem 3.16. *Let F be a field and $f(X), g(X) \in F[X]$ not both zero. Then there is a unique gcd $d(X)$ of $f(X)$ and $g(X)$. Furthermore, there exists $u(X), v(X) \in F[X]$ such that*

$$d(X) = f(X)u(X) + g(X)v(X).$$

Proof. Consider the set

$$S := \{f(X)u(X) + g(X)v(X) : u(X), v(X) \in F[X]\}.$$

(Observe that S is an ideal in the ring $F[X]$.)

Let $T(X)$ be a monix polynomial of smallest degree in S. So $T(X) = f(X)u(X) + g(X)v(X)$ for some $u(X), v(X) \in F[X]$.

Claim: $T(X) = \gcd(f(X), g(X))$.

We first show that $T(X)|f(X)$: By the Division Algorithm, there exist $q(X)$ and $r(X)$ in $F[X]$ such that $f(X) = T(X)q(X) + r(X)$, with $r(X) = 0$ or $\deg r(X) < \deg T(X)$. So

$$\begin{aligned} r(X) &= f(X) - T(X)q(X) \\ &= f(X) - [f(X)u(X) + g(X)v(X)]q(X) \\ &= f(X)[1 - u(X)q(X)] + g(X)[-q(X)v(X)]. \end{aligned}$$

Thus, $r(X) \in S$. So we cannot have $\deg r(X) < \deg T(X)$ (by choice of $T(X)$), which means that we must hav e$r(X) = 0$. So $f(X) = T(X)q(X)$. Similarly, $T(X)|g(X)$.

We now show that if $c(X)$ is another common divisor of $f(X)$ and $g(X)$, then $\deg c(X) \leq \deg T(X)$.

We have $f(X) = c(X)w(X)$ and $g(X) = c(X)s(X)$, so

$$\begin{aligned} T(X) &= f(X)u(X) + g(X)v(X) \\ &= c(X)w(X)u(X) + c(X)s(X)v(X) \\ &= c(X)[w(X)u(X) + s(X)v(X)]. \end{aligned}$$

Thus, $c(X)|T(X)$ and $\deg c(X) \leq \deg T(X)$.

We now prove the uniqueness.

Suppose $d(X)$ is another gcd of $f(X)$ and $g(X)$. By the above observation we have that $d(X)$ divides $T(X)$ and $T(X)$ divides $d(X)$. Therefore $\deg T(X) = \deg d(X)$ and $T(X) = cd(X)$ for some $c \in F \setminus \{0\}$.

But both $T(X)$ are $c(X)$ are monic, so $c = 1$ and $T(X) = d(X)$. □

Observe that we proved a little bit more than the definition of gcd, and that we obtained

Exercise 3.17. *$d(X)$ is the gcd of $f(X)$ and $g(X)$ if and only if*

1. *$d(X)$ divides $f(X)$ and $g(X)$, and*
2. *every divisor of both $f(X)$ and $g(X)$ divides $d(X)$.*

Definition 3.18. *Two polynomials $f(X)$ and $g(X)$ in $F[X]$ are relatively prime if* $\gcd(f(X), g(X)) = 1$.

Note that by the previous theorem, if $f(X)$ and $g(X)$ are relatively prime, then there exist $u(X), v(X) \in F[X]$ such that

$$f(X)u(X) + g(X)v(X) = 1.$$

Question: Given two polynomials f_1 and f_2, how do we find their gcd?

Answer: Use the division algorithm algorithm until we get a zero remainder: Say for instance that $\deg f_1 \geq \deg f_2$.

- Divide f_1 by f_2, you get a remainder f_3.
- Divide f_2 by f_3, you get a remainder f_4.
- Keep doing this, dividing f_i by f_{i+1}, obtaining a remainder f_{i+2}.
- Let $d(X)$ be the last non-zero remainder.

If necessary: Multiply $d(X)$ by a non-zero constant to get a monic polynomial which is the gcd.

Example 3.19. *Find the gcd of $f(X) = 3X^4 + X^3 + 2X^2 + 1$ and $g(X) = X^2 + 4X + 2$ in $\mathbb{Z}_5[X]$.*

We have seen

$$3X^4 + X^3 + 2X^2 + 1 = (X^2 + 4X + 2)(\underbrace{3X^2 + 4X}_{q(X)}) + \underbrace{2X + 1}_{r(X)}.$$

We now divide $X^2 + 4X + 2$ by $2X + 1$:

$$\begin{array}{rrrr} & & 3X & +3 \\ \cline{2-4} 2X+1 \mid & X^2 & +4X & +2 \\ & X^2 & +3X & \\ \cline{2-4} & & X & +2 \\ & & X & +3 \\ \cline{2-4} & & & 4 \end{array}$$

So $X^2+4X+2=(2X+1)(3X+3)+4$. *We now divide* $3X+4$ *by* 4*:*

$$\begin{array}{rll} & 3X & +4 \\ \cline{2-3} 4\,| & 2X & +1 \\ & 2X & \\ \cline{2-3} & & 1 \\ & & 1 \\ \cline{3-3} & & 0 \end{array}$$

So $2X+1=4(3X+4)+0$.

The last non-zero remainder is 4. *We make it monic:* $4(4)=1$, *so*

$$\gcd(3X^4+X^3+2X^2+1, X^2+4X+2)=1,$$

$f(X)$ *and* $g(X)$ *are relatively prime.*

Remark 3.20. *If* $f(X)=c_nX^n+\cdots+c_0$ *with* $c_n\neq 0_F$, *what is* $\gcd(f(X),0_F)$*?*

Since $u(X).0_F=0_F$, *we have* $u(X)|0_F$ *for every* $u(X)\in F[X]$. *Thus,* $u(X)$ *is a common divisor of* $f(X)$ *and* 0_F *if and only if* $u(X)|f(X)$. *Taking* $u(X)=f(X)$ *we have a common divisor of maximal degree. Making it monic, we get* $\gcd(f(X),0_F)=c_n^{-1}f(X)$.

3.3 Irreducibles

Note: In $\mathbb{Z}$ the units are 1 and -1. We have a different situation in a polynomial ring.

Theorem 3.21. *Let* R *be an integral domain. Then* $f(X)$ *is a unit in* $R[X]$ *if and only if* f *is a constant polynomial that is a unit in* R.

Proof. ($\Rightarrow$) By hypothesis $f(X)g(X)=1_R$ for some $g(X)\in R[X]$. Thus $\deg f(X)+\deg g(X)=\deg 1_R=0$. This implies that $\deg f(X)=\deg g(X)=0$ and so $f(X)=a$ and $g(X)=b$ for some $a,b\in R\setminus\{0\}$. As $f(X)g(X)=1$ we get that $f(X)$ is a unit in R.

($\Leftarrow$) If $f(X)=b$, b a unit in R. Take $g(X)=b^{-1}$. Then $f(X)g(X)=1$ proving that $f(X)$ is a unit in $R[X]$. □

Example 3.22. *Be careful if* R *is not an integral domain:*

In $\mathbb{Z}_4$, $2X+1$ *is a unit, as*

$$(2X+1)(2X+1)=4X^2+4X+1=1,$$

but $2X+1$ *is not a constant polynomial.*

Corollary 3.23. *Let F be a field. Then $f(X)$ is a unit in $F[X]$ if and only if $f(X)$ is a non-zero constant polynomial.*

Definition 3.24. *An element a in a commutative ring R with identity is an associate of an element $b \in R$ if $a = bu$ for some unit $u \in R$.*

Remark 3.25.

1. *If a is an associate of b, then b is an associate of a (as $a = bu$ implies $b = au^{-1}$ and u^{-1} is a unit).*

2. *In $\mathbb{Z}$, the only associates of a are a and $-a$.*

3. *By the previous Corollary, $f(X)$ is an associate of $g(X)$ in $F[X]$ if and only if $f(X) = cg(X)$ for some $c \in F \setminus \{0\}$.*

We are interested in developping a notion similar to prime numbers, but in $F[X]$. Recall:

In $\mathbb{Z}$, p is prime if $p \neq \pm 1$ (p is not a unit) and the only divisors of p are ± 1 (units) and $\pm p$ (associates of p).

Definition 3.26. *Let F be a field. A non-constant polynomial $p(X) \in F[X]$ is irreducible if its only divisors are its associates and the non-zero constant polynomials (units). (In other words: If $p(X) = r(X)s(X)$ then one of $r(X)$ and $s(X)$ is a unit.)*

A non-constant polynomial that is not irreducible is called reducible.

Example 3.27.

1. *$X - 1$ is irreducible in $\mathbb{Q}[X]$.*

2. *$X^2 - 3$ is irreducible in $\mathbb{Q}[X]$ (why? We will come back to this).*

3. *$X^2 + 1$ is reducible in $\mathbb{Z}_5[X]$, as $X^2 + 1 = (X + 2)(X + 3)$.*

Theorem 3.28. *Let F be a field and let $f(X) \in F[X] \setminus \{0\}$. Then $f(X)$ is reducible in $F[X]$ if and only if $f(X)$ can be written as the product of two polynomials in $F[X]$ of (strictly) lower degree.*

Proof. ($\Rightarrow$) There exist $g(X) \in F[X]$ that is not an associate of $f(X)$ or a non-zero constant such that $f(X) = g(X)h(X)$ for some $h(X) \in F[X]$.

If either $\deg h(X) = \deg f(X)$ or $\deg g(X) = \deg f(X)$, then $\deg g(X) = 0$ or $\deg h(X) = 0$. So either $g(X) = c$ for some $c \in F$ (contradiction), or $h(X) = b$ for some $b \in F$. This gives that $f(X) = g(X)b$, so $g(X)$ is an associate of $f(X)$, contradiction.

Therefore $\deg g(X), \deg h(X) < \deg f(X)$.

($\Leftarrow$) Suppose $f(X) = g(X)h(X)$ where $\deg g(X), \deg h(X) < \deg f(X)$. Then $g(X)|f(X)$ and $g(X)$ is not an associate of $f(X)$ (associates have the same degree), and $g(X)$ is not a unit (if $g(X)$ is a unit, then $\deg g(X) = 0$, so $\deg h(X) = \deg f(X)$, contradiction). Thus f is reducible. □

3.4 Roots and Reducibility

Question: How can we determine if a given polynomial is irreducible in a polynomial ring $F[X]$? (We often say "over F".)

Definition 3.29. *Let R be a commutative ring and $f(X) \in R[X]$. An element a of R is a root of $f(X)$ is $f(a) = 0_R$.*

Example 3.30.

- *$f(X) = X^2 - X \in \mathbb{Z}_2[X]$ has roots 0 and 1, as $f(0) = 0^2 - 0 = 0$ and $f(1) = 1^2 - 1 = 0$.*
- *$X^4 + 2X^2 + 1$ has no root in $\mathbb{Q}$ since it is equal to $(x^2+1)^2$, which only takes positive values.*

Theorem 3.31 (Remainder Theorem)**.** *Let F be a field, $f(X) \in F[X]$ and $a \in F$. Then $f(a)$ is the remainder in the division of $f(X)$ by $X - a$.*

Proof. By the Division Algorithm, $f(X) = (X - a)q(X) + r(X)$, where $\deg r(X) < \deg(X - a) = 1$ or $r(X) = 0$. Thus, $\deg r(X) = 0$ or $r(X) = 0$. In either case, $r(X) = c$ for some $c \in F$. Hence $f(X) = (X - a)q(X) + c$, so $f(a) = (a - a)q(a) + c = c$. □

Theorem 3.32 (Factor Theorem)**.** *Let F be a field, $f(X) \in F[X]$ and $a \in F$. Then a is a root of $f(X)$ if and only if $X - a$ is a factor of $f(X)$ in $F[X]$.*

Proof. $X - a$ is a factor of $f(X)$ if and only if the remainder in the division of $f(X)$ by $X - a$ is 0, if and only if $f(a) = 0$ (by the previous Theorem), if and only if a is a root of $f(X)$. □

Example 3.33. *$X^4 + X^2 + 1 \in \mathbb{Z}_3[X]$ has 1 as a root: $1^4 + 1^2 + 1 = 3 = 0$. Thus, $X - 1 (= X + 2)$ is a factor of $X^4 + X^2 + 1$ in $\mathbb{Z}_3[X]$.*

Corollary 3.34. *Let F be a field and $f(X)$ a non-zero polynomial of degree n in $F[X]$. Then $f(X)$ has at most n roots in F.*

Proof. Suppose by contradiction that $f(X)$ has k roots $a_1, \ldots, a_k$ with $k > n$. Therefore $f(X) = (X - a_1)f_1(X)$, and $a_2, \ldots, a_k$ are roots of $f_1(X)$. Then $f_1(X) = (X - a_2)f_2(X)$, $a_3, \ldots, a_k$ are roots of $f_2(X)$. Observe that $f(X) = (X - a_1)(X - a_2)f_2(X)$. And so on (induction, really).

We obtain: $f(X) = (X - a_1)\cdots(X - a_k)f_k(X)$ Therefore $\deg f(X) \geq k > n$, contradiction. □

This result is not necessarily true is F is not a field. Think about $\mathbb{Z}_6[X]$.

Corollary 3.35. *Let F be a field and $f(X) \in F[X]$ with $\deg f(X) \geq 2$. If $f(X)$ is irreducible in $F[X]$ then $f(X)$ has no roots in F.*

Remark 3.36. 1. *The converse is false in general if* $\deg f(X) > 3$*. Consider for instance* $X^4 + 2X^2 + 1 = (X^2 + 1)^2 \in \mathbb{Q}[X]$*. It is reducible but has no roots in* $\mathbb{Q}$*.*

2. *If* $\deg f(X) = 2$ *or* 3*, then the converse is true.*

Exercise 3.37. *Prove that* $p(X)$ *is irreducible is equivalent to:*
If $p(X) = r(X)s(X)$ *then* $r(X)$ *or* $s(X)$ *is a nonzero constant polynomial.*

Corollary 3.38. *Let* F *be a field and let* $f(X) \in F[X]$ *be a polynomial of degree* 2 *or* 3*. Then* $f(X)$ *is irreducible in* $F[X]$ *if and only if* $f(X)$ *has no roots in* F*.*

Proof. ($\Rightarrow$) See previous Corollary.

($\Leftarrow$) We first observe that no polynomial of degee 1 is a factor of $f(X)$. Why Not? If $f(X) = (cX + d)g(X)$ (with $c \in F \setminus \{0\}$), then $f(-c^{-1}d) = 0$, impossible.

Thus, if $f(X) = r(X)s(X)$, then $\deg r(X), \deg s(X) \neq 1$, and so

$$2 \text{ or } 3 = \deg f(X) = \underbrace{\deg r(X)}_{\neq 1} + \underbrace{\deg s(X)}_{\neq 1}.$$

Therefore, either $\deg r(X) = 0$ or $\deg s(X) = 0$, i.e., $r(X)$ or $s(X)$ is a nonzero constant. So $f(X)$ is irreducible by the previous exercise. □

Example 3.39. 1. X^2+1 *is irreducible in* $\mathbb{Z}_3$*, as* $0^2+1 = 1 \neq 0$*,* $1^2+1 = 2 \neq 0$*,* $2^2 + 1 = 2 \neq 0$*.*

2. $X^2 + 1$ *is reducible in* $\mathbb{Z}_5$ *as* $2^2 + 1 = 5 = 0$*, thus* 2 *is a root.*

Observe that for $f(X) \in \mathbb{Z}_p[X]$, p prime (so that $\mathbb{Z}_p$ is a field), and $\deg f(X) = 2$ or 3, we can check irreducibility of $f(X)$ by checking that $f(X) \neq 0$ for $X = 0, 1, \ldots, p-1$.

3.5 Irreducibility in $\mathbb{Q}$

Idea: Consider $f(X) \in \mathbb{Q}[X]$. How can we determine if $f(X)$ is irreducible in $\mathbb{Q}[X]$? Factoring in $\mathbb{Q}[X]$ can be reduced to factoring in $\mathbb{Z}[X]$.

Remark 3.40. 1. *If* $f(X) \in \mathbb{Q}[X]$*, there there exists* $c \in \mathbb{Z} \setminus \{0\}$ *such that* $cf(X) \in \mathbb{Z}[X]$*.*

2. $f(X)$ *and* $cf(X)$ *have the same roots, and (clearly)* $f(X)$ *is reducible in* $\mathbb{Q}[X]$ *if and only if* $cf(X)$ *is reducible in* $\mathbb{Q}[X]$*. Thus, we consider polynomials with integer coefficients.*

Recall that a is a root of $f(X)$ if and only if $X - a$ is a factor of $f(X)$. So, to find linear factors, we search for roots. How can we find roots?

Theorem 3.41 (Rational Root Test). *Let* $f(X) = a_nX^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0 \in \mathbb{Z}[X]$*. If* $r \neq 0$ *and the rational number* $\frac{r}{s}$, $\gcd(r, s) = 1$ *is a root of* $f(X)$*, then* $r|a_0$ *and* $s|a_n$*.*

Proof. As $\frac{r}{s}$ is a root of $f(X)$, we have

$$a_n\frac{r^n}{s^n} + a_{n-1}\frac{r^{n-1}}{s^{n-1}} + \cdots + a_1\frac{r}{s} + a_0 = 0.$$

Multiply both sides by s^n and rearrange to get

$$\begin{aligned} a_0s^n &= -(a_nr^n + a_{n-1}sr^{n-1} + \cdots + a_1s^{n-1}r \\ &= -r(a_nr^{n-1} + a_{n-1}sr^{n-2} + \cdots + a_1s^{n-1}). \end{aligned}$$

So, $r|a_0s^n$. As $\gcd(r, s) = 1$, then $\gcd(r, s^n) = 1$ and thus $r|a_0$. A similar argument implies that $s|a_n$. □

Example 3.42. $f(X) = 2X^4 - 5X^3 + 3X^2 + 4X - 6 \in \mathbb{Q}[X]$*. The possible roots in* $\mathbb{Q}$ *are given by*

$$r = \pm1, \pm2, \pm3, \pm6 \text{ and } s = \pm1, \pm2.$$

We check $\frac{r}{s} = \pm1, \pm2, \pm3, \pm6, \pm\frac{1}{2}, \pm\frac{3}{2}$ *to see which are roots, and get that the roots are* $-1, \frac{3}{2}$*.*

So $f(X) = (X + 1)(2X - 3)q(X)$*, where* $q(X)$ *has degree* 2*. We find* $q(X) = X^2 - 2X + 2$*, which is then irreducible in* $\mathbb{Q}[X]$ *(we can also check with the quadratic formula to make sure that the roots of* $q(X)$ *are not in* $\mathbb{Q}$*).*

Let us insist on the following:

$$a \text{ is a root of } f(X) \Leftrightarrow X - a \text{ is a factor of } f(X).$$

So:

$f(X)$ has no roots in $F \Leftrightarrow f(X)$ has no linear factors with coefficients in F.

But this does not necessarily imply that $f(X)$ is irreducible in $F[X]$ (it is only the case when $\deg f = 2$ or 3).

Example 3.43. *$f(X) = X^4 + X^3 + X + 2 \in \mathbb{Z}_3[X]$ has no roots in $\mathbb{Z}_3$ but $f(X) = (X^2 + 1)(X^2 + X + 2) \in \mathbb{Z}_3[X]$.*

Definition 3.44. *The content of a non-zero polynomial $f(X) = a_nX^n + \cdots + a_0 \in \mathbb{Z}[X]$ is $\gcd(a_n, a_{n-1}, \ldots, a_0)$. A primitive polynomial in $\mathbb{Z}[X]$ is a polynomial with content 1.*

Theorem 3.45 (Gauss' Lemma). *The product of two primitive polynomials is primitive.*

Proof. Let $f(X), g(X)$ be primitive. Suppose $f(X)g(X)$ is not primitive, and let p be a prime divisor of the content of $f(X)g(X)$. Let $\bar{f}(X), \bar{g}(X)$ and $\overline{f(X)g(X)}$ be the polynomials obtained from f, g and fg by reducing the coefficients modulo p.

Then $\bar{f}, \bar{g} \in \mathbb{Z}_p[X]$ and $\bar{f}(X)\bar{g}(X) = \overline{f(X)g(X)} = 0$ in $\mathbb{Z}_p[X]$ (check this). As $\mathbb{Z}_p[X]$ is an integral domain, $\bar{f}(X) = 0$ or $\bar{g}(X) = 0$. Thus, either p divides every coefficient of $f(X)$ or p divides every coefficient of $g(X)$. Then either f or g is not primitive, contradiction. □

Theorem 3.46 (Reducibility over $\mathbb{Q}$ $\Leftrightarrow$ Reducibility over $\mathbb{Z}$). *Let $f(X) \in \mathbb{Z}[X]$. If $f(X)$ is reducible in $\mathbb{Q}[X]$ then $f(X)$ is reducible in $\mathbb{Z}[X]$.*

Proof. Suppose $f(X) = g(X)h(X)$ where $g(X), h(X) \in \mathbb{Q}[X]$. We can assume that $f(X)$ is primitive (dividing f by its content if necessary). Let a be the least common multiple of the denominators of the coefficients of $g(X)$, and b be the least common multiple of the denominators of the coefficients of $h(X)$. Then

$$abf(X) = ag(X) \cdot bh(X),$$

where $ag(X), bh(X) \in \mathbb{Z}[X]$. Let c_1 be the content of $ag(X)$ and c_2 be the content of $bh(X)$. Then $ag(X) = c_1g_1(X)$ and $bh(X) = c_2h_1(X)$, where both g_1 and h_1 are primitive. Thus

$$abf(X) = c_1c_2g_1(X)h_1(X).$$

Since f is primitive, the content of $abf(X)$ is ab. By Gauss' Lemma, g_1h_1 is primitive and so the content of $c_1c_2g_1(X)h_1(X)$ is c_1c_2. Thus $ab = c_1c_2$ and $f(X) = g_1(X)h_1(X)$ where $g_1(X), h_1(X) \in \mathbb{Z}[X]$. □

We look now at more general results on irreducibility.

Theorem 3.47 (Eisenstein's Criterion). *Let $f(X) = a_nX^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0$ be a non-constant polynomial in $\mathbb{Z}[X]$. If there is a prime p such that*

- *$p|a_i$ for $i = 0, 1, \dots, n-1$;*
- *$p \nmid a_n$;*
- *$p^2 \nmid a_0$,*

then $f(X)$ is irreducible in $\mathbb{Q}[X]$.

Proof. We procede by contradiction. Suppose that $f(X)$ is reducible in $\mathbb{Q}[X]$. Then by the previous theorem, $f(X)$ is reducible in $\mathbb{Z}[X]$, say

$$f(X) = (b_0 + b_1X + \cdots + b_rX^r)(c_0 + c_1X + \cdots + c_sX^s),$$

where $b_i, c_j \in \mathbb{Z}$, $r \geq 1$, $s \geq 1$. Note that $a_0 = b_0c_0$.

As $p|a_0$, then $p|b_0$ or $p|c_0$. Without loss of generality we assume that $p|b_0$. Since $p^2 \nmid a_0$, we have $p \nmid c_0$ (otherwise $p|c_0$, so $p^2|a_0$, contradiction). We also have $a_n = b_rc_s$. Thus $p \nmid b_r$ (since otherwise $p|a_n$, contradiction). There may exist another i such that $p \nmid b_i$. Let b_k be the first of the b_i's such that $p \nmid b_k$.

Then $0 < k \leq r < n$ and $p|b_i$ for $i < k$, and $p \nmid b_k$. Since

$$a_k = b_0c_k + b_1c_{k-1} + \cdots + b_{k-1}c_1 + b_kc_0,$$

we have

$$b_kc_0 = a_k - b_0c_k - b_1c_{k-1} - \cdots - b_{k-1}c_1.$$

As $p|a_k$ and $p|b_i$ for $i < k$, then p divides the right hand side. Thus, $p|b_kc_0$ and so $p|b_k$ or $p|c_0$, a contradiction in either case. Thus $f(X)$ is irreducible in $\mathbb{Q}[X]$. □

Example 3.48. 1. *$f(X) = 3X^5 + 15X^4 - 20X^2 + 10X + 20$ is irreducible in $\mathbb{Q}[X]$:*

By Eisenstein's criterion with $p = 5$: $5 \nmid 3$, $25 \nmid 20$, but $5|15, 20, 10, 20$.

2. *Is $f(X) = 21X^3 - 3X^2 + 2X + 9$ irreducible in $\mathbb{Q}[X]$?*

Use Eisenstein? $p = 3$ does not work ($p \nmid 2$) and no other prime seems to work.

Now what?

Notation: $\overline{f(X)}$ means reduce the coefficients of $f(X)$ modulo p.

Theorem 3.49 (Mod p Test). *Let $f(X) = a_kX^k + \cdots + a_1X + a_0 \in \mathbb{Z}[X]$ and let p be a prime such that $p \nmid a_k$. If $\overline{f(X)}$ is irreducible in $\mathbb{Z}_p[X]$, then $f(X)$ is irreducible in $\mathbb{Q}[X]$.*

Proof. We use contradiction. Suppose f is reducible in $\mathbb{Q}[X]$. By Theorem 3.46, $f(X) = g(X)h(X)$ with $g(X), h(X)$ non-constant polynomials in $\mathbb{Z}[X]$. Since $p \not| a_k$, p does not divide the leading coefficients of $g(X)$ and $h(X)$. Thus, $\deg g(X) = \deg \overline{g(X)}$ and $\deg h(X) = \deg \overline{h(X)}$. So neither $\overline{g(X)}$ not $\overline{h(X)}$ are constant polynomials in $\mathbb{Z}_p[X]$.

Check that $f(X) = g(X)h(X)$ in $\mathbb{Z}[X]$ implies $\overline{f(X)} = \overline{g(X)h(X)}$ in $\mathbb{Z}_p[X]$.

Thus, $\overline{f(X)}$ is reducible in $\mathbb{Z}_p[X]$, a contradiction. □

Example 3.50. *Is $f(X) = 21X^3 - 3X^2 + 2X + 9$ irreducible in $\mathbb{Q}[X]$?*

Take $p = 2$. Then $\overline{f(X)} = X^3 + X^2 + 1$ in $\mathbb{Z}_2[X]$. Is $\overline{f(X)}$ irreducible in $\mathbb{Z}_2[X]$? As $\deg \overline{f(X)} = 3$, we need only check roots in $\mathbb{Z}_2$.

As $\bar{f}(0) = 1$ and $\bar{f}(1) = 1$, $\overline{f(X)}$ is irreducible in $\mathbb{Z}_2[X]$. By the "Mod p test", f is irreducible in $\mathbb{Q}[X]$.

Remark 3.51. *1. Given $f(X) \in \mathbb{Z}[X]$, if $\overline{f(X)}$ is reducible in $\mathbb{Z}_p[X]$, then we have nothing (try another p).*

2. What if $\deg \overline{f(X)} > 3$? How do we show that $\overline{f(X)}$ is irreducible in $\mathbb{Z}_p[X]$?

Example 3.52. *Is $f(X) = X^5 + 4X^4 + 2X^3 + 3X^2 - X + 5$ irreducible in $\mathbb{Q}[X]$?*

Take $p = 3$. Then $\overline{f(X)} = X^5 + X^4 + 2X^3 + 2X + 2$ in $\mathbb{Z}_3[X]$. If $\overline{f(X)}$ irreducible in $\mathbb{Z}_3[X]$?

We check for roots in $\mathbb{Z}_3$: $\bar{f}(0) = 2$, $\bar{f}(1) = 2$, $\bar{f}(2) = 1$. There are no roots, i.e., no linear factors. So if $\overline{f(X)}$ is reducible, it must be a product of two monic irreducible polynomials of degrees 2 and 3 in $\mathbb{Z}_3[X]$.

We list all the monic irreducible polynomials of degree 2 in $\mathbb{Z}_3[X]$ (check this):

$$X^2 + 1,\ X^2 + X + 2,\ X^2 + 2X + 2$$

Then we check that each one is not a factor of $\overline{f(X)}$ (how? Use the division algorithm and check that the remainder is not 0).

Hence, $\overline{f(X)}$ is irreducible in $\mathbb{Z}_3[X]$. By the "Mod 3 Test", $f(X)$ is irreducible in $\mathbb{Q}[X]$.

Idea: We can now connect irreducible polynomials to ideals.

Definition 3.53. *A principal ideal domain (PID) is an integral domain R in which every ideal is principal (i.e., is generated by one element, cf. Remark 2.34).*

Example 3.54. 1. *$\mathbb{Z}$ is a PID. Quick proof (check the details): Let I be an ideal of $\mathbb{Z}$ and let a be the smallest positive elements of I. Recall that $\langle a\rangle = \{na : n \in \mathbb{Z}\}$. Clearly $\langle a\rangle \subseteq I$. To show that every element of I is in $\langle a\rangle$, take an element $t \in I$ and consider the division of t by a, $t = aq + r$. Show that $r \in I$ and that it follows that we must have $r = 0$.*

2. *Let F be a field. Prove that $F[X]$ is a PID.*

Hint: Let I be an ideal of $F[X]$. If $I = F[X]$, then $I = \langle 1\rangle$, so we assume that $I \neq F[X]$. Let $f(X)$ be a non-constant monic polynomial of minimal degree in I. Show that $I = \langle f(X)\rangle$ (use the Division Algorithm, and the same strategy as in the case of $\mathbb{Z}$ above).

Observe that if $c \in F \setminus \{0\}$, then

$$\langle f(X)\rangle = \langle cf(X)\rangle,$$

in particular, we can always take $f(X)$ monic (multiply it by the inverse of its leading coefficient to make it monic).

Theorem 3.55. *Let F be a field and let $p(X) \in F[X]$. Then $\langle p(X)\rangle$ is a maximal ideal in $F[X]$ if and only if $p(X)$ is irreducible in $F[X]$.*

Proof. ($\Rightarrow$) Note that $p(X) \neq 0$ and $p(X)$ is not a unit in $F[X]$ (since $\langle 0\rangle = \{0\}$ in the first case, and $\langle p(X)\rangle = F[X]$ in the second case, which are not maximal ideals in $F[X]$).

If $p(X) = g(X)h(X)$, then

$$\langle p(X)\rangle \subseteq \langle g(X)\rangle \subseteq F[X].$$

Thus (since $\langle p(X)\rangle$ is a maximal ideal), $\langle p(X)\rangle = \langle g(X)\rangle$ or $\langle g(X)\rangle = F[X]$. In the first case, we have $\deg p(X) = \deg g(X)$ (why?). In the second case, it follows that $\deg g(X) = 0$ (why?) and so $\deg h(X) = \deg p(X)$. Thus, $p(X)$ cannot be written as a product of two polynomials in $F[X]$ of lower degree, i.e., $p(X)$ is irreducible in $F[X]$.

($\Leftarrow$) Let I be an ideal of $F[X]$ such that

$$\langle p(X)\rangle \subseteq I \subseteq F[X].$$

As $F[X]$ is a PID (see exercise above), $I = \langle g(X)\rangle$ for some $g(X) \in F[X]$. So $p(X) \in \langle g(X)\rangle$ and thus $p(X) = g(X)h(X)$ for some $h(X) \in F[X]$. As $p(X)$ is irreducible, either $g(X)$ or $h(X)$ is a constant polynomial. In the first case, we have $I = F[X]$. In the second case, we have $\langle p(X)\rangle = \langle g(X)\rangle$. Thus $\langle p(X)\rangle$ is maximal. □

Corollary 3.56. *If $p(X)$ is irreducible in $F[X]$, then $F[X]/\langle p(X)\rangle$ is a field. Moreover, it contains a subfield $F^\star$ that is isomorphic to F.*

Proof. By the previous Theorem, $\langle p(X)\rangle$ is maximal. Thus, by Theorem 2.48, $F[X]/\langle p(X)\rangle$ is a field. Consider

$$F^\star := \{a + \langle p(X)\rangle : a \in F\}.$$

Check that $F^\star$ is a subring of $F[X]/\langle p(X)\rangle$ that is a field (hint: the inverse of $a + \langle p(X)\rangle$ is $a^{-1} + \langle p(X)\rangle$ for $a + \langle p(X)\rangle \neq 0$).

Claim: F and $F^\star$ are isomorphic.

Why? Define $\Phi : F \to F^\star$ by $\Phi(a) = a + \langle p(X)\rangle$. We have

$$\begin{aligned}\Phi(a+b) &= a + b + \langle p(X)\rangle \\ &= (a + \langle p(X)\rangle) + (b + \langle p(X)\rangle) \\ &= \Phi(a) + \Phi(b).\end{aligned}$$

Similarly (do it!) $\Phi(a \cdot b) = \Phi(a) \cdot \Phi(b)$. thus, Φ is a homomorphism. It is surjective by definition, so we only have to show that it is injective. Suppose $\Phi(a) = \Phi(b)$, i.e., $a + \langle p(X)\rangle = b + \langle p(X)\rangle$, so $a - b \in \langle p(X)\rangle$. Therefore $p(X)|a - b$. As $\deg p(X) \geq 1$ and $a - b \in F$, we must have $a - b = 0$, so $a = b$. □

Remark 3.57. *1. We can now construct finite fields:*

Let p be a prime number, and let $f(X)$ be irreducible in $\mathbb{Z}_p[X]$, $\deg f(x) = n$.

Then $\mathbb{Z}_p[X]/\langle f(X)\rangle$ is a field with p^n elements.

We know that it is a field since $\langle f(X)\rangle$ is a maximal ideal in $F[X]$. But why does it have p^n elements?

$$\mathbb{Z}_p[X]/\langle f(X)\rangle = \{g(X) + \langle f(X)\rangle : g(X) \in \mathbb{Z}_p[X]\}.$$

By the Division Algorithm, $g(X) = f(X)q(X) + r(X)$, where $r(X) = 0$ or $0 < \deg r(X) < \deg f(X) = n$. And

$$g(X) - r(X) = f(X)q(X) \in \langle f(X)\rangle.$$

Thus

$$g(X) + \langle f(X)\rangle = r(X) + \langle f(X)\rangle,$$

and

$$\begin{aligned}\mathbb{Z}_p[X]/\langle f(X)\rangle &= \{r(X) + \langle f(X)\rangle : r(X) \in F[X],\ \deg r(X) < n\} \\ &= \{a_{n-1}X^{n-1} + \cdots + a_1X + a_0 + \langle f(X)\rangle : a_i \in \mathbb{Z}_p\}.\end{aligned}$$

There are p choices for each of the a_i, so altogether p^n elements. Provided that they are all different, which we now check:

Assume that $(a_0, \ldots, a_{n-1}) \neq (b_0, \ldots, b_{n-1})$, i.e., the polynomials $a(X) := a_0 + a_1X + \cdots + a_{n-1}X^{n-1}$ and $b(X) := b_0 + b_1X + \cdots + b_{n-1}X^{n-1}$ are different in $\mathbb{Z}_p[X]$. We show that $a(X)+\langle f(X)\rangle \neq b(X)+\langle f(X)\rangle$: If we had equality, then $a(X) - b(X) \in \langle p(X)\rangle$, so $p(X)$ divides $a(X) - b(X)$ which has degree less than n. This is only possible if $a(X) - b(X) = 0$, i.e., $a(X) = b(X)$, a contradiction.

2. *By the previous result, we have*

$$F \cong F^\star \subseteq F[X]/\langle p(X)\rangle.$$

So we may view F has a subfield of $F[X]/\langle p(x)\rangle$. We say that $F[X]/\langle p(X)\rangle$ and an "extension field" of F.

Chapter 4

Field Theory (a brief glimpse. . .)

Overview: We saw a glimpse of “field extensions”, namely

$$\begin{tikzcd} K \arrow[d, dash] & := & F[X]/\langle p(X)\rangle,\ p(X) \text{ irreducible in F[X]} \\ F & & \end{tikzcd}$$

(where the field that is “up” is larger than the field “below”)

Can we get a general picture?

Recall the definition:

Definition 4.1. *If F and K are fields and F is a subring of K that contains 1_K, we say that K is an extension field of F (or a field extension of F, or that F is a subfield of K).*

Remark 4.2. *(Very important) In this case, K is a vector space over F: We have addition of elements of K and multiplication by elements of F. The properties of vector space are satisfied because K is a ring.*

Definition 4.3. *Let K be an extension field of F. We denote by $[K : F]$ the number of elements of any basis of K over F (the dimension of K over F), and call it the degree of K over F.*

Proposition 4.4. *Let K be a field and let L_i, $i \in I$, be subfields of K. Then $\bigcap_{i\in I} L_i$ is a subfield of K.*

Proof. We first show that $L := \bigcap_{i\in I} L_i$ is a subring of K by using the subring test (Exercise 2.18): Let $a, b \in L$. Then $a, b \in L_i$ for every i, and since L_i

is a subfield, we get $a - b, a \cdot b \in L_i$. Thus $a - b, a \cdot b \in L$. Clearly L is commutative, since K is.

To check that L is a subfield, we now have to check that if $a \in L \setminus \{0\}$, then a has an inverse in L. Since K is a field, a has an inverse a^{-1} in K. Since $a \in L_i$ for every i, a has an inverse in L_i, and this inverse must be equal to a^{-1} (do it). So $a^{-1} \in L_i$ for every i, which gives $a^{-1} \in L$. □

Corollary 4.5. *Let K be a field and let S be a subset of K. Then there is a smallest subfield L of K containing S (in the sense that if N is a subfield of K containing S, then $L \subseteq N$).*

Proof. By Proposition 4.4, any intersection of subfields of K is a subfield of K. Take for L the intersection of all subfields of K containing S. It is necessarilly the smallest subfield of K containing S. □

Definition 4.6. *Let K be an extension field of F and $u \in K$. Then $F(u)$ ("F adjoin u") is the smallest subfield of K containing F and u and is called a simple extension of F.*

We can give an explicit description of $F(u)$:

Proposition 4.7. *With notation as in the previous definition,*

$$F(u) = \left\{ \frac{f(u)}{g(u)} : f(X), g(X) \in F[X],\ g(u) \neq 0 \right\}.$$

Proof. We check both inclusions.

($\supseteq$) Let $f(X) = a_0 + \cdots + a_n X^n \in F[X]$. Since $F(u)$ is a field containing F and u, we have $a_i u^i \in F[u]$ for every i. Therefore $f(u) \in F(u)$. Similarly $g(u) \in F(u)$ and since $g(u) \neq 0$ and $F(u)$ is a field, $g(u)^{-1} \in F(u)$. Thus $\dfrac{f(u)}{g(u)} \in F(u)$.

($\subseteq$) Let

$$N := \left\{ \frac{f(u)}{g(u)} : f(X), g(X) \in F[X],\ g(u) \neq 0 \right\}.$$

It is easy to check that N is a subfield of K that contains F and u. Therefore, by definition of $F(u)$, we have $F(u) \subseteq N$. □

In picture:

$$\begin{tikzcd} K \arrow[d, dash] \\ F(u) \arrow[d, dash] \\ F \end{tikzcd}$$

Example:

$$\begin{tikzcd} \mathbb{C} \arrow[d, dash] \\ \mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\} \arrow[d, dash] \\ \mathbb{Q} \end{tikzcd}$$

Observe that the description of $\mathbb{Q}(\sqrt{2})$ is simpler than what is given in Proposition 4.7. We will come back to this.

Definition 4.8. *Let K be an extension field of F and let $u \in K$. We say that u is algebraic over F if u is the root of a polynomial in $F[X]$. If u is not algebraic over F, we say that u is transcendental over F.*

Example 4.9. *1. $\sqrt{2}$ is algebraic over $\mathbb{Q}$, as it is a root of $X^2 - 2 \in \mathbb{Q}[X]$.*

2. $\sqrt{-3} + \sqrt{2}$ is algebraic over $\mathbb{Q}$.

Why? Let $\alpha = \sqrt{-3} + \sqrt{2}$. Then $\alpha^2 = -3 + 2\sqrt{-6} + 2$. So $\alpha^2 + 1 = 2\sqrt{-6}$ and taking squares we get $\alpha^4 + 2\alpha^2 + 1 = 4(-6)$, so $\alpha^4 + 2\alpha^2 + 25 = 0$. So, α is a root of $X^4 + 2X^2 + 25$.

3. π is transcendental over $\mathbb{Q}$ (Lindemann, 1882).

e is transcendental over $\mathbb{Q}$ (Hermite, 1873).

Theorem 4.10. *Let K be an extension of F and let $u \in K$ be algebraic over F. There there exists a unique monic irreducible polynomial $m_{u/F}(X) \in F[X]$ that has u as a root.*

This polynomial is called the minimal polynomial of u over F, and satisfies

$$\langle m_{u/F} \rangle = \{f(X) \in F[X] : f(u) = 0\}.$$

Proof. Since u is algebraic over F, the set

$$I := \{f(X) \in F[X] : f(u) = 0\}$$

is a non-zero ideal of $F[X]$. As seen in Example 3.54, $I = \langle p(X) \rangle$ with $p(X)$ polynomial of minimal degree in I, and we can choose $p(X)$ monic. So $m_{u/F}(X) := p(X)$ is a monic polynomial of minimal degree that has u as root and it satisfies the final statement.

We check that $m_{u/F}(X)$ is irreducible: Supppose that $m_{u/X}(X) = f(X)g(X)$ with $f(X)$ and $g(X)$ of degree less than $m_{u/F}(X)$. Since $m_{u/F}(u) = 0$, we get $f(u) = 0$ or $g(u) = 0$, impossible since they both have degree less than $m_{u/F}(X)$.

We check that $m_{u/F}(X)$ is unique: Let $t(X)$ be another monic irreducible polynomial such that $t(u) = 0$. By definition of $m_{u/F(X)}$ we have $\deg m_{u/F}(X) \leq \deg t(X)$. We divide $t(X)$ by $m_{u/F}(X)$: $t(X) = m_{u/F}(X)q(X) + r(X)$ with $r(X) = 0$ or $\deg r(X) < \deg m_{u/F}(X)$. Replacing X by u we get $r(u) = 0$, which means that we must have $r(X) = 0$ (otherwise we would get a monic polynomial with u as root, of degree less than $m_{u/X}$). So $m_{u/F}(X)$ divides $t(X)$, and since $t(X)$ is irreducible, $q(X)$ is a constant. Since they are both monic, we must have $q(X) = 1$, so $t(X) = m_{u/F}(X)$. □

Example 4.11. *The minimal polynomial of $\alpha := \sqrt{-3} + \sqrt{2}$ over $\mathbb{Q}$ is $f(X) := X^4 + 2X^2 + 25$.*

Why? We already checked that α is a root of $f(X)$. We check that $f(X)$ is irreducible in $\mathbb{Q}[X]$. We use the mod 3 test ($p = 3$). Then $\overline{f(X)} = X^4 + 2X^2 + 1$ has no roots in $\mathbb{Z}_3$, so $\overline{f(X)}$ has no linear factors. We check all the monic quadratic irreducible polynomials in $\mathbb{Z}_3[X]$ (how many are there? See exercise set 5) using the Division Algorithm, and we see that none of them divides $\overline{f(X)}$. So by the Mod 4 Test, $f(X)$ is irreducible in $\mathbb{Q}[X]$.

Final question: What is the connection between $F(u)$, u algebraic, the minimal polynomial $m_{u/F}(X)$ of u over F, and $F[X]/\langle m_{u/F}(u)\rangle$?

Theorem 4.12. *Let K be an extension field of F and let $u \in K$. The following are equivalent.*

1. *u is algebraic over F;*
2. *$F[X]/\langle m_{u/F}(X)\rangle \cong F[u] := \{f(u) : f(X) \in F[X]\}$;*
3. *$F(u) = F[u]$;*
4. *$[F(u) : F]$ is finite.*

In this case, and if $\deg m_{u/F}(X) = n$, then $\{1_F, u, u^2, \dots, u^{n-1}\}$ is a basis of $F(u)$ over F. In particular $[F(u) : F] = \deg m_{u/F}$ and

$$F(u) = \{a_0 + a_1 u + \cdots + a_{n-1}u^{n-1} : a_i \in F\}.$$

Proof. (1)⇒(2): Consider the map

$$\lambda : F[X] \to F[u],\ \lambda(f(X)) = f(u)$$

(the evaluation at u). A direct computation shows that $(f(X)+g(X))(u) = f(u)+g(u)$ and $(f(X)g(X))(u) = f(u)g(u)$, i.e., $\lambda(f(X)+g(X)) = \lambda(f(X))+\lambda(g(X))$ and $\lambda(f(X)g(X)) = \lambda(f(X))\lambda(g(X))$, so that λ is a ring homomorphism.

The map λ is clearly surjective, so by the first isomorphism theorem, we have

$$F[X]/\operatorname{Ker}\lambda \cong F[u],$$

and the result follows since by Theorem 4.10

$$\operatorname{Ker}\lambda = \{f \in F[X] : f(u) = 0\} = \langle m_{u/F}(X)\rangle.$$

(2)⇒(3): Clearly, $F[u] \subseteq F(u)$, and we know that $F(u)$ is the smallest subfield of K containing F and u. But $F[X]/\langle m_{u/F}(X)\rangle$ is a field since

$m_{u/F}(X)$ is irreducible. So $F[u]$ is also a field. Since it contains F and u, it cannot be smaller than $F(u)$, so we have $F[u] = F(u)$.

(3)⇒(4): It suffices to show that $[F[u] : F]$ is finite, i.e., that $\dim_F F[u]$ is finite. By definition

$$\begin{aligned} F[u] &= \{f(u) : f \in F[X]\} \\ &= \operatorname{Span}_F\{1, u, u^2, \ldots\}, \end{aligned}$$

and, writing $m_{u/F}(X) = a_0 + a_1X + \cdots + a_{n-1}X^{n-1} + X^n$, we have

$$a_0 + a_1u + \cdots + a_{n-1}u^{n-1} + u^n = 0.$$

Therefore:

$$\begin{aligned} &u^n \in \operatorname{Span}_F\{1, u, \ldots, u^{n-1}\}, \\ &u^{n+1} \in \operatorname{Span}_F\{1, u, \ldots, u^{n-1}, u^n\} = \operatorname{Span}_F\{1, u, \ldots, u^{n-1}\}, \\ &\cdots \\ &u^i \in \operatorname{Span}_F\{1, u, \ldots, u^{n-1}\} \quad \forall i \in \mathbb{N}. \end{aligned}$$

and we obtain

$$F[u] = \operatorname{Span}_F\{1, u, u^2, \ldots\} = \operatorname{Span}_F\{1, u, \ldots, u^{n-1}\}.$$

(4)⇒(1): Consider the elements $1, u, u^2, u^3, \ldots$ in $F(u)$. Since $[F(u) : F] = \dim_F F(u)$ is finite, they cannot be linearly independent, so there are $m \in \mathbb{N}$ and $a_0, \ldots, a_m \in F$ such that

$$a_0 + a_1u + \cdots + a_mu^m = 0,$$

so u is algebraic over F.

Suppose now that one (=all) of the above conditions hold, with $n = \deg m_{u/F}(X)$. As seen in the proof of (3)⇒(4), $\{1, u, \ldots, u^{n-1}\}$ generates $F[u]$:

$$\begin{aligned} F[u] &= \{g(u) : g(X) \in F[X],\ \deg g(X) < \deg m_{u/F}(X)\} \\ &= \{a_0 + a_1u + \cdots + a_{n-1}u^{n-1} : a_i \in F\}. \end{aligned}$$

We check that $\{1, u, \ldots, u^{n-1}\}$ is linearly independent over F: Assume that $b_0 + b_1u + \cdots + b_{n-1}u^{n-1} = 0$ for some $b_0, \ldots, b_{n-1} \in F$ not all 0, and let $g(X) = b_0 + b_1X + \cdots + b_{n-1}X^{n-1}$. Then $g(u) = 0$ but $\deg g(X) < \deg m_{u/F}(X)$, impossible. □

Example 4.13.

$$\begin{aligned}\mathbb{Q}(\sqrt{2}) = \mathbb{Q}[\sqrt{2}] &= \{a + b\sqrt{2} : a, b \in \mathbb{Q}\} \\ &\cong \mathbb{Q}[X]/\langle X^2 - 2\rangle.\end{aligned}$$

Note that $X^2 - 2$ is irreducible in $\mathbb{Q}$ (for instance by Eisenstein with $p = 2$, or using that it has degree 2 and no root in $\mathbb{Q}$).

$\{1, \sqrt{2}\}$ *is a basis over $\mathbb{Q}$ of the field $\mathbb{Q}(\sqrt{2})$, which is the smallest subfield of $\mathbb{C}$ (or $\mathbb{R}$) containing both $\mathbb{Q}$ and $\sqrt{2}$.*

$[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2.$