Algebraic Methods

F. Oggier

November 11, 2011

These notes were written to suit the contents of the course "Algebraic methods" given at NTU from August to October 2009, 2010 and 2011.

The main structure of the notes comes from the book by Robert Ash [1], a reference for this course.

The discussions on permutations were inspired by the notes of Peter Hendrikus Kropholler (`http://www.maths.gla.ac.uk/~phk/3H_GRF_Chapter_2.pdf`).

The proof of Jordan-Holder Theorem was inspired by the one given by Stuart Rankin (`www.math.uwo.ca/~srankin`).

The presentation of the Primitive Element Theorem is based on the one given by Ken Brown (`www.math.cornell.edu/~kbrown/4340/primitive.pdf`).

Many parts in the Chapter on Galois Theory are based on the book by Stewart [6].

For the history comments, they are taken from [2, 3, 4, 5].

Exercises have been collected during these past years from different sources. Are included a couple of exercises from Lam's book, and a couple of exercises from G. Berhuy.

Finally, pictures are coming from Wikipedia.

Contents

Chapter 1

Group Theory

Most lectures on group theory actually start with the definition of what is a group. It may be worth though spending a few lines to mention how mathematicians came up with such a concept.

Around 1770, Lagrange initiated the study of permutations in connection with the study of the solution of equations. He was interested in understanding solutions of polynomials in several variables, and got this idea to study the behaviour of polynomials when their roots are permuted. This led to what we now call Lagrange's Theorem, though it was stated as [5] *If a function $f(x_1, \ldots, x_n)$ of n variables is acted on by all $n!$ possible permutations of the variables and these permuted functions take on only r values, then r is a divisior of $n!$.* It is Galois (1811-1832) who is considered by many as the founder of group theory. He was the first to use the term "group" in a technical sense, though to him it meant a collection of permutations closed under multiplication. Galois theory will be discussed much later in these notes. Galois was also motivated by the solvability of polynomial equations of degree n. From 1815 to 1844, Cauchy started to look at permutations as an autonomous subject, and introduced the concept of permutations generated by certain elements, as well as several notations still used today, such as the cyclic notation for permutations, the product of permutations, or the identity permutation. He proved what we call today Cauchy's Theorem, namely that if p is prime divisor of the cardinality of the group, then there exists a subgroup of cardinality p. In 1870, Jordan gathered all the applications of permutations he could find, from algebraic geometry, number theory, function theory, and gave a unified presentation (including the work of Cauchy and Galois). Jordan made explicit the notions of homomorphism, isomorphism (still for permutation groups), he introduced solvable groups, and proved that the indices in two composition series are the same (now called Jordan-Hölder Theorem). He also gave a proof that the alternating group A_n is simple for $n > 4$.

In 1870, while working on number theory (more precisely, in generalizing

Kummer's work on cyclotomic fields to arbitrary fields), Kronecker described in one of his papers a finite set of arbitrary elements on which he defined an abstract operation on them which satisfy certain laws, laws which now correspond to axioms for finite abelian groups. He used this definition to work with ideal classes. He also proved several results now known as theorems on abelian groups. Kronecker did not connect his definition with permutation groups, which was done in 1879 by Frobenius and Stickelberger.

Apart permutation groups and number theory, a third occurence of group theory which is worth mentioning arose from geometry, and the work of Klein (we now use the term Klein group for one of the groups of order 4), and Lie, who studied transformation groups, that is transformations of geometric objects. The work by Lie is now a topic of study in itself, but Lie theory is beyond the scope of these notes.

The abstract point of view in group theory emerged slowly. It took something like one hundred years from Lagrange's work of 1770 for the abstract group concept to evolve. This was done by abstracting what was in commun to permutation groups, abelian groups, transformation groups... In 1854, Cayley gave the modern definition of group for the first time:
"*A set of symbols all of them different, and such that the product of any two of them (no matter in what order), or the product of any one of them into itself, belongs to the set, is said to be a group. These symbols are not in general convertible [commutative], but are associative.*"

Let us start from there.

1.1 Groups and subgroups

We start by introducing the object that will interest us for the whole chapter.

Definition 1.1. A group is a non-empty set G on which there is a binary operation $(a, b) \mapsto ab$ such that

- if a and b belong to G then ab is also in G (*closure*),
- $a(bc) = (ab)c$ for all a, b, c in G (*associativity*),
- there is an element $1 \in G$ such that $a1 = 1a = a$ for all $a \in G$ (*identity*),
- if $a \in G$, then there is an element $a^{-1} \in G$ such that $aa^{-1} = a^{-1}a = 1$ (*inverse*).

One can easily check that this implies the unicity of the identity and of the inverse.

A group G is called abelian if the binary operation is commutative, i.e., $ab = ba$ for all $a, b \in G$.

Remark. There are two standard notations for the binary group operation: either the additive notation, that is $(a, b) \mapsto a + b$ in which case the identity is

denoted by 0, or the multiplicative notation, that is $(a, b) \mapsto ab$ for which the identity is denoted by 1.

Examples 1.1. 1. $\mathbb{Z}$ with the addition and 0 as identity is an abelian group.

2. $\mathbb{Z}$ with the multiplication is not a group since there are elements which are not invertible in $\mathbb{Z}$.

3. The set of $n \times n$ invertible matrices with real coefficients is a group for the matrix product and identity the matrix $\mathbf{I}_n$. It is denoted by $GL_n(\mathbb{R})$ and called the general linear group. It is not abelian for $n \geq 2$.

The above examples are the easiest groups to think of. The theory of algebra however contains many examples of famous groups that one may discover, once equipped with more tools (for example, the Lie groups, the Brauer group, the Witt group, the Weyl group, the Picard group,...to name a few).

Definition 1.2. The order of a group G, denoted by $|G|$, is the cardinality of G, that is the number of elements in G.

We have only seen infinite groups so far. Let us look at some examples of finite groups.

Examples 1.2. 1. The trivial group $G = \{0\}$ may not be the most exciting group to look at, but still it is the only group of order 1.

2. The group $G = \{0, 1, 2, \ldots, n-1\}$ of integers modulo n is a group of order n. It is sometimes denoted by $\mathbb{Z}_n$ (this should not be confused with p-adic integers though!).

Definition 1.3. A subgroup H of a group G is a non-empty subset of G that forms a group under the binary operation of G.

Examples 1.3. 1. If we consider the group $G = \mathbb{Z}_4 = \{0, 1, 2, 3\}$ of integers modulo 4, $H = \{0, 2\}$ is a subgroup of G.

2. The set of $n \times n$ matrices with real coefficients and determinant of 1 is a subgroup of $GL_n(\mathbb{R})$, denoted by $SL_n(\mathbb{R})$ and called the special linear group.

At this point, in order to claim that the above examples are actually subgroups, one has to actually check the definition. The proposition below gives an easier criterion to decide whether a subset of a group G is actually a subgroup.

Proposition 1.1. *Let G be a group. Let H be a non-empty subset of G. The following are equivalent:*

1. *H is a subgroup of G.*
2. *(a) $x, y \in H$ implies $xy \in H$ for all x, y.*
   (b) $x \in H$ implies $x^{-1} \in H$.

3. *$x, y \in H$ implies $xy^{-1} \in H$ for all x, y.*

Proof. We prove that 1. $\Rightarrow$ 3. $\Rightarrow$ 2. $\Rightarrow$ 1.

1. $\Rightarrow$ 3. This part is clear from the definition of subgroup.

3. $\Rightarrow$ 2. Since H is non-empty, let $x \in H$. By assumption of 3., we have that $xx^{-1} = 1 \in H$ and that $1x^{-1} \in H$ thus x is invertible in H. We now know that for $x, y \in H$, x and y^{-1} are in H, thus $x(y^{-1})^{-1} = xy$ is in H.

2. $\Rightarrow$ 1. To prove this direction, we need to check the definition of group. Since closure and existence of an inverse are true by assumption of 2., and that associativity follows from the associativity in G, we are left with the existence of an identity. Now, if $x \in H$, then $x^{-1} \in H$ by assumption of 2., and thus $xx^{-1} = 1 \in H$ again by assumption of 2., which completes the proof.

□

We will often use the last equivalence to check that a subset of a group G is a subgroup.

Now that we have these structures of groups and subgroups, let us introduce a map that allows to go from one group to another and that respects the respective group operations.

Definition 1.4. Given two groups G and H, a group homomorphism is a map $f : G \to H$ such that

$$f(xy) = f(x)f(y) \text{ for all } x, y \in G.$$

Note that this definition immediately implies that the identity 1_G of G is mapped to the identity 1_H of H. The same is true for the inverse, that is $f(x^{-1}) = f(x)^{-1}$.

Example 1.4. The map $\exp : (\mathbb{R}, +) \to (\mathbb{R}^*, \cdot),\ x \mapsto \exp(x)$ is a group homomorphism.

Definition 1.5. Two groups G and H are isomorphic if there is a group homomorphism $f : G \to H$ which is also a bijection.

Roughly speaking, isomorphic groups are "essentially the same".

Example 1.5. If we consider again the group $G = \mathbb{Z}_4 = \{0, 1, 2, 3\}$ of integers modulo 4 with subgroup $H = \{0, 2\}$, we have that H is isomorphic to $\mathbb{Z}_2$, the group of integers modulo 2.

A crucial definition is the definition of the order of a group element.

Definition 1.6. The order of an element $a \in G$ is the least positive integer n such that $a^n = 1$. If no such integer exists, the order of a is infinite. We denote it by $|a|$.

Note that the critical part of this definition is that the order is the *least* positive integer with the given property. The terminology *order* is used both for groups and group elements, but it is usually clear from the context which one is considered.

1.2 Cyclic groups

Let us now introduce a first family of groups, the cyclic groups.

Definition 1.7. A group G is cyclic if it is generated by a single element, which we denote by $G = \langle a \rangle$. We may denote by C_n a cyclic group of n elements.

Example 1.6. A finite cyclic group generated by a is necessarily abelian, and can be written (multiplicatively)

$$\{1, a, a^2, \dots, a^{n-1}\} \text{ with } a^n = 1$$

or (additively)

$$\{0, a, 2a, \dots, (n-1)a\} \text{ with } na = 0.$$

A finite cyclic group with n elements is isomorphic to the additive group $\mathbb{Z}_n$ of integers modulo n.

Example 1.7. An nth root of unity is a complex number z which satisfies the equation $z^n = 1$ for some positive integer n. Let $\zeta_n = e^{2i\pi/n}$ be an nth root of unity. All the nth roots of unity form a group under multiplication. It is a cyclic group, generated by ζ_n, which is called a primitive root of unity. The term "primitive" exactly refers to being a generator of the cyclic group, namely, an nth root of unity is primitive when there is no positive integer k smaller than n such that $\zeta_n^k = 1$.

Here are some properties of cyclic groups and its generators.

Proposition 1.2. *If G is a cyclic group of order n generated by a, the following conditions are equivalent:*

1. $|a^k| = n$.
2. *k and n are relatively prime.*
3. *k has an inverse modulo n, that is there exists an integer s such that $ks \equiv 1$ modulo n.*

Proof. Before starting the proof, recall that since a generates G of order n, we have that the order of a is n and in particular $a^n = 1$. The fact that $|a^k| = n$ means in words that the order of a^k is also n, that is, a^k is also a generator of G. We first prove that 1. $\iff$ 2., while 2. $\iff$ 3. follows from Bezout identity.

1. $\Rightarrow$ 2. Suppose by contradiction that k and n are not relatively prime, that is, there exists $s > 1$ such that $s|k$ and $s|n$. Thus $n = ms$ and $k = sr$ for some $m, r \geq 1$ and we have
$$(a^k)^m = a^{srm} = a^{nr} = 1.$$
Now since $s > 1$, $m < n$, which contradicts that n is the order of a^k.

2. $\Rightarrow$ 1. Suppose that the order of a^k is not n, then there exists $m < n$ such that $(a^k)^m = 1$ and thus $n|km$ since n is the order of a. If k and n were to be relatively prime, then n would divide m, which is a contradiction since $m < n$.

2. $\Rightarrow$ 3. If k and n are relatively prime, then by Bezout identity, there exist r, s such that $1 = kr + ns$ and thus $kr \equiv 1$ modulo n.

3. $\Rightarrow$ 2. If $kr \equiv 1$ modulo n then $1 = kr + ns$ for some s and the greatest common divisor of k and n must divide 1, which shows k and n are relatively prime.

□

Corollary 1.3. *The set of invertible elements modulo n forms a group under multiplication, whose order is the Euler function $\varphi(n)$, which by definition counts the number of positive integers less than n that are relatively prime to n.*

Example 1.8. Consider the group $\mathbb{Z}_6 = \{0, 1, 2, 3, 4, 5\}$, the group $\mathbb{Z}_6^*$ of invertible elements in $\mathbb{Z}_6$ is $\mathbb{Z}_6^* = \{1, 5\}$. We have that $\varphi(6) = \varphi(2)\varphi(3) = 2$.

1.3 Cosets and Lagrange's Theorem

Definition 1.8. Let H be a subgroup of a group G. If $g \in G$, the right coset of H generated by g is
$$Hg = \{hg,\ h \in H\}$$
and similarly the left coset of H generated by g is
$$gH = \{gh,\ h \in H\}.$$

In additive notation, we get $H + g$ (which usually implies that we deal with a commutative group where we do not need to distinguish left and right cosets).

Example 1.9. If we consider the group $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ and its subgroup $H = \{0, 2\}$ which is isomorphic to $\mathbb{Z}_2$, the cosets of H in G are
$$0 + H = H,\ 1 + H = \{1, 3\},\ 2 + H = H,\ 3 + H = \{1, 3\}.$$
Clearly $0 + H = 2 + H$ and $1 + H = 3 + H$.

We see in the above example that while an element of $g \in G$ runs through all possible elements of the group G, some of the left cosets gH (or right cosets Hg) may be the same. It is easy to see when this exactly happens.

Lemma 1.4. *We have that $Ha = Hb$ if and only if $ab^{-1} \in H$ for $a, b \in G$. Similarly, $aH = bH$ if and only if $a^{-1}b \in H$ for $a, b \in G$.*

Proof. If two right cosets are the same, that is $Ha = Hb$, since H is a subgroup, we have $1 \in H$ and $a = hb$ for some $h \in H$, so $ab^{-1} = h \in H$.

Conversely, if $ab^{-1} = h \in H$, then $Ha = Hhb = Hb$, again since H is a subgroup. □

While one may be tempted to define a coset with a subset of G which is not a subgroup, we see that the above characterization really relies on the fact that H is actually a subgroup.

Example 1.10. It is thus no surprise that in the above example we have $0+H = 2+H$ and $1+H = 3+H$, since we have modulo 4 that $0-2 \equiv 2 \in H$ and $1-3 \equiv 2 \in H$.

Saying that two elements $a, b \in G$ generate the same coset is actually an equivalence relation in the following sense. We say that a is equivalent to b if and only if $ab^{-1} \in H$, and this relation satisfies the three properties of an equivalence relation:

- *reflexivity*: $aa^{-1} = 1 \in H$.
- *symmetry*: if $ab^{-1} \in H$ then $(ab^{-1})^{-1} = ba^{-1} \in H$.
- *transitivity*: if $ab^{-1} \in H$ and $bc^{-1} \in H$ then $(ab^{-1})(bc^{-1}) = ac^{-1} \in H$.

The equivalence class of a is the set of elements in G which are equivalent to a, namely

$$\{b,\ ab^{-1} \in H\}.$$

Since $ab^{-1} \in H \iff (ab^{-1})^{-1} = ba^{-1} \in H \iff b \in Ha$, we further have that

$$\{b,\ ab^{-1} \in H\} = Ha,$$

and a coset is actually an equivalence class.

Example 1.11. Let us get back to our example with the group $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ and its subgroup $H = \{0, 2\}$. We compute the first coset $0 + H = H$, and thus we now know that the equivalence class of 0 is H, and thus there is no need to compute the coset generated by 2, since it will give the same coset. We then compute the coset $1 + H = \{1, 3\}$ and again there is no need to compute the one of 3 since it is already in the coset of 1. We thus get 2 cosets, and clearly they partition $\mathbb{Z}_4$:

$$\mathbb{Z}_4 = \{0, 2\} \sqcup \{1, 3\} = H \sqcup (1 + H).$$

It is important to notice that the right (resp. left) cosets partition the group G (that the union of all cosets is G is clear since we run through all elements of G and H contains 1, and it is easy to see that if $x \in Ha$ and $x \in Hb$ then $Ha = Hb$).

Example 1.12. Consider $\mathbb{R}$ as an additive group with subgroup $\mathbb{Z}$. Every real number up to addition by an integer looks like a number in $[0, 1)$. Thus

$$\mathbb{R} = \cup_{0 \leq x < 1}(x + \mathbb{Z}),$$

and the cosets of $\mathbb{Z}$ partition $\mathbb{R}$.

Furthermore, since the map $h \mapsto ha$, $h \in H$, is a one-to-one correspondence, each coset has $|H|$ elements.

Definition 1.9. The index of a subgroup H in G is the number of right (left) cosets. It is a positive number or ∞ and is denoted by $[G : H]$.

If we think of a group G as being partitioned by cosets of a subgroup H, then the index of H tells how many times we have to translate H to cover the whole group.

Example 1.13. In Example 1.12, the index $[\mathbb{R} : \mathbb{Z}]$ is infinite, since there are infinitely many cosets of $\mathbb{Z}$ in $\mathbb{R}$.

Theorem 1.5. (Lagrange's Theorem). *If H is a subgroup of G, then $|G| = |H|[G : H]$. In particular, if G is finite then $|H|$ divides $|G|$ and $[G : H] = |G|/|H|$.*

Proof. Let us start by recalling that the left cosets of H forms a partition of G, that is

$$G = \sqcup gH,$$

where g runs through a set of representatives (one for each coset). Let us look at the cardinality of G:

$$|G| = |\sqcup\, gH| = \sum |gH|$$

since we have a disjoint union of cosets, and the sum is again over the set of representatives. Now

$$\sum |gH| = \sum |H|$$

since we have already noted that each coset contains $|H|$ elements. We then conclude that

$$|G| = \sum |H| = [G : H]|H|.$$

□

Example 1.14. Consider $G = \mathbb{Z}$, $H = 3\mathbb{Z}$, then $[G : H] = 3$.

Of course, Lagrange did not prove Lagrange's theorem! The modern way of defining groups did not exist yet at his time. Lagrange was interested in polynomial equations, and in understanding the existence and nature of the roots (does every equation has a root? how many roots?...). What he actually proved was that if a polynomial in n variables has its variables permuted in all $n!$ ways, the number of different polynomials that are obtained is always a

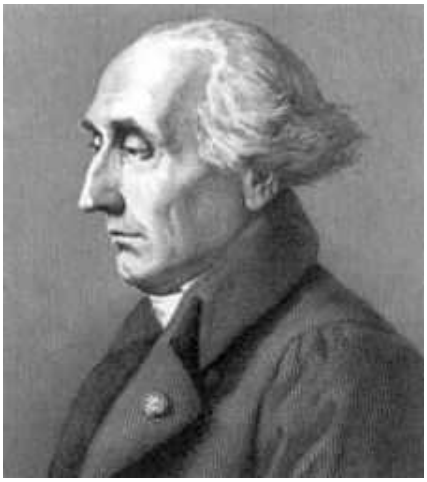


Figure 1.1: Joseph-Louis Lagrange (1736-1813)

factor of $n!$. Since all the permutations of n elements are actually a group, the number of such polynomials is actually the index in the group of permutations of n elements of the subgroup H of permutations which preserve the polynomial. So the size of H divides $n!$, which is exactly the number of all permutations of n elements. This is indeed a particular case of what we call now Lagrange's Theorem.

Corollary 1.6. *1. Let G be a finite group. If $a \in G$, then $|a|$ divides $|G|$. In particular, $a^{|G|} = 1$.*

2. If G has prime order, then G is cyclic.

Proof. 1. If $a \in G$ has order say m, then the subgroup $H = \{1, a, \ldots, a^{m-1}\}$ is a cyclic subgroup of G with order $|H| = m$. Thus m divides $|G|$ by the theorem.

2. Since $|G|$ is prime, we may take $a \neq 1$ in G, and since the order of a has to divide $|G|$, we have $|a| = |G|$. Thus the cyclic group generated by a coincides with G.

□

Example 1.15. Using Lagrange's Theorem and its corollaries, we can already determine the groups of order from 1 to 5, up to isomorphism (see Table 1.1). If $|G|$ is prime, we now know that G is cyclic.

Let us look at the case where G is of order 4. Let $g \in G$. We know that the order of g is either 1,2 or 4. If the order of g is 1, this is the identity. If G contains an element g of order 4, then that means that g generates the whole group, thus G is cyclic. If now G does not contain an element of order 4, then

$\lvert G\rvert$	G
1	$\{1\}$
2	C_2
3	C_3
4	C_4, $C_2 \times C_2$
5	C_5

Table 1.1: Groups of order from 1 to 5. C_n denotes the cyclic group of order n.

apart the identity, all the elements have order 2. From there, it is easy to obtain a multiplication table for G, and see that it coincides with the one of the group

$$\mathbb{Z}_2 \times \mathbb{Z}_2 = \{(x, y) \mid x, y \in \mathbb{Z}_2\}$$

with binary operation $(x, y) + (x', y') = (x + x', y + y')$. This group is called the Klein group, and it has several interpretations, for example, it is the group of isometries fixing a rectangle.

Remark. The above example also shows that the converse of Lagrange's Theorem is not true. If we take the group $G = C_2 \times C_2$, then 4 divides the order of G, however there is no element of order 4 in G.

Once Lagrange's Theorem and its corollaries are proven, we can easily deduce Euler's and Fermat's Theorem.

Theorem 1.7. (Euler's Theorem). *If a and n are relatively prime positive integers, with $n \geq 2$, then*

$$a^{\varphi(n)} \equiv 1 \mod n.$$

Proof. Since a and n are relatively prime, we know from Proposition 1.2 that a has an inverse modulo n, and by its corollary that the group of invertible elements has order $\varphi(n)$. Thus

$$a^{\varphi(n)} \equiv 1 \mod n$$

by Lagrange's Theorem first corollary. □

Corollary 1.8. (Fermat's Little Theorem). *If p is a prime and a is a positive integer not divisible by p, then*

$$a^{p-1} \equiv 1 \mod p.$$

This is particular case of Euler's Theorem when n is a prime, since then $\varphi(n) = p - 1$.

1.4 Normal subgroups and quotient group

Given a group G and a subgroup H, we have seen how to define the cosets of H, and thanks to Lagrange's Theorem, we already know that the number of cosets $[G : H]$ is related to the order of H and G by $|G| = |H|[G : H]$. A priori, the set of cosets of H has no structure. We are now interested in a criterion on H to give the set of its cosets a structure of group.

In what follows, we may write $H \leq G$ for H is a subgroup of G.

Definition 1.10. Let G be a group and $H \leq G$. We say that H is a normal subgroup of G, or that H is normal in G, if we have

$$cHc^{-1} = H, \text{ for all } c \in G.$$

We denote it $H \trianglelefteq G$, or $H \triangleleft G$ when we want to emphasize that H is a proper subgroup of G.

The condition for a subgroup to be normal can be stated in many slightly different ways.

Lemma 1.9. *Let $H \leq G$. The following are equivalent:*

1. *$cHc^{-1} \subseteq H$ for all $c \in G$.*
2. *$cHc^{-1} = H$ for all $c \in G$, that is $cH = Hc$ for all $c \in G$.*
3. *Every left coset of H in G is also a right coset (and vice-versa, every right coset of H in G is also a left coset).*

Proof. Clearly 2. implies 1., now $cHc^{-1} \subseteq H$ for all $c \in G$ if and only if $cH \subseteq Hc$. Let $x \in Hc$, that is $x = hc$ for some $h \in H$, so that

$$x = (cc^{-1})hc = c(c^{-1}hc) = ch'$$

for some $h' \in H$ since $cHc^{-1} \subset H$ for all c and thus in particular for c^{-1}. This shows that Hc is included in cH or equivalently that $H \subseteq cHc^{-1}$.

Also 2. clearly implies 3. Now suppose that $cH = Hd$. This means that c belongs to Hd by assumption and to Hc by definition, which means that $Hd = Hc$. □

Example 1.16. Let $GL_n(\mathbb{R})$ be the group of $n \times n$ real invertible matrices, and let $SL_n(\mathbb{R})$ be the subgroup formed by matrices whose determinant is 1. Let us see that $SL_n(\mathbb{R}) \triangleleft GL_n(\mathbb{R})$.

For that, we have to check that $ABA^{-1} \in SL_n(\mathbb{R})$ for all $B \in SL_n(\mathbb{R})$ and $A \in GL_n(\mathbb{R})$. This is clearly true since

$$\det(ABA^{-1}) = \det(B) = 1.$$

Proposition 1.10. *If H is normal in G, then the cosets of H form a group.*

Proof. Let us first define a binary operation on the cosets: $(aH, bH) \mapsto (aH)(bH) = \{(ah)(bh'),\ ah \in aH,\ bh' \in bH\}$. We need to check that the definition of group is satisfied.

- **closure.** This is the part which asks a little bit of work. Since $cH = Hc$ for all $c \in G$, then

$$(aH)(bH) = a(Hb)H = a(bH)H = abHH = abH.$$

  Note that this product does not depend on the choice of representatives.

- **Associativity** comes from G being associative.
- The **identity** is given by the coset $1H = H$.
- The **inverse** of the coset aH is $a^{-1}H$.

□

Definition 1.11. The group of cosets of a normal subgroup N of G is called the quotient group of G by N. It is denoted by G/N.

Let us finish this section by discussing the connection between normal subgroups and homomorphisms. The first normal subgroup of interest will be the kernel of a group homomorphism.

Recall that if $f : G \to H$ is a group homomorphism, the kernel of f is defined by

$$\mathrm{Ker}(f) = \{a \in G,\ f(a) = 1\}.$$

It is easy to see that $\mathrm{Ker}(f)$ is a normal subgroup of G, since

$$f(aba^{-1}) = f(a)f(b)f(a)^{-1} = f(a)f(a)^{-1} = 1$$

for all $b \in \mathrm{Ker}(f)$ and all $a \in G$.

The converse is more interesting.

Proposition 1.11. *Let G be a group. Every normal subgroup of G is the kernel of a homomorphism.*

Proof. Suppose that $N \trianglelefteq G$ and consider the map

$$\pi : G \to G/N,\ a \mapsto aN.$$

To prove the result, we have to show that π is a group homomorphism whose kernel is N. First note that π is indeed a map from group to group since G/N is a group by assuming that N is normal. Then we have that

$$\pi(ab) = abN = (aN)(bN) = \pi(a)\pi(b)$$

where the second equality comes from the group structure of G/N. Finally

$$\mathrm{Ker}(\pi) = \{a \in G \mid \pi(a) = N\} = \{a \in G \mid aN = N\} = N.$$

□

Definition 1.12. Let $N \trianglelefteq G$. The group homomorphism

$$\pi : G \to G/N,\ a \mapsto aN$$

is called the natural or canonical map or projection.

Recall for further usage that for f a group homomorphism, we have the following characterization of injectivity: a homomorphism f is injective if and only if its kernel is trivial (that is, contains only the identity element). Indeed, if f is injective, then $\mathrm{Ker}(f) = \{a,\ f(a) = 1\} = \{1\}$ since $f(1) = 1$. Conversely, if $\mathrm{Ker}(f) = \{1\}$ and we assume that $f(a) = f(b)$, then

$$f(ab^{-1}) = f(a)f(b)^{-1} = f(a)f(a)^{-1} = 1$$

and $ab^{-1} = 1$ implying that $a = b$ and thus f is injective.

Terminology.

monomorphism=injective homomorphism

epimorphism=surjective homomorphism

isomorphism=bijective homomorphism

endomorphism=homomorphism of a group to itself

automorphism=isomorphism of a group with itself

We have looked so far at the particular subgroup of G which is its kernel. The proposition below describes more generally subgroups of G and H.

Proposition 1.12. *Let $f : G \to H$ be a homomorphism.*

1. *If K is a subgroup of G, then $f(K)$ is a subgroup of H. If f is an epimorphism and K is normal, then $f(K)$ is normal.*
2. *If K is a subgroup of H, then $f^{-1}(K) = \{x \in G,\ f(x) \in K\}$ is a subgroup of G. If K is normal, so is $f^{-1}(K)$.*

Proof. 1. To prove that $f(K)$ is a subgroup of H, it is enough to show that $f(a)f(b)^{-1} \in f(K)$ by Proposition 1.1, which is clear from

$$f(a)f(b)^{-1} = f(ab^{-1}) \in f(K).$$

If K is normal, we have to show that $cf(K)c^{-1} = f(K)$ for all $c \in H$. Since f is an epimorphism, there exists $d \in G$ such that $f(d) = c$, so that

$$cf(K)c^{-1} = f(d)f(K)f(d)^{-1} = f(dKd^{-1}) = f(K)$$

using that K is normal.

2. As before, to prove that $f^{-1}(K)$ is a subgroup of G, it is enough to showing that $ab^{-1} \in f^{-1}(K)$ for $a, b \in f^{-1}(K)$, which is equivalent to show that $f(ab^{-1}) \in K$. This is now true since $f(ab^{-1}) = f(a)f(b)^{-1}$ with $a, b \in f^{-1}(K)$ and K a subgroup.

   For the second claim, we have to show that

$$cf^{-1}(K)c^{-1} = f^{-1}(K)$$

   or equivalently

$$f(cf^{-1}(K)c^{-1}) = K,\ c \in G.$$

   For $c \in G$ and $a \in f^{-1}(K)$, then

$$f(cac^{-1}) = f(c)f(a)f(c)^{-1} \in K$$

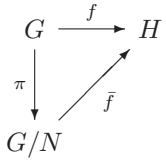
   since K is normal.

□

1.5 The isomorphism theorems

This section presents different isomorphism theorems which are important tools for proving further results. The first isomorphism theorem, that will be the second theorem to be proven after the factor theorem, is easier to motivate, since it will help us in computing quotient groups.

But let us first start with the so-called factor theorem. Assume that we have a group G which contains a normal subgroup N, another group H, and $f : G \to H$ a group homomorphism. Let π be the canonical projection (see Definition 1.12) from G to the quotient group G/N:



We would like to find a homomorphism $\bar{f} : G/N \to H$ that makes the diagram commute, namely

$$f(a) = \bar{f}(\pi(a))$$

for all $a \in G$.

Theorem 1.13. (Factor Theorem). *Any homomorphism f whose kernel K contains N can be factored through G/N. In other words, there is a unique homomorphism $\bar{f} : G/N \to H$ such that $\bar{f} \circ \pi = f$. Furthermore*

1. *$\bar{f}$ is an epimorphism if and only if f is.*
2. *$\bar{f}$ is a monomorphism if and only if $K = N$.*

3. $\bar{f}$ is an isomorphism if and only if f is an epimorphism and $K = N$.

Proof. **Unicity.** Let us start by proving that if there exists $\bar{f}$ such that $\bar{f} \circ \pi = f$, then it is unique. Let $\tilde{f}$ be another homomorphism such that $\tilde{f} \circ \pi = f$. We thus have that

$$(\bar{f} \circ \pi)(a) = (\tilde{f} \circ \pi)(a) = f(a)$$

for all $a \in G$, that is

$$\bar{f}(aN) = \tilde{f}(aN) = f(a).$$

This tells us that for all $bN \in G/N$ for which there exists an element b in G such that $\pi(b) = bN$, then its image by either $\bar{f}$ or $\tilde{f}$ is determined by $f(b)$. This shows that $\bar{f} = \tilde{f}$ by surjectivity of π.

Existence. Let $aN \in G/N$ such that $\pi(a) = aN$ for $a \in G$. We define

$$\bar{f}(aN) = f(a).$$

This is the most natural way to do it, however, we need to make sure that this is indeed well-defined, in the sense that it should not depend on the choice of the representative taken in the coset. Let us thus take another representative, say $b \in aN$. Since a and b are in the same coset, they satisfy $a^{-1}b \in N \subset K$, where $K = \mathrm{Ker}(f)$ by assumption. Since $a^{-1}b \in K$, we have $f(a^{-1}b) = 1$ and thus $f(a) = f(b)$.

Now that $\bar{f}$ is well defined, let us check this is indeed a group homomorphism. First note that G/N is indeed a group since $N \trianglelefteq G$. Then, we have

$$\bar{f}(aNbN) = \bar{f}(abN) = f(ab) = f(a)f(b) = \bar{f}(aN)\bar{f}(bN)$$

and $\bar{f}$ is a homomorphism.

1. The fact that $\bar{f}$ is an epimorphism if and only if f is comes from the fact that both maps have the same image.

2. First note that the statement $\bar{f}$ is a monomorphism if and only if $K = N$ makes sense since $K = \mathrm{Ker}(f)$ is indeed a normal subgroup, as proved earlier.

   To show that $\bar{f}$ is a monomorphism is equivalent to show that $\mathrm{Ker}(\bar{f})$ is trivial. By definition, we have

   $$\begin{aligned} \mathrm{Ker}(\bar{f}) &= \{aN \in G/N,\ \bar{f}(aN) = 1\} \\ &= \{aN \in G/N,\ \bar{f}(\pi(a)) = f(a) = 1\} \\ &= \{aN \in G/N,\ a \in K = \mathrm{Ker}(f)\}. \end{aligned}$$

   So the kernel of $\bar{f}$ is exactly those cosets of the form aN with $a \in K$, but for the kernel to be trivial, we need it to be equal to N, that is we need $K = N$.

3. This is just a combination of the first two parts.

□

We are now ready to state the first isomorphism theorem.

Theorem 1.14. (1st Isomorphism Theorem). *If $f : G \to H$ is a homomorphism with kernel K, then the image of f is isomorphic to G/K:*

$$\mathrm{Im}(f) \simeq G/\mathrm{Ker}(f).$$

Proof. We know from the Factor Theorem that

$$\bar{f} : G/\mathrm{Ker}(f) \to H$$

is an isomorphism if and only if f is an epimorphism, and clearly f is an epimorphism on its image, which concludes the proof. □

Example 1.17. We have seen in Example 1.16 that $SL_n(\mathbb{R}) \lhd GL_n(\mathbb{R})$. Consider the map

$$\det : GL_n(\mathbb{R}) \to (\mathbb{R}^*, \cdot),$$

which is a group homomorphism. We have that $\mathrm{Ker}(\det) = SL_n(\mathbb{R})$. The 1st Isomorphism Theorem tells that

$$\mathrm{Im}(\det) \simeq GL_n(\mathbb{R})/SL_n(\mathbb{R}).$$

It is clear that det is surjective, since for all $a \in \mathbb{R}^*$, one can take the diagonal matrix with all entries at 1, but one which is a. Thus we conclude that

$$\mathbb{R}^* \simeq GL_n(\mathbb{R})/SL_n(\mathbb{R}).$$

The 1st Isomorphism Theorem can be nicely illustrated in terms of exact sequences.

Definition 1.13. Let $F, G, H, I, \ldots$ be groups, and let $f, g, h, \ldots$ be group homomorphisms. Consider the following sequence:

$$\cdots \qquad F \xrightarrow{f} G \xrightarrow{g} H \xrightarrow{h} I \qquad \cdots$$

We say that this sequence is exact in one point (say G) if $\mathrm{Im}(f) = \mathrm{Ker}(g)$. A sequence is exact if it is exact in all points.

A short exact sequence of groups is of the form

$$1 \xrightarrow{i} F \xrightarrow{f} G \xrightarrow{g} H \xrightarrow{j} 1$$

where i is the inclusion and j is the constant map 1.

Proposition 1.15. *Let*

$$1 \xrightarrow{i} F \xrightarrow{f} G \xrightarrow{g} H \xrightarrow{j} 1$$

be a short exact sequence of groups. Then $\mathrm{Im}(f)$ *is normal in* G *and we have a group isomorphism*

$$G/\mathrm{Im}(f) \simeq H,$$

or equivalently

$$G/\mathrm{Ker}(g) \simeq H.$$

Proof. Since the sequence is exact, we have that $\mathrm{Im}(f) = \mathrm{Ker}(g)$ thus $\mathrm{Im}(f)$ is a normal subgroup of G. By the first Isomorphism Theorem, we have that

$$G/\mathrm{Ker}(g) \simeq \mathrm{Im}(g) = H.$$

since $\mathrm{Im}(g) = \mathrm{Ker}(j) = H$. □

The formulation in terms of exact sequences is useful to know, since it happens very often in the literature that an exact sequence is given exactly to be able to compute such quotient groups.

Let us state the second and third isomorphism theorem.

Theorem 1.16. (2nd Isomorphism Theorem). *If* H *and* N *are subgroups of* G*, with* N *normal in* G*, then*

$$H/(H \cap N) \simeq HN/N.$$

There are many things to discuss about the statement of this theorem.

- First we need to check that HN is indeed a subgroup of G. To show that, notice that $HN = NH$ since N is a normal subgroup of G. This implies that for $hn \in HN$, its inverse $(hn)^{-1} = n^{-1}h^{-1} \in G$ actually lives in HN, and so does the product $(hn)(h'n') = h(nh')n'$.
- Note that by writing HN/N, we insist on the fact that there is no reason for N to be a subgroup of H. On the other hand, N is a normal subgroup of HN, since for all $hn \in HN$, we have

  $$hnNn^{-1}h^{-1} = hNh^{-1} \subseteq N$$

  since N is normal in G.
- We now know that the right hand side of the isomorphism is a quotient group. In order to see that so is the left hand side, we need to show that $H \cap N$ is a normal subgroup of H. This comes by noticing that $H \cap N$ is the kernel of the canonical map $\pi : G \to G/N$ restricted to H.

Now that all these remarks have been done, it is not difficult to see that the 2nd Isomorphism Theorem follows from the 1st Isomorphism Theorem, as does the 3rd Isomorphism Theorem.

Theorem 1.17. (3rd Isomorphism Theorem). *If* N *and* H *are normal subgroups of* G*, with* N *contained in* H*, then*

$$G/H \simeq (G/N)/(H/N).$$

1.6 Direct and semi-direct products

So far, we have seen how given a group G, we can get smaller groups, such as subgroups of G or quotient groups. We will now do the other way round, that is, starting with a collection of groups, we want to build larger new groups.

Let us start with two groups H and K, and let $G = H \times K$ be the cartesian product of H and K, that is

$$G = \{(h, k),\ h \in H,\ k \in K\}.$$

We define a binary operation on this set by doing componentwise multiplication (or addition if the binary operations of H and K are denoted additively) on G:

$$(h_1, k_1)(h_2, k_2) = (h_1 h_2, k_1 k_2) \in H \times K.$$

Clearly G is closed under multiplication, its operation is associative (since both operations on H and K are), it has an identity element given by $1_G = (1_H, 1_K)$ and the inverse of (h, k) is (h^{-1}, k^{-1}). In summary, G is a group.

Definition 1.14. Let H, K be two groups. The group $G = H \times K$ with binary operation defined componentwise as described above is called the external direct product of H and K.

Examples 1.18.

1. Let $\mathbb{Z}_2$ be the group of integers modulo 2. We can build a direct product of $\mathbb{Z}_2$ with itself, namely $\mathbb{Z}_2 \times \mathbb{Z}_2$ with additive law componentwise. This is actually the Klein group, also written $C_2 \times C_2$. This group is not isomorphic to $\mathbb{Z}_4$!

2. Let $\mathbb{Z}_2$ be the group of integers modulo 2, and $\mathbb{Z}_3$ be the group of integers modulo 3. We can build a direct product of $\mathbb{Z}_2$ and $\mathbb{Z}_3$, namely $\mathbb{Z}_2 \times \mathbb{Z}_3$ with additive law componentwise. This group is actually isomorphic to $\mathbb{Z}_6$!

3. The group $(\mathbb{R}, +) \times (\mathbb{R}, +)$ with componentwise addition is a direct product.

Note that G contains isomorphic copies $\bar{H}$ and $\bar{K}$ of respectively H and K, given by

$$\bar{H} = \{(h, 1_K),\ h \in H\},\ \bar{K} = \{(1_H, k),\ k \in K\},$$

which furthermore are normal subgroups of G. Let us for example see that $\bar{H}$ is normal in G. By definition, we need to check that

$$(h, k)\bar{H}(h^{-1}, k^{-1}) \subseteq \bar{H},\ (h, k) \in G.$$

Let $(h', 1_K) \in \bar{H}$, we compute that

$$(h, k)(h', 1_k)(h^{-1}, k^{-1}) = (hh'h^{-1}, 1_k) \in \bar{H},$$

since $hh'h^{-1} \in H$. The same computation holds for $\bar{K}$.

If we gather what we know about $G, \bar{H}$ and $\bar{K}$, we get that

- by definition, $G = \bar{H}\bar{K}$ and $\bar{H} \cap \bar{K} = \{1_G\}$,
- by what we have just proved, $\bar{H}$ and $\bar{K}$ are two normal subgroups of G.

This motivates the following definition.

Definition 1.15. If a group G contains normal subgroups H and K such that $G = HK$ and $H \cap K = \{1_G\}$, we say that G is the internal direct product of H and K.

Examples 1.19.

1. Consider the Klein group $\mathbb{Z}_2 \times \mathbb{Z}_2$, it contains the two subgroups $H = \{(h, 0),\ h \in \mathbb{Z}_2\}$ and $K = \{(0, k),\ k \in \mathbb{Z}_2\}$. We have that both H and K are normal, because the Klein group is commutative. We also have that $H \cap K = \{(0, 0)\}$, so the Klein group is indeed an internal direct product. On the other hand, $\mathbb{Z}_4$ only contains as subgroup $H = \{0, 2\}$, so it is not an internal direct product!
2. Consider the group $\mathbb{Z}_2 \times \mathbb{Z}_3$, it contains the two subgroups $H = \{(h, 0),\ h \in \mathbb{Z}_2\}$ and $K = \{(0, k),\ k \in \mathbb{Z}_3\}$. We have that both H and K are normal, because the group is commutative. We also have that $H \cap K = \{(0, 0)\}$, so this group is indeed an internal direct product. Also $\mathbb{Z}_6$ contains the two subgroups $H = \{0, 3\} \simeq \mathbb{Z}_2$ and $K = \{0, 2, 4\} \simeq \mathbb{Z}_3$. We have that both H and K are normal, because the group is commutative. We also have that $H \cap K = \{0\}$, so this group is indeed an internal direct product, namely the internal product of $\mathbb{Z}_2$ and $\mathbb{Z}_3$. This is in fact showing that $\mathbb{Z}_6 \simeq \mathbb{Z}_2 \times \mathbb{Z}_3$.

The next result makes explicit the connection between internal and external products.

Proposition 1.18. *If G is the internal direct product of H and K, then G is isomorphic to the external direct product $H \times K$.*

Proof. To show that G is isomorphic to $H \times K$, we define the following map

$$f : H \times K \to G,\ f(h, k) = hk.$$

First remark that if $h \in H$ and $k \in K$, then $hk = kh$. Indeed, we have using that both K and H are normal that

$$(hkh^{-1})k^{-1} \in K,\ h(kh^{-1}k^{-1}) \in H$$

implying that

$$hkh^{-1}k^{-1} \in K \cap H = \{1\}.$$

We are now ready to prove that f is a group isomorphism.

1. This is a group homomorphism since

$$f((h, k)(h', k')) = f(hh', kk') = h(h'k)k' = h(kh')k' = f(h, k)f(h', k').$$

2. The map f is injective. This can be seen by checking that its kernel is trivial. Indeed, if $f(h,k)=1$ then

$$hk=1 \Rightarrow h=k^{-1} \Rightarrow h\in K \Rightarrow h\in H\cap K=\{1\}.$$

We have then that $h=k=1$ which proves that the kernel is $\{(1,1)\}$.

3. The map f is surjective since by definition $G=HK$.

□

Note that the definitions of external and internal product are surely not restricted to two groups. One can in general define them for n groups $H_1,\dots,H_n$. Namely

Definition 1.16. If $H_1,\dots,H_n$ are arbitrary groups, the external direct product of $H_1,\dots,H_n$ is the cartesian product

$$G=H_1\times H_2\times\cdots\times H_n$$

with componentwise multiplication.

If G contains normal subgroups $H_1,\dots,H_n$ such that $G=H_1\cdots H_n$ and each g can be represented as $h_1\cdots h_n$ uniquely, we say that G is the internal direct product of $H_1,\dots,H_n$.

We can see a slight difference in the definition of internal product, since in the case of two subgroups, the condition given was not that each g can be represented uniquely as h_1h_2, but instead that the intersection of the two subgroups is $\{1\}$. The next proposition shows the connection between these two points of view.

Proposition 1.19. *Suppose that $G=H_1\cdots H_n$ where each H_i is a normal subgroup of G. The following conditions are equivalent.*

1. *G is the internal direct product of the H_i.*
2. *$H_1H_2\cdots H_{i-1}\cap H_i=\{1\}$, for all $i=1,\dots,n$.*

Proof. Let us prove 1. $\iff$ 2.

1. $\Rightarrow$ 2. Let us assume that G is the internal direct product of the H_i, which means that every element in G can be written uniquely as a product of elements in H_i. Now let us take $g\in H_1H_2\cdots H_{i-1}\cap H_i=\{1\}$. We have that $g\in H_1H_2\cdots H_{i-1}$, which is uniquely written as $g=h_1h_2\cdots h_{i-1}1_{H_i}\cdots 1_{H_n}$, $h_j\in H_j$. On the other hand, $g\in H_i$ thus $g=1_{H_1}\cdots 1_{H_{i-1}}g$ and by unicity of the representation, we have $h_j=1$ for all j and $g=1$.

2. $\Rightarrow$ 1. Conversely, let us assume that $g\in G$ can be written either

$$g=h_1h_2\cdots h_n,\ h_j\in H_j,$$

or

$$g = k_1 k_2 \cdots k_n,\ k_j \in H_j.$$

Recall that since all H_j are normal subgroups, then

$$h_i h_j = h_j h_i,\ h_i \in H_i,\ h_j \in H_j.$$

(If you cannot recall the argument, check out the proof of Proposition 1.18). This means that we can do the following manipulations:

$$\begin{aligned} & h_1 h_2 \cdots h_n = k_1 k_2 \cdots k_n \\ \iff \quad & h_2 \cdots h_n = \underbrace{(h_1^{-1} k_1)}_{\in H_1} k_2 \cdots k_n \\ \iff \quad & h_3 \cdots h_n = (h_1^{-1} k_1)(h_2^{-1} k_2) k_3 \cdots k_n \end{aligned}$$

and so on and so forth till we reach

$$h_n k_n^{-1} = (h_1^{-1} k_1) \cdots (h_{n-1}^{-1} k_{n-1}).$$

Since the left hand side belongs to H_n while the right hand side belongs to $H_1 \cdots H_{n-1}$, we get that

$$h_n k_n^{-1} \in H_n \cap H_1 \cdots H_{n-1} = \{1\},$$

implying that $h_n = k_n$. We conclude the proof by iterating this process.

□

Let us get back to the case of two groups. We have seen above that we can endow the cartesian product of two groups H and K with a group structure by considering componentwise binary operation

$$(h_1, k_1)(h_2, k_2) = (h_1 h_2, k_1 k_2) \in H \times K.$$

The choice of this binary operation of course determines the structure of $G = H \times K$, and in particular we have seen that the isomorphic copies of H and K in G are normal subgroups. Conversely in order to define an internal direct product, we need to assume that we have two normal subgroups.

We now consider a more general setting, where the subgroup K does not have to be normal (and will not be in general), for which we need to define a new binary operation on the cartesian product $H \times K$. This will lead us to the definition of internal and external semi-direct product.

Recall that an automorphism of a group H is a bijective group homomorphism from H to H. It is easy to see that the set of automorphisms of H forms a group with respect to the composition of maps and identity element the identity map Id_H. We denote it by $\mathrm{Aut}(H)$.

Proposition 1.20. *Let H and K be groups, and let*

$$\rho : K \to \mathrm{Aut}(H),\ k \mapsto \rho_k$$

be a group homomorphism. Then the binary operation

$$(H \times K) \times (H \times K) \to (H \times K),\ ((h,k),(h',k')) \mapsto (h\rho_k(h'), kk')$$

endows $H \times K$ with a group structure, with identity element $(1,1)$.

Proof. First notice that the closure property is satisfied.

(Identity). Let us show that $(1,1)$ is the identity element. We have

$$(h,k)(1,1) = (h\rho_k(1), k) = (h,k)$$

for all $h \in H$, $k \in K$, since ρ_k is a group homomorphism. We also have

$$(1,1)(h',k') = (\rho_1(h'), k') = (h',k')$$

for all $h' \in H$, $k' \in K$, since ρ being a group homomorphism, it maps 1_K to $1_{\mathrm{Aut}(K)} = \mathrm{Id}_H$.

(Inverse). Let $(h,k) \in H \times K$ and let us show that $(\rho_k^{-1}(h^{-1}), k^{-1})$ is the inverse of (h,k). We have

$$(h,k)(\rho_k^{-1}(h^{-1}), k^{-1}) = (h\rho_k(\rho_k^{-1}(h^{-1})), 1) = (hh^{-1}, 1) = (1,1).$$

We also have

$$\begin{aligned}(\rho_k^{-1}(h^{-1}), k^{-1})(h,k) &= (\rho_k^{-1}(h^{-1})\rho_{k^{-1}}(h), 1)\\ &= (\rho_{k^{-1}}(h^{-1})\rho_{k^{-1}}(h), 1)\end{aligned}$$

using that $\rho_k^{-1} = \rho_{k^{-1}}$ since ρ is a group homomorphism. Now

$$(\rho_{k^{-1}}(h^{-1})\rho_{k^{-1}}(h), 1) = (\rho_{k^{-1}}(h^{-1}h), 1) = (\rho_{k^{-1}}(1), 1) = (1,1)$$

using that $\rho_{k^{-1}}$ is a group homomorphism for all $k \in K$.

Associativity. This is the last thing to check. On the one hand, we have

$$\begin{aligned}[(h,k)(h',k')](h'',k'') &= (h\rho_k(h'), kk')(h'',k'')\\ &= (h\rho_k(h')\rho_{kk'}(h''), (kk')k''),\end{aligned}$$

while on the other hand

$$\begin{aligned}(h,k)[(h',k')(h'',k'')] &= (h,k)(h'\rho_{k'}(h''), k'k'')\\ &= (h\rho_k(h'\rho_{k'}(h'')), k(k'k'')).\end{aligned}$$

Since K is a group, we have $(kk')k'' = k(k'k'')$. We now look at the first component. Note that $\rho_{kk'} = \rho_k \circ \rho_{k'}$ using that ρ is a group homomorphism, so that

$$h\rho_k(h')\rho_{kk'}(h'') = h\rho_k(h')\rho_k(\rho_{k'}(h'')).$$

Furthermore, ρ_k is a group homomorphism, yielding

$$h\rho_k(h')\rho_k(\rho_{k'}(h'')) = h\rho_k(h'\rho_{k'}(h''))$$

which concludes the proof. □

We are now ready to define the first semi-direct product.

Definition 1.17. Let H and K be two groups, and let

$$\rho : K \to \mathrm{Aut}(H)$$

be a group homomorphism. The set $H \times K$ endowed with the binary operation

$$((h,k),(h',k')) \mapsto (h\rho_k(h'), kk')$$

is a group G called an external semi-direct product of H and K by ρ, denoted by $G = H \times_\rho K$.

Example 1.20. Let us consider the group $\mathbb{Z}_2$ of integers modulo 2. Suppose we want to compute the semi-direct product of $\mathbb{Z}_2$ with itself, then we need to first determine $\mathrm{Aut}(\mathbb{Z}_2)$. Since an automorphism of $\mathbb{Z}_2$ must send 0 to 0, it has no other choice than send 1 to 1, and thus $\mathrm{Aut}(\mathbb{Z}_2)$ is only the identity map, in which case the semi-direct product is just the direct product, since ρ_k is the identity for every k. To have a bigger automorphism group, let us consider $H = \mathbb{Z}_3$. In that case, apart the identity map, we also have the map $x \mapsto x^{-1}$, that is $0 \mapsto 0,\ 1 \mapsto 2,\ 2 \mapsto 1$. Thus $\rho(0) = \rho_0$ is the identity, $\rho(1) = \rho_1$ is the inverse map, and we can form the external semi-direct product $G = \mathbb{Z}_3 \times_\rho \mathbb{Z}_2$.

In fact, this example holds for $\mathbb{Z}_n$, $n \geq 3$.

Example 1.21. Let $H = \mathbb{Z}_n$ be the group of integers mod n, $K = \mathbb{Z}_2$ be the group of integers mod 2, and let $\rho : K \to \mathrm{Aut}(H)$ be the homomorphism that sends 0 to the identity, and 1 to the inverse map of H, given by $x \mapsto x^{-1}$, which is indeed a group homomorphism of H since H is abelian. Since the subgroup of $\mathrm{Aut}(H)$ generated by the inverse map is of order 2, it is isomorphic to K. We can thus define the external semi-direct product $G = \mathbb{Z}_n \times_\rho \mathbb{Z}_2$.

We can make observations similar to what we did for direct products. Namely, we can identify two isomorphic copies $\bar{H}$ and $\bar{K}$ of respectively H and K, given by

$$\bar{H} = \{(h, 1_K),\ h \in H\},\ \bar{K} = \{(1_H, k),\ k \in K\},$$

and look at the properties of these subgroups.

- The subgroup $\bar{H} = \{(h,1),\ h \in H\}$ is normal in $H \times_\rho K$, this can be seen by writing down the definition of normal subgroup. (We cannot claim the same for $\bar{K}$!).
- We have $\bar{H}\bar{K} = H \times_\rho K$, since every element $(h,k) \in H \times_\rho K$ can be written as $(h,1)(1,k)$ (indeed $(h,1)(1,k) = (h\rho_1(1),k) = (h,k)$).
- We have $\bar{H} \cap \bar{K} = \{1_G\}$.

This motivates the definition of internal semi-direct products.

Definition 1.18. Let G be a group with subgroups H and K. We say that G is the internal semi-direct product of H and K if H is a normal subgroup of G, such that $HK = G$ and $H \cap K = \{1_G\}$. It is denoted by

$$G = H \rtimes K.$$

Example 1.22. The dihedral group D_n is the group of all reflections and rotations of a regular polygon with n vertices centered at the origin. It has order $2n$. Let a be a rotation of angle $2\pi/n$ and let b be a reflection. We have that

$$D_n = \{a^i b^j,\ 0 \leq i \leq n-1,\ j = 0, 1\},$$

with

$$a^n = b^2 = (ba)^2 = 1.$$

We thus have that $\langle a \rangle = C_n$ and $\langle b \rangle = C_2$, where C_n denotes the cyclic group of order n.

The geometric interpretation of D_n as symmetries of a regular polygon with n vertices holds for $n \geq 3$, however, note that when $n = 2$, we can still look at the relations defined above: we then have $a^2 = b^2 = (ba)^2 = 1$, thus D_2 contains only 4 elements, the identity and 3 elements of order 2, showing that it is isomorphic to the Klein group $C_2 \times C_2$.

To prove, for $n \geq 3$, that

$$D_n \simeq C_n \rtimes C_2,$$

we are left to check that $\langle a \rangle \cap \langle b \rangle = \{1\}$ and that $\langle a \rangle$ is normal in D_n. The former can be seen geometrically (a reflection cannot be obtained by possibly successive rotations of angle $2\pi/n$, $n \geq 3$). For the latter, we first show that

$$bab^{-1} \in \langle a \rangle,$$

which can be easily checked, since $(ba)^2 = baba = 1$, thus $bab = a^{-1} = bab^{-1}$ using that $b^2 = 1$. This also shows that $ba = a^{-1}b$ from which we have:

$$ba^2b^{-1} = baab^{-1} = a^{-1}(bab^{-1}) \in \langle a \rangle,$$

similarly

$$ba^3b^{-1} = baa^2b^{-1} = a^{-1}(ba^2b^{-1}) \in \langle a \rangle.$$

Again similarly to the case of direct products, these assumptions guarantee that we can write uniquely elements of the internal semi-direct product. Let us repeat things explicitly.

Lemma 1.21. *Let G be a group with subgroups H and K. Suppose that $G = HK$ and $H \cap K = \{1_G\}$. Then every element g of G can be written uniquely in the form hk, for $h \in H$ and $k \in K$.*

Proof. Since $G = HK$, we know that g can be written as hk. Suppose it can also be written as $h'k'$. Then $hk = h'k'$ so $h'^{-1}h = k'k^{-1} \in H \cap K = \{1\}$. Therefore $h = h'$ and $k = k'$. □

The internal and external direct products were two sides of the same objects, so are the internal and external semi-direct products. If $G = H \times_\rho K$ is the external semi-direct product of H and K, then $\bar{H} = H \times \{1\}$ is a normal subgroup of G and it is clear that G is the internal semi-direct product of $H \times \{1\}$ and $\{1\} \times K$. This reasoning allows us to go from external to internal semi-direct products. The result below goes in the other direction, from internal to external semi-direct products.

Proposition 1.22. *Suppose that G is a group with subgroups H and K, and G is the internal semi-direct product of H and K. Then $G \simeq H \times_\rho K$ where $\rho : K \to \mathrm{Aut}(H)$ is given by $\rho_k(h) = khk^{-1}$, $k \in K$, $h \in H$.*

Proof. Note that ρ_k belongs to $\mathrm{Aut}(H)$ since H is normal.

By the above lemma, every element g of G can be written uniquely in the form hk, with $h \in H$ and $k \in K$. Therefore, the map

$$\varphi : H \times_\rho K \to G,\ \varphi(h,k) = hk$$

is a bijection. It only remains to show that this bijection is a homomorphism.

Given (h,k) and (h',k') in $H \times_\rho K$, we have

$$\varphi((h,k)(h',k')) = \varphi((h\rho_k(h'), kk')) = \varphi(hkh'k^{-1}, kk') = hkh'k' = \varphi(h,k)\varphi(h',k').$$

Therefore φ is a group isomorphism, which concludes the proof. □

In words, we have that every internal semi-direct product is isomorphic to some external semi-direct product, where ρ is the conjugation.

Example 1.23. Consider the dihedral group D_n from the previous example:

$$D_n \simeq C_n \rtimes C_2.$$

According to the above proposition, D_n is isomorphic to an external semi-direct product

$$D_n \simeq C_n \times_\rho C_2,$$

where

$$\rho : C_2 \to \mathrm{Aut}(C_n),$$

maps to the conjugation in $\mathrm{Aut}(C_n)$. We have explicitly that

$$1 \mapsto \rho_1 = \mathrm{Id}_{C_n},\ b \mapsto \rho_b,\ \rho_b(a) = bab^{-1} = a^{-1}.$$

In fact, we are back to Example 1.21!

Before finishing this section, note the following distinction: the external (semi-)direct product of groups allows to construct new groups starting from different abstract groups, while the internal (semi-)direct product helps in analyzing the structure of a given group.

$\lvert G\rvert$	G abelian	G non-abelian
1	$\{1\}$	-
2	C_2	-
3	C_3	-
4	C_4, $C_2 \times C_2$	-
5	C_5	-
6	$C_6 = C_3 \times C_2$	$D_3 = C_3 \rtimes C_2$
7	C_7	-
8	C_8, $C_4 \times C_2$, $C_2 \times C_2 \times C_2$	$D_4 = C_4 \rtimes C_2$

Table 1.2: C_n denotes the cyclic group of order n, D_n the dihedral group

Example 1.24. Thanks to the new structures we have seen in this section, we can go on our investigation of groups of small orders. We can get two new groups of order 6 and 4 of order 8:

- $C_3 \times C_2$ is the direct product of C_3 and C_2. You may want to check that it is actually isomorphic to C_6.
- The dihedral group $D_3 = C_3 \rtimes C_2$ is the semi-direct product of C_3 and C_2. We get similarly $D_4 = C_4 \rtimes C_2$.
- The direct product $C_4 \times C_2$ and the direct product of the Klein group $C_2 \times C_2$ with C_2.

The table actually gives an exact classification of groups of small order (except the missing non-abelian quaternion group of order 8), though we have not proven it. The reason why the quaternion group of order 8 is missing is exactly because it cannot be written as a semi-direct product of smaller groups (see Exercises).

1.7 Permutations and Group action

Since we introduced the definition of group as a set with a binary operation which is closed, we have been computing things internally, that is inside a group structure. This was the case even when considering cartesian products of groups, where the first thing we did was to endow this set with a group structure.

In this section, we wonder what happens if we have a group and a set, which may or may not have a group structure. We will define a group action, that is a way to do computations with two objects, one with a group law, not the other one.

As a first result, we will prove the so-called Cayley's theorem, whose proof will motivate the introduction of group action. Since the statement of this theorem uses permutation groups, we start by recalling the necessary definitions. We will give enough background on permutations to define the alternating group, a group which is useful as an illustration of many concepts!

Definition 1.19. A permutation of a set S is a bijection on S. The set of all such functions (with respect to function composition) is a group called the symmetric group on S. We denote by S_n the symmetric group on n elements.

Example 1.25. Consider the symmetric group S_3 of permutations on 3 elements. It is given by (note here that by ab we mean that we first apply the permutation b, then a)

$$\begin{aligned} e &: 123 \to 123 \text{ or } () \\ a &: 123 \to 213 \text{ or } (12) \\ b &: 123 \to 132 \text{ or } (23) \\ ba &: 123 \to 312 \text{ or } (132) \\ ab &: 123 \to 231 \text{ or } (123) \\ aba &: 123 \to 321 \text{ or } (13) \end{aligned}$$

One can check that this is indeed a group. The notation (132) means that the permutation sends 1 to 3, 3 to 2, and 2 to 1.

We can generally write a permutation on m elements as $(i_1, \ldots, i_m)$, which is called a cycle notation. The permutation $(i_1, \ldots, i_m)$ is called an m-cycle. When $m = 2$, a 2-cycle is called a transposition. Note that several different cycles can represent the same permutation (e.g., $(132) = (321) = (213)$), and not every permutation is a cycle. For example, if we consider $1 \mapsto 2$, $2 \mapsto 1$, $3 \mapsto 4$, $4 \mapsto 3$, this permutation is not a cycle. However, it is clearly the product of two disjoint cycles, namely $(12)(34)$. Formally, we say that two cycles $(i_1, \ldots, i_s)$ and $(j_1, \ldots, j_t)$ are disjoint if and only if $\{i_1, \ldots, i_s\} \cap \{j_1, \ldots, j_t\}$ is empty. Such a decomposition of a permutation into product of disjoint cycles is true in general.

Proposition 1.23. *Every element of S_n can be expressed uniquely as a product of disjoint cycles, up to ordering of the cycles, and notational redundancy within each cycle. Furthermore, every cycle can be written as a product of transpositions.*

Proof. Let σ be an element of S_n. Choose any index $i_1 \in \{1, \ldots, n\}$. By applying σ repeatedly on i_1, we construct a sequence of elements of $\{1, \ldots, n\}$: $i_1, i_2, \ldots$, where $i_j = \sigma(i_{j-1})$ for $j \geq 2$. If we let j grow, this sequence necessarily contains repetitions: suppose that the kth term is the first one which is repeated, that is $i_k = i_j$ with $j < k$. But this means that both i_{k-1} and i_{j-1} are mapped to i_k by σ, and since σ is a bijection, two elements cannot be mapped to i_k, thus it must be that $i_k = i_1$ (the only element which has not yet a preimage). We then see that σ defines a $k-1$-cycle

$$(i_1, \ldots, i_{k-1}).$$

If $k - 1 = n$, we are done. If not, we take i_1' another index not covered by the first cycle, and iterate. This second cycle has to be disjoint from the first one, since σ is a bijection. We then obtain a disjoint cycle representation for σ.

Now take an l-cycle $(i_1, \dots, i_l)$. It can be rewritten as

$$(i_1, \dots, i_l) = (i_1\ i_2)(i_2\ i_3)\cdots(i_{l-1}\ i_l).$$

Indeed, start with the right hand-side: i_l and i_{l-1} are swapped. Thus $(i_1, \dots, i_l)$ is mapped to an l-tuple whose last 2 terms are i_l, i_{l-1}. The next transposition is (i_{l-2}, i_{l-1}), thus both terms are swapped, and we now have as last 3 terms i_{l-1}, i_l, i_{l-2}. The next swap will yield

$$i_{l-2}, i_{l-1}, i_l, i_{l-3}$$

and by iterating this process, we reach the last swap (i_1, i_2), that is

$$i_2, i_3, \dots, i_l, i_1$$

as we wanted to prove. □

The representation has a product of transpositions is not unique, for example

$$(2,5,3,6) = (2,6)(2,3)(2,5) = (5,2)(3,5)(6,3) = (1,7)(2,6)(2,3)(2,5)(1,7).$$

We can however define an invariant of a permutation, called the parity.

Definition 1.20. An element of S_n is said to be even if it can be expressed as a product of an even number of transpositions. It is said odd otherwise.

For this definition to make sense, parity of an element of S_n should be unique, which it is.

Theorem 1.24. *For $n \geq 2$, every element of S_n has a unique parity, even or odd.*

Proof. To prove this, we need to introduce some ordering on the permutations. We call the switching number of a permutation σ the number of ordered pairs (i, j) with $i < j$ but $\sigma(i) > \sigma(j)$. The switching number is an invariant of a permutation. Let t be the switching number of σ, and let τ be an arbitrary transposition, say $\tau = (i\ j)$. Without loss of generality, we may assume that i comes before j in the permutation

$$\sigma(1), \dots, \sigma(n).$$

By applying τ to σ, we switch i and j, and the picture now looks like

$$(1, 2, \dots, \sigma^{-1}(i), \dots, \sigma^{-1}(j), \dots, n) \xrightarrow{\sigma} (\sigma(1), \sigma(2), \dots, i, \dots, j, \dots, \sigma(n))$$

$$\xrightarrow{\tau} (\sigma(1), \sigma(2), \dots, \tau(i), \dots, \tau(j), \dots, \sigma(n))$$

(where the first vector is ordered, but not the second and the third).

To understand the effect of the transposition τ on the switching number of σ (that is we are computing the switching number of $\tau\sigma$ and see how it differs from that of σ), we need to remember that we are looking at all the ordered pairs (k, l), $k < l$, in $(1, 2, \dots, n)$:

1. For the ordered pair $(\sigma^{-1}(i), \sigma^{-1}(j))$, when applying σ, either (a) the ordering is preserved (i.e., $i < j$), the switching number thus does not change, however when applying τ, the ordering is reversed, and thus t increases by 1, or (b) the ordering is changed, but τ changes again the ordering, so that t decreases by 1.

2. Let us now assume that $i < j$ (if not do the same with $j > i$). Then for every index l such $i < l < j$, we can look at the non-ordered pairs (i, l) and (l, j). It might be that $\sigma^{-1}(i)$ is either greater or smaller than $\sigma^{-1}(l)$, yielding one ordered pair or the other, and similarly for $\sigma^{-1}(j)$ and $\sigma^{-1}(l)$. Thus each ordered pair might or not contribute to the switching number of σ, but after τ is applied, i and j are reversed, and thus both (i, l) and (l, j) are changed at once. Thus the switching number increases by 2, decreases by 2, or does not change. We can write down the cases explicitly:

$$(\sigma^{-1}(i), \sigma^{-1}(l)), (\sigma^{-1}(l), \sigma^{-1}(j)) \xrightarrow{\sigma} (i, l), (l, j) \xrightarrow{\tau} (j, l), (l, i), \ i < l < j$$

thus the switching number of σ is t including no switch for these 2 pairs, and that of $\tau\sigma$ has two switches for these 2 paires, thus is of $t + 2$.

$$(\sigma^{-1}(i), \sigma^{-1}(l)), (\sigma^{-1}(j), \sigma^{-1}(l)) \xrightarrow{\sigma} (i, l), (j, l) \xrightarrow{\tau} (j, l), (i, l), \ i < l < j$$

and the switching number of σ is here t including one switch for the second pair, and that of $\tau\sigma$ has one switch for the first pair, but none for the second, thus a total of t. The case $(\sigma^{-1}(l), \sigma^{-1}(i)), (\sigma^{-1}(l), \sigma^{-1}(j))$ also gives t, and finally

$$(\sigma^{-1}(l), \sigma^{-1}(i)), (\sigma^{-1}(j), \sigma^{-1}(l)) \xrightarrow{\sigma} (l, i), (j, l) \xrightarrow{\tau} (l, j), (i, l), \ i < l < j$$

has a switching number of t for σ including two switches for these 2 pairs, and $\tau\sigma$ has no switch, thus a total of $t - 2$.

3. All the non-ordered pairs (k, l), where $l < i < j$ and $k < l$ or $k > l$, or $l > j > i$ and $k < l$ or $k > l$ (that is all the cases not conisdered so far) do not induce any change in the switching number, since by swapping i and j, we do not change the ordering of the pairs.

This shows that given a permutation σ with switching number t, composing with one transposition always changes the parity of the switching number. Since the switching number is invariant, this means that it always takes an even number of transpositions applied to σ to have a chance to keep the same switching number. This establishes that the parity of a permutation is either even or odd, but not both. □

Definition 1.21. The set of even permutations forms a subgroup of S_n called the alternating group, denoted by A_n.

Note that if τ is any odd permutation, then the coset τA_n consists entirely of odd permutations, and conversely, if σ is an odd permutation, then $\tau^{-1}\sigma$ is even, so $\sigma \in \tau A_n$. This shows that $|A_n| = |S_n|/2$.

We will encounter the alternating group in the future, but for now, we only need to recall the definition of the symmetric group S_n to prove Cayley's Theorem.

Theorem 1.25. (Cayley's Theorem.) *Every group is isomorphic to a group of permutations.*

Proof. Let S_G be the group of permutations of G. We will prove that every group is isomorphic to a subgroup of S_G. The idea is that each element $g \in G$ corresponds to a permutation of G, namely we need to define a map from G to S_G:

$$\lambda : G \to S_G,\ g \mapsto \lambda(g) = \lambda_g$$

and since λ_g is a bijection on G, we need to define what λ_g does:

$$\lambda_g : G \to G,\ \lambda_g(x) = gx.$$

For justifying that λ_g is indeed a bijection, it is enough to see that g^{-1} exists since G is a group (try to write down the definition of injection and surjection).

We are left to check that λ is an injective group homomorphism. Injectivity again comes from G being a group, for if $\lambda_g(x) = \lambda_h(x)$ for all $x \in G$, then it has to be true that $gx = hx$ when $x = 1$.

Now

$$(\lambda(a) \circ \lambda(b))(x) = (\lambda_a \circ \lambda_b)(x) = a(bx) = \lambda_{ab}(x) = \lambda(ab)(x)$$

for all x, so that $\lambda(a) \circ \lambda(b) = \lambda(ab)$ which concludes the proof. □

Examples 1.26.

1. Consider the group $\{0, 1\}$ of integers modulo 2. The group element 0 corresponds to the identity permutation, while the group element 1 corresponds to the permutation (12).

2. Let us consider the group $\{0, 1, 2\}$ of integers modulo 3 to get a less simple example. Again 0 corresponds to the identity permutation, 1 corresponds to the permutation (123), and 2 to the permutation (132). To see that it makes sense, you may want to check that the arithmetic works similarly on both sides. For example, we can say that $1 + 1 = 2$ on the one hand, now on the other hand, we have $(123)(123) = (132)$.

3. One can check that the dihedral group D_3 of order 6 is isomorphic to S_3 (this can be done for example by working out the multiplication table for each group).

The key point in the proof of Cayley's Theorem is the way the function λ_g is defined. We see that for $x \in G$, g "acts" (via λ_g) on x by multiplication.

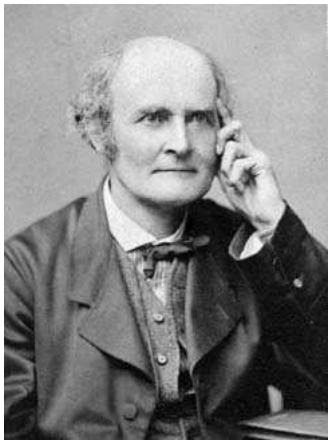


Figure 1.2: Arthur Cayley (1821-1895): he was the first to define the concept of a group in the modern way. Before him, groups referred to permutation groups.

Definition 1.22. The group G acts on the set X if for all $g \in G$, there is a map

$$G \times X \to X,\ (g, x) \mapsto g \cdot x$$

such that

1. $h \cdot (g \cdot x) = (hg) \cdot x$ for all $g, h \in G$, for all $x \in X$.
2. $1 \cdot x = x$ for all $x \in X$.

The first condition says that we have two laws, the group law between elements of the group, and the action of the group on the set, which are compatible.

Examples 1.27. Let us consider two examples where a group G acts on itself.

1. Every group acts on itself by left multiplication. This is called the regular action.
2. Every group acts on itself by conjugation. Let us write this action as

   $$g \cdot x = gxg^{-1}.$$

   Let us check the action is actually well defined. First, we have that

   $$h \cdot (g \cdot x) = h \cdot (gxg^{-1}) = hgxg^{-1}h^{-1} = (hg)xg^{-1}h^{-1} = (hg) \cdot x.$$

   As for the identity, we get

   $$1 \cdot x = 1x1^{-1} = x.$$

Similarly to the notion of kernel for a homomorphism, we can define the kernel of an action.

Definition 1.23. The kernel of an action $G \times X \to X,\ (g,x) \mapsto g \cdot x$ is given by

$$\text{Ker} = \{g \in G,\ g \cdot x = x \text{ for all } x\}.$$

This is the set of elements of G that fix everything in X. When the group G acts on itself, that is $X = G$ and the action is the conjugation, we have

$$\text{Ker} = \{g \in G,\ gxg^{-1} = x \text{ for all } x\} = \{g \in G,\ gx = xg \text{ for all } x\}.$$

This is called the center of G, denoted by $Z(G)$.

Definition 1.24. Suppose that a group G acts on a set X. The orbit $B(x)$ of x under the action of G is defined by

$$B(x) = \{g \cdot x,\ g \in G\}.$$

This means that we fix an element $x \in X$, and then we let g act on x when g runs through all the elements of G. By the definition of an action, $g \cdot x$ belongs to X, so the orbit gives a subset of X.

It is important to notice that orbits partition X. Clearly, one has that $X = \cup_{x \in X} B(x)$. But now, assume that one element x of X belongs to two orbits $B(y)$ and $B(z)$, then it means that $x = g \cdot y = g' \cdot z$, which in turn implies, due to the fact that G is a group, that

$$y = g^{-1}g' \cdot z,\ z = (g')^{-1}g \cdot y.$$

In words, that means that y belongs to the orbit of z, and vice-versa, z belongs to the orbit of y, and thus $B(y) = B(z)$. We can then pick a set of representatives for each orbit, and write that

$$X = \sqcup B(x),$$

where the disjoint union is taken over a set of representatives.

Definition 1.25. Suppose that a group G acts on a set X. We say that the action is transitive, or that G acts transitively on X if there is only one orbit, namely, for all $x, y \in X$, there exists $g \in G$ such that $g \cdot x = y$.

Definition 1.26. The stabilizer of an element $x \in X$ under the action of G is defined by

$$\text{Stab}(x) = \{g \in G,\ g \cdot x = x\}.$$

Given x, the stabilizer $\text{Stab}(x)$ is the set of elements of G that leave x fixed. One may check that this is a subgroup of G. We have to check that if $g, h \in \text{Stab}(x)$, then $gh^{-1} \in \text{Stab}(x)$. Now

$$(gh^{-1}) \cdot x = g \cdot (h^{-1} \cdot x)$$

by definition of action. Since $h \in \mathrm{Stab}(x)$, we have $h \cdot x = x$ or equivalently $x = h^{-1} \cdot x$, so that

$$g \cdot (h^{-1} \cdot x) = g \cdot x = x,$$

which shows that $\mathrm{Stab}(x)$ is a subgroup of G.

Examples 1.28.

1. The regular action (see the previous example) is transitive, and for all $x \in X = G$, we have $\mathrm{Stab}(x) = \{1\}$, since x is invertible and we can multiply $g \cdot x = x$ by x^{-1}.

2. Let us consider the action by conjugation, which is again an action of G on itself ($X = G$): $g \cdot x = gxg^{-1}$. The action has no reason to be transitive in general, and for all $x \in X = G$, the orbit of x is given by

$$B(x) = \{gxg^{-1},\ g \in G\}.$$

This is called the conjugacy class of x. Let us now consider the stabilizer of an element $x \in X$:

$$\mathrm{Stab}(x) = \{g \in G,\ gxg^{-1} = x\} = \{g \in G,\ gx = xg\},$$

which is the centralizer of x, that we denote by $C_G(x)$.

Note that we can define similarly the centralizer $C_G(S)$ where S is an arbitrary subset of G as the set of elements of G which commute with everything in S. The two extreme cases are: if $S = \{x\}$, we get the centralizer of one element, if $S = G$, we get the center $Z(G)$.

Theorem 1.26. (The Orbit-Stabilizer Theorem). *Suppose that a group G acts on a set X. Let $B(x)$ be the orbit of $x \in X$, and let $\mathrm{Stab}(x)$ be the stabilizer of x. Then the size of the orbit is the index of the stabilizer, that is*

$$|B(x)| = [G : \mathrm{Stab}(x)].$$

If G is finite, then

$$|B(x)| = |G|/|\mathrm{Stab}(x)|.$$

In particular, the size of an orbit divides the order of the group.

Proof. Recall first that $[G : \mathrm{Stab}(x)]$ counts the number of left cosets of $\mathrm{Stab}(x)$ in G, that is the cardinality of

$$G/\mathrm{Stab}(x) = \{g\mathrm{Stab}(x),\ g \in G\}.$$

Note that cosets of $\mathrm{Stab}(x)$ are well-defined since we saw that $\mathrm{Stab}(x)$ is a subgroup of G. The idea of the proof is to build a function between the sets $B(x)$ and $G/\mathrm{Stab}(x)$ which is a bijection. That the cardinalities are the same will then follow.

Take $y \in B(x)$, that is $y = g \cdot x$ for some $g \in G$. We define a map

$$f : B(x) \to G/\mathrm{Stab}(x),\ y = g \cdot x \mapsto g\mathrm{Stab}(x).$$

Before checking that this map is a bijection, we need to check that it is well defined. Indeed, for a given y, there is no reason for the choice of g to be unique (there is in general no bijection between G and $B(x)$). Suppose that

$$y = g_1 \cdot x = g_2 \cdot x$$

then

$$g_2^{-1} g_1 \cdot x = x \iff g_1 \mathrm{Stab}(x) = g_2 \mathrm{Stab}(x).$$

The equivalence is the characterization of having two equal cosets. This is exactly what we wanted: the image by f does not depend on the choice of g, and if we choose two different g's, their image falls into the same coset.

The surjectivity of f is immediate.

We conclude the proof by showing the injectivity. Let us assume that $f(y_1) = f(y_2)$ for $y_1 = g_1 \cdot x \in B(x)$, $y_2 = g_2 \cdot x \in B(x)$. Thus

$$g_1 \mathrm{Stab}(x) = g_2 \mathrm{Stab}(x) \iff g_2^{-1} g_1 \in \mathrm{Stab}(x) \iff g_2^{-1} g_1 \cdot x = x \iff g_1 \cdot x = g_2 \cdot x.$$

□

Let G be a finite group. We consider again as action the conjugation ($X = G$), given by: $g \cdot x = gxg^{-1}$. Recall that orbits under this action are given by

$$B(x) = \{gxg^{-1},\ g \in G\}.$$

Let us notice that x always is in its orbit $B(x)$ (take $g = 1$). Thus if we have an orbit of size 1, this means that

$$gxg^{-1} = x \iff gx = xg$$

and we get an element x in the center $Z(G)$ of G. In words, elements that have an orbit of size 1 under the action by conjugation are elements of the center.

Recall that the orbits $B(x)$ partition X:

$$X = \sqcup B(x)$$

where the disjoint union is over a set of representatives. We get

$$\begin{aligned} |G| &= \sum |B(x)| \\ &= |Z(G)| + \sum |B(x)| \\ &= |Z(G)| + \sum [G : \mathrm{Stab}(x)], \end{aligned}$$

where the second equality comes by splitting the sum between orbits with 1 element and orbits with at least 2 elements, while the third follows from the Orbit-Stabilizer Theorem. By remembering that $\mathrm{Stab}(x) = C_G(x)$ when the action is the conjugation, we can alternatively write

$$|G| = |Z(G)| + \sum [G : C_G(x)].$$

This formula is called the class equation.

Example 1.29. Consider the dihedral D_4 of order 8, given by

$$D_4 = \{1, s, r, r^2, r^3, rs, r^2s, r^3s\},$$

with $s^2 = 1$, $r^4 = 1$ and $srs = r^{-1}$. We have that the center $Z(D_4)$ of D_4 is $\{1, r^2\}$ (just check that $r^2s = sr^2$). There are three conjugacy classes given by

$$\{r, r^3\},\ \{rs, r^3s\}, \{s, r^2s\}.$$

Thus

$$|D_4| = 8 = |Z(D_4)| + |B(r)| + |B(rs)| + |B(s)|.$$

The following result has many names: Burnside's lemma, Burnside's counting theorem, the Cauchy-Frobenius lemma or the orbit-counting theorem. This result is not due to Burnside himself, who only quoted it. It is attributed to Frobenius.

Theorem 1.27. (Orbit-Counting Theorem). *Let the finite group G act on the finite set X, and denote by X^g the set of elements of X that are fixed by g, that is $X^g = \{x \in X,\ g \cdot x = x\}$. Then*

$$\textit{number of orbits} = \frac{1}{|G|} \sum_{g \in G} |X^g|,$$

that is the number of orbits is the average number of points left fixed by elements of G.

Proof. We have

$$\begin{aligned} \sum_{g \in G} |X^g| &= |\{(g, x) \in G \times X,\ g \cdot x = x\}| \\ &= \sum_{x \in X} |\mathrm{Stab}(x)| \\ &= \sum_{x \in X} |G|/|B(x)| \end{aligned}$$

by the Orbit-Stabilizer Theorem. We go on:

$$\begin{aligned} \sum_{x \in X} |G|/|B(x)| &= |G| \sum_{x \in X} 1/|B(x)| \\ &= |G| \sum_{B \in \text{ set of orbits}} \sum_{x \in B} \frac{1}{|B|} \\ &= |G| \sum_{B \in \text{ set of orbits}} 1 \end{aligned}$$

which concludes the proof. Note that the second equality comes from the fact that we can write X as a disjoint union of orbits. □

1.8 The Sylow theorems

We look at orders of groups again, but this time paying attention to the occurrence of prime factors. More precisely, we will fix a given prime p, look at the partial factorization of the group order n as $n = p^r m$ where p does not divide m, and study the existence of subgroups of order p or a power of p. In a sense, this is trying to establish some kind of converse for Lagrange's Theorem. Recall that Lagrange's Theorem tells that the order of a subgroup divides the order of the group. Here we conversely pick a divisor of the order of the group, and we try to find a subgroup with order the chosen divisor.

Definition 1.27. Let p be a prime. The group G is said to be a p-group if the order of each element of G is a power of p.

Examples 1.30. We have already encountered several 2-groups.

1. We have seen in Example 1.15 that the cyclic group C_4 has elements of order 1,2 and 4, while the direct product $C_2 \times C_2$ has elements of order 1 and 2.

2. The dihedral group D_4 is also a 2-group.

Definition 1.28. If $|G| = p^r m$, where p does not divide m, then a subgroup P of order p^r is called a Sylow p-subgroup of G. Thus P is a p-subgroup of G of maximum possible size.

The first thing we need to check is that such a subgroup of order p^r indeed exists, which is not obvious. This will be the content of the first Sylow theorem. Once we have proven the existence of a subgroup of order p^r, it has to be a p-group, since by Lagrange's Theorem the order of each element must divide p^r.

We need a preliminary lemma.

Lemma 1.28. *If $n = p^r m$ where p is prime, then $\binom{n}{p^r} \equiv m \mod p$. Thus if p does not divide m, then p does not divide $\binom{n}{p^r}$.*

Proof. We have to prove that

$$\binom{n}{p^r} \equiv m \mod p,$$

after which we have that if p does not divide m, the $m \not\equiv 0 \mod p$ implying that $\binom{n}{p^r} \not\equiv 0 \mod p$ and thus p does not divide $\binom{n}{p^r}$.

Let us use the binomial expansion of the following polynomial

$$(x+1)^{p^r} = \sum_{k=0}^{p^r} \binom{p^r}{k} x^{p^r-k} 1^k \equiv x^{p^r} + 1 \mod p$$

where we noted that all binomial coefficients but the first and the last are divisible by p. Thus

$$(x+1)^{p^r m} \equiv (x^{p^r} + 1)^m \mod p$$

Figure 1.3: Ludwig Sylow (1832-1918)

which we can expand again into

$$\sum_{k=0}^{p^r m} \binom{p^r m}{k} x^{p^r m-k} \equiv \sum_{k=0}^{m} \binom{m}{k} (x^{p^r})^{m-k} \mod p.$$

We now look at the coefficient of x^{p^r} on both sides:

- on the left, take $k = p^r(m-1)$, to get $\binom{p^r m}{p^r}$,
- on the right, take $k = m-1$, to get $\binom{m}{m-1} = m$.

The result follows by identifying the coefficients of x^{p^r}. □

We are ready to prove the first Sylow Theorem.

Theorem 1.29. (1st Sylow Theorem). *Let G be a finite group of order $p^r m$, p a prime such that p does not divide m, and r some positive integer. Then G has at least one Sylow p-subgroup.*

Proof. The idea of the proof is to actually exhibit a subgroup of G of order p^r. For that, we need to define a clever action of G on a carefully chosen set X. Take the set

$$X = \{\text{subsets of } G \text{ of size } p^r\}$$

and for action that G acts on X by left multiplication. This is clearly a well-defined action. We have that

$$|X| = \binom{p^r m}{p^r}$$

which is not divisible by p (by the previous lemma). Recall that the action of G on X induces a partition of X into orbits:

$$X = \sqcup B(S)$$

where the disjoint union is taken over a set of representatives. Be careful that here S is an element of X, that is S is a subset of size p^r. We get

$$|X| = \sum |B(S)|$$

and since p does not divide $|X|$, it does not divide $\sum |B(S)|$, meaning that there is at least one S for which p does not divide $|B(S)|$. Let us pick this S, and denote by P its stabilizer.

The subgroup P which is thus by choice the stabilizer of the subset $S \in X$ of size p^r whose orbit size is not divisible by p is our candidate: we will prove it has order p^r.

$|P| \geq p^r$. Let us use the Orbit-Stabilizer Theorem, which tells us that

$$|B(S)| = |G|/|P| = p^r m/|P|.$$

By choice of the S we picked, p does not divide $|B(S)|$, that is p does not divide $p^r m/|P|$ and $|P|$ has to be a multiple of p^r, or equivalently p^r divides $|P|$.

$|P| \leq p^r$. Let us define the map λ_x, $x \in S$, by

$$\lambda_x : P \to S,\ g \mapsto \lambda_x(g) = gx.$$

In words, this map goes from P, which is a subgroup of G, to S, which is an element of X, that is a subset of G with cardinality p^r. Note that this map is well-defined since $gx \in S$ for any $x \in S$ and any $g \in P$ by definition of P being the stabilizer of S. It is also clearly injective ($gx = hx$ implies $g = h$ since x is an element of the group G and thus is invertible). If we have an injection from P to S, that means $|P| \leq |S| = p^r$.

□

Example 1.31. Consider the general group $G = GL_n(\mathbb{F}_p)$ of $n \times n$ invertible matrices with coefficients in $\mathbb{F}_p$, which denotes integers mod p, p a prime. Let us compute a Sylow p-subgroup of G. For that, we first need to know the cardinality of G. This is a classical combinatorial computation: to build an invertible matrix with coefficients in $\mathbb{F}_p$, the first column can be anything but the whole zero vector, thus $p^n - 1$ choices, the 2nd column can be anything but a multiple of the first column, thus $p^n - p$ choices, the 3rd column can be anything but a linear combination of the first 2 columns, thus $p^n - p^2$ choices,..., thus the cardinality is

$$\begin{aligned} |G| = |GL_n(\mathbb{F}_p)| &= (p^n - 1)(p^n - p)(p^n - p^2)\cdots(p^n - p^{n-1}) \\ &= (p^n - 1)p(p^{n-1} - 1)p^2(p^{n-2} - 1)\cdots p^{n-1}(p - 1) \\ &= p^{n(n-1)/2}\prod_{k=1}^{n}(p^k - 1). \end{aligned}$$

Figure 1.4: Augustin Louis Cauchy (1789-1857)

Clearly $p^{n(n-1)/2}$ is the highest power of p that divides G, we thus have to find a subgroup of G of that order. Consider the set $U(p,n)$ of $n \times n$ upper triangular matrices with every diagonal coefficient at 1 and elements of $\mathbb{F}_p$ above. This is clearly a subgroup of G, since such matrices are invertible and form a group. Its cardinality is $p^{n(n-1)/2}$ as wanted (there are $n(n-1)/2$ coefficients above the diagonal, which can take any value mod p).

Corollary 1.30. (Cauchy Theorem). *If the prime p divides the order of G, then G has an element of order p.*

Proof. Let P be a Sylow p-subgroup of G (which exists by the 1st Sylow Theorem), and pick $x \neq 1$ in P. The order $|x|$ of x is a power of p by definition of a p-group, say $|x| = p^k$. Then $x^{p^{k-1}}$ has order p. □

The above corollary gives some converse to Lagrange's Theorem. The one below gives an alternative definition of a finite p-group. It is tempting to use it as a definition of p-group, however it cannot be used for infinite groups.

Corollary 1.31. *A finite group G is a p-group if and only if the order of G is a power of p.*

Proof. If the order of G is not a power of p, then it is divisible by some other prime q, in which case G contains an element of order q by Cauchy's Theorem, which contradicts the definition of p-group.

The converse is clear using Lagrange's Theorem. □

Example 1.32. Let us consider again the group $H = U(p,n)$ of $n \times n$ upper triangular matrices with every diagonal coefficient at 1 and elements of $\mathbb{F}_p$ above computed in Example 1.31. We know that $|H| = p^{n(n-1)/2}$, thus it is a p-group. Now let us consider the union

$$G = \bigcup_{n \geq 1} U(p,n).$$

This forms an infinite p-group, since every element has order a power of p.

Now that we know that at least one Sylow p-subgroup exists, let us derive a result on the number n_p of Sylow p-subgroups in a group G.

We need again a preliminary lemma.

Lemma 1.32. *Let H and K be arbitrary finite subgroups of a group G. Then*

$$|HK| = \frac{|H||K|}{|H \cap K|}.$$

Note that if K is further assumed to be normal in G, then this result can be deduced from the second isomorphism Theorem.

Proof. Consider the map

$$f : H \times K \to HK,\ (h,k) \mapsto hk.$$

Since f is surjective, $|HK| \leq |H \times K| < \infty$ since H and K are finite, and thus HK is finite. Let $h_1k_1, \ldots, h_dk_d$ be the distinct elements of HK. Then $H \times K$ is the disjoint union of the $f^{-1}(h_ik_i)$, $i = 1, \ldots, d$. Now we can check that

$$f^{-1}(hk) = \{(hg, g^{-1}k),\ g \in H \cap K\}$$

and this set has cardinality

$$|f^{-1}(hk)| = |H \cap K|.$$

Thus

$$|H \times K| = d|H \cap K|$$

which concludes the proof. □

Theorem 1.33. (2nd Sylow Theorem). *Let G be a finite group of order $p^r m$, p a prime such that p does not divide m, and r some positive integer. Denote by n_p the number of Sylow p-subgroups of G. Then*

$$n_p \equiv 1 \mod p.$$

Proof. Consider the set

$$X = \{ \text{ all Sylow } p - \text{subgroups of } G \ \}$$

whose cardinality $|X|$ is denoted by n_p. By the 1st Sylow Theorem, this set is non-empty and there exists at least one Sylow p-subgroup P in X, whose order is p^r. We can thus let P act on X by conjugation, i.e., $g \cdot Q = gQg^{-1}$, $g \in P$, $Q \in X$. Note that in the case where P is the only Sylow p-subgroup, then we can take $Q = P$.

By the Orbit-Stabilizer Theorem, we have that the orbit $B(Q)$ of any Sylow p-subgroup Q in X has cardinality

$$|B(Q)| = |P|/|\text{Stab}(Q)| = p^r/|\text{Stab}(Q)|.$$

In particular, the size of an orbit of any Sylow p-subgroups divides p^r, meaning it has to be either 1 or a power of p.

Let us recall that the set X is partitioned by the orbits $B(Q)$ under the action of P, so that the cardinality of X is:

$$|X| = \sum |B(Q)| = \sum |B(Q')| + \sum |B(Q'')|$$

where Q' and Q'' denote subgroups whose orbit has respectively one element or at least two elements. Since p divides the second sum, we have

$$|X| \equiv \text{ number of orbits of size 1} \mod p.$$

To conclude the proof, we thus have to show that there is only one Sylow p-subgroup whose orbit has size 1, namely P itself (it is clear that P has an orbit of size 1, since conjugating P by itself will not give another subgroup).

Let us assume there is another Sylow p-subgroup Q whose orbit has only one element, namely (recall that one element is always in its orbit):

$$gQg^{-1} = Q,\ g \in P,$$

which translates into

$$gQ = Qg \text{ for all } g \in P \iff PQ = QP.$$

This easily implies that PQ is a subgroup of G, and by the previous lemma

$$|PQ| = \frac{p^r p^r}{|P \cap Q|}$$

implying that $|PQ|$ is a power of p, say p^c for some c which cannot be bigger than r, since $|G| = p^r m$. Thus

$$p^r = |P| \leq |PQ| \leq p^r$$

so that $|PQ| = p^r$ and thus $|P| = |PQ|$, saying that Q is included in P. But both Q and P have same cardinality of p^r, so $Q = P$. □

The third of the Sylow Theorems tells us that all Sylow p-subgroups are conjugate.

Theorem 1.34. (3rd Sylow Theorem). *Let G be a finite group of order $p^r m$, p a prime such that p does not divide m, and r some positive integer. Then all Sylow p-subgroups are conjugate.*

Proof. Let P be a Sylow p-subgroup of G and let R be a p-group of G. We will prove that R (being a p-group in general) is contained in a conjugate of P.

Let R act by multiplication on the set Y of left cosets of P:

$$Y = \{gP,\ g \in G\}.$$

It is a well-defined action (it is multiplication in the group G).

We want to prove that there is an orbit of size 1 under this action. By Lagrange's Theorem, we know that

$$|Y| = |G|/|P| = \frac{p^r m}{p^r} = m$$

and thus p does not divide $|Y|$ by assumption on m. By writing that we have a partition of Y by its orbits, we get

$$|Y| = \sum |B(y)|$$

and there exists one orbit $B(y)$ whose size is not divisible by p. By the Orbit-Stabilizer Theorem, we have that the size of every orbit divides $|R|$, which has order a power of p (by a corollary of the 1st Sylow Theorem), so every orbit size must divide p, which gives as only possibility that there is an orbit of size 1.

Let $gP \in Y$ be the element whose orbit size is 1. We have

$$h \cdot gP = gP$$

for $h \in R$, since gP belongs to its orbit. Thus

$$g^{-1}hg \in P \iff h \in gPg^{-1}$$

for all h in R. We have just proved that the p-group R is contained in a conjugate of P.

All we needed for the proof is that R is a p-group, so the same proof holds for the case of a Sylow p-subgroup, for which we get that R is contained in a conjugate of P, and both have same cardinality, which concludes the proof.

We will use the fact that the proof works for R a p-group in general for proving one corollary. □

Corollary 1.35. *1. Every p-subgroup of G is contained in a Sylow p-subgroup.*

2. The number n_p of Sylow p-subgroups divides m.

Proof. 1. Now we know that if P is a Sylow p-subgroup, then so is gPg^{-1}, $g \in G$, by the above theorem. The proof of the theorem itself shows that any p-group is included in gPg^{-1} and we are done.

2. Let the group G act by conjugation on the set of its subgroups. In particular, G acts on the Sylow p-subgroup P, and the orbit of P has size the number of Sylow p-subgroups in G, denoted by n_p. By the Orbit-Stabilizer Theorem, n_p divides $|G| = p^r m$. But p cannot be a prime factor of n_p since $n_p \equiv 1 \mod p$, from which it follows that n_p must divide m.

□

1.9 Simple groups

We will now see a few applications of the Sylow Theorems, in particular to determine the structure of so-called simple groups.

Definition 1.29. A group G is simple if $G \neq \{1\}$ and the only normal subgroups of G are G itself and $\{1\}$.

Finite simple groups are important because in a sense they are building blocks of all finite groups, similarly to the way prime numbers are building blocks of the integers. This will be made clearer in the coming section by the Jordan-Hölder Theorem. Infinite simple groups exist and can be found for example among Lie groups, but we will concentrate here on finite groups.

The case of simple abelian groups is easy to understand. Suppose that G is a simple abelian group. Note that G cannot be $\{1\}$. Now G being abelian, all its subgroups are normal. On the other hand, G being simple, its only normal subgroups are $\{1\}$ and itself, leaving as only solution that G has only two subgroups, namely $\{1\}$ and G. Thus G has to be a cyclic group of prime order.

We now start looking at non-abelian simple groups. We start with some preliminary results.

Proposition 1.36. *If P is a non-trivial finite p-group, then P has a non-trivial center.*

Proof. Let P act on itself by conjugation. The orbits of this action are the conjugacy classes of P, and we have that x belongs to an orbit of size 1 if and only if x belongs to the center $Z(P)$.

By the Orbit-Stabilizer, the size of any orbit must divide $|P|$, which is a power of p by a corollary of the 1st Sylow Theorem.

If it were true that the center is trivial, that is $Z(P) = \{1\}$, then that means there is only one orbit of size 1, and thus all the other orbits must have size that divides p, namely they are congruent to $0 \mod p$. Thus

$$|P| = |Z(P)| + \sum |B| \equiv 1 \mod p,$$

where the sum is over orbits of size at least 2. This is clearly a contradiction, which concludes the proof. □

Lemma 1.37. *The group P is a normal Sylow p-subgroup of a group G if and only if P is the unique Sylow p-subgroup of G.*

Proof. We know from the 3rd Sylow Theorem that the Sylow p-subgroups form a single conjugacy class. Then P is the unique Sylow p-subgroup means that P is the only element in the conjugacy class, and thus it satisfies

$$gPg^{-1} = P,$$

for every $g \in G$, which exactly means that P is a normal subgroup of G. Conversely, if P is normal, then $gPg^{-1} = P$ for all $g \in G$, which means that

when we conjugate P by any element of G, P is the only Sylow p-subgroup in its conjugacy class. Since all the Sylow p-subgroups belong to the same conjugacy class, P is the only Sylow p-subgroup. □

Thanks to the two above results, we can now prove that a non-abelian simple group must have more than one Sylow p-subgroup.

Proposition 1.38. *Let G be a finite group which is non-abelian and simple. If the prime p divides $|G|$, then the number n_p of Sylow p-subgroups is strictly bigger than 1.*

Proof. Let us look at the prime factors appearing in the order of G.

- If p is the only prime factor of $|G|$, then $|G|$ must be a power of p, that is G is a non-trivial p-group (it is non-trivial by definition of simple and a p-group by a corollary of the 1st Sylow Theorem). Now the above proposition tells us that its center $Z(G)$ is non-trivial as well. Since $Z(G)$ is a normal subgroup of G and G is simple, it must be that $G = Z(G)$, which contradicts the assumption that G is non-abelian.
- We then know that $|G|$ is divisible by at least two distinct primes. So if P is a Sylow p-subgroup, then

  $$\{1\} < P < G,$$

  where the second inclusion is strict since the order of G is divisible by two primes.

  If there were only one Sylow p-subgroup, namely $n_p = 1$, then this Sylow p-subgroup would be normal by the above lemma, which contradicts the simplicity of G.

□

Let us see if we can be more precise by refining the assumptions on the order of the group G we consider. The group G can be either abelian or non-abelian, though the results on simplicity are more interesting for non-abelian groups.

Proposition 1.39. *Let G be a group of order pq, where p and q are distinct primes.*

1. *If $q \not\equiv 1 \mod p$, then G has a normal Sylow p-subgroup.*
2. *If both $q \not\equiv 1 \mod p$ and $p \not\equiv 1 \mod q$, then G is cyclic.*
3. *G is not simple.*

Proof. 1. By Lemma 1.37, saying that G has a normal Sylow p-subgroup is the same as saying that there is a unique Sylow p-subgroup. This is now indeed the case, since the number n_p of Sylow p-subgroups has to satisfy both $n_p \equiv 1 \mod p$ and $n_p \mid q$ by the Sylow Theorems. Since q is prime, n_p is either 1 or q. It cannot be that $n_p = q$, since it would imply that $q \equiv 1 \mod p$ which contradicts the assumption.

2. By the previous point, the group G has a normal Sylow p-subgroup P and a normal Sylow q-subgroup Q, both of them cyclic (since they are of prime order). Let us write them respectively $P = \langle x \rangle$ and $Q = \langle y \rangle$. Since both P and Q are normal, with $P \cap Q = \{1\}$, we have that $xy = yx$ (we have seen that before, but the argument goes like that: take the element $xyx^{-1}y^{-1}$ and show that it belongs to $P \cap Q$ by normality of P and Q). Thanks to this commutativity property, we have that $(xy)^n = x^n y^n$ and the order of xy is pq, showing that G is cyclic with generator xy.

3. Without loss of generality, we can assume that $p > q$ so that p does not divide $q - 1$ which can be rewritten as

$$q \not\equiv 1 \mod p.$$

   By the first point, we know that G has a normal Sylow p-group, and thus G cannot be simple.

□

Here is another family of groups which are not simple. The proof contains an interesting combinatorial argument!

Proposition 1.40. *Let G be a group of order $|G| = p^2 q$ where p and q are two distinct primes. Then G contains either a normal Sylow p-subgroup or a normal Sylow q-subgroup. In particular, G is not simple.*

Proof. Recall that having a normal Sylow p-subgroup (resp. q-subgroup) is the same as saying there is a unique Sylow p-subgroup (resp. q-subgroup). Suppose that the claim is not true, that is both the number of Sylow p-subgroups n_p and the number of Sylow q-subgroups n_q are bigger than 1. Let us start this proof by counting the number of elements of order q in G.

If a Sylow q-subgroup has order q, it is cyclic and can be generated by any of its elements which is not 1. This gives $q - 1$ elements of order q per Sylow q-subgroup of G. Conversely, if y has order q, then the cyclic group it generates is a Sylow q-subgroup, and any two distinct Sylow q-subgroups have trivial intersection. Thus

$$\text{number of elements of order } q = n_q(q - 1).$$

Now we know from the Sylow Theorems that $n_q \mid p^2$, thus n_q is either p of p^2 ($n_q = 1$ is ruled out by the fact that we do a proof by contradiction).

- $n_q = p^2$: then the number of elements of order NOT q is

$$p^2 q - p^2(q - 1) = p^2.$$

  On the other hand, if P is a Sylow p-subgroup, then it also contains p^2 elements, and all of them have order not q, so that we can conclude that P actually contains all elements of order not q, which implies that we have only one Sylow p-subgroup, yielding the wanted contradiction.

$\|G\|$	abelian	non-abelian
p	C_p simple	not possible
p^r	not simple	not simple since $\|Z(G)\| > 1$
pq	not simple	not simple
p^2q	not simple	not simple

Table 1.3: C_p refers to a cyclic group of prime order.

- $n_q = p$: We know from Sylow Theorems that

  $$n_q \equiv 1 \mod q \Rightarrow p \equiv 1 \mod q \Rightarrow p > q,$$

  but also that

  $$n_p \mid q$$

  and since q is prime, that leaves $n_p = 1$ or $n_p = q$ and thus $n_p = q$. As before

  $$n_p \equiv 1 \mod p \Rightarrow q \equiv 1 \mod p \Rightarrow q > p.$$

  This concludes the proof.

□

We have thus shown that the situation is easy for simple abelian groups. For non-abelian groups, we have seen two cases ($|G| = pq$ and $|G| = p^2q$) where groups are not simple. To find a non-abelian group which is simple, one has to go to groups of order at least 60. Indeed, it has been proven that the smallest non-abelian simple group is the alternating group A_5 of order 60, this is the group of even permutations of a finite set. This result is attributed to Galois (1831). It is not an easy task to determine the list of simple groups, and in fact, the classification of finite simple groups was only accomplished in 1982 (there has been some controversy as to whether the proof is correct, given its length - tens of thousands of pages - and complexity).

1.10 The Jordan-Hölder Theorem

We have mentioned when introducing simple groups in the previous section that they can be seen as building blocks for decomposing arbitrary groups. This will be made precise in this section.

Definition 1.30. Let G be a group, and let $G_0, \ldots, G_n$ be subgroups of G such that

1. $G_n = \{1\}$ and $G_0 = G$,
2. $G_{i+1} \trianglelefteq G_i$, $i = 0, \ldots, n-1$.

Then the series

$$\{1\} = G_n \trianglelefteq G_{n-1} \trianglelefteq \cdots \trianglelefteq G_0 = G$$

is called a subnormal series for G.

Suppose that G_{i+1} is not a maximal normal subgroup of G_i, then we can refine the subnormal series by inserting a group H such that $G_{i+1} \lhd H \lhd G_i$, and we can repeat this process hoping it will terminate (it will if G is finite, it may not otherwise).

Definition 1.31. Let G be a group, and let $G_0, \ldots, G_n$ be subgroups of G such that

1. $G_n = \{1\}$ and $G_0 = G$,
2. $G_{i+1} \lhd G_i$, $i = 0, \ldots, n-1$, such that G_{i+1} is a maximal normal subgroup of G_i.

Then the series

$$\{1\} = G_n \lhd G_{n-1} \lhd \cdots \lhd G_0 = G$$

is called a composition series for G. The factor groups G_i/G_{i+1} are called the factors of the composition series, whose length is n.

Another way of stating the condition G_{i+1} is a maximal normal subgroup of G_i is to say that G_i/G_{i+1} is simple, $i = 0, \ldots, n-1$. To see that asks a little bit of work. This result is sometimes called the 4rth isomorphism theorem.

Theorem 1.41. (Correspondence Theorem). *Let N be a normal subgroup of G and let H be a subgroup of G containing N. Then the map*

$$\psi : \{\textit{subgroups of } G \textit{ containing } N\} \to \{\textit{subgroups of } G/N\},\ H \mapsto \psi(H) = H/N$$

is a bijection. Furthermore, H is a normal subgroup of G if and only if H/N is a normal subgroup of G/N.

Proof. We first prove that ψ is a bijection.

Injectivity. If $H_1/N = H_2/N$, then cosets in each subgroup are the same, that is for any $h_1 \in H_1$, we have $h_1 N = h_2 N$ for some $h_2 \in H_2$, implying that $h_2^{-1} h_1 \in N \subset H_2$ and thus $h_1 \in H_2$, showing that $H_1 \subseteq H_2$. By repeating the same argument but reverting the role of H_1 and H_2, we get $H_2 \subseteq H_1$ and thus $H_1 = H_2$.

Surjectivity. Let Q be a subgroup of G/N and let $\pi : G \to G/N$ be the canonical projection. Then

$$\pi^{-1}(Q) = \{a \in G,\ aN \in Q\}.$$

This is a subgroup of G containing N and

$$\psi(\pi^{-1}(Q)) = \{aN,\ aN \in Q\} = Q.$$

We are left to prove that $H \trianglelefteq G \iff H/N \trianglelefteq G/N$. Assume thus that $H \trianglelefteq G$. For any $a \in G$, we have to show that

$$(aN)(H/N)(aN)^{-1} = H/N.$$

Now for any $hN \in H/N$, we have

$$(aN)(hN)(aN)^{-1} = (aha^{-1})N \in H/N$$

and we are done.

Conversely, suppose that $H/N \trianglelefteq G/N$. Consider the homomorphism

$$a \mapsto (aN)(H/N)$$

which is the composition of the canonical projection π of G onto G/N, and the canonical projection of G/N onto $(G/N)/(H/N)$ (the latter makes sense since $H/N \trianglelefteq G/N$). We now want to show that H is the kernel of this map, which will conclude the proof since the kernel of a group homomorphism is normal.

An element a is in the kernel if and only if $(aN)(H/N) = H/N$, that is if and only if $aN \in H/N$, or equivalently $aN = hN$ for some $h \in H$. Since N is contained in H, this means aN is in H and thus so is a, which is what we wanted to prove. □

Let us now go back to the composition series of G. If G/N is simple, then by definition it has only trivial normal subgroups, namely N and G/N. Now using the Correspondence Theorem, the normal subgroups N and G/N exactly correspond to the normal subgroups N and G in G, which shows that N is the maximal normal subgroup of G.

The Jordan-Hölder Theorem will tell us that if G has a composition series, then the resulting composition length n and the simple composition factors G_i/G_{i+1} are unique up to isomorphism and rearrangement. This for example shows that if G_1 and G_2 are two groups with different composition factors, then they cannot be isomorphic.

Lemma 1.42. *Let G be a group with composition series*

$$\{1\} = G_n \lhd G_{n-1} \lhd \cdots \lhd G_0 = G.$$

Then for any normal subgroup K of G, if we remove the duplicates from the series

$$\{1\} = K \cap G_n \trianglelefteq K \cap G_{n-1} \trianglelefteq \cdots \trianglelefteq K \cap G_0 = K,$$

the result is a composition series for K of length at most n.

Proof. We need to show that $K \cap G_{i+1} \lhd K \cap G_i$ and that the group $(K \cap G_i)/(K \cap G_{i+1})$ is simple for all i.

Let $x \in K \cap G_i$ and $g \in K \cap G_{i+1}$. Then $xgx^{-1} \in K$ since by assumption K is a normal subgroup of G, and $xgx^{-1} \in G_{i+1}$ since $G_{i+1} \lhd G_i$. Thus $xgx^{-1} \in K \cap G_{i+1}$ which proves that $K \cap G_{i+1} \lhd K \cap G_i$.

We now look at the quotient group $(K \cap G_i)/(K \cap G_{i+1})$. Since G_i/G_{i+1} is simple, G_{i+1} is a maximal normal subgroup of G_i, and thus the only normal subgroups of G_i that contain G_{i+1} are G_i and G_{i+1}.

Recall that $K \cap G_i$ is normal in G_i (it is the kernel of the canonical projection of G to G/K restricted to G_i), so that we get

$$G_{i+1} \lhd (K \cap G_i)G_{i+1} \lhd G_i.$$

For the first normal inclusion, compute that for $kg \in (K \cap G_i)G_{i+1}$ we have

$$kgG_{i+1}g^{-1}k^{-1} = kG_{i+1}k^{-1} \subseteq G_{i+1}$$

since $k \in G_i$ and G_{i+1} is normal in G_i. For the second normal inclusion, we have for $g \in G_i$ that

$$g(K \cap G_i)G_{i+1}g^{-1} = (K \cap G_i)gG_{i+1}g^{-1}$$

since $K \cap G_i$ is normal in G_i and

$$(K \cap G_i)gG_{i+1}g^{-1} \subseteq (K \cap G_i)G_{i+1}$$

since $G_{i+1} \lhd G_i$.

Thus either $G_{i+1} = (K \cap G_i)G_{i+1}$ or $(K \cap G_i)G_{i+1} = G_i$. Using the second isomorphism theorem (with $G_{i+1} \lhd G_i$ and $(K \cap G_i) \leq G_i$), we have

$$(K \cap G_i)G_{i+1}/G_{i+1} \simeq (K \cap G_i)/(K \cap G_i \cap G_{i+1}) = (K \cap G_i)/(K \cap G_{i+1}).$$

We can see that if $G_{i+1} = (K \cap G_i)G_{i+1}$, then $K \cap G_i = K \cap G_{i+1}$ and we have a duplicate to remove. If $(K \cap G_i)G_{i+1} = G_i$, then

$$G_i/G_{i+1} \simeq (K \cap G_i)/(K \cap G_{i+1})$$

and thus $(K \cap G_i)/(K \cap G_{i+1})$ is simple. □

Theorem 1.43. (Jordan-Hölder Theorem). *Let G be a group that has a composition series. Then any two composition series for G have the same length. Moreover, if*

$$\{1\} = G_n \lhd G_{n-1} \lhd \cdots \lhd G_0 = G$$

and

$$\{1\} = H_n \lhd H_{n-1} \lhd \cdots \lhd H_0 = G$$

are two composition series for G, there exists a permutation τ such that $G_i/G_{i+1} \simeq H_{\tau(i)}/H_{\tau(i)+1}$.

Proof. The proof will be on induction on the length of a composition series. Suppose that G is a group with a composition series of length 1. Then the subnormal series

$$G \rhd \{1\}$$

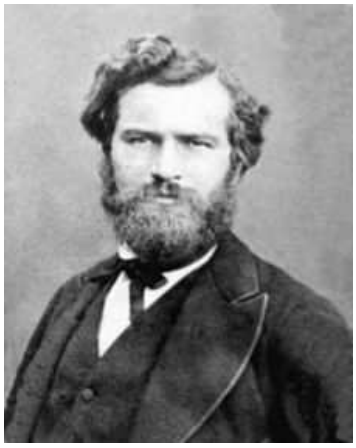



Figure 1.5: Camille Jordan (1838–1922) and Otto Hölder (1859–1937)

cannot be refined, so it must be a composition series. In particular $G \simeq G/\{1\}$ is simple. This is also the only composition series for G and so all the assertions are true for length 1.

Suppose now that $n > 1$ and that the claims are true for composition series of length up till $n-1$. Let G be a group with composition series of length n, say

$$\{1\} = G_n \lhd G_{n-1} \lhd \cdots \lhd G_0 = G$$

(so that $G_i \neq G_{i+1}$ for each i). Now let

$$\{1\} = H_m \lhd H_{m-1} \lhd \cdots \lhd H_0 = G$$

be a composition series for G (again $H_i \neq H_{i+1}$ for each i).

We first have to show that $m = n$ after which we discuss the unicity of the decomposition.

(Proof that $m = n$). The idea of the proof goes as follows: to use the induction hypothesis, we need to get a composition series of length smaller than n, that is, we need to identify the first composition factors, which we will using the above lemma. Concretely, we first exclude the case when $G_1 = H_1$, then compute a composition series of length $n-2$ for $H_1 \cap G_1$, which will indeed be the second composition factor. We then use the second composition series of G to get another composition series for $H_1 \cap G_1$ whose length depends on m, that we can compare to the known one.

If $G_1 = H_1$, then by the induction hypothesis applied to G_1, we have $n-1 = m-1$, we have a suitable permutation τ of the $n-1$ factors, and we are done.

Suppose then that $H_1 \neq G_1$. Since both G_1 and H_1 are maximal normal in G, we see that $H_1 \lhd G_1H_1 \lhd G$ with $H_1 \neq G_1H_1$ since we assumed $H_1 \neq G_1$. Thus $G_1H_1 = G$, from which we conclude by the 2nd isomorphism theorem that

$$G_1H_1/H_1 \simeq G/H_1 \simeq G_1/(H_1 \cap G_1).$$

Since G/H_1 is simple, we get that $G_1/(H_1 \cap G_1)$ is simple as well. Now by the above lemma, upon removing duplicates from the series

$$\{1\} = H_1 \cap G_n \trianglelefteq \cdots \trianglelefteq H_1 \cap G_0 = H_1,$$

we get a composition series for H_1 of length at most n and thus upon removing duplicates

$$\{1\} = H_1 \cap G_n \trianglelefteq \cdots \trianglelefteq H_1 \cap G_1$$

is a composition series for $H_1 \cap G_1$ of length at most $n-1$. Since $G_1/(H_1 \cap G_1)$ is simple, it follows that upon removing duplicates

$$\{1\} = H_1 \cap G_n \trianglelefteq \cdots \trianglelefteq H_1 \cap G_1 \triangleleft G_1$$

is a composition series for G_1. But then

$$G_1 \rhd G_2 \rhd \cdots \rhd G_n = \{1\}$$

and

$$G_1 \rhd H_1 \cap G_1 \trianglerighteq H_1 \cap G_2 \trianglerighteq \cdots \trianglerighteq H_1 \cap G_n = \{1\}$$

are both composition series for G_1, with the first series of length $n-1$. By induction hypothesis, both series have the same length. Since $G_1 \neq H_1 \cap G_1$ (recall that we assumed $H_1 \neq G_1$), any duplication must occur later in the series. Let

$$G_1 = K_1 \rhd K_2 = H_1 \cap G_1 \rhd K_3 \rhd \cdots \rhd K_n = \{1\}$$

denote the composition series for G_1 of length $n-1$ that results from removing the duplicates. By hypothesis, there exists a permutation α such that $G_i/G_{i+1} \simeq K_{\alpha(i)}/K_{\alpha(i)+1}$ for each $i = 1, \ldots, n-1$. Set α not to move the index 0, then

$$G = G_0 \rhd G_1 \rhd G_2 \rhd \cdots \rhd G_n = \{1\}$$

and

$$G = K_0 \rhd G_1 = K_1 \rhd K_2 = H_1 \cap G_1 \rhd K_3 \rhd \cdots \rhd K_n = \{1\}$$

are composition series of length n for G and α is a permutation such that $G_i/G_{i+1} \simeq K_{\alpha(i)}/K_{\alpha(i)+1}$ for each $i = 0, \ldots, n-1$. Moreover, we have found a composition series for $H_1 \cap G_1$ of length $n-2$.

Let us now repeat similar computations for the composition series

$$G = H_0 \rhd H_1 \rhd \cdots \rhd H_m = \{1\}$$

and the normal subgroup G_1 of G. Again by the above lemma, upon removing the duplicates from the series

$$G_1 = H_0 \cap G_1 \trianglerighteq H_1 \cap G_1 \trianglerighteq \cdots \trianglerighteq H_m \cap G_1 = \{1\}$$

we obtain a composition series for G_1, so that upon removing the duplicates

$$H_1 \cap G_1 \trianglerighteq \cdots \trianglerighteq H_m \cap G_1 = \{1\}$$

yields a composition series for $H_1 \cap G_1$. Now since $H_1 \cap G_1$ has a composition series of length $n-2$, namely

$$K_2 = H_1 \cap G_1 \rhd \cdots \rhd K_n = \{1\},$$

we apply the induction hypothesis to $H_1 \cap G_1$ to conclude that all composition series of $H_1 \cap G_1$ have length $n-2$, and so in particular the preceding composition series

$$H_1 \cap G_1 \trianglerighteq \cdots \trianglerighteq H_m \cap G_1 = \{1\}$$

has length $n-2$. We cannot conclude yet, since we do not know how many terms there are in function of m in the above composition series (we need to get rid of the duplicates).

Since we know from the 2nd isomorphism theorem that $H_1/(H_1 \cap G_1) \simeq H_1G_1/G_1 = G_0/G_1$, which is a simple group, it follows that $H_1/(H_1 \cap G_1)$ is simple. Thus upon the removal of the duplicates from

$$H_1 \rhd H_1 \cap G_1 \trianglerighteq \cdots \trianglerighteq H_m \cap G_1 = \{1\}$$

the result is a composition series for H_1 of length $n-1$ (we added the term H_1 to the composition series for $H_1 \cap G_1$ of length $n-2$). Also

$$H_1 \rhd H_2 \rhd \cdots \rhd H_m = \{1\}$$

is another composition series for H_1. Since the first series has length $n-1$, by our induction hypothesis, the second series must also have length $n-1$. Since its length is $m-1$, it follows that $m = n$.

(Unicity of the composition factors). Again by induction hypothesis on H_1, we have a permutation β of the $n-1$ composition factors (which can be extended to n factors by setting $\beta(0) = 0$.) Namely, let L_i, $i = 1, 2, \ldots, n$ denote the distinct terms in the series

$$H_1 \rhd H_1 \cap G_1 \rhd H_2 \cap G_1 \rhd \cdots \rhd H_n \cap G_1 = \{1\}$$

so that $L_1 = H_1$ and $L_2 = H_1 \cap G_1$. Then we have composition series

$$G = H_0 \rhd H_1 \rhd \cdots H_n = \{1\} \text{ and } G = L_0 \rhd L_1 \rhd \cdots L_n = \{1\}$$

of length n for G and there exists a permutation β of $\{0, 1, \ldots, n-1\}$ such that $H_i/H_{i+1} \simeq L_{\beta(i)}/L_{\beta(i)+1}$ for each $i = 0, 1, \ldots, n-1$.

We are almost done but for the fact that we need an isomorphism between H_i/H_{i+1} and $G_{\beta(i)}/G_{\beta(i)+1}$ instead of having $H_i/H_{i+1} \simeq L_{\beta(i)}/L_{\beta(i)+1}$. We recall that we already have a permutation α such that $G_i/G_{i+1} \simeq K_{\alpha(i)}/K_{\alpha(i)+1}$. We are thus left to find one between L_i/L_{i+1} and K_i/K_{i+1}.

Finally, since $K_2 = L_2 = H_1 \cap G_1$, we have two composition series for G:

$$\begin{array}{ccccccccc} G\rhd & G_1\rhd & H_1 \cap G_1\rhd & K_3\rhd & \cdots & K_{n-1}\rhd & K_n = \{1\} \\ G\rhd & H_1\rhd & H_1 \cap G_1\rhd & L_3\rhd & \cdots & L_{n-1}\rhd & L_n = \{1\}. \end{array}$$

We may apply the induction hypothesis to $H_1 \cap G_1$ to obtain the existence of a permutation γ of $\{2, 3, \ldots, n-1\}$ such that for each i in this set we have $K_i/K_{i+1} \simeq L_{\gamma(i)}/L_{\gamma(i)+1}$. We have already seen that $G/G_1 \simeq H_1/(H_1 \cap G_1)$ and $G/H_1 \simeq G_1/(H_1 \cap G_1)$, so we may extend γ to a permutation of $\{0, 1, \ldots, n-1\}$ by setting $\gamma(0) = 1$ and $\gamma(1) = 0$. Then since

$$K_0 = G = L_0,\ K_1 = G_1,\ L_1 = H_1,\ K_2 = L_2 = H_1 \cap G_1,$$

we have

$$K_i/K_{i+1} \simeq L_{\gamma(i)}/L_{\gamma(i)+1},\ i = 0, \ldots, n-1.$$

In summary, we have $m = n$, and for $\tau = \beta^{-1}\gamma\alpha$, we have

$$G_i/G_{i+1} \simeq H_{\tau(i)}/H_{\tau(i)+1},\ i = 0, \ldots, n-1.$$

This concludes the proof. □

Example 1.33. The cyclic group C_{12} has three composition series

$$C_1 \lhd C_2 \lhd C_6 \lhd C_{12},\ C_1 \lhd C_2 \lhd C_4 \lhd C_{12},\ C_1 \lhd C_3 \lhd C_6 \lhd C_{12}$$

and all of them have the same length. Furthermore, the factor groups appearing are

$$\{C_2, C_3, C_2\}, \{C_2, C_2, C_3\}, \{C_3, C_2, C_2\}$$

which are indeed the same up to permutation.

Corollary 1.44. (Fundamental Theorem of arithmetic). *Let $n > 1$ be a positive integer. Then there exist unique primes $p_1 < p_2 < \cdots < p_k$ and unique positive integers $r_1, \ldots, r_k$ such that $n = p_1^{r_1} p_2^{r_2} \cdots p_k^{r_k}$.*

Proof. Let $G = \langle g \rangle$ be a cyclic group of order n. Then every subgroup of G is normal, and there is a unique subgroup of size d for each positive divisor d of n. Let d be the largest proper divisor of n, and let G_1 be the unique subgroup of G of size d. Then G/G_1 is simple and cyclic, hence of prime order. We may repeat this construction on the cyclic subgroup G_1, so by induction, we obtain a composition series

$$G = G_0 \lhd G_1 \lhd G_2 \lhd \cdots \lhd G_m = \{1\}$$

for G with G_i/G_{i+1} of prime order p_i for each i. Thus

$$\begin{aligned} n &= |G| \\ &= |G/G_1||G_1| \\ &= |G/G_1||G_1/G_2| \cdots |G_{m-1}/G_m||G_m| \\ &= p_1 p_2 \cdots p_{m-1}. \end{aligned}$$

The uniqueness of the prime decomposition of n follows from the Jordan-Hölder Theorem applied to G. □

1.11 Solvable and nilpotent groups

Let us start by introducing a notion stronger than normality.

Definition 1.32. A subgroup H of the group G is called characteristic in G if for each automorphism f of G, we have

$$f(H) = H.$$

We may write H char G.

This is stronger than normal since normality corresponds to choose for f the conjugation by an element of g.

Note that f restricted to H a characteristic subgroup (denoted by $f|_H$) is an automorphism of H (it is an endomorphism by definition of H being characteristic).

Here are a few immediate properties of characteristic subgroups.

Lemma 1.45. *Let G be a group, and let H, K be subgroups of G.*

1. *If H is characteristic in K and K is characteristic in G, then H is characteristic in G (being characteristic is transitive).*
2. *If H is characteristic in K, and K is normal in G, then H is normal in G.*

Proof.

1. Note that by assumption $H \leq K \leq G$. Let ϕ be an automorphism of G. Since K is characteristic in G, then $\phi(K) = K$ by definition, and thus $\phi|_K$ is an automorphism of K. Now $\phi|_K(H) = H$ since H is characteristic in K. But $\phi|_K$ is just the restriction of ϕ (recall $H \leq K$), so $\phi(H) = H$.
2. Consider the automorphism of K given by $k \mapsto gkg^{-1}$, $g \in G$, which is well defined since K is normal in G. For any choice of g, we get a different automorphism of K, which will always preserve H since H is characteristic in K, and thus $gHg^{-1} \subset H$ which proves that H is normal in G.

□

Let us introduce a new definition, that will give us an example of characteristic subgroup.

Definition 1.33. The commutator subgroup G' of a group G is the subgroup generated by all commutators

$$[x, y] = xyx^{-1}y^{-1}.$$

It is also called the derived subgroup of G.

Let us make a few remarks.

- By the subgroup generated by all commutators, we mean that by definition we take the smallest subgroup of G containing all the commutators. It is thus closed by construction.

- The product of two or more commutators need not be a commutator. It is known that the least order of a finite group for which there exists two commutators whose product is not a commutator is 96.
- Note that the inverse $[x,y]^{-1}$ of $[x,y]$ is given by $[x,y]^{-1} = [y,x] = yxy^{-1}x^{-1}$.

Here are a list of properties of the commutator subgroup G'.

Lemma 1.46. *Let G' be the commutator subgroup of G.*

1. *G' is characteristic in G.*
2. *G is abelian if and only if G' is trivial.*
3. *G/G' is abelian.*
4. *If N is normal in G, then G/N is abelian if and only if $G' \leq N$.*

Proof. 1. To show that G' is characteristic in G, we have to show that $f(G') = G'$ for f any automorphism of G. Now

$$f([x,y]) = f(xyx^{-1}y^{-1}) = f(x)f(y)f(x)^{-1}f(y)^{-1} = [f(x), f(y)].$$

2. We have that G' is trivial if and only if $xyx^{-1}y^{-1} = 1$ which exactly means that $xy = yx$.

3. Since G' is characteristic, it is also normal in G, and G/G' is a group. We are left to prove it is an abelian group. Take two elements (that is two cosets) $G'x$ and $G'y$ in G/G'. We have that $G'xG'y = G'yG'x \iff G'xy = G'yx$ by definition of the law group on G/G'. Now

$$G'xy = G'yx \iff xy(yx)^{-1} \in G' \iff xyx^{-1}y^{-1} \in G',$$

which holds by definition.

4. Let us assume that N is normal in G. We have that G/N is a group, and G/N is abelian if and only if for Nx, Ny two cosets we have

$$NxNy = NyNx \iff Nxy = Nyx \iff xy(yx)^{-1} \in N \iff xyx^{-1}y^{-1} \in N$$

which exactly tells that each commutator must be in N.

□

We can iterate the process of taking commutators:

$$G^{(0)} = G,\ G^{(1)} = G',\ G^{(2)} = (G')', \ldots, G^{(i+1)} = (G^{(i)})', \ldots$$

The process may or may not reach $\{1\}$.

Definition 1.34. The group G is said to be solvable if $G^{(r)} = 1$ for some r. We then have a normal series

$$\{1\} = G^{(r)} \trianglelefteq G^{(r-1)} \trianglelefteq \cdots \trianglelefteq G^{(0)} = G$$

called the derived series of G. The term"solvable" historically refers to Galois theory and the question of "solvability" of quintic equations, as we will see later.

We have already seen the notion of subnormal series in the previous section. By normal series, we mean a serie where not only each group is normal in its successor, but also each group is normal in the whole group, namely each $G^{(i)}$ is normal in G. We have indeed such series here using the fact that the commutator subgroup is a characteristic subgroup, which is furthermore a transitivity property.

Let us make a few remarks about the definition of solvable group.

Lemma 1.47. *1. Every abelian group is solvable.*

2. A group G both simple and solvable is cyclic of prime order.

3. A non-abelian simple group G cannot be solvable.

Proof. 1. We know that G is abelian if and only if G' is trivial. We thus get the normal series

$$G^{(0)} = G \rhd G^{(1)} = \{1\}.$$

2. If G is simple, then its only normal subgroups are $\{1\}$ and G. Since G' is characteristic and thus normal, we have either $G' = \{1\}$ or $G' = G$. The latter cannot possibly happen, since then the derived serie cannot reach $\{1\}$ which contradicts the fact that G is solvable. Thus we must have that $G' = \{1\}$, which means that G is abelian. We conclude by remembering that an abelian simple group must be cyclic of order a prime p.

3. If G is non-abelian, then G' cannot be trivial, thus since G is simple, its only normal subgroups can be either $\{1\}$ or $\{G\}$, thus G' must be either one of the other, and it cannot be $\{1\}$, so it must be G. Thus the derived series never reaches $\{1\}$ and G cannot be solvable.

□

There are several ways to define solvability.

Proposition 1.48. *The following conditions are equivalent.*

1. G is solvable, that is, it has a derived series

$$\{1\} = G^{(r)} \trianglelefteq G^{(r-1)} \trianglelefteq \cdots \trianglelefteq G^{(0)} = G.$$

2. G has a normal series

$$\{1\} = G_r \trianglelefteq G_{r-1} \trianglelefteq \cdots \trianglelefteq G_0 = G$$

where all factors, that is all quotient groups G_i/G_{i+1} are abelian.

3. G has a subnormal series

$$\{1\} = G_r \trianglelefteq G_{r-1} \trianglelefteq \cdots \trianglelefteq G_0 = G$$

where all factors, that is all quotient groups G_i/G_{i+1} are abelian.

Proof. That 1. $\Rightarrow$ 2. is clear from Lemma 1.46 where we proved that G/G' is abelian, where G' is the commutator subgroup of G.

That 2. $\Rightarrow$ 3. is also clear since the notion of normal series is stronger than subnormal series.

What we need to prove is thus that 3. $\Rightarrow$ 1. Starting from G, we can always compute G', then $G^{(2)}, \ldots$.To prove that G has a derived series, we need to check that $G^{(s)} = \{1\}$ for some s. Suppose thus that G has a subnormal series

$$1 = G_r \trianglelefteq G_{r-1} \trianglelefteq \cdots \trianglelefteq G_0 = G$$

where all quotient groups G_i/G_{i+1} are abelian. For $i = 0$, we get $G_1 \trianglelefteq G$ and G/G_1 is abelian. By Lemma 1.46, we know that G/G_1 is abelian is equivalent to $G' \leq G_1$. By induction, let us assume that $G^{(i)} \leq G_i$, that is taking i times the derived subgroup of G is a subgroup which is contained in the ith term G_i of the subnormal series, and see what happens with $G^{(i+1)}$. We have that $G^{(i+1)} = (G^{(i)})' \leq G_i'$ by induction hypothesis (and noting that if $H \subset G$ then $H' \subset G'$, since all the commutators in H surely belong to those of G). Furthermore, $G_i' \leq G_{i+1}$ since G_i/G_{i+1} is abelian. Thus $G^{(r)} \leq G_r = \{1\}$. $\square$

Let us see what are the properties of subgroups and quotients of solvable groups.

Proposition 1.49. *Subgroups and quotients of a solvable group are solvable.*

Proof. Let us first consider subgroups of a solvable groups. If H is a subgroup of a solvable group G, then H is solvable because $H^{(i)} \leq G^{(i)}$ for all i, and in particular for r such that $H^{(r)} \leq G^{(r)} = \{1\}$ which proves that the derived series of H terminates.

Now consider N a normal subgroup of a solvable group G. The commutators of G/N are cosets of the form $xNyNx^{-1}Ny^{-1}N = xyx^{-1}y^{-1}N$, so that the commutator subgroup $(G/N)'$ of G/N satisfies $(G/N)' = G'N/N$ (we cannot write G'/N since there is no reason for N to be a subgroup of G'). Inductively, we have $(G/N)^{(i)} = G^{(i)}N/N$. Since G is solvable, $G^{(r)} = \{1\}$ and thus $(G/N)^{(r)} = N/N = \{1\}$ which shows that G/N is solvable. $\square$

Example 1.34. Consider the symmetric group S_4. It has a subnormal series

$$\{1\} \lhd C_2 \times C_2 \lhd A_4 \lhd S_4,$$

where A_4 is the alternating group of order 12 (given by the even permutations of 4 elements) and $C_2 \times C_2$ is the Klein group of order 4 (corresponding to the permutations $1, (12)(34), (13)(24), (14)(23)$). The quotient groups are

$$\begin{array}{rcl} C_2 \times C_2/\{1\} & \simeq C_2 \times C_2 & \text{abelian of order 4} \\ A_4/C_2 \times C_2 & \simeq C_3 & \text{abelian of order 3} \\ S_4/A_4 & \simeq C_2 & \text{abelian of order 2.} \end{array}$$

We finish by introducing the notion of a nilpotent group. We will skip the general definition, and consider only finite nilpotent groups, for which the following characterization is available.

Proposition 1.50. *The following statements are equivalent.*

1. *G is the direct product of its Sylow subgroups.*
2. *Every Sylow subgroup of G is normal.*

Proof. If G is the direct product of its Sylow subgroups, that every Sylow subgroup of G is normal is immediate since the factors of a direct product are normal subgroups.

Assume that every Sylow subgroup of G is normal, then by Lemma 1.37, we know that every normal Sylow p-subgroup is unique, thus there is a unique Sylow p_i-subgroup P_i for each prime divisor p_i of $|G|$, $i = 1, \dots, k$. Now by Lemma 1.32, we have that $|P_1P_2| = |P_1||P_2|$ since $P_1 \cap P_2 = \{1\}$, and thus $|P_1 \cdots P_k| = |P_1| \cdots |P_k| = |G|$ by definition of Sylow subgroups. Since we work with finite groups, we deduce that G is indeed the direct product of its Sylow subgroups, having that $G = P_1 \cdots P_k$ and $P_i \cap \prod_{j \neq i} P_j$ is trivial. □

Definition 1.35. A finite group G which is the product of its Sylow subgroups, or equivalently by the above proposition satisfies that each of its Sylow subgroup is normal is called a nilpotent group.

Corollary 1.51. *Every finite abelian group and every finite p-group is nilpotent.*

Proof. A finite abelian group surely has the property that each of its Sylow subgroup is normal, so it is nilpotent.

Now consider P a finite p-group. Then by definition P has only one Sylow subgroup, namely itself, so it is the direct product if its Sylow subgroups and thus is nilpotent. □

Finite nilpotent groups are also nicely described with respect to their normalizer.

Proposition 1.52. *If G is a finite nilpotent group, then no proper subgroup H of G is equal to its normalizer $N_G(H) = \{g \in G,\ gH = Hg\}$.*

Proof. Let H be a proper subgroup of G, and let n be the largest index such that $G_n \subseteq H$ (such index exists since G is nilpotent). There exists $a \in G_{n+1}$ such that $a \notin G$ (since H is a proper subgroup). Now for every $h \in H$, the cosets aG_n and hG_n commute (since $G_{n+1}/G_n \subseteq Z(G/G_n)$), namely:

$$G_nah = (G_na)(G_nh) = (G_nh)(G_na) = G_nha$$

and thus there is some $h' \in G_n \subseteq H$ for which

$$ah = h'ha$$

that is

$$aha^{-1} = h'h \in H.$$

Thus $a \in N_G(H)$ and $a \notin H$. □

Here is the definition for possibly infinite nilpotent groups.

Definition 1.36. A central series for a group G is a normal series

$$\{1\} = G_n \trianglelefteq G_{n-1} \trianglelefteq \cdots \trianglelefteq G_0 = G$$

such that $G_i/G_{i+1} \subseteq Z(G/G_{i+1})$ for every $i = 0, \ldots, n-1$. An arbitrary group G is said to be nilpotent if it has a central series. The smallest n such that G has a central series of length n is called the nilpotency class of G, and G is said to be nilpotent of class n.

Example 1.35. Abelian groups are nilpotent of class 1, since

$$\{1\} = G_1 \trianglelefteq G_0 = G$$

is a normal series for G and for $i = 0$ we have $G/\{1\} \simeq G \subseteq Z(G)$.

Nilpotent groups in general are discussed with solvable groups since they can be described with normal series, and one can prove that they are solvable. Indeed, if $G_i/G_{i+1} \subseteq Z(G/G_{i+1})$, then the elements of G_i/G_{i+1} commute with each other, since they commute with everything in G/G_{i+1}, thus G_i/G_{i+1} is abelian. It is not true that solvable groups are necessarily nilpotent (see Exercises for an example).

The main definitions and results of this chapter are

- **(1.1-1.2).** Definitions of: group, subgroup, group homomorphism, order of a group, order of an element, cyclic group.
- **(1.3-1.4).** Lagrange's Theorem. Definitions of: coset, normal subgroup, quotient group
- **(1.5).** 1st, 2nd and 3rd Isomorphism Theorems.
- **(1.6).** Definitions of: external (semi-)direct product, internal (semi-)direct product.
- **(1.7).** Cayley's Theorem, the Orbit-Stabilizer Theorem, the Orbit-Counting Theorem. Definitions of: symmetric group, group action, orbit, transitive action, stabilizer, centralizer. That the orbits partition the set under the action of a group
- **(1.8).** Definition: p-group, Sylow p-subgroup. The 3 Sylow Theorems, Cauchy Theorem
- **(1.9).** Definition: simple group. Applications of the Sylow Theorems.
- **(1.10).** Definitions: subnormal series, composition series. Jordan-Hölder Theorem.
- **(1.11).** Definitions: characteristic subgroup, commutator subgroup, normal and derived series, solvable group, finite nilpotent group.

Chapter 2

Exercises on Group Theory

Exercises marked by (*) are considered difficult.

2.1 Groups and subgroups

Exercise 1. Let G be a group and let H be a nonempty subset of G. We have seen that the two following statements are equivalent:

a) H is a subgroup of G,

b) b_1) $x, y \in H \Rightarrow xy \in H$
b_2) $x \in H \Rightarrow x^{-1} \in H$.

1. Show that b_1) is not sufficient to show that H is a subgroup of G.
2. Show that however, if G is a finite group, then b_1) is sufficient.

Answer.

1. Consider for example the group $G = \mathbb{Q}^*$ with multiplication. Then the set $\mathbb{Z}$ with multiplication satisfies that if $x, y \in \mathbb{Z}$ then $xy \in \mathbb{Z}$. However, $\mathbb{Z}$ is not a group with respect to multiplication since $2 \in \mathbb{Z}$ but $1/2$ is not in $\mathbb{Z}$.
2. Let $x \in H$. Then take the powers $x, x^2, x^3, \ldots$ of x. Since G is finite, there is some n such that $x^n = 1$, and by b_1), $x^n \in H$ thus $1 \in H$, and $x^{n-1} = x^{-1} \in H$.

Exercise 2. Let G be a finite group of order n such that all its non-trivial elements have order 2.

1. Show that G is abelian.

2. Let H be a subgroup of G, and let $g \in G$ but not in H. Show that $H \cup gH$ is a subgroup of G.

3. Show that the subgroup $H \cup gH$ has order twice the order of H.

4. Deduce from the previous steps that the order of G is a power of 2.

Answer.

1. Let $x, y \in G$, x, y not 1. By assumption, $x^2 = y^2 = 1$, which also means that x, y and xy are their own inverse. Now
$$(xy)(xy) = 1 \Rightarrow xy = (xy)^{-1} = y^{-1}x^{-1} = yx.$$

2. First note that $H \cup gH$ contains 1 since $1 \in H$. Let $x, y \in H \cup gH$. Then $x \in H$ or $x \in gH$, and $y \in H$ or $y \in gH$. If both $x, y \in H$, then clearly $xy \in H$ since H is a subgroup. If both $x, y \in gH$, then $x = gh, y = gh'$ and $xy = ghgh' = hh' \in H$ since G is commutative and $g^2 = 1$. If say $x \in H$ and $y \in gH$ (same proof vice-versa), then $xy = xgh = g(xh) \in gH$ since G is commutative. For the inverse, if $x \in H$, then $x^{-1} \in H$ since H is a subgroup. If $x \in gH$, then $x = gh$, and $x^{-1} = h^{-1}g^{-1} = gh$ since G is commutative and all elements have order 2.

3. It is enough to show that the intersection of H and gH is empty. Let $x \in H$ and $x \in gH$. Then $x = gh$ for $h \in H$, so that $xh = gh^2 = g$, which is a contradiction, since $xh \in H$ and g is not in H by assumption.

4. Take h an element of order 2 in G, and take $H = \{1, h\}$. If $G = H$ we are done. If not, there is a g not in H, and by the previous point $H \cup gH$ has order 4. We can now iterate. If $G = H \cup gH$ we are done. Otherwise, $H \cup gH = H'$ is a subgroup of G, and there exists a g' not in H', so that $H' \cup g'H'$ has order 8. One can also write a nice formal proof by induction.

Exercise 3. Let G be an abelian group, and let $x, y \in G$ of finite order. Show that $|xy|$ divides $lcm(|x|, |y|)$, where lcm stands for "least common multiple". Give an example to illustrate that $|xy| \neq lcm(|x|, |y|)$ in general.

Answer. Let $x \in G$ be of order n and let $y \in G$ be of order m. Since G is an abelian group, we have that
$$(xy)^k = x^k y^k$$
for any k. Thus by definition, the order $|xy|$ is the smallest positive k such that $x^k y^k = 1$. Also by definition, the $lcm(|x|, |y|) = lcm(n, m) = N$ satisfies that $N = nn' = mm'$, so that
$$x^N y^N = (x^n)^{n'} (y^m)^{m'} = 1.$$
Finally $(xy)^k = (xy)^N = 1$ and since k is the smallest such positive integer with this property, it must divide N. (If you are not yet convinced, you can add that k is smaller than N, thus you can divide N by k and write $N = kq + r$, $r < k$, which implies that $1 = (xy)^{kq+r} = (xy)^r$, a contradiction to the definition of k.) An easy counter-example is: take $y = x^{-1}$, $x \neq 1$, then $|1| \neq lcm(|x|, |x^{-1}|)$.

Exercise 4. Let G be a group and let H and K be two subgroups of G.

1. Is $H \cap K$ a subgroup of G? If your answer is yes, prove it. If your answer is no, provide a counterexample.
2. Is $H \cup K$ a subgroup of G? If your answer is yes, prove it. If your answer is no, provide a counterexample.

Answer.

1. This is true. It is enough to check that $xy^{-1} \in H \cap K$ for $x, y \in H \cap K$. But since $x, y \in H$, we have $xy^{-1} \in H$ since H is a subgroup, and likewise, $xy^{-1} \in K$ for $x, y \in K$ since K is a subgroup.
2. This is false. For example, take the groups of integers modulo 3 and 2, namely $\mathbb{Z}/3\mathbb{Z}$ and $\mathbb{Z}/2\mathbb{Z}$. Then 2 and 3 are in their union, but 5 is not.

Exercise 5. Show that if G has only one element of order 2, then this element is in the center of G (that is the elements of G which commute with every element in G).

Answer. Let x be the element of order 2. Then yxy^{-1} has also order 2. Thus it must be either 1 or x. If $yxy^{-1} = 1$, then $x = 1$ a contradiction. Thus $yxy^{-1} = x$.

Exercise 6. Let G be a group and H be a subgroup of G. Show that

$$N_G(H) = \{g \in G,\ gH = Hg\}$$

and

$$C_G(H) = \{g \in G,\ gh = hg \text{ for all } h \in H\}$$

are subgroups of G.

Answer. Take $x, y \in N_G(H)$. We have to check that $xy^{-1} \in N_G(H)$, that is, that $xy^{-1}H = Hxy^{-1}$. But $Hxy^{-1} = xHy^{-1}$ since $x \in N_G(H)$, and $xHy^{-1} = xy^{-1}H$ since $yH = Hy \iff y^{-1}H = Hy^{-1}$.

Now take $x, y \in C_G(H)$. We have to check that $xy^{-1}h = hxy^{-1}$ for all $h \in H$. But $hxy^{-1} = xhy^{-1}$ because $x \in C_G(H)$, and $xhy^{-1} = xy^{-1}h$ since $yh = hy \iff y^{-1}h = hy^{-1}$.

2.2 Cyclic groups

Exercise 7. Let $G = \mathbb{Z}_{24}^*$ be the group of invertible elements in $\mathbb{Z}_{24}$. Find all cyclic subgroups of G.

Answer. We have that the size of G is

$$|G| = \varphi(24) = \varphi(3)\varphi(2^3) = 2 \cdot 4 = 8.$$

G is given by the elements that are invertible mod 24, that is, those that are coprime to 24:

$$G = \{1, 5, 7, 11, 13, 17, 19, 23\}.$$

Now

$$\begin{aligned}
\langle 1\rangle &= \{1\},\\
\langle 5\rangle &= \{5, 5^2 = 1\},\\
\langle 7\rangle &= \{7, 7^2 = 49 = 1\},\\
\langle 11\rangle &= \{11, 11^2 = 121 = 1\}\\
\langle 13\rangle &= \{13, 13^2 = 169 = 1\}\\
\langle 17\rangle &= \{17, 17^2 = (-7)^2 = 1\}\\
\langle 19\rangle &= \{19, 19^2 = (-5)^2 = 1\}\\
\langle 23\rangle &= \{23, 23^2 = (-1)^2 = 1
\end{aligned}$$

and there are 8 cyclic subgroups of G, including the trivial subgroup $\{1\}$.

Exercise 8. Let $G = \mathbb{Z}_{20}^*$ be the group of invertible elements in $\mathbb{Z}_{20}$. Find two subgroups of order 4 in G, one that is cyclic and one that is not cyclic.

Answer. As in the exercise above, G contains

$$|G| = \varphi(20) = \varphi(4)\varphi(5) = 2 \cdot 4 = 8.$$

These 8 elements are coprime to 20, that is

$$G = \{1, 3, 7, 9, 11, 13, 17, 19\}.$$

The subgroup

$$\langle 3\rangle = \{3, 3^2 = 9, 3^3 = 7, 3^4 = 21 = 1\}$$

is cyclic of order 4. We have that

$$11, 11^2 = 121 = 1, 19, 19^2 = (-1)^2 = 1, 11 \cdot 19 = (-11) = 9, 9^2 = 81 = 1$$

and

$$\{1, 11, 19, 9\}$$

is a group of order 4 which is not cyclic.

Exercise 9. Let φ be the Euler totient function. Let G be a cyclic group of order n.

1. First show that the order of g^k is
$$|g^k| = n/gcd(k, n).$$
2. Show that if $m|n$, then $\langle g^{n/m}\rangle$ is the unique subgroup of G of order m.
3. Prove that for every factor m of n, the number of elements in G with order m is exactly $\varphi(m)$.

4. Furthermore, show that $\sum_{m|n} \varphi(m) = n$.

Answer. Let $G = \langle g \rangle$ be a cyclic group of order n, so that every element in G is of the form

$$g^k,\ 1 \leq k \leq n.$$

1. Set $m = gcd(k, n)$, so that $k = mk'$, $n = mn'$. If $(g^k)^r = 1$, then $n|kr$, and
$$\frac{n}{m} | \frac{kr}{m}.$$
By definition of m, n/m and k/m are coprime, so that n/m divides r. Hence n/m is the smallest power of g^k such that $(g^k)^{n/m} = 1$ showing that $n/m = |g^k|$.

2. Let H be a subgroup of order m, then $H = \langle g^k \rangle$ with $|H| = m$ and some $k > 0$. We will show first that H can be also generated by an element g^d where $d = gcd(k, n)$, and in particular, we can always write
$$H = \langle g^d \rangle,\ d|n.$$
Since $d|k$, $k = dq$ and $g^k = g^{dq} \in \langle g^d \rangle$ and $\langle g^k n \rangle \subseteq \langle g^d \rangle$. Conversely, $d = gcd(k, n) = kr + ns$ for some r, s and
$$g^d = g^{kr+ns} = g^{kr} \in \langle g^k \rangle$$
and $\langle g^d \rangle \subseteq \langle g^k \rangle$. Now $m = |H| = |g^k| = |g^d| = n/gcd(d, n)$ by the above, and since $d|n$, we get that $m = n/d$, or $d = m/n$.

3. Now for $m|n$, an element is of order m if and only if it is the generator of the only subgroup of G of order m. Now there are as many generators for this subgroup as elements coprime to m, that is $\varphi(m)$.

4. To show that $\sum_{m|n} \varphi(m) = n$, we can sort the elements of G according to their order. Since the order of each element divides n, we have
$$n = \sum_{m|n} \text{nb of elements of order } m = \sum_{m|n} \varphi(m).$$

1.3 Cosets and Lagrange's Theorem

Exercise 10. Let $G = S_3$ be the group of permutations of 3 elements, that is

$$G = \{(1), (12), (13), (23), (123), (132)\}$$

and let $H = \{(1), (12)\}$ be a subgroup. Compute the left and right cosets of H.

Answer. We have

g	gH	Hg
(1)	$\{(1),(12)\}$	$\{(1),(12)\}$
(12)	$\{(1),(12)\}$	$\{(1),(12)\}$
(13)	$\{(13),(123)\}$	$\{(13),(132)\}$
(23)	$\{(23),(132)\}$	$\{(23),(123)\}$
(123)	$\{(13),(123)\}$	$\{(23),(123)\}$
(132)	$\{(23),(132)\}$	$\{(13),(132)\}$

For example, $H(23)$ is $\{(1)(23),(12)(23)\}$. Clearly $(1)(23)=(23)$. Now $(12)(23)$ sends $123 \mapsto 132$ via (23), and then sends $132 \mapsto 231$ via (12), so that finally we have $123 \mapsto 231$ which can be written (123).

Exercise 11. Let G be a finite group and let H and K be subgroups with relatively prime order. Then $H \cap K = \{1\}$.

Answer. Since $H \cap K$ is a subgroup of both H and K, we have

$$|H \cap K| \mid |H|,\ |H \cap K| \mid |K|$$

by Lagrange's Theorem. Since $(|H|,|K|) = 1$, it must be that $|H \cap K| = 1$ implying that $H \cap K = \{1\}$.

2.3 Normal subgroups and quotient group

Exercise 12. Consider the following two sets:

$$T = \left\{ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix},\ a,c \in \mathbb{R}^*,\ b \in \mathbb{R} \right\},\ U = \left\{ \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix},\ b \in \mathbb{R} \right\}.$$

1. Show that T is a subgroup of $GL_2(\mathbb{R})$.
2. Show that U is a normal subgroup of T.

Answer.

1. It is enough to show that if $X, Y \in T$, then $XY^{-1} \in T$. Let

$$X = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix},\ Y = \begin{pmatrix} a' & b' \\ 0 & c' \end{pmatrix}$$

then

$$XY^{-1} = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \frac{1}{a'c'} \begin{pmatrix} c' & -b' \\ 0 & a' \end{pmatrix} = \frac{1}{a'c'} \begin{pmatrix} ac' & -ab'+a'b \\ 0 & a'c \end{pmatrix} \in T$$

2. We have to show that $XYX^{-1} \in U$ when $Y \in U$ and $X \in T$. We have

$$\begin{aligned} XYX^{-1} &= \begin{pmatrix} a' & b' \\ 0 & c' \end{pmatrix} \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \frac{1}{a'c'} \begin{pmatrix} c' & -b' \\ 0 & a' \end{pmatrix} \\ &= \begin{pmatrix} a' & a'b+b' \\ 0 & c' \end{pmatrix} \frac{1}{a'c'} \begin{pmatrix} c' & -b' \\ 0 & a' \end{pmatrix} \\ &= \frac{1}{a'c'} \begin{pmatrix} a'c' & -b'a'+a'(a'b+b') \\ 0 & a'c' \end{pmatrix} \in U. \end{aligned}$$

Exercise 13. Let G be a group, and let H be a subgroup of index 2. Show that H is normal in G.

Answer. If H is of index 2, that means by definition that there are only 2 cosets, say H and g_1H for some g_1 not in H. Note that if $g_1 \neq g_2 \in G$ are not in H,then $g_1g_2 \in H$. Indeed, we have that either $g_1g_2 \in H$ or $g_1g_2 \in g_1H$ (recall that the cosets partition the group), and $g_1g_2 \in g_1H$ is not possible since g_2 is not in H. In other words, if both g_1, g_2 are not in H, then $(g_1g_2)H(g_1g_2)^{-1} \in H$.

Now let $h \in H$, $g \in G$. If $g \in H$, then $ghg^{-1} \in H$ and we are done. If g is not in H, then gh is not in H and by the above remark we have that $ghg^{-1} = (gh)g^{-1} \in H$ (take $g_1 = gh, g_2 = g^{-1}$). Alternatively by the same above remark, since $(g_1g_2)H(g_1g_2)^{-1} \in H$ for every g_1, g_2 not in H, it is enough to wrote g as g_1g_2, say $g_1 = g$ (g is not in H) and $g_2 = g^{-1}h$ (which is not in H either).

Exercise 14. If G_1 is normal in G_2 and G_2 is normal in G_3, then G_1 is normal in G_3. True or false?

Answer. This is wrong (we need the notion of characteristic to get transitivity, this is introduced in the section "Solvable and Nilpotent Groups"). An example is the dihedral group D_4:

$$D_4 = \langle r, f | f^2 = 1, r^4 = 1, fr = r^{-1}f \rangle.$$

The subgroup

$$H = \langle rf, fr \rangle = \{1, rf, r^2, fr\} \simeq C_2 \times C_2$$

is isomorphic to the Klein group. We have that $H \lhd G$. Finally

$$K = \langle rf \rangle = \{1, rf\} \lhd H$$

but K is not normal in G, since $f \cdot rf \cdot f^{-1} = f \cdot rf \cdot f = fr$ which is not in K.

Exercise 15. Let G be a group and let $Z(G)$ be its center (that is the elements of G which commute with every element in G). Show that if $G/Z(G)$ is cyclic then G is abelian. Give an example to show that if $G/Z(G)$ is only abelian, then G does not have to be abelian.

Answer. If $G/Z(G)$ is cyclic, then $G/Z(G) = \langle gZ(G)\rangle$. Let $x, y \in G$, then their corresponding cosets are $xZ(G), yZ(G)$ which can be written

$$xZ(G) = (gZ(G))^k = g^k Z(G),\ yZ(G) = (gZ(G))^l = g^l Z(G)$$

and

$$x = g^k z_1,\ y = g^l z_2,\ z_1, z_2 \in Z(G).$$

Now

$$xy = g^k z_1 g^l z_2 = yx$$

since $z_1, z_2 \in Z(G)$. For example, consider the dihedral group $D_4 = \langle r, f | f^2 = 1, r^4 = 1, fr = r^{-1}f\rangle$. Its center is $Z(D_4) = \{1, r^2\}$. Thus $D_4/Z(D_4)$ is a group of order 4, it contains 4 cosets: $Z(D_4), rZ(D_4), fZ(D_4), rfZ(D_4)$, which is isomorphic to the Klein group, which is abelian but not cyclic.

Exercise 16.
1. Let G be a group. Show that if H is a normal subgroup of order 2, then H belongs to the center of G.
2. Let G be a group of order 10 with a normal subgroup H or order 2. Prove that G is abelian.

Answer.

1. Since H is of order 2, then $H = \{1, h\}$. It is furthermore normal, so that $gHg^{-1} = \{1, ghg^{-1}\}$ is in H, thus $ghg^{-1} = h$ and we are done, since this is saying that h commutes with every $g \in G$.
2. Since H is normal in G, G/H has a group structure, and $|G/H| = |G|/|H| = 10/2 = 5$. Thus the quotient group G/H is a group of order 5, implying that it is cyclic. Now take x, y in G, with respective coset xH, yH. Since the quotient group is cyclic, there exists a coset gH such that $xH = (gH)^k = g^k H$, and $yH = (gH)^l = g^l H$ for some k, l. Thus $x = g^k h$, $y = g^l h'$ for some $h, h' \in H$. We are left to check that $xy = yx$, that is $g^k h g^l h' = g^l h' g^k h$, which is true since we know that $h, h' \in H$ which is contained in the center of G (by the part above).

2.4 The isomorphism theorems

Exercise 17. Consider A the set of affine maps of $\mathbb{R}$, that is

$$A = \{f : x \mapsto ax + b,\ a \in \mathbb{R}^*,\ b \in \mathbb{R}\}.$$

1. Show that A is a group with respect to the composition of maps.
2. Let
$$N = \{g : x \mapsto x + b,\ b \in \mathbb{R}\}.$$
Show that N is a normal subgroup of A.
3. Show that the quotient group A/N is isomorphic to $\mathbb{R}^*$.

Answer.

1. Let $f, g \in A$. Then

$$(f \circ g)(x) = f(ax+b) = a'(ax+b) + b' = a'ax + a'b + b',$$

where $a'a \in \mathbb{R}^*$ thus the closure property is satisfied. The composition of maps is associative. The identity element is given by the identity map since

$$\mathrm{Id} \circ f = f \circ \mathrm{Id} = f.$$

Finally, we need to show that every $f \in A$ is invertible. Take $f^{-1}(x) = a^{-1}x - a^{-1}b$. Then

$$f^{-1} \circ f(x) = f^{-1}(ax+b) = a^{-1}(ax+b) - a^{-1}b = x.$$

2. Let $g \in N$ and let $f \in A$. We have to show that

$$f \circ g \circ f^{-1} \in N.$$

We have

$$f \circ g(a^{-1}x - a^{-1}b) = f(a^{-1}(x) - a^{-1}b + b') = x - b + ab' + b \in N.$$

3. Define the map

$$\varphi : A \to \mathbb{R}^*,\ f(x) = ax + b \mapsto a.$$

It is a group homomorphism since

$$\varphi(f \circ g) = a'a = \varphi(f)\varphi(g).$$

The kernel of φ is N and its image is $\mathbb{R}^*$. By the 1st isomorphism theorem, we thus have that

$$A/N \simeq \mathbb{R}^*.$$

Exercise 18. Use the first isomorphism theorem to

1. show that

$$GL_n(\mathbb{R})/SL_n(\mathbb{R}) \simeq \mathbb{R}^*.$$

2. show that

$$\mathbb{C}^*/U \simeq \mathbb{R}_+^*,$$

where

$$U = \{z \in \mathbb{C}^* \mid |z| = 1\}.$$

3. compute

$$\mathbb{R}/2\pi\mathbb{Z}.$$

Answer.

1. Consider the map:

$$\det : GL_n(\mathbb{R}) \to \mathbb{R}^*,\ X \mapsto \det(X).$$

It is a group homomorphism. Its kernel is $SL_n(\mathbb{R})$, its image is $\mathbb{R}^*$ and thus by the 1st isomorphism theorem, we have

$$GL_n(\mathbb{R})/SL_n(\mathbb{R}) \simeq \mathbb{R}^*.$$

2. Consider the map

$$\exp : \mathbb{C}^* \to \mathbb{R}_+^*.$$

It is a group homomorphism. Its kernel is U, and its image is $\mathbb{R}_+^*$ and thus by the 1st isomorphism theorem, we have

$$\mathbb{C}^*/U \simeq \mathbb{R}_+^*.$$

3. Define the map

$$f : \mathbb{R} \to \mathbb{C}^*,\ x \mapsto e^{ix}.$$

It is a group homomorphism. Its kernel is $2\pi\mathbb{Z}$. Its image is $\{e^{ix},\ x \in \mathbb{R}\} = U$. Thus by the 1st isomorphism theorem

$$\mathbb{R}/2\pi\mathbb{Z} \simeq U.$$

Exercise 19. Let $G = \langle x \rangle$ be a cyclic group of order $n \geq 1$. Let $h_x : \mathbb{Z} \to G$, $m \mapsto x^m$.

- Show that h_x is surjective and compute its kernel.
- Show that $G \simeq \mathbb{Z}/n\mathbb{Z}$.

Answer.

- Let $g \in G$. Since $G = \langle x \rangle$, $g = x^k$ for some $0 \leq k \leq n-1$ and thus h_x is surjective. Its kernel is the set of m such that $x^m = 1$, thus m must be a multiple of n and $Ker(h_x) = n\mathbb{Z}$.
- By the 1st isomorphism theorem, since h_x is a group homomorphism, we have

$$G \simeq \mathbb{Z}/n\mathbb{Z}.$$

Exercise 20. Prove the second isomorphism theorem for groups, namely that if H and N are subgroups of G, with N normal in G, then

$$H/(H \cap N) \simeq HN/N.$$

Answer. Let π be the canonical epimorphism from G to G/N, and let π_0 be the restriction of π to H. Then the kernel of π_0 is $H \cap N$, so by the 1st isomorphism theorem for groups, we have that $H/(H \cap N)$ is isomorphic to the image of π_0 which is $\{hN,\ h \in H\} = HN/N$.

Exercise 21. Prove the third isomorphism theorem for groups, namely that if N and H are normal subgroups of G, with N contained in H, then

$$G/H \simeq (G/N)/(H/N).$$

Answer. This follows from the 1st isomorphism theorem for groups, if we can find an epimorphism of G/N into G/H with kernel H/N: take $f(aN) = aH$. Now f well-defined, since if $aN = bN$, then $a^{-1}b \in N \subset H$ so $aH = bH$. Since a is arbitrary in G, f is surjective. By definition of coset multiplication, f is a homomorphism. The kernel is

$$\{aN,\ aH = H\} = \{aN,\ a \in H\} = H/N.$$

Exercise 22. Consider the short exact sequence of groups

$$1 \xrightarrow{i} A \xrightarrow{u} B \xrightarrow{v} C \xrightarrow{j} 1$$

where i is the inclusion and j is the constant map 1.

1. Show that $Im(u) = Ker(v) \iff v \circ u = 1$, and $Ker(v) \subset Im(u)$ (1 denotes the constant map here).
2. Show that in the short exact sequence, we have that u is injective and v is surjective.
3. Show that $u(A)$ is normal in B and that we have a group isomorphism

$$B/u(A) \simeq C.$$

Answer.

1. If $Im(u) = Ker(v)$, then clearly $Ker(v) \subset Im(u)$. Then $v \circ u(x) = v(u(x)) = 1$ since $u(x)$ is in the kernel of v. Conversely, we have to show that $Im(u) \subset Ker(v)$. Let $u(x) \in Im(u)$. Now $v(u(x)) = 1$ and thus $u(x)$ is in $Ker(v)$.
2. To show that u is injective, we compute its kernel. Now $Ker(u) = Im(i) = \{1\}$ and u is injective. To show that v is surjective, we have to see that $Im(v) = C$, but $Im(v) = Ker(j) = C$.
3. Since $Im(u) = Ker(v)$, $u(A)$ is normal in B, and we conclude by the 1st isomorphism theorem.

2.5 Direct and semi-direct products

Exercise 23. The *quaternion group* Q_8 is defined by

$$Q_8 = \{1, -1, i, -i, j, -j, k, -k\}$$

with product $\cdot$ computed as follows:

$$\begin{aligned} 1 \cdot a = a \cdot 1 = a, \ \forall\, a \in Q_8 \\ (-1) \cdot (-1) = 1, \ (-1) \cdot a = a \cdot (-1) = -a, \ \forall\, a \in Q_8 \\ i \cdot i = j \cdot j = k \cdot k = -1 \\ i \cdot j = k, \ j \cdot i = -k, \\ j \cdot k = i, \ k \cdot j = -i, \\ k \cdot i = j, \ i \cdot k = -j. \end{aligned}$$

Show that Q_8 cannot be isomorphic to a semi-direct product of smaller groups.

Answer. By definition, a semi direct product must contain two smaller subgroups of trivial intersection $\{1\}$. Now the smaller subgroups of Q_8 are $\{1, -1\}$, $\{1, i, -i, -1\}$, $\{1, j, -j, -1\}$, $\{1, k, -k, -1\}$, and each contains -1 so that it is not possible that Q_8 is a semi-direct product.

Exercise 24. Consider the set of matrices

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

(where $\mathbb{F}_p$ denotes the integers mod p).

1. Show that G is a subgroup of $SL_2(\mathbb{F}_p)$.
2. Write G as a semi-direct product.

Answer.

1. That G is a subset of $SL_2(\mathbb{F}_p)$ is clear because the determinant of every matrix in G is 1. We have to show that for $X, Y \in G$, $XY^{-1} \in G$. This is a straightforward computation:

$$\begin{pmatrix} a & b \\ 0 & a^{-1} \end{pmatrix} \begin{pmatrix} c^{-1} & -d \\ 0 & c \end{pmatrix} = \begin{pmatrix} ac^{-1} & -da + bc \\ 0 & a^{-1}c \end{pmatrix} \in G.$$

2. Take

$$K = \left\{ \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}, \ a \neq 0, \ a \in \mathbb{F}_p \right\}$$

and

$$H = \left\{ \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}, \ b \in \mathbb{F}_p \right\}.$$

Both K and H are subgroups of G. Their intersection is the 2-dimensional identity matrix, and $HK = G$, since

$$\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} = \begin{pmatrix} a & ba^{-1} \\ 0 & a^{-1} \end{pmatrix}$$

and ba^{-1} runs through every possible element of $\mathbb{F}_p$ (since b does). Also H is normal in G, since

$$\begin{pmatrix} a & b \\ 0 & a^{-1} \end{pmatrix} \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a^{-1} & -b \\ 0 & a \end{pmatrix} = \begin{pmatrix} 1 & a^2 b \\ 0 & 1 \end{pmatrix} \in H.$$

Note that K is not normal, which can be seen by doing the same computation. Thus G is the semi-direct product of H and K.

Exercise 25. Show that the group $\mathbb{Z}_n \times \mathbb{Z}_m$ is isomorphic to $\mathbb{Z}_{mn}$ if and only if m and n are relatively prime. Here $\mathbb{Z}_n$ denotes the integers modulo n.

Answer. If m and n are relatively prime, then for a multiple of $(1,0)$ to be zero, it must be a multiple of n, and for a multiple of $(0,1)$ to be zero, it must be a multiple of m. Thus for a multiple k of $(1,1)$ to be zero, it must be a multiple of both n and m, and since they are coprime, the smallest possible value of k is mn. Hence $\mathbb{Z}_n \times \mathbb{Z}_m$ contains an element of order mn, showing that $\mathbb{Z}_m \times \mathbb{Z}_n$ is isomorphic to $\mathbb{Z}_{mn}$. Conversely, suppose that $\gcd(m,n) > 1$. Then the least common multiple of m and n is smaller than mn, let us call it d. This shows that every element of $\mathbb{Z}_m \times \mathbb{Z}_n$ has order at most d and thus none of them can generate the whole group, so that it cannot be cyclic, and thus cannot be isomorphic to $\mathbb{Z}_{mn}$.

Note that one can also prove this result by the definition of direct product: we have that $\mathbb{Z}_m$ and $\mathbb{Z}_n$ are both normal subgroups of $\mathbb{Z}_{mn}$ because this is an abelian group. We are thus left to look at the intersection of $\mathbb{Z}_m$ and $\mathbb{Z}_n$. Recall that $\mathbb{Z}_m$ and $\mathbb{Z}_n$ are embedded into $\mathbb{Z}_{mn}$ as respectively

$$\mathbb{Z}_m = \{0, n, 2n, \ldots, (m-1)n\},\ \mathbb{Z}_n = \{0, m, 2m, \ldots, (n-1)m\}.$$

If m and n are coprime, then $\mathbb{Z}_m \cap \mathbb{Z}_n = \{0\}$. Conversely, if x belongs to the intersection and is non-zero, then x must be a multiple of both n and m which is not congruent to 0 modulo mn, and thus m and n cannot be coprime.

Exercise 26. Let $\mathbb{Z}_3$ denote the group of integers modulo 3.

1. Show that the map

$$\sigma : \mathbb{Z}_3 \times \mathbb{Z}_3 \to \mathbb{Z}_3 \times \mathbb{Z}_3, (x,y) \mapsto (x+y, y)$$

is an automorphism of $\mathbb{Z}_3 \times \mathbb{Z}_3$ of order 3.

2. Show that the external semi-direct product of $\mathbb{Z}_3 \times \mathbb{Z}_3$ and $\mathbb{Z}_3$ by ρ, $\rho : \mathbb{Z}_3 \to Aut(\mathbb{Z}_3 \times \mathbb{Z}_3),\ i \mapsto \sigma^i$, is a non-abelian group G satisfying that

$$a^3 b^3 = (ab)^3$$

for any a, b in G.

Answer.

1. So to be an automorphism, σ has to be a group homomorphism, but

   $$\sigma((x+x', y+y')) = (x+x'+y+y', y+y') = (x+y, y)+(x'+y', y') = \sigma(x,y)+\sigma(x',y').$$

   It clearly goes from the group to itself, and it is a bijection. It is an injection

   $$\sigma(x,y) = \sigma(x',y') \Rightarrow (x+y, y) = (x'+y', y') \Rightarrow y = y', x = x',$$

   and thus it is a surjection since the group is finite. It is of order 3, since

   $$\sigma(x,y) = (x+y, y),\ \sigma^2(x,y) = (x+2y, y),\ \sigma^3(x,y) = (x+3y, y) = (x,y).$$

2. An element in the external semi-direct product is of the form $((x,y),i)$, and we have

   $$((x,y),i)((x,y),i) = ((x,y) + \sigma^i(x,y), 2i),$$

   $$\begin{aligned} ((x,y),i)^3 &= ((x,y) + \sigma^i(x,y) + \sigma^{2i}(x,y), 3i) \\ &= ((x,y) + (x+iy, y) + (x+2iy, y), 3i) \\ &= ((3x+3iy, 3y), 3i) \\ &= ((0,0),0). \end{aligned}$$

   This shows that for any element a of the semi-direct product $a^3 = 0$, thus $b^3 = 0$, ab is another element of the group thus $(ab)^3 = 0$ which shows that $a^3b^3 = 0 = (ab)^3$, though the group is non-abelian (because σ is not the identity).

2.6 Permutations and Group action

Exercise 27. (*) In a group G of order n, for all divisors d of n, there exists at least one subgroup of order d. True or false? [Though the statement only involves the order of a group, there were not enough examples of groups seen in the lecture notes earlier to come up with a counter-example.]

Answer. This is false. The smallest counterexample is the alternating group $G = A_4$ of even permutation on 4 elements, given explicitly by

$$A_4 = \{e, (12)(34), (13)(24), (14)(23), (123),$$

$$(132), (124), (142), (134), (143), (234), (243)\}.$$

It has 12 elements. We will now show that it has no subgroup of order 6. Let H denote a subgroup of order 6, that is, H has index 2 in A_4 and there are only two cosets, satisfying

$$A_4 = H \cup Ha,$$

for all $a \in A_4$ but not in H. Consider the coset Ha^2. We have that either $Ha^2 = H$ or $Ha^2 = Ha$. If $Ha^2 = Ha$, then $Ha = H$ and $a \in H$, a contradiction. Thus

$$Ha^2 = H,$$

for all $a \in A_4$ but not in H. But those $a \in H$, because H is a subgroup, also satisfy that $Ha^2 = H$, so that we deduce that

$$Ha^2 = H, \ \forall a \in A_4.$$

This in turn implies that

$$a^2 \in H \ \forall a \in A_4.$$

Let now $b \in A_4$ be an element of order 3, that is $b^3 = e$. Then $b^2 = b^{-1}$ and $b^2 \in H$ showing that $b^{-1} \in H$ and finally $b \in H$. We have just shown that every element of order 3 in A_4 are in H, which is contradiction, since A_4 contains 8 elements of order 3. (If you are not satisfied with this proof, please check "Variations on a Theme: A_4 Definitely Has No Subgroup of Order Six!" by M. Brennan and D. Machale, available online, where 12 different proofs are provided.)

Exercise 28. 1. Let $G = GL_n(\mathbb{C})$ and $X = \mathbb{C}^n - \{\mathbf{0}\}$. Show that G acts on X by $G \times X \to X$, $(M, \nu) \mapsto M\nu$.

2. Show that the action is transitive.

Answer.

1. We have to show that

$$M \cdot (M' \cdot \nu) = (MM') \cdot \nu, \ 1_G \cdot \nu = \nu.$$

The first point is clear by properties of matrix vector multiplication. The second is also clear since 1_G is the identity matrix.

2. We have to show that there is only one orbit (which is why we have to remove the whole zero vector from $\mathbb{C}^n$). For that, we need to show that for any two vectors $\nu, \nu' \in X$, there is a matrix $M \in G$ such that $M\nu = \nu'$. We thus have a system of n linear equations for n^2 unknowns, so that we have enough degrees of freedom to find such a matrix. Alternatively, if $\nu = (a_1, \ldots, a_n)$, $\nu' = (b_1, \ldots, b_n)$, where a_i, b_i are all non-zero, take the matrix

$$\mathrm{diag}(a_1^{-1}, \ldots, a_n^{-1})$$

and notice that

$$\mathrm{diag}(b_1, \ldots, b_n)\mathrm{diag}(a_1^{-1}, \ldots, a_n^{-1})\nu = \nu'.$$

The case where some a_i, b_j are zero can be done similarly.

Exercise 29. Let G be group, and H be a subgroup of G. Show that

$$g \cdot g'H = gg'H$$

defines an action of G on the set G/H of cosets of H. Find the stabilizer of gH.

Answer. To show that the action is well defined we have to check that it does not depend on the choice of the representative, and that it satisfies the definition of group action. First suppose that $g'H = g''H$. We have to show that $g \cdot g''H = gg'H$. But $g'H = g''H \iff (g'')^{-1}g' \in H \iff (gg'')^{-1}(gg') \in H \iff gg'H = gg''H$. The definition of group action can be checked easily:

$$g_1 \cdot (g_2 \cdot g'H) = g_1 \cdot g_2g'H = g_1g_2g'H = g_1g_2 \cdot g'H,\ 1 \cdot g'H = g'H.$$

The stabilizer of gH is formed by g' such that $g'gH = gH$ that is $g^{-1}g'g \in H$. Thus $g^{-1}g'g = h$, for some $h \in H$, or equivalently $g' = ghg^{-1}$, thus the stabilizer is gHg^{-1}.

Exercise 30. Consider the *dihedral group* D_8 given by

$$D_8 = \{1, s, r, r^2, r^3, rs, r^2s, r^3s\}$$

(that is $s^2 = 1$, $r^4 = 1$ and $(rs)^2 = 1$).

1. Divide the elements of the dihedral group D_8 into conjugacy classes.
2. Verify the class equation.

Answer.

1. There are 5 conjugacy classes

$$\{1\}, \{r^2\}, \{r, r^3\}, \{s, sr^2\}, \{sr, sr^3\}.$$

2. We have that $\{1\}$ and $\{r^2\}$ are in the center. Thus

$$|D_4| = 8 = |Z(D_4)| + |B(r)| + |B(rs)| + |B(s)|.$$

Exercise 31. The *quaternion group* Q_8 is defined by

$$Q_8 = \{1, -1, i, -i, j, -j, k, -k\}$$

with product $\cdot$ computed as follows:

$$\begin{aligned} 1 \cdot a = a \cdot 1 = a,\ \forall\ a \in Q_8 \\ (-1) \cdot (-1) = 1,\ (-1) \cdot a = a \cdot (-1) = -a,\ \forall\ a \in Q_8 \\ i \cdot i = j \cdot j = k \cdot k = -1 \\ i \cdot j = k,\ j \cdot i = -k, \\ j \cdot k = i,\ k \cdot j = -i, \\ k \cdot i = j,\ i \cdot k = -j. \end{aligned}$$

1. Show that if $x \notin Z(Q_8)$, then $|C_{Q_8}(x)| = 4$.

2. Show that as a consequence, the class of conjugacy of $x \notin Z(D_8)$ has only two elements.

Answer.

1. The center $Z(Q_8)$ is $Z(Q_8) = \{1, -1\}$. We have by definition that

$$C_{Q_8}(x) = \{g \in Q_8,\ gx = xg\}.$$

Thus

$$C_{Q_8}(i) = \{1, -1, i, -i\},\ C_{Q_8}(j) = \{1, -1, j, -j\},\ C_{Q_8}(k) = \{1, -1, k, -k\}.$$

2. When the action is defined by conjugation, we have that $\mathrm{Stab}(x) = C_{Q_8}(x)$. Thus by the Orbit-Stabilizer, the size of an orbit, which is a conjugacy class, is

$$|B(x)| = |Q_8|/|C_{Q_8}(x)| = 8/4 = 2.$$

Exercise 32. Let G be a group and let H and K be two subgroups of G.

1. Show that the subgroup H acts on the set of left cosets of K by multiplication.

2. Consider the coset $1K = K$. Compute its orbit $B(K)$ and its stabilizer $\mathrm{Stab}(K)$.

3. Compute the union of the cosets in $B(K)$ and deduce how many cosets are in the orbit.

4. Use the Orbit-Stabilizer Theorem to get another way of counting the number of cosets in $B(K)$. By comparing the two expressions to count the cardinality of $B(K)$, can you recover a known result on the cardinality of HK?

Answer.

1. Let $X = \{gK,\ g \in G\}$ be the set of left cosets of K. We have to check that $h' \cdot (h \cdot gK) = (h'h) \cdot gK$ which clearly holds, as does $1_H \cdot gK = gK$.

2. We have that $B(K) = \{h \cdot K,\ h \in H\}$ and $\mathrm{Stab}(K) = \{h \in H,\ h \cdot K = K\} = H \cap K$.

3. The union of the cosets in $B(K)$ is HK, the cosets in $B(K)$ are disjoint and each has cardinality K, so that we have $|HK|/|K|$ cosets in $B(K)$.

4. By the Orbit-Stabilizer Theorem, we have

$$|B(K)| = |H|/|\text{Stab}(K)| \Rightarrow |HK|/|K| = |H|/|H \cap K|$$

and thus

$$|HK| = \frac{|H||K|}{|H \cap K|}.$$

Exercise 33. Let G be a finite group, and let p be the smallest prime divisor of the order of G.

1. Let H be a normal subgroup of G. Show that G acts on H by conjugation.
2. Let H be a normal subgroup of G of order p.
   - Show that the orbits of H under the action of G are all of size 1.
   - Conclude that a normal subgroup H of order p is contained in the center of G.

Answer.

1. We check the definition, that is, the group G acts on H if for the map $(g, x) \mapsto g \cdot x = gxg^{-1}$, $x \in H$, defined from $G \times H \to H$ (note that we need here H normal to guarantee that $gxg^{-1} \in H$!), we have
   - $h \cdot (g \cdot x) = h \cdot (gxg^{-1}) = h(gxg^{-1})h^{-1} = (hg) \cdot x$
   - $1 \cdot x = x$ for all $x \in H$
2. - By the orbit stabilizer theorem, the size of an orbit $B(x), x \in H$ divides the size of G, the group that acts on H, thus if $|B(x)|$ is not 1, it must be at least p, since p is the smallest divisor of the order of G. Now the orbits partition H, that is $H = \cup B(x)$ and thus $|H| = \sum |B(x)|$, that is the sum of the cardinals of the orbits is $|H| = p$. Among all the $B(x)$, we can take $x = 1 \in H$ since H is a subgroup. The orbit $B(1) = \{g \cdot 1, g \in G\} = \{g1g^{-1} = 1\}$ has only 1 element, there is at least one orbit of size 1, and thus no orbit can have size greater or equal to p, since then $p + 1 > p$. Thus all orbits of H are of size 1.
   - We have that $B(x) = \{g \cdot x, g \in G\} = \{gxg^{-1}, g \in G\}$ is always of size 1, and since for $g = 1 \in G$ we have $x \in B(x)$, we deduce that $B(x) = \{x\}$, that is $gxg^{-1} = x$, or $gx = xg$ showing that for all $x \in H$, x actually commutes with every $g \in G$, that is, H is contained in the center.

Exercise 34. Let G be a group acting on a finite set X.

1. We assume that every orbit contains at least 2 elements, that $|G| = 15$, and that $|X| = 17$. Find the number of orbits and the cardinality of each of them.

2. We assume that $|G| = 33$ and $|X| = 19$. Show that there exists at least one orbit containing only 1 element.

Answer.

1. The cardinal of every orbit divides the order of G. Furthermore, the sum of the orbit cardinalities is equal to the cardinality of X. If $|G| = 15$, $|X| = 17$, and there is no orbit of size 1, there is only one possibility: 4 orbits of length 3 and 1 of length 5. Indeed, we are looking for integers such that their sum is 17, but each integer must divide 15, that is we need to realize 17 as a sum of integers belonging to $\{3, 5, 15\}$ (1 is excluded by assumption). Then 15 is not possible, and we can use only 3 and 5: 15+2 is not possible, 10+7 is not possible, so only 5+12 works.

2. Now $|G| = 33$ and $|X| = 19$. The divisors of 33 are 1,3,11 and 33. We need to obtain as above 19 as a sum of these divisors. 33 is too big, and we cannot possibly use only 11 and 3. Thus there must be at least one orbit of size 1.

Exercise 35. Let G be a finite group of order $n \geq 1$ and let p be a prime. Consider the set

$$X = \{x = (g_1, g_2, \ldots, g_p) \ \in G^p \mid g_1 \cdot g_2 \cdots g_p = 1_G\}.$$

1. Compute the cardinality $|X|$ of the set X.

2. Show that if $(g_1, \ldots, g_p) \in X$, then $(g_2, \ldots, g_p, g_1) \in X$. Denote by σ the corresponding permutation. Show that $< \sigma >$ acts on X as follows:

$$\sigma^k \cdot (g_1, \ldots, g_p) = (g_{\sigma^k(1)}, \ldots, g_{\sigma^k(p)}), \ k \in \mathbb{Z}$$

3. What is the cardinal of one orbit of X?

4. What are the orbits with one element? Show that there is at least one such orbit.

5. Deduce that if p does not divide n, then

$$n^{p-1} \equiv 1 \mod p.$$

6. Deduce Cauchy Theorem from the above, namely, if $p \mid n$ then G has at least one element of order p.

Answer.

1. Since $g_1, \ldots, g_{p-1}$ can take any value in G (only g_p is constrained so as to have $g_1 \cdot g_2 \cdots g_p = 1_G$), we have $|X| = |G|^{p-1} = n^{p-1}$.

2. Since $(g_1, \ldots, g_p) \in X$, then $g_1 \cdot g_2 \cdots g_p = 1_G$ and $g_2 \cdots g_p \cdot g_1 = g_1^{-1} \cdot 1_G \cdot g_1$ showing that $(g_2, \ldots, g_p, g_1) \in X$. To show that $< \sigma >$ acts on X, check the definition, namely $\sigma^l \cdot (\sigma^k \cdot (g_1, \ldots, g_p)) = \sigma^l \sigma^k \cdot (g_1, \ldots, g_p)$ and $\sigma^0 \cdot (g_1, \ldots, g_p) = (g_1, \ldots, g_p)$.

3. The answer is either 1 or p. There are two ways to do it: one can notice that $< \sigma >$ has order p, and thus by the Orbit-Stabilizer Theorem the size of the orbit divides p, so it can be either 1 or p. Also one can just write down the definition of one orbit: the orbit of $(g_1, \ldots, g_p)$ is formed by all the shifts of the components, and thus since p is prime, there will be p distinct shifts, apart if all the components are all the same, in which case there is only one element in the orbit.

4. Since an element always belongs to its orbit, we have that orbits with one element are of the form $B(x) = \{x\}$, and if there is only one element, that means that the shifts are doing nothing on $x = (g_1, \ldots, g_p)$ thus $x = (g, \ldots, g)$ and since $x \in X$, that further means that $g^p = 1_G$. To show one such orbit exists, take the orbit of $(1, \ldots, 1)$.

5. Since the orbits partition X, we have

$$|X| = \sum |B(x)| + \sum |B(x')|$$

where the first sum is over orbits of size 1, and the second over orbits of size greater or equal to 2. By the above, if the size is at least 2, it is p, and thus $|B(x')| \equiv 0 \mod p$. Then if there were more than $(1, \ldots, 1)$ with orbit of size 1, that means an element g such that $g^p = 1$, which would mean $p|n$, a contradiction. Thus only there is only one orbit of size 1, and

$$|X| = n^{p-1} \equiv 1 \mod p.$$

6. Again, we have that

$$n^{p-1} = |X| = \sum |B(x)| + \sum |B(x')|$$

and if $p|n$ then $0 \equiv \sum |B(x)|$ and there must be at least another element with orbit size 1, that is an element g of order p.

2.7 The Sylow theorems

Exercise 36. Let G be a group of order 399.

1. Show that G has a unique Sylow 19-subgroup P which is normal in G.

2. Let Q be a Sylow 7-subgroup. Show that $N = PQ$ is a subgroup of order 133 of G.

Answer.

1. The number n_{19} of Sylow 19-subgroups is $\equiv 1 \mod 19$ and divides 21, thus it must be 1. Since it is unique, it has to be normal.

2. Since P is normal in G, we have that $N = PQ$ is a subgroup of G. (The fact that P is normal can be used to check directly the definition of subgroup). By the 2nd isomorphism theorem for groups, we have

$$Q/(Q \cap P) \simeq PQ/P \Rightarrow Q \simeq PQ/P$$

since $Q \cap P = \{1\}$ so that $|PQ| = |P||Q| = 19 \cdot 7 = 133$.

Exercise 37. Let G be a simple group of order 168.

1. Compute its number of Sylow 7-subgroups.
2. Deduce the number of elements of order 7 in G.

Answer.

1. Since $168 = 2^3 \cdot 3 \cdot 7$, the number of Sylow 7-subgroups must be $\equiv 1 \mod 7$ and must divide 24. The only possibilities are thus 1 and 8. Since G is simple, it cannot be 1.
2. Elements of order 7 correspond to generators of the cyclic Sylow 7-subgroups, and there are 6 of them per Sylow 7-subgroup, that is $6 \times 8 = 48$ elements of order 7.

Exercise 38. (*) This exercise aims at classifying groups of order up to 8.

1. For p prime, show that any group G with cardinality p^2 is abelian.
2. For p an odd prime, show that any non-abelian group G of order $2p$ is isomorphic to the dihedral group D_p.
3. Determine all the finite groups of order at most 8 up to isomorphism.

Answer.

1. Let $Z(G)$ be the center of G. We have by Lagrange Theorem that $|Z(G)|$ divides p^2, thus we have 3 cases:
   - $|Z(G)| = 1$: we know that the center of a p-group cannot be trivial, thus this case cannot happen.
   - $|Z(G)| = p^2$: then clearly G is abelian.
   - $|Z(G)| = p$: then $|G/Z(G)| = p^2/p = p$, and then quotient group is cyclic, and we already showed (see Exercise 15) that in this case it implies that G is abelian.
2. Let G be a group of order $2p$. Then G contains an element α of order p, and an element β of order 2, by Cauchy Theorem. We will prove

$$\langle \alpha, \beta \rangle \simeq D_p.$$

- First we have that $|\langle\alpha,\beta\rangle| = 2p$. Indeed, it must divide $2p$, and it must be great than p, since $\langle\alpha\rangle$ has already cardinality p and does not contain β.
- We now show that
$$\beta\alpha\beta^{-1} = \alpha^{-1}.$$
Since $\langle\alpha\rangle$ is normal in G, then
$$\beta\alpha\beta^{-1} = \alpha^{i}$$
for some i. Since $\beta^2 = 1$
$$\alpha = \beta^2\alpha\beta^{-2} = \beta(\beta\alpha\beta^{-1})\beta^{-1} = \beta\alpha^{i}\beta^{-1} = (\beta\alpha\beta^{-1})^{i} = \alpha^{i^2}$$
and $i^2 \equiv 1 \mod p$. If $i \equiv 1 \mod p$, then α and β commute, so G cannot be non-abelian. Thus $i \equiv -1 \mod p$ as wanted.
- Finally, it is enough to conclude to show that
$$f(a^m b^e) = \alpha^m\beta^e$$
is an isomorphism, where a is a rotation of angle $2\pi/p$ and b is a reflection. This map is surjective, and both sets have same size, so it is injective. It is also clear that f is a group homomorphism.

3. - For prime order $|G|$, we have that G is cyclic, so that gives C_2, C_3, C_5, C_7.
   - For $|G| = 4$, we already know that either G contains an element of order 4, and $G = C_4$, or G has only elements of order 2 (except of course the identity), and then $G = C_2 \times C_2$.
   - If $|G| = 6$, if G contains an element of order 6, then $G = C_6$. If not, then G must contain 1 element of order 3, and one of order 2, and by the above computations, we have $G = D_6$.
   - If $|G| = 8$, then elements in G can have order 2,4 and 8. If there is an element of order 8, then $G = C_8$. If all elements have order 2, then G is abelian (namely $C_2 \times C_2 \times C_2$), so for G to be non-abelian, we must have an element of order 4, say g. Now we have $1, g, g^2, g^3 \in G$. If there is an element h of order 2, h not in $\langle g\rangle$, then hah^{-1} has order 4, and repeating the above computations, we can see that $G = D_8$. If such an h does not exist, then all elements not in $\langle g\rangle$ have order 4. Let k be such an element of order 4, then k^2 has order 2 and must then be g^2. In this case, we obtain $G = Q_8$. Finally, if G has an element of order 4 and is abelian but not cyclic, then $G = C_2 \times C_4$.

Exercise 39. Consider the set of matrices of the form
$$\begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}$$

where a, b, c are integers modulo 2. Show that this set forms a group and compute its cardinality. Can you identify this group?

Answer. This set forms a group under matrix multiplication. Matrix multiplication is clearly associative. One can check that the product of two such matrices still belongs to the set. The identity element is given by the identity matrix. Every matrix is clearly invertible with inverse

$$\begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -a & ac-b \\ 0 & 1 & -c \\ 0 & 0 & 1 \end{pmatrix}.$$

Since a, b, c are integers modulo 2, there are 8 possible such matrices. Thus it is a group of order 8. It is not commutative, thus it is either D_4 or Q_8. We note that

$$\begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & ac \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

thus

$$\begin{pmatrix} 1 & 1 & b \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

is an element of order 4, and the subgroup it generates is

$$\begin{pmatrix} 1 & 1 & b \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 & b+1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, I_3.$$

So when $a = 0$ and $c = 1$, or $a = 1$ and $c = 0$, we see that we get elements of order 2 not in this subgroup, thus it must be D_4. (This line of argument comes from the above exercise where we did the classification of groups of order 8.) Alternatively, we can notice that

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

is of order 2, and the intersection of this subgroup of order 2 with that above of order 4 is the identity. Again this cannot happen with Q_8 since we know the intersection of any two subgroups cannot be trivial (this was the line of argument used to show that Q_8 cannot be a semi-direct product).

Exercise 40. Let G be a finite p-group, and let H be a normal subgroup of G. Show that $H \cap Z(G)$ cannot be trivial (where $Z(G)$ denotes the center of G). Is it still true when G is an infinite p-group?[this is harder!]

Answer. The subgroup H is normal, thus G acts by conjugation on H. Since $|G| = p^r$ for some r, the size of the non-trivial orbits is divisible by p. Since

G is a p-group, H is also a p-group, with say $|H| = p^s$ for some $s < r$. We deduce that the union of orbits has size p^s, and since the orbits of size ≥ 2 are divisible by p, the union of trivial orbits, that is of orbits of fixed points, is also of cardinality divisible by p. But the union of orbits of fixed points is $H \cap Z(G)$. Since it contains the identity element, it at least contains p elements and cannot be trivial.

This is not true anymore when the group is infinite. Take for example the group

$$G = \bigcup_{n \geq 1} U(p, n),$$

where $U(p, n)$ is the group of $n \times n$ upper triangular matrices with every diagonal coefficient at 1 and elements of $\mathbb{F}_p$ discussed in Example 1.31. This is an infinite p-group. Now the center $Z(U(p, n))$ of $U(p, n)$ is a cyclic group of order p, formed by matrices with every diagonal coefficient at 1, one element of $\mathbb{F}_p$ in position $(1, n)$ and zeros elsewhere. Thus the center of the infinite p-group G which is itself the limit of the $U(p, n)$ when n grows will be contained in the limit of these centers, which is in fact the trivial group (that is the limit of the identity matrix when n tends to infinity).

Exercise 41. Let G be a group of order 57 which is not cyclic.

1. Compute the number of its Sylow 19-subgroups.
2. Deduce the number of elements of order 3 in G.

Answer.

1. Since $57 = 3 \cdot 19$, $n_{19}|3$ and $n_{19} \equiv 1 \mod 19$, showing that $n_{19} = 1$.
2. Let $a \neq 1$ be in G. Then $|a| = 3$ or $|a| = 19$ ($|a| \neq 57$ since G is not cyclic and $|a| \neq 1$ since $a \neq 1$). Since there is only one Sylow 19-subgroup, which is cyclic, it contains 18 elements of order 19, which are all the elements of order 19 of G. Thus there are exactly 38 elements of order 3 in G.

Exercise 42. Let G be a group of order 231.

1. Show that G as a unique Sylow 11-subgroup M.
2. Compute the number n_7 of its Sylow 7-subgroups.
3. Let P be a Sylow 3-subgroup and L be a Sylow 7-subgroup. Show that $N := PL$ forms a subgroup of order 21.
4. Show that N is furthermore normal.(*)
5. Deduce from the above that G can be written as MN.
6. Prove that the Sylow 11-subgroup M belongs to the center of G.

Answer.

1. We have that
$$231 = 11 \cdot 3 \cdot 7.$$
Thus $n_{11} \equiv 1 \mod 11$ and $n_{11} \mid 21$. Thus it must be 1.

2. Again, $n_7 \equiv 1 \mod 7$ and n_7 divides 33. Thus it must be 1.

3. Notice first that since L is a unique Sylow 7-subgroup, it is normal. We have to see that PL is a subgroup. Let pl be an element of PL. Its inverse is $(pl)^{-1} = l^{-1}p^{-1}$ and since L is normal, we know that $gL = Lg$ for every g in G. Thus in particular $l^{-1}p^{-1} \in Lp^{-1} = p^{-1}L$ and $l^{-1}p^{-1} = p^{-1}l'$ showing that the inverse belongs to PL. The same argument works to show that $plp'l' \in PL$. We know that $lp' = p'l''$ by normality of L, thus $p(lp')l' = p(p'l'')l' \in PL$ and 1 is in PL. Now about the order of PL, since $|P| = 3$ and $|L| = 7$, and $P \cap L = \{1\}$ because they are groups of order respectively 3 and 7, we have that $|PL| = |P||L|/|P \cap L| = 21$.

4. This is the most difficult part. The most likely quickest way to do it is by using the normalizer $N_G(N)$ of N. It is a subgroup of G, thus its order divides $|G|$. Since N is contained in its normalized, $|N_G(N)| \geq 21$. So if we can prove that $|N_G(N)| > 21$, we are done, because then $|N_G(N)| = 231$ and the normalizer is G, which yields the desired conclusion. To show that $|N_G(N)| > 21$, it is enough to exhibit at least one element in the normalizer which is not in N. This is not the most economical way of doing it, but one way is to just get an element in the center. By Cauchy Theorem, we know G contains an element g of order 3, and since M is the unique Sylow 11-subgroup, it is a normal subgroup of G, which is furthermore cyclic, generated by say m. We thus have that $gM = Mg$ that is $gm = m^l g$ and we are left to show that l is one. Since g is of order 3, then $m = g^3 m g^{-3} = g^2(gmg^{-1})g^{-2} = g^2 m^l g^{-2} = g(gm^l g^{-1})g^{-1} = g(m^l)^l g^{-1} = m^{l^3}$. This shows that $l^3 \equiv 1 \mod 11$. But this is possible only if $l = 1$ ($\gcd(3, 10) = 1$) and m belongs to the center.

5. That MN is a group works as above, because N is normal. That the cardinality is right also works as above, since $|M| = 11$ and $|N| = 21$, thus their intersection is 1, and thus $|MN| = 231$, from which it follows that $MN = G$.

6. We first observe that M commutes with N, since if $m \in M, n \in N$, then by normality of M (it is a unique Sylow 11-subgroup) we have that $mnm^{-1}n^{-1} \in M \cap N = \{1\}$ implying that $mn = nm$. To show that M commutes with G, we use the fact that $G = MN$, and thus every element $g \in G$ can be written as $g = m'n'$. Now $gm = (m'n')m = m'mn'$ since M and N commute, and $m'mn' = mm'n' = mg$ since M is abelian (it is in fact cyclic).

2.8 Simple groups

Exercise 43. Show that no group of order 200 is simple.

Answer. The number of Sylow 5-subgroups of a group of order $200 = 5^2 \cdot 2^3$ is $\equiv 1 \mod 5$ and divides 8, thus it must be 1. Thus the unique Sylow 5-subgroup is normal and thus the group cannot be simple.

Exercise 44. (*) Let G be a group such that $|G| \leq 59$. Show that if G is simple, then $|G|$ is prime.

Answer. We know that if $|G| = pq$, p, q two distinct primes, then G is not simple. This can be extended by noting that if $|G| = p^k m$, where $k > 0$ and $(m, p) = 1$, $m < p$. Indeed, in this case, n_p must divide mp^k, and thus must divide m. But also, $n_p \equiv 1 \mod p$, so that if $n_p \neq 1$ then $n_p \geq p + 1 > m$, which contradicts that $n_p | m$. We are thus left to check that groups of order

$$12, 24, 30, 36, 40, 45, 48, 56$$

are not simple. We know G with $|G| = 45 = 5 \cdot 3^2$ is not simple since G is not simple when $|G| = p^2 q$. If $|G| = 40$, then the number n_5 of Sy- low 5-subgroups is congruent to 1 modulo 5 and divides 8. The only possibility is $n_5 = 1$, and G has a normal Sylow 5-subgroup. If $|G| = 56 = 2^3 \cdot 7$, then the number n_7 of Sylow 7-subgroups must divide 8 and $\equiv 1 \mod 7$. If $n_7 = 1$, G is not simple. If $n_7 = 8$, then we get $6 \cdot 8 = 48$ elements of order 7, and only 8 elements are left in the group not of order 7, which correspond to the Sylow 2-subgroup of size 8. Similarly if $|G| = 12$, there are $n_3 = 1 \mod 3$ Sylow 3-subgroups, where $n_3 | 12$, so this can be 1 or 4. If $n_3 = 1$ G is not simple. If $n_3 = 4$, we get $4 \cdot 2 = 8$ elements of order 3. The other 4 elements must be part of the Sylow 2-subgroup which is of order 4. We are thus left with

$$24, 30, 36, 48$$

One way to take care of 24, 36 and 48 at once is to prove that the order of G divides $n_p!/2$. Otherwise it can be done case by case. $|G| = 30$ is done individually.

Exercise 45. Let G be a group of order 105. Prove that it is impossible that $|Z(G)| = 7$.

Suppose to the contrary that $|Z(G)| = 7$, then $|G|/|Z(G)| = 15$ and thus $G/Z(G)$ is a group of order 15. Since $15 = 3 \cdot 5$, that is $p = 3$, $q = 5$, with p which does not divide $q - 1 = 4$, then $G/Z(G)$ is cyclic (by Proposition 1.39) and thus G is abelian (by Exercise 15).

Exercise 46. Let G be a group, H a subgroup, and consider $N_G(H) = \{g \in G,\ gH = Hg\}$.

1. Show that the number of conjugates of H in G is equal to the index of $N_G(H)$ in G.
2. Deduce a formula for the number of Sylow p-subgroups of G.
3. Use the above to show that a simple group G of order 60 cannot have a subgroup of order 20.

Answer. First recall that $N_G(H)$ is a subgroup of G (see Exercise 6).

1. Let G act by conjugation on the set X of its subgroups. By the orbit-stabilizer theorem, we have that $|B(H)| = |G|/|N_G(H)|$ since $N_G(H)$ is the stabilizer of H. Now $|B(H)|$ is the number of conjugates of H, which proves the claim.
2. Let H be a Sylow p-subgroup of G. We know by the Sylow theorems that all Sylow p-subgroups are conjugate, thus $n_p = |B(H)|$ and n_p is the index of $N_G(H)$.
3. Assume by contradiction that K is a subgroup of order 20, then K has a unique Sylow 5-subgroup L which is then normal, and thus $K \subset N_G(L)$. Now the order of K must divide the order of $N_G(L)$ and the order of G, and since $|K| = 20$, we have that $|N_G(L)|$ is 20 or 60:
   - if it is 60, then the index of $N_G(L)$ is 1, and there is a unique Sylow 5-subgroup in G, which is then normal, and contradicts the simplicity of G.
   - if it is 20, then the index is 3, and there are 3 Sylow 5-subgroups in G, that is $n_5 = 3$ which contradicts the Sylow theorems: $n_5 \equiv 1 \mod 5$

Exercise 47. (*) Let G be a simple group of order 60, which thus cannot contain a subgroup of order 20.

1. Show that if G contains a subgroup K of order 12, then K contains 4 Sylow 3-subgroups.
2. Show that if H and K are two distinct subgroups of order 4 of G, then $H \cap K = \{1\}$.

Answer.

1. K is of order 12, thus its number n_3 of Sylow 3-subgroups is congruent to 1 mod 3, so
$$n_3 = 1 \text{ or } n_3 = 4$$
(it cannot be more since K is of order 12). To show that it must be 4, we show that it cannot be 1. Let us assume that K has a unique Sylow 3-subgroup L, then L must be normal in K, then K must be contained in $N_G(L) = \{g \in G,\ gL = Lg\}$, and since $N_G(L)$ is a subgroup of G, its

order must divide 60, and also be divisible by the order of K which is 12. Thus

$$|N_G(L)| = 12 \text{ or } 60.$$

Its index in G is then respectively 5 or 1. Since its index is also the number of Sylow 3-subgroups (because they are all conjugate, and using the Orbit Stabilizer theorem), that means that we have either 1 or 5 Sylow 3-subgroups in G, which is impossible: 1 is impossible since G is simple, and 5 is impossible since $|G| = 60$, and thus its number of Sylow 3-subgroups is congruent to 1 mod 3.

2. The intersection of two subgroups is a subgroup, thus its order is either 1,2 or 4. It cannot be 4 since they are distinct, to prove that it is 1, let us assume by contradiction that it is 2, thus a cyclic group of order 2:

$$H \cap K = \langle a \rangle$$

and $a^2 = 1$. Since the order of H is 4, the index of $\langle a \rangle$ in H is 2, and $\langle a \rangle$ is normal in H, that is $hah^{-1} = a$ and H is contained in $C_G(a) = \{g \in G,\ ga = ag\}$. The same holds for K and since both H and K are in $C_G(a)$, so must be their union $H \cup K$. Now the size of $H \cup K$ is 6, and thus $|C_G(a)| \geq 6$ and must be divisible by 4. The possibilities are

$$12, 20, 60.$$

Now 20 is impossible, by what is mentioned in the statement of the exercise. 12 is also impossible by the previous question: then $C_G(a)$ would have 4 Sylow 3-subgroups, showing that there should be only 1 Sylow 2-subgroup, but both H and K are in $C_G(a)$. This leaves 60: in that case, we would get that $\langle a \rangle$ is a normal subgroup, and thus it should be the unique Sylow 2-subgroup, which is also a contradiction.

2.9 The Jordan-Hölder Theorem

Exercise 48. Prove that every finite group has a composition series.

Answer. Take the longest possible subnormal series of G, say

$$\{1\} = G_n \lhd G_{n-1} \lhd \cdots \lhd G_0 = G$$

which is possible since G is finite. Then the composition factors are all simple. Indeed, if there exists a composition factor G_i/G_{i+1} which is not simple, then it has a non-trivial normal subgroup, which by the correspondence theorem gives a normal subgroup between G_i and G_{i+1}. This extends the subnormal series assumed to be the longest possible, thus a contradiction. The longest possible subnormal series is then a composition series.

Exercise 49. Prove that the infinite cyclic group G has no composition series.

Answer. Let $G = \langle g \rangle$ where $|g| = \infty$. Suppose there exists a composition series for G, that is

$$\{1\} = G_n \lhd G_{n-1} \lhd \cdots \lhd G_0 = G.$$

It cannot be that $n = 1$, that is

$$\{1\} = G_0 \lhd G$$

because that would mean that G is simple, and clearly G is not simple, because for example $\langle g^2 \rangle$ is a proper normal subgroup of G. Thus $n \geq 2$, and the composition series contains a subgroup G_{n-1} different than $\{1\}$ and G. Then G_{n-1} is a non-trivial subgroup of G, which means it is of the form $G_{n-1} = \langle g^k \rangle$ for some positive integer k. But then G_{n-1} is an infinite cyclic group, so it cannot be simple, which contradicts the definition of composition series.

Exercise 50. Show that the group $GL_n(\mathbb{R})$ has a subnormal series, but no composition series.

Answer. The series

$$\{1\} \lhd Z(GL_n(\mathbb{R})) \lhd SL_n(\mathbb{R}) \lhd GL_n(\mathbb{R})$$

is a subnormal series, however $GL_n(\mathbb{R})$ cannot have a composition series, because it contains a normal subgroup isomorphic to the infinite cyclic group (look at matrices of the form a scalar times the identity matrix, which clearly commute with every matrix in $GL_n(\mathbb{R})$).

2.10 Solvable and nilpotent groups

Exercise 51. Consider the general dihedral group

$$D_{2n} = \{a, b \mid a^n = b^2 = 1, b^{-1}ab = a^{-1}\}.$$

Is D_{2n} solvable? Prove your answer.

Answer. We have seen several equivalent definitions of G solvable. One of them is that G is solvable if and only if there exists a normal series

$$\{1\} = G_0 \lhd G_1 \lhd G_2 \lhd \cdots \lhd G_n = G$$

such that G_{i+1}/G_i is abelian. The rotations $\langle a \rangle$ form a subgroup of order n, this is not the case of the reflections (composing two reflections gives a rotation). We have that $\langle a \rangle$ is normal. So the series

$$\{1\} \lhd \langle a \rangle \lhd D_{2n}$$

is normal and the quotient $D_{2n}/\langle a \rangle$ has order two hence is the cyclic group of order 2 which is abelian.

Exercise 52. Consider thel dihedral group

$$D_6 = \{a, b \mid a^3 = b^2 = 1, b^{-1}ab = a^{-1}\},$$

which was shown to solvable in the previous exercise. Is D_6 nilpotent? Prove your answer.

Answer. It is not, because the number n_2 of its Sylow 2-subgroups is 3. Indeed, we know that $n_2 \equiv 1 \bmod 2$, and $n_2 \equiv 0 \bmod 3$. This contradicts the fact that a finite nilpotent group must have a unique Sylow p-subgroup for every p.

Exercise 53. Show that the nilpotency class of the quaternion group Q_8 is 2.

Answer. A central series for Q_8 is

$$\{1\} = G_2 \lhd G_1 = \{1, -1\} \lhd G_0 = Q_8.$$

Indeed, we notice that G_1 is the center of Q_8, thus it is normal in Q_8, which shows that this is a normal series. Now for $i = 1$, we have that $G_1/G_2 \simeq \{1, -1\} \subseteq Z(G/G_2) \simeq \{1, -1\}$. For $i = 0$, we have that $G_0/G_1 \simeq G/Z(G)$, since $G/Z(G)$ is of order 4 and cannot be cyclic (otherwise Q_8 would be abelian) it must be the Klein group, which is commutative (anyway knowing that the quotient group is of order 4 is enough to conclude the commutativity). Thus $G/Z(G) \subseteq Z(G/Z(G))$. Finally, the central series cannot be shortened, since Q_8 is not abelian, thus the nilpotency class of Q_8 is 2.

Exercise 54. (*) Let G be a group of order 16, which contains an element g of order 4. Show that $\langle g^2 \rangle$ is normal in G.

Answer. There are two cases:

1. If $\langle g \rangle$ is normal in G, then $\langle g^2 \rangle$ is also normal in G, since it is characteristic in G (this can be checked by the definition, take any automorphism f of $\langle g \rangle$, it thus maps some g^i to some g^j and thus $f((g^2)^i) = f(g^i)^2 = (g^j)^2 \in \langle g^2 \rangle$).

2. If $\langle g \rangle$ is not normal in G, then the subgroup

   $$H = \{h \in G,\ h\langle g \rangle = \langle g \rangle h\}$$

   (called the normalizer of $\langle g \rangle$) will be order 8. Indeed H is a subgroup of G, thus $|H|$ divides 16, that is $|H| = 1, 2, 4, 8, 16$, and $|H| \geq |\langle g \rangle|$, so that $|H| = 4, 8, 16$. It cannot be 16, since otherwise $\langle g \rangle$ would be normal in G. Finally, if $|H| = 4$, then that would mean that $\langle g \rangle = H$. This is not possible either (see result on nilpotent groups (currently Prop 1.50) in the lecture notes). If we can now prove that $\langle g^2 \rangle$ is characteristic in H, then we are done, since then $\langle g^2 \rangle$ is characteristic in H which is normal in G (it is of index 2 in G) thus $\langle g^2 \rangle$ is normal in G.

We are thus left to prove that if H is a group of order 8, with a cyclic group $\langle g \rangle$ of order 4, then $\langle g^2 \rangle$ is characteristic in H.

1. If H is abelian, then H is isomorphic to C_8 or to $\langle g \rangle \times C_2$. In both cases it is characteristic (in the first case, $\langle g^2 \rangle$ is a subgroup of a cyclic group, thus it is characteristic).

2. If H is not abelian, then its center $Z(H)$ is of order 2 (the order of the center has to divide $|H|$, if it were 4 or 8, then $H/Z(H)$ would be of order 2 or 1, thus cyclic, which implies that H is abelian). Furthermore, the center intersects non trivially every non-trivial normal subgroup of H. Thus $Z(H) = \langle g^2 \rangle$ since this is the only subgroup of $\langle g \rangle$ of order 2 and $\langle g \rangle$ is normal in H (it is of index 2). Finally $Z(H)$ is a characteristic subgroup of H which concludes the proof.

Exercise 55. True/False.

Q1. There are 3 kinds of groups of order 4, up to isomorphism.

Q2. Let H and K be two subgroups of G. Then HK is a subgroup of G.

Q3. Let G be a group, and let X be a set. Then the orbit B(x) of x in X under the action of G is a subgroup of G.

Q4. The dihedral group D_{10} of order 20 is simple.

Q5. The dihedral group D_3 is isomorphic to the symmetric group S_3.

Q6. Let H and N be two subgroups of G, with N normal in G. Then the following two quotient groups are isomorphic: $HN/N \simeq H/N$.

Q7. Every simple p-group G is abelian.

Q8. The number of elements in any conjugacy class of a finite group G divides the order of G.

Q9. The Klein group is a 2-group.

Q10. Let G be a cyclic group of order n. Then there is a subgroup of size d for each positive divisor d of n.

Answer.

Q1. False. There are only the Klein group $C_2 \times C_2$ and the cyclic group C_4.

Q2. False. Indeed, you need the extra condition that $HK = KH$ for it be true! If you try to find an inverse for hk, you will see you cannot find it. Of course, in G it should be $k^{-1}h^{-1}$, but without the assumption that $HK = KH$, this element has no reason to live in HK.

Q3. False. For $B(x)$ to be a subgroup of G, then $B(x)$ at least need to contain the identity element 1. However gx has no reason to be 1 in general, since x belongs to X which is an arbitrary set.

Q4. False. Remember that D_{10} has order $2 \cdot 10 = 20$. Now $20 = (2^2) \cdot 5 = p^2 q$ and we have seen that groups of such order cannot be simple.

Q5. True. One can for example check out the multiplication table for both groups.

Q6. False. In fact H/N is not even properly defined since N has no reason to be included in H to start with.

Q7. True. G is a p-group, thus its center is non-trivial (result proved in the lecture notes). The center of G is always normal in G. Now G is simple, thus its normal subgroups are only $\{1\}$ and G. Thus either the center is $\{1\}$ or it is G. It cannot be one since it's non-trivial, thus it is G and G is abelian.

Q8. True. Use the Orbit-Stabilizer theorem to deduce that the number of elements in an orbit divides the order of the group, and now notice that a conjugacy class is nothing else than an orbit when G acts on itself by conjugation.

Q9. True. The Klein group is of order 4, and is actually isomorphic to $C_2 \times C_2$. All its elements are order 2, so it is indeed a 2-group.

Q10. True. It is not true in general for an arbitrary group, but it is true for cyclic groups. Indeed, take g to be the generator of G of order n. Now if d divides n, then $n = kd$ for some k. Take the subgroup generated by g^k. Clearly $(g^k)^d = 1$ since $g^n = 1$. There cannot be a $d' < d$ such that $(g^k)^{d'} = 1$, otherwise this would mean the order of G is $< n$.

Chapter 3

Ring Theory

In the first section below, a ring will be defined as an abstract structure with a commutative addition, and a multiplication which may or may not be commutative. This distinction yields two quite different theories: the theory of respectively commutative or non-commutative rings. These notes are mainly concerned about commutative rings.

Non-commutative rings have been an object of systematic study only quite recently, during the 20th century. Commutative rings on the contrary have appeared though in a hidden way much before, and as many theories, it all goes back to Fermat's Last Theorem.

In 1847, the mathematician Lamé announced a solution of Fermat's Last Theorem, but Liouville noticed that the proof depended on a unique decomposition into primes, which he thought was unlikely to be true. Though Cauchy supported Lamé, Kummer was the one who finally published an example in 1844 (in an obscure journal, rediscovered in 1847) to show that the uniqueness of prime decompositions failed. Two years later, he restored the uniqueness by introducing what he called "ideal complex numbers" (today, simply "ideals") and used it to prove Fermat's Last Theorem for all $n < 100$ except $n = 37$, 59, 67 and 74.

It is Dedekind who extracted the important properties of "ideal numbers", defined an "ideal" by its modern properties: namely that of being a subgroup which is closed under multiplication by any ring element. He further introduced prime ideals as a generalization of prime numbers. Note that today we still use the terminology "Dedekind rings" to describe rings which have in particular a good behavior with respect to factorization of prime ideals. In 1882, an important paper by Dedekind and Weber developed the theory of rings of polynomials. At this stage, both rings of polynomials and rings of numbers (rings appearing in the context of Fermat's Last Theorem, such as what we call now the Gaussian integers) were being studied. But it was separately, and no one made connection between these two topics. Dedekind also introduced the term

"field" (Körper) for a commutative ring in which every non-zero element has a multiplicative inverse but the word "ring" is due to Hilbert, who, motivated by studying invariant theory, studied ideals in polynomial rings proving his famous "Basis Theorem" in 1893.

It will take another 30 years and the work of Emmy Noether and Krull to see the development of axioms for rings. Emmy Noether, about 1921, is the one who made the important step of bringing the two theories of rings of polynomials and rings of numbers under a single theory of abstract commutative rings.

In contrast to commutative ring theory, which grew from number theory, non-commutative ring theory developed from an idea of Hamilton, who attempted to generalize the complex numbers as a two dimensional algebra over the reals to a three dimensional algebra. Hamilton, who introduced the idea of a vector space, found inspiration in 1843, when he understood that the generalization was not to three dimensions but to four dimensions and that the price to pay was to give up the commutativity of multiplication. The quaternion algebra, as Hamilton called it, launched non-commutative ring theory.

Other natural non-commutative objects that arise are matrices. They were introduced by Cayley in 1850, together with their laws of addition and multiplication and, in 1870, Pierce noted that the now familiar ring axioms held for square matrices.

An early contributor to the theory of non-commutative rings was the Scottish mathematician Wedderburn, who in 1905, proved "Wedderburn's Theorem", namely that every finite division ring is commutative and so is a field.

It is only around the 1930's that the theories of commutative and non-commutative rings came together and that their ideas began to influence each other.

3.1 Rings, ideals and homomorphisms

Definition 3.1. A ring R is an abelian group with a multiplication operation

$$(a, b) \mapsto ab$$

which is associative, and satisfies the distributive laws

$$a(b + c) = ab + ac, \ (a + b)c = ac + bc$$

with identity element 1.

There is a group structure with the addition operation, but not necessarily with the multiplication operation. Thus an element of a ring may or may not be invertible with respect to the multiplication operation. Here is the terminology used.

Definition 3.2. Let a, b be in a ring R. If $a \neq 0$ and $b \neq 0$ but $ab = 0$, then we say that a and b are zero divisors. If $ab = ba = 1$, we say that a is a unit or that a is invertible.

While the addition operation is commutative, it may or not be the case with the multiplication operation.

Definition 3.3. Let R be ring. If $ab = ba$ for any a, b in R, then R is said to be commutative.

Here are the definitions of two particular kinds of rings where the multiplication operation behaves well.

Definition 3.4. An integral domain is a commutative ring with no zero divisor. A division ring or skew field is a ring in which every non-zero element a has an inverse a^{-1}. A field is a commutative ring in which every non-zero element is invertible.

Let us give two more definitions and then we will discuss several examples.

Definition 3.5. The characteristic of a ring R, denoted by char$R$, is the smallest positive integer such that

$$n \cdot 1 = \underbrace{1 + 1 + \ldots + 1}_{n\text{times}} = 0.$$

If there is no such positive integer, we say that the ring has characteristic 0.

We can also extract smaller rings from a given ring.

Definition 3.6. A subring of a ring R is a subset S of R that forms a ring under the operations of addition and multiplication defined in R.

Examples 3.1.

1. $\mathbb{Z}$ is an integral domain but not a field.
2. The integers modulo n form a commutative ring, which is an integral domain if and only if n is prime.
3. For $n \geq 2$, the $n \times n$ matrices $\mathcal{M}_n(\mathbb{R})$ with coefficients in $\mathbb{R}$ are a non-commutative ring, but not an integral domain.
4. The set
$$\mathbb{Z}[i] = \{a + bi,\ a, b \in \mathbb{Z}\},\ i^2 = -1,$$
is a commutative ring. It is also an integral domain, but not a field.
5. Let us construct the smallest and also most famous example of division ring. Take $1, i, j, k$ to be basis vectors for a 4-dimensional vector space over $\mathbb{R}$, and define multiplication by
$$i^2 = j^2 = k^2 = -1,\ ij = k,\ jk = i,\ ki = j,\ ji = -ij,\ kj = -jk,\ ik = -ki.$$
Then
$$\mathbb{H} = \{a + bi + cj + dk,\ a, b, c, d \in \mathbb{R}\}$$

	commutative	non-commutative
has zero divisor	integers mod n, n not a prime	matrices over a field
has no zero divisor	$\mathbb{Z}$	$\{a+bi+cj+dk,\ a,b,c,d \in \mathbb{Z}\}$
non-zero element invertible	$\mathbb{R}$	$\mathbb{H}$

forms a division ring, called the Hamilton's quaternions. So far, we have only seen the ring structure. Let us now discuss the fact that every non-zero element is invertible. Define the conjugate of an element $h = a+bi+cj+dk \in \mathbb{H}$ to be $\bar{h} = a-bi-cj-dk$ (yes, exactly the same way you did it for complex numbers). It is an easy computation (and a good exercise if you are not used to the non-commutative world) to check that

$$q\bar{q} = a^2 + b^2 + c^2 + d^2.$$

Now take q^{-1} to be

$$q^{-1} = \frac{\bar{q}}{q\bar{q}}.$$

Clearly $qq^{-1} = q^{-1}q = 1$ and the denominator cannot possibly be 0, but if $a = b = c = d = 0$.

6. If R is a ring, then the set $R[X]$ of polynomials with coefficients in R is a ring.

Similarly to what we did with groups, we now define a map from a ring to another which has the property of carrying one ring structure to the other.

Definition 3.7. Let R, S be two rings. A map $f : R \to S$ satisfying

1. $f(a+b) = f(a) + f(b)$ (this is thus a group homomorphism)
2. $f(ab) = f(a)f(b)$
3. $f(1_R) = 1_S$

for $a, b \in R$ is called ring homomorphism.

We do need to mention that $f(1_R) = 1_S$, otherwise, since a ring is not a group under multiplication, strange things can happen. For example, if $\mathbb{Z}_6$ denotes the integers mod 6, the map $f : \mathbb{Z}_6 \to \mathbb{Z}_6$, $n \mapsto 3n$ satisfies that $f(m+n) = 3(m+n) = 3m+3n = f(m)+f(n)$, and $f(n)f(m) = 3m3n = 3mn = f(mn)$ but $f(1) \neq 1$ and f is not a ring homomorphism. Notice the difference with group homomorphism: from $f(a+b) = f(a)+f(b)$, we deduce that $f(a+0) = f(a)+f(0)$, that is $f(a) = f(a)+f(0)$. Now because $f(a)$ is invertible, it must be that $f(0) = 0$! Once we reach $f(a) = f(a)f(1)$, because $f(a)$ does not have to be invertible, we cannot conclude!

The notion of "ideal number" was introduced by the mathematician Kummer, as being some special "numbers" (well, nowadays we call them groups) having the property of unique factorization, even when considered over more

general rings than $\mathbb{Z}$ (a bit of algebraic number theory would be good to make this more precise). Today only the name "ideal" is left, and here is what it gives in modern terminology:

Definition 3.8. Let $\mathcal{I}$ be a subset of a ring R. Then an additive subgroup of R having the property that

$$ra \in \mathcal{I} \text{ for } a \in \mathcal{I},\ r \in R$$

is called a left ideal of R. If instead we have

$$ar \in \mathcal{I} \text{ for } a \in \mathcal{I},\ r \in R$$

we say that we have a right ideal of R. If an ideal happens to be both a right and a left ideal, then we call it a two-sided ideal of R, or simply an ideal of R.

Example 3.2. The even integers $2\mathbb{Z} = \{2n,\ n \in \mathbb{Z}\}$ form an ideal of $\mathbb{Z}$. The set of polynomials in $\mathbb{R}[X]$ with constant coefficient zero form an ideal of $\mathbb{R}[X]$.

Of course, for any ring R, both R and $\{0\}$ are ideals. We thus introduce some terminology to precise whether we consider these two trivial ideals.

Definition 3.9. We say that an ideal $\mathcal{I}$ of R is proper if $\mathcal{I} \neq R$. We say that is it non-trivial if $\mathcal{I} \neq R$ and $\mathcal{I} \neq 0$.

If $f : R \to S$ is a ring homomorphism, we define the kernel of f in the most natural way:

$$\mathrm{Ker} f = \{r \in R,\ f(r) = 0\}.$$

Since a ring homomorphism is in particular a group homomorphism, we already know that f is injective if and only if $\mathrm{Ker} f = \{0\}$. It is easy to check that $\mathrm{Ker} f$ is a proper two-sided ideal:

- $\mathrm{Ker} f$ is an additive subgroup of R.
- Take $a \in \mathrm{Ker} f$ and $r \in R$. Then

  $$f(ra) = f(r)f(a) = 0 \text{ and } f(ar) = f(a)f(r) = 0$$

  showing that ra and ar are in $\mathrm{Ker} f$.
- Then $\mathrm{Ker} f$ has to be proper (that is, $\mathrm{Ker} f \neq R$), since $f(1) = 1$ by definition.

We can thus deduce the following (extremely useful) result.

Lemma 3.1. *Suppose $f : R \to S$ is a ring homomorphism and the only two-sided ideals of R are $\{0\}$ and R. Then f is injective.*

Proof. Since $\mathrm{Ker} f$ is a two-sided ideal of R, then either $\mathrm{Ker} f = \{0\}$ or $\mathrm{Ker} f = R$. But $\mathrm{Ker} f \neq R$ since $f(1) = 1$ by definition (in words, $\mathrm{Ker} f$ is a proper ideal). □

At this point, it may be worth already noticing the analogy between on the one hand rings and their two-sided ideals, and on the other hand groups and their normal subgroups.

- Two-sided ideals are stable when the ring acts on them by multiplication, either on the right or on the left, and thus

$$rar^{-1} \in \mathcal{I},\ a \in \mathcal{I},\ r \in R,$$

  while normal subgroups are stable when the groups on them by conjugation

$$ghg^{-1} \in H,\ h \in H,\ g \in G\ (H \leq G).$$

- Groups with only trivial normal subgroups are called simple. We will not see it formally here, but rings with only trivial two-sided ideals as in the above lemma are called simple rings.

- The kernel of a group homomorphism is a normal subgroup, while the kernel of a ring homomorphism is an ideal.

- Normal subgroups allowed us to define quotient groups. We will see now that two-sided ideals will allow to define quotient rings.

3.2 Quotient rings

Let $\mathcal{I}$ be a proper two-sided ideal of R. Since $\mathcal{I}$ is an additive subgroup of R by definition, it makes sense to speak of cosets $r + \mathcal{I}$ of $\mathcal{I}$, $r \in R$. Furthermore, a ring has a structure of abelian group for addition, so $\mathcal{I}$ satisfies the definition of a normal subgroup. From group theory, we thus know that it makes sense to speak of the quotient group

$$R/\mathcal{I} = \{r + \mathcal{I},\ r \in R\},$$

group which is actually abelian (inherited from R being an abelian group for the addition).

We now endow $R/\mathcal{I}$ with a multiplication operation as follows. Define

$$(r + \mathcal{I})(s + \mathcal{I}) = rs + \mathcal{I}.$$

Let us make sure that this is well-defined, namely that it does not depend on the choice of the representative in each coset. Suppose that

$$r + \mathcal{I} = r' + \mathcal{I},\ s + \mathcal{I} = s' + \mathcal{I},$$

so that $a = r' - r \in \mathcal{I}$ and $b = s' - s \in \mathcal{I}$. Now

$$r's' = (a + r)(b + s) = ab + as + rb + rs \in rs + \mathcal{I}$$

since ab, as and rb belongs to $\mathcal{I}$ using that $a, b \in \mathcal{I}$ and the definition of ideal. This tells us $r's'$ is also in the coset $rs + \mathcal{I}$ and thus multiplication does not depend on the choice of representatives. Note though that this is true only because we assumed a two-sided ideal $\mathcal{I}$, otherwise we could not have concluded, since we had to deduce that both as and rb are in $\mathcal{I}$.

Definition 3.10. The set of cosets of the two-sided ideal $\mathcal{I}$ given by

$$R/\mathcal{I} = \{r + \mathcal{I},\ r \in R\}$$

is a ring with identity $1_R + \mathcal{I}$ and zero element $0_R + \mathcal{I}$ called a quotient ring.

Note that we need the assumption that $\mathcal{I}$ is a proper ideal of R to claim that $R/\mathcal{I}$ contains both an identity and a zero element (if $R = \mathcal{I}$, then $R/\mathcal{I}$ has only one element).

Example 3.3. Consider the ring of matrices $\mathcal{M}_2(\mathbb{F}_2[i])$, where $\mathbb{F}_2$ denotes the integers modulo 2, and i is such that $i^2 = -1 \equiv 1 \mod 2$. This is thus the ring of 2×2 matrices with coefficients in

$$\mathbb{F}_2[i] = \{a + ib,\ a, b \in \{0, 1\}\}.$$

Let $\mathcal{I}$ be the subset of matrices with coefficients taking values 0 and $1 + i$ only. It is a two-sided ideal of $\mathcal{M}_2(\mathbb{F}_2[i])$. Indeed, take a matrix $U \in \mathcal{I}$, a matrix $M \in \mathcal{M}_2(\mathbb{F}_2[i])$, and compute UM and MU. An immediate computation shows that all coefficients are of the form $a(1+i)$ with $a \in \mathbb{F}_2[i]$, that is all coefficients are in $\{0, 1 + i\}$. Clearly $\mathcal{I}$ is an additive group.

We then have a quotient ring

$$\mathcal{M}_2(\mathbb{F}_2[i])/\mathcal{I}.$$

We have seen that $\mathrm{Ker} f$ is a proper two-sided ideal when f is a ring homomorphism. We now prove the converse.

Proposition 3.2. *Every proper two-sided ideal $\mathcal{I}$ is the kernel of a ring homomorphism.*

Proof. Consider the canonical projection π that we know from group theory. Namely

$$\pi : R \to R/\mathcal{I},\ r \mapsto \pi(r) = r + \mathcal{I}.$$

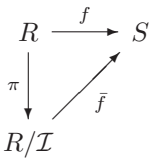
We already know that π is group homomorphism, and that its kernel is $\mathcal{I}$. We are only left to prove that π is a ring homomorphism:

- since $\mathcal{I}$ is two-sided, then $R/\mathcal{I}$ is a ring.
- $\pi(rs) = rs + \mathcal{I} = (r + \mathcal{I})(s + \mathcal{I}) = \pi(r)\pi(s)$.
- $\pi(1_R) = 1_R + \mathcal{I}$ which is indeed the identity element of $R/\mathcal{I}$.

□

We are now ready to state a factor theorem and a 1st isomorphism theorem for rings, the same way we did for groups. It may help to keep in mind the analogy between two-sided ideals and normal subgroups mentioned above.

Assume that we have a ring R which contains a proper two-sided ideal $\mathcal{I}$, another ring S, and $f : R \to S$ a ring homomorphism. Let π be the canonical projection from R to the quotient group $R/\mathcal{I}$:



We would like to find a ring homomorphism $\bar{f} : R/\mathcal{I} \to S$ that makes the diagram commute, namely

$$f(a) = \bar{f}(\pi(a))$$

for all $a \in R$.

Theorem 3.3. (Factor Theorem for Rings). *Any ring homomorphism f whose kernel K contains $\mathcal{I}$ can be factored through $R/\mathcal{I}$. In other words, there is a unique ring homomorphism $\bar{f} : R/\mathcal{I} \to S$ such that $\bar{f} \circ \pi = f$. Furthermore*

1. *$\bar{f}$ is an epimorphism if and only if f is.*
2. *$\bar{f}$ is a monomorphism if and only if $K = \mathcal{I}$.*
3. *$\bar{f}$ is an isomorphism if and only if f is an epimorphism and $K = \mathcal{I}$.*

Proof. Since we have already done the proof for groups with many details, here we will just mention a few important points in the proof.

Let $a + \mathcal{I} \in R/\mathcal{I}$ such that $\pi(a) = a + \mathcal{I}$ for $a \in R$. We define

$$\bar{f}(a + \mathcal{I}) = f(a).$$

This is the most natural way to do it, however, we need to make sure that this is indeed well-defined, in the sense that it should not depend on the choice of the representative taken in the coset. Let us thus take another representative, say $b \in a + \mathcal{I}$. Since a and b are in the same coset, they satisfy $a - b \in \mathcal{I} \subset K$, where $K = \mathrm{Ker}(f)$ by assumption. Since $a - b \in K$, we have $f(a - b) = 0$ and thus $f(a) = f(b)$.

Now that $\bar{f}$ is well defined, it is an easy computation to check that $\bar{f}$ inherits the property of ring homomorphism from f.

The rest of the proof works exactly the same as for groups. □

The first isomorphism theorem for rings is similar to the one for groups.

Theorem 3.4. (1st Isomorphism Theorem for Rings). *If $f : R \to S$ is a ring homomorphism with kernel K, then the image of f is isomorphic to R/K:*

$$\mathrm{Im}(f) \simeq R/\mathrm{Ker}(f).$$

Proof. We know from the Factor Theorem that

$$\bar{f} : R/\mathrm{Ker}(f) \to S$$

is an isomorphism if and only if f is an epimorphism, and clearly f is an epimorphism on its image, which concludes the proof. □

Example 3.4. Let us finish Example 3.3. We showed there that $\mathcal{M}_2(\mathbb{F}_2[i])/\mathcal{I}$ is a quotient ring, where $\mathcal{I}$ is the ideal formed of matrices with coefficients in $\{0, 1+i\}$. Consider the ring homomorphism:

$$f : \mathcal{M}_2(\mathbb{F}_2[i]) \to \mathcal{M}_2(\mathbb{F}_2), M = (m_{k,l}) \mapsto f(M) = (m_{k,l} \mod 1+i)$$

that is f looks at the coefficients of $M \mod 1+i$. Its kernel is $\mathcal{I}$ and it is surjective. By the first isomorphism for rings, we have

$$\mathcal{M}_2(\mathbb{F}_2[i])/\mathcal{I} \simeq \mathcal{M}_2(\mathbb{F}_2).$$

Example 3.5. A less exotic example, which we will study in more details later on, is the following. Consider the map $f : \mathbb{R}[X] \to \mathbb{C}$, $f(p(X)) = p(i)$, that is, f takes a polynomial $p(X)$ with real coefficients, and evaluate this polynomial in i ($i^2 = -1$). This map is surjective (take the polynomial $p(X) = X + (z - i)$, $z \in \mathbb{C}$) and its kernel is formed by polynomials which, when evaluated in i, are giving 0, meaning that i is a root of the polynomial, or equivalently that (X^2+1) is a factor of the polynomial. Thus $\mathrm{Ker}(f) = (X^2+1)\mathbb{R}[X] = \{p(X) = (X^2+1)q(X),\ q(X) \in \mathbb{R}[X]\}$. Using the first isomorphism for rings, we have

$$\mathbb{R}[X]/(X^2+1)\mathbb{R}[X] \simeq \mathbb{C}.$$

3.3 The Chinese Remainder Theorem

The name "Chinese Remainder Theorem" supposely comes from the following question: How many soldiers were part of Han Xing's army if, sorted by 3 columns, 2 soldiers were left, sorted by 5 columns, 3 soldiers were left, and sorted by 7 columns, 2 soldiers were left.

The Chinese Remained Theorem is attributed to Sun Zi (in the 3rd century), and was later published by Qin Jiushao (around 1247).

We will prove a "general" Chinese Remainder Theorem, rephrased in terms of rings and ideals.

For that let us start by introducing some new definitions about ideals, that will collect some of the manipulations one can do on ideals. Let us start with the sum.

Definition 3.11. Let $\mathcal{I}$ and $\mathcal{J}$ be two ideals of a ring R. The sum of $\mathcal{I}$ and $\mathcal{J}$ is the ideal

$$\mathcal{I} + \mathcal{J} = \{x + y,\ x \in \mathcal{I},\ y \in \mathcal{J}\}.$$

If $\mathcal{I}$ and $\mathcal{J}$ are right (resp. left) ideals, so is their sum.

Note that the intersection $\mathcal{I} \cap \mathcal{J}$ of two (resp. right, left, two-sided) ideals of R is again a (resp. right, left, two-sided) ideal of R.

Definition 3.12. The product of two left (resp. right) ideals $\mathcal{I}$ and $\mathcal{J}$ is the left (resp. right) ideal

$$\mathcal{I}\mathcal{J} = \{\sum_{i=1}^{n} x_i y_i,\ x_i \in \mathcal{I},\ y_i \in \mathcal{J}\}.$$

Example 3.6. Take $\mathcal{I} = 2\mathbb{Z}$ and $\mathcal{J} = 3\mathbb{Z}$ which are both two-sided ideals of $\mathbb{Z}$. We have

$$\mathcal{I} + \mathcal{J} = \{2x + 3y,\ x, y \in \mathbb{Z}\} = \mathbb{Z},$$

using Bezout identity (since $\gcd(2, 3) = 1$). Also

$$\mathcal{I} \cap \mathcal{J} = 6\mathbb{Z},\ \mathcal{I}\mathcal{J} = \{\sum_{i=1}^{n} 2x_i 3y_i,\ x, y \in \mathbb{Z}\} = 6\mathbb{Z}.$$

We can define a notion of being co-prime for ideals as follows.

Definition 3.13. The two-sided ideals $\mathcal{I}$ and $\mathcal{J}$ of a ring R are relatively prime if

$$\mathcal{I} + \mathcal{J} = R.$$

In a sense, this definition generalizes Bezout identity for rings.

Notice that for a commutative ring, if $\mathcal{I}$ and $\mathcal{J}$ are relatively prime then

$$\mathcal{I}\mathcal{J} = \mathcal{I} \cap \mathcal{J}.$$

(This is also illustrated in the above example.) Indeed, we clearly have that

$$\mathcal{I}\mathcal{J} \subset \mathcal{I} \cap \mathcal{J}$$

since $\mathcal{I}\mathcal{J}$ contains by definition sums of elements xy, $x \in \mathcal{I}, y \in \mathcal{J}$, with $xy \in \mathcal{I}$ and $xy \in \mathcal{J}$ by definition of two-sided ideal. Conversely

$$\mathcal{I} \cap \mathcal{J} \subset \mathcal{I}\mathcal{J}$$

since there exist $x \in \mathcal{I}, y \in \mathcal{J}$ such that $x + y = 1$ by definition of relatively prime, and for every element $a \in \mathcal{I} \cap \mathcal{J}$, we have that

$$a = a(x + y) = ax + ay = xa + ay \in \mathcal{I}\mathcal{J}.$$

For R a non-commutative ring, where $\mathcal{I}, \mathcal{J}$ are two-sided and co-prime, all we can say is that

$$\mathcal{I} \cap \mathcal{J} = \mathcal{I}\mathcal{J} + \mathcal{J}\mathcal{I}.$$

Indeed, $a(x + y) = ax + ay \in \mathcal{J}\mathcal{I} + \mathcal{I}\mathcal{J}$ since $ax \neq xa$.

Finally, let us extend the notion of "modulo" to ideals.

Definition 3.14. If $a, b \in R$ and $\mathcal{I}$ is an ideal of R, we say that a is congruent to b modulo $\mathcal{I}$ if

$$a - b \in \mathcal{I}.$$

A last definition this time about rings is needed before we can state the theorem.

Definition 3.15. If $R_1, \ldots, R_n$ are rings, the direct product of $R_1, \ldots, R_n$, denoted by $\prod_{i=1}^{n} R_i$, is defined as the ring of n-tuples $(a_1, \ldots, a_n)$, $a_i \in R_i$, with componentwise addition and multiplication. The zero element is $(0, \ldots, 0)$ and the identity is $(1, \ldots, 1)$ where 1 means 1_{R_i} for each i.

This definition is an immediate generalization of the direct product we studied for groups.

Theorem 3.5. (Chinese Remainder Theorem). *Let R be a commutative ring, and let $\mathcal{I}_1, \ldots, \mathcal{I}_n$ be ideals in R, such that*

$$\mathcal{I}_i + \mathcal{I}_j = R, \ i \neq j.$$

1. *If $a_1, \ldots, a_n$ are elements of R, there exists an element $a \in R$ such that*

$$a \equiv a_i \mod \mathcal{I}_i, \ i = 1, \ldots, n.$$

2. *If b is another element of R such that $b \equiv a_i \mod \mathcal{I}_i$, $i = 1, \ldots, n$, then*

$$b \equiv a \mod \cap_{i=1}^{n} \mathcal{I}_i.$$

Conversely, if b satisfies the above congruence, then $b \equiv a_i \mod \mathcal{I}_i$, $i = 1, \ldots, n$.

3. *We have that*

$$R/\cap_{i=1}^{n} \mathcal{I}_i \simeq \prod_{i=1}^{n} R/\mathcal{I}_i.$$

Proof. 1. For $j > 1$, we have by assumption that $\mathcal{I}_1 + \mathcal{I}_j = R$, and thus there exist $b_j \in \mathcal{I}_1$ and $d_j \in \mathcal{I}_j$ such that

$$b_j + d_j = 1, \ j = 2, \ldots, n.$$

This yields that

$$\prod_{j=2}^{n} (b_j + d_j) = 1. \tag{3.1}$$

Now if we look at the left hand side of the above equation, we have

$$(b_2 + d_2)(b_3 + d_3) \cdots (b_n + d_n) = (\underbrace{b_2 b_3 + b_2 d_3 + d_2 b_3}_{\in \mathcal{I}_1} + d_2 d_3) \cdots (b_n + d_n)$$

and all the terms actually belong to $\mathcal{I}_1$, but $c_1 := \prod_{j=2}^n d_j \in \prod_{j=2}^n \mathcal{I}_j$. Thus

$$c_1 \equiv 1 \mod \mathcal{I}_1$$

from (3.1). On the other hand, we also have

$$c_1 \equiv 0 \mod \mathcal{I}_j$$

for $j > 1$ since $c_1 \in \prod_{j=2}^n \mathcal{I}_j$.

More generally, for all i, we can find c_i with

$$c_i \equiv 1 \mod \mathcal{I}_i,\ c_i \equiv 0 \mod \mathcal{I}_j,\ j \neq i.$$

Now take arbitrary elements $a_1, \ldots, a_n \in R$, and set

$$a = a_1 c_1 + \ldots + a_n c_n.$$

Let us check that a is the solution we are looking for. Since $c_j \equiv 0 \mod \mathcal{I}_j$,$j \neq i$, we have for a given i that

$$a \equiv a_i c_i \equiv a_i \mod \mathcal{I}_i$$

using that $c_i \equiv 1 \mod \mathcal{I}_i$.

2. We have just shown the existence of a solution a modulo $\mathcal{I}_i$ for $i = 1, \ldots, n$. We now discuss the question of unicity, and show that the solution is actually not unique, but any other solution than a is actually congruent to $a \mod \cap_{i=1}^n \mathcal{I}_i$.

   We have for all $i = 1, \ldots, n$ that

   $$b \equiv a_i \mod \mathcal{I}_i \iff b \equiv a \mod \mathcal{I}_i \iff b - a \equiv 0 \mod \mathcal{I}_i$$

   which finally is equivalent to

   $$b - a \in \cap_{i=1}^n \mathcal{I}_i.$$

3. Define the ring homomorphism $f : R \to \prod_{i=1}^n R/\mathcal{I}_i$, sending

   $$a \mapsto f(a) = (a + \mathcal{I}_1, \ldots, a + \mathcal{I}_n).$$

   - This map is surjective: take any $(a_1 + \mathcal{I}_1, \ldots, a_n + \mathcal{I}_n) \in \prod_{i=1}^n R/\mathcal{I}_i$, then we must find an $a \in R$ such that $f(a) = (a_1 + \mathcal{I}_1, \ldots, a_n + \mathcal{I}_n)$, that is $a + \mathcal{I}_i = a_i + \mathcal{I}_i$, or equivalently $a_i \equiv a \mod \mathcal{I}_i$, which is true by the first point.
   - Its kernel is given by

   $$\begin{aligned} \mathrm{Ker} f &= \{a \in R,\ f(a) = (\mathcal{I}_1, \ldots, \mathcal{I}_n)\} \\ &= \{a \in R,\ a \in \mathcal{I}_i,\ i = 1, \ldots, n\} \\ &= \prod_{i=1}^n \mathcal{I}_i. \end{aligned}$$

   We conclude using the first isomorphism Theorem for rings.

□

Example 3.7. If $R = \mathbb{Z}$, the Chinese Remainder Theorem simplifies to say that if $n = \prod_i n_i$ where the n_i are coprime, then

$$\mathbb{Z}/n\mathbb{Z} \simeq \prod_i \mathbb{Z}/n_i\mathbb{Z}.$$

In the particular case of Example 3.6, we have

$$\mathbb{Z}/6\mathbb{Z} \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}.$$

This version of the Chinese remainder Theorem does not hold in the non-commutative case, because the property that $\mathcal{I}\mathcal{J} = \mathcal{I} \cap \mathcal{J}$ does not hold anymore, as pointed out earlier. There is though a commutative version if all the co-prime ideals are assumed to be two-sided.

3.4 Maximal and prime ideals

Here are a few special ideals.

Definition 3.16. The ideal generated by the non-empty set X of R is the smallest ideal of R that contains X. It is denoted by $\langle X \rangle$. It is the collection of all finite sums of the form $\sum_i r_i x_i s_i$.

Definition 3.17. An ideal generated by a single element a is called a principal ideal, denoted by $\langle a \rangle$.

Definition 3.18. A maximal ideal in the ring R is a proper ideal that is not contained in any strictly larger proper ideal.

One can prove that every proper ideal is contained in a maximal ideal, and that consequently every ring has at least one maximal ideal. We skip the proof here, since it heavily relies on set theory, requires many new definitions and the use of Zorn's lemma.

Instead, let us mention that a correspondence Theorem exists for rings, the same way it exists for groups, since we will need it for characterizing maximal ideals.

Theorem 3.6. (Correspondence Theorem for rings). *If $\mathcal{I}$ is a two-sided ideal of a ring R, then the canonical map*

$$\pi : R \to R/\mathcal{I}$$

sets up a one-to-one correspondence between

- *the set of all subrings of R containing $\mathcal{I}$ and the set of all subrings of $R/\mathcal{I}$,*
- *the set of all ideals of R containing $\mathcal{I}$ and the set of all ideals of $R/\mathcal{I}$.*

Here is a characterization of maximal ideals in commutative rings.

Theorem 3.7. *Let M be an ideal in the commutative ring R. We have*

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

Proof. Let us start by assuming that M is maximal. Since R/M is a ring, we need to find the multiplicative inverse of $a+M \in R/M$ assuming that $a+M \neq 0$ in R/M, that is $a \notin M$. Since M is maximal, the ideal $Ra + M$ has to be R itself, since $M \subset Ra + M$. Thus $1 \in Ra + M = R$, that is

$$1 = ra + m,\ r \in R,\ m \in M.$$

Then

$$(r + M)(a + M) = ra + M = (1 - m) + M = 1 + M$$

proving that $r + M$ is $(a + M)^{-1}$.

Conversely, let us assume that R/M is a field. First we notice that M must be a proper ideal of R, since if $M = R$, then R/M contains only one element and $1 = 0$.

Let N be an ideal of R such that $M \subset N \subset R$ and $N \neq R$. We have to prove that $M = N$ to conclude that M is maximal.

By the correspondence Theorem for rings, we have a one-to-one correspondence between the set of ideals of R containing M, and the set of ideals of R/M. Since N is such an ideal, its image $\pi(N) \in R/M$ must be an ideal of R/M, and thus must be either $\{0\}$ or R/M (since R/M is a field). The latter yields that $N = R$, which is a contradiction, letting as only possibility that $\pi(N) = \{0\}$, and thus $N = M$, which completes the proof. □

Definition 3.19. A prime ideal in a commutative ring R is a proper ideal P of R such that for any $a, b \in R$, we have that

$$ab \in P \Rightarrow a \in P \text{ or } b \in P.$$

Here is again a characterization of a prime ideal P of R in terms of its quotient ring R/P.

Theorem 3.8. *If P is an ideal in the commutative ring R*

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

Proof. Let us start by assuming that P is prime. It is thus proper by definition, and R/P is a ring. We must show that the definition of integral domain holds, namely that

$$(a + P)(b + P) = 0 + P \Rightarrow a + P = P \text{ or } b + P = P.$$

Since

$$(a + P)(b + P) = ab + P = 0 + P,$$

we must have $ab \in P$, and thus since P is prime, either $a \in P$ or $b \in P$, implying respectively that either $a + P = P$ or $b + P = P$.

Conversely, if R/P is an integral domain, then P must be proper (otherwise $1 = 0$). We now need to check the definition of a prime ideal. Let us thus consider $ab \in P$, implying that

$$(a + P)(b + P) = ab + P = 0 + P.$$

Since R/P is an integral domain, either $a + P = P$ or $b + P = P$, that is

$$a \in P \text{ or } b \in P,$$

which concludes the proof. □

Corollary 3.9. *In a commutative ring, a maximal ideal is prime.*

Proof. If M is maximal, then R/M is a field, and thus an integral domain, so that M is prime. □

Corollary 3.10. *Let $f : R \to S$ be an epimorphism of commutative rings.*

1. *If S is a field, then* $\mathrm{Ker} f$ *is a maximal ideal of R.*
2. *If S is an integral domain, then* $\mathrm{Ker} f$ *is a prime ideal of R.*

Proof. By the first isomorphism theorem for rings, we have that

$$S \simeq R/\mathrm{Ker} f.$$

□

Example 3.8. Consider the ring $\mathbb{Z}[X]$ of polynomials with coefficients in $\mathbb{Z}$, and the ideal generated by the indeterminate X, that is $\langle X \rangle$ is the set of polynomials with constant coefficient 0. Clearly $\langle X \rangle$ is a proper ideal. To show that it is prime, consider the following ring homomorphism:

$$\varphi : \mathbb{Z}[X] \to \mathbb{Z},\ f(X) \mapsto \varphi(f(X)) = f(0).$$

We have that $\langle X \rangle = \mathrm{Ker}\varphi$ which is prime by the above corollary.

3.5 Polynomial rings

For this section, we assume that R is a commutative ring. Set $R[X]$ to be the set of polynomials in the indeterminate X with coefficients in R. It is easy to see that $R[X]$ inherits the properties of ring from R.

We define the evaluation map E_x, which evaluates a polynomial $f(X) \in R[X]$ in $x \in R$, as

$$E_x : R[X] \to R,\ f(X) \mapsto f(X)|_{X=x} = f(x).$$

We can check that E_x is a ring homomorphism.

The degree of a polynomial is defined as usual, that is, if $p(X) = a_0 + a_1 X + \ldots + a_n X^n$ with $a_n \neq 0$, then $\deg(p(X)) = \deg p = n$. By convention, we set $\deg(0) = -\infty$.

Euclidean division will play an important role in what will follow. Let us start by noticing that there exists a polynomial division algorithm over $R[X]$, namely: if $f, g \in R[X]$, with g monic, then there exist unique polynomials q and r in $R[X]$ such that

$$f = qg + r, \ \deg r < \deg g.$$

The requirement that g is monic comes from R being a ring and not necessarily a field. If R is a field, g does not have to be monic, since one can always multiply g by the inverse of the leading coefficient, which is not possible if R is not a field.

Example 3.9. Take $f(X) = X^2 - 2$ and $g(X) = 2X - 1$. It is not possible to divide $f(X)$ by $g(X)$ in $\mathbb{Z}[X]$. If it were, then

$$f(X) = X^2 - 2 = (q_0 + q_1 X)(2X - 1) + r_0$$

and the coefficient of X^2 is 1 on the left hand side, and $2q_1$ on the right hand side. Now in $\mathbb{Z}$, there is no solution to the equation $2q_1 = 1$. Of course, this is possible in $\mathbb{Q}$, by taking $q_1 = 1/2$!

This gives the following:

Theorem 3.11. (Remainder Theorem). *If $f \in R[X]$, $a \in R$, then there exists a unique polynomial $q(X) \in R[X]$ such that*

$$f(X) = q(X)(X - a) + f(a).$$

Hence $f(a) = 0 \iff X - a \mid f(X)$.

Proof. Since $(X - a)$ is monic, we can do the division

$$f(X) = q(X)(X - a) + r(X).$$

But now since $\deg r < \deg(X - a)$, $r(X)$ must be a constant polynomial, which implies that

$$f(a) = r(X)$$

and thus

$$f(X) = q(X)(X - a) + f(a)$$

as claimed. Furthermore, we clearly have that

$$f(a) = 0 \iff X - a \mid f(X).$$

□

The following result sounds well known, care should be taken not to generalize it to rings which are not integral domain!

Theorem 3.12. *If R is an integral domain, then a non-zero polynomial f in $R[X]$ of degree n has at most n roots in R, counting multiplicity.*

Proof. If f has no root in $R[X]$, then we are done. Let us thus assume that f has a root a_1 in R, that is $f(a_1) = 0$. Then

$$X - a_1 \mid f(X)$$

by the remainder Theorem above, meaning that

$$f(X) = q_1(X)(X - a_1)^{n_1}$$

where $q_1(a_1) \neq 0$ and $\deg q_1 = n - n_1$ since R is an integral domain. Now if a_1 is the only root of f in R, then $n_1 \leq n$ and we are done. If not, consider similarly $a_2 \neq a_1$ another root of f, so that

$$0 = f(a_2) = q_1(a_2)(a_2 - a_1)^{n_1}.$$

Since R is an integral domain, we must have that $q_1(a_2) = 0$, and thus a_2 is a root of $q_1(X)$. We can repeat the process with $q_1(X)$ instead of $f(X)$: since a_2 is a root of $q_1(X)$, we have

$$q_1(X) = q_2(X)(X - a_2)^{n_2}$$

with $q_2(a_2) \neq 0$ and $\deg q_2 = n - n_1 - n_2$. By going on iterating the process, we obtain

$$\begin{aligned} f(X) &= q_1(X)(X - a_1)^{n_1} \\ &= q_2(X)(X - a_2)^{n_2}(X - a_1)^{n_1} \\ &= \dots \\ &= (X - a_1)^{n_1}(X - a_2)^{n_2} \cdots (X - a_k)^{n_k} \cdot c(X) \end{aligned}$$

where $c(X)$ is a polynomial with no root in R, possibly constant, and

$$n \geq n_1 + n_2 + \cdots + n_k.$$

Since R is an integral domain, the only possible roots of f are $a_1, \dots, a_k$, $k \leq n$, and the number of roots counting multiplicity is less than n. □

Example 3.10. Take $R = \mathbb{Z}_8$ the ring of integers modulo 8. Consider the polynomial

$$f(X) = X^3.$$

It is easy to check that is has 4 roots: $0, 2, 4, 6$. This comes from the fact that $\mathbb{Z}_8$ is not an integral domain.

3.6 Unique factorization and Euclidean division

In this section, all rings are assumed to be integral domains.

Let us start by defining formally the notions of irreducible and prime. The elements a, b, c, u in the definitions below all belong to an integral domain R.

Definition 3.20. The elements a, b are called associate if $a = ub$ for some unit u.

Definition 3.21. Let a be a non-zero element which is not a unit. Then a is said to be irreducible if $a = bc$ implies that either b or c must be a unit.

Definition 3.22. Let a be a non-zero element which is not a unit. Then a is called prime if whenever $a \mid bc$, then $a \mid b$ or $a \mid c$.

Between prime and irreducible, which notion is the stronger? The answer is in the proposition below.

Proposition 3.13. *If a is prime, then a is irreducible.*

Proof. Suppose that a is prime, and that $a = bc$. We want to prove that either b or c is a unit. By definition of prime, we must have that a divides either b or c. Let us say that a divides b. Thus

$$b = ad \Rightarrow b = bcd \Rightarrow b(1 - cd) = 0 \Rightarrow cd = 1$$

using that R is an integral domain, and thus c is a unit. The same argument works if we assume that a divides c, and we conclude that a is irreducible. □

Example 3.11. Consider the ring

$$R = \mathbb{Z}[\sqrt{-3}] = \{a + ib\sqrt{3},\ a, b \in \mathbb{Z}\}.$$

We want to see that 2 is irreducible but not prime.

- Let us first check that 2 is indeed irreducible. Suppose that

  $$2 = (a + ib\sqrt{3})(c + id\sqrt{3}).$$

  Since 2 is real, it is equal to its conjugate, and thus

  $$2\overline{2} = (a + ib\sqrt{3})(c + id\sqrt{3})(a - ib\sqrt{3})(c - id\sqrt{3})$$

  implies that

  $$4 = (a^2 + 3b^2)(c^2 + 3d^2).$$

  We deduce that $a^2 + 3b^2$ must divide 4, and it cannot possibly be 2, since we have a sum of squares in $\mathbb{Z}$. If $a^2 + 3b^2 = 4$, then $c^2 + 3d^2 = 1$ and $d = 0$, $c = \pm 1$. Vice versa if $c^2 + 3d^2 = 4$ then $a^2 + 3b^2 = 1$, and $b = 0$, $a = \pm 1$. In both cases we get that one of the factors of 2 is unit, namely ± 1.

- We now have to see that 2 is not a prime. Clearly

$$2 \mid (1+i\sqrt{3})(1-i\sqrt{3}) = 4.$$

  But 2 divides neither $1+i\sqrt{3}$ nor $1-i\sqrt{3}$.

We can see from the above example that the problem which arises is the lack of unique factorization.

Definition 3.23. A unique factorization domain (UFD) is an integral domain R satisfying that

1. every element $0 \neq a \in R$ can be written as a product of irreducible factors $p_1, \dots p_n$ up to a unit u, namely:

$$a = up_1 \dots p_n.$$

2. The above factorization is unique, that is, if

$$a = up_1 \dots p_n = vq_1 \dots q_m$$

   are two factorizations into irreducible factors p_i and q_j with units u, v, then $n = m$ and p_i and q_i are associate for all i.

We now prove that the distinction between irreducible and prime disappear in a unique factorization domain.

Proposition 3.14. *In a unique factorization domain R, we have that a is irreducible if and only if a is prime.*

Proof. We already know that prime implies irreducible. Let us show that now, we also have irreducible implies prime.

Take a to be irreducible and assume that $a \mid bc$. This means that $bc = ad$ for some $d \in R$. Using the property of unique factorization, we decompose d, b and c into products of irreducible terms (resp. d_i, b_i, c_i up to units u, v, w):

$$a \cdot ud_1 \cdots d_r = vb_1 \cdots b_s \cdot wc_1 \dots c_t.$$

Since the factorization is unique, a must be associate to some either b_i or c_i, implying that a divides b or c, which concludes the proof. □

We now want to connect the property of unique factorization to ideals.

Definition 3.24. Let $a_1, a_2, \dots$ be elements of an integral domain R. If the sequence of principal ideals

$$(a_1) \subseteq (a_2) \subseteq (a_3) \subseteq \dots$$

stabilizes, i.e., we have

$$(a_n) = (a_{n+1}) = \dots$$

for some n, then we say that R satisfies the ascending chain condition on principal ideals.

If the same condition holds but for general ideals, not necessarily principal, we call R a Noetherian ring, in honor of the mathematician Emmy Noether.

Figure 3.1: Amalie Emmy Noether (1882-1935)

Examples 3.12. 1. Consider the polynomial ring in infinitely many indeterminates $X_1, X_2, \ldots$ over $\mathbb{R}$. The chain

$$(X_1) \subset (X_1, X_2) \subset (X_1, X_2, X_3) \subset \ldots$$

of non-principal ideals is ascending and does not terminate. The ideal generated by all indeterminates is maximal.

2. Consider the polynomial ring $\mathbb{Z} + X\mathbb{Q}[X]$ of all rational polynomials with integral constant term. The chain

$$(X) \subset (X/2) \subset (X/4) \subset \ldots$$

of principal ideals is ascending and does not terminate.

Theorem 3.15. *Let R be an integral domain.*

1. *If R is a UFD, then R satisfies the ascending chain condition on principal ideals.*
2. *If R satisfies the ascending chain condition on principal ideals, then every non-zero element of R can be factored into irreducible (this says nothing about the unicity of the factorization).*
3. *If R is such that every non-zero element of R can be factored into irreducible, and in addition every irreducible element is prime, then R is a UFD.*

Thus R is a UFD if and only if it satisfies the ascending chain condition on principal ideals and every irreducible element of R is prime.

Proof. 1. Recall that in a UFD, prime and irreducible are equivalent. Consider an ascending chain of principal ideals

$$(a_1) \subseteq (a_2) \subseteq (a_3) \subseteq \dots$$

We have that $a_{i+1} \mid a_i$ for all i. Thus the prime factors of a_{i+1} consist of some (possibly all) prime factors of a_i. Since a_1 has a unique factorization into finitely many prime factors, the prime factors will end up being the same, and the chain will stabilize.

2. Take $0 \neq a_1 \in R$. If a_1 is irreducible, we are done. Let us thus assume that a_1 is not irreducible, that is

$$a_1 = a_2 b_2$$

where a_2 and b_2 are not unit. Since $a_2 \mid a_1$, we have $(a_1) \subseteq (a_2)$, and actually

$$(a_1) \subsetneq (a_2).$$

Indeed, if $(a_1) = (a_2)$, then a_2 would be a multiple of a_1, namely $a_2 = ca_1$ and thus

$$a_1 = a_2 b_2 \Rightarrow a_1 = c a_1 b_2 \Rightarrow a_1(1 - cb_2) = 0$$

implying that $cb_2 = 1$ and thus b_2 is a unit. This contradicts the assumption that a_1 is not irreducible. This computation has shown us that whenever we get a factor which is not irreducible, we can add a new principal ideal to the chain of ideals. Thus, if $a_2 b_2$ is a product of irreducible, we are done. Otherwise, we have that say a_2 is not irreducible, and $a_2 = a_3 b_3$, yielding

$$(a_1) \subsetneq (a_2) \subsetneq (a_3).$$

Since R satisfies the ascending chain condition on principal ideals, this process cannot go on and must stop, showing that we have a factorization into irreducible.

3. We now know that R allows a factorization into irreducible. We want to prove that this factorization is unique, under the assumption that every irreducible is prime. Suppose thus that

$$a = u p_1 p_2 \cdots p_n = v q_1 q_2 \cdots q_m$$

where u, v are units and p_i, q_j are irreducible. p_1 is an irreducible but also a prime by assumption, thus it must divide one of the q_j, say q_1, and we have $q_1 = p_1 d$. Since q_1 is irreducible, d must be a unit, and q_1 and p_1 are associate. We can iterated the process to find that q_i and p_i are associate for all i.

□

We now introduce a notion stronger than being a unique factorization domain.

Definition 3.25. A principal ideal domain (PID) is an integral domain in which every ideal is principal.

Theorem 3.16. *A principal ideal domain R is a unique factorization domain.*

Proof. What we will prove is that if R is a principal ideal domain, then

- R satisfies the ascending chain condition on principal ideals.
- every irreducible in R is also prime.

Having proved these two claims, we can conclude using the above theorem.

Let us first prove that R satisfies the ascending chain condition on principle ideals. Consider the following sequence of principal ideals

$$(a_1) \subseteq (a_2) \subseteq (a_3) \dots$$

and let $\mathcal{I} = \cup_{i=1}^{\infty}(a_i)$. Note that $\mathcal{I}$ is an ideal of R (be careful, a union of ideals is not an ideal in general!). Indeed, we have that $\mathcal{I}$ is closed under addition: take $a, b \in \mathcal{I}$, then there are ideals (a_j) and (a_k) in the chain with $a \in (a_j)$ and $b \in (a_k)$. If $m \geq \max(j, k)$, then both $a, b \in (a_m)$ and so do $a+b$. To check that $\mathcal{I}$ is closed under multiplication by an element of R, take again $a \in \mathcal{I}$. Then $a \in (a_j)$ for some j. If $r \in R$, then $ra \in (a_j)$ implying that $ra \in \mathcal{I}$.

Now by assumption, $\mathcal{I}$ is a principal ideal, generated by, say b: $\mathcal{I} = (b)$. Since b belongs to $\cup_{i=1}^{\infty}(a_i)$, it must belongs to some (a_n). Thus $\mathcal{I} = (b) \subseteq (a_n)$. For $j \geq n$, we have

$$(a_j) \subseteq \mathcal{I} \subseteq (a_n) \subseteq (a_j)$$

which proves that the chain of ideal stabilizes.

We are left to prove that every irreducible element is also prime. Let thus a be an irreducible element. Consider the principal ideal (a) generated by a. Note that (a) is a proper ideal: if $(a) = R$, then $1 \in (a)$ and thus a is a unit, which is a contradiction.

We have that (a) is included in a maximal ideal $\mathcal{I}$ (this can be deduced from either the ascending chain condition or from the theorem (Krull's theorem) that proves that every ideal is contained in a maximal ideal). Since R is a principal ideal domain, we have that $\mathcal{I} = (b)$. Thus

$$(a) \subseteq (b) \Rightarrow b \mid a \Rightarrow a = bd$$

where a is irreducible, b cannot be a unit (since $\mathcal{I}$ is by definition of maximal ideal a proper ideal), and thus d has to be a unit of R. In other words, a and b are associate. Thus

$$(a) = \mathcal{I} = (b).$$

Since $\mathcal{I}$ is a maximal ideal, it is prime implying that a is prime, which concludes the proof. □

Determining whether a ring is a principal ideal domain is in general quite a tough question. It is still an open conjecture (called Gauss's conjecture) to decide whether there are infinitely many real quadratic fields which are principal (we use the terminology "principal" for quadratic fields by abuse of notation, it actually refers to their ring of integers, that is rings of the form either $\mathbb{Z}[\sqrt{d}]$ if $d \equiv 2$ or $3 \mod 4$ or $\mathbb{Z}[\frac{1+\sqrt{d}}{2}]$ else).

One way mathematicians have found to approach this question is to actually prove a stronger property, namely whether a ring R is Euclidean.

Definition 3.26. Let R be an integral domain. We say that R is a Euclidean domain if there is a function Ψ from $R\backslash\{0\}$ to the non-negative integers such that

$$a = bq + r,\ a, b \in R,\ b \neq 0,\ q, r \in R$$

where either $r = 0$ or $\Psi(r) < \Psi(b)$.

When the division is performed with natural numbers, it is clear what it means that $r < b$. When we work with polynomials instead, we can say that $\deg r < \deg b$. The function Ψ generalizes these notions.

Theorem 3.17. *If R is a Euclidean domain, then R is a principal ideal domain.*

Proof. Let $\mathcal{I}$ be an ideal of R. If $\mathcal{I} = \{0\}$, it is principal and we are done. Let us thus take $\mathcal{I} \neq \{0\}$. Consider the set

$$\{\Psi(b),\ b \in \mathcal{I},\ b \neq 0\}.$$

It is included in the non-negative integers by definition of Ψ, thus it contains a smallest element, say n. Let $0 \neq b \in \mathcal{I}$ such that $\Psi(b) = n$.

We will now prove that $\mathcal{I} = (b)$. Indeed, take $a \in \mathcal{I}$, and compute

$$a = bq + r$$

where $r = 0$ or $\Psi(r) < \Psi(b)$. This yields

$$r = a - bq \in \mathcal{I}$$

and $\Psi(r) < \Psi(b)$ cannot possibly happen by minimality of n, forcing r to be zero. This concludes the proof. □

Example 3.13. Consider the ring

$$\mathbb{Z}[\sqrt{d}] = \{a + b\sqrt{d},\ a, b \in \mathbb{Z}\}$$

with

$$\Psi(a + b\sqrt{d}) = |a^2 - b^2 d|.$$

We will show that we have a Euclidean domain for $d = -2, -1, 2$.

Note that $\mathbb{Z}[\sqrt{d}]$ is an integral domain. Take $\alpha, \beta \neq 0$ in $\mathbb{Z}[\sqrt{d}]$. Now we would like to perform the division of α by β to get something of the form

$$\alpha = \beta q + r,\ q, r \in \mathbb{Z}[\sqrt{d}].$$

Since $\mathbb{Z}[\sqrt{d}]$ is not a field, there is no reason for this division to give a result in $\mathbb{Z}[\sqrt{d}]$ (that is, $q, r \in \mathbb{Z}[\sqrt{d}]$), however, we can compute the division in $\mathbb{Q}(\sqrt{d})$:

$$\alpha/\beta = q',$$

with $q' = x + \sqrt{d}y$ with x, y rational. Let us now approximate x, y by integers x_0, y_0, namely take x_0, y_0 such that

$$|x - x_0| \leq 1/2, \ |y - y_0| \leq 1/2.$$

Take

$$q = x_0 + y_0\sqrt{d}, \ r = \beta((x - x_0) + (y - y_0)\sqrt{d}),$$

where clearly $q \in \mathbb{Z}[\sqrt{d}]$, then

$$\begin{aligned} \beta q + r &= \beta(x_0 + y_0\sqrt{d}) + \beta((x - x_0) + (y - y_0)\sqrt{d}) \\ &= \beta(x + y\sqrt{d}) = \beta q' = \alpha, \end{aligned}$$

which at the same time shows that $r \in \mathbb{Z}[\sqrt{d}]$. We are left to show that $\Psi(r) < \Psi(\beta)$. We have

$$\begin{aligned} \Psi(r) &= \Psi(\beta)\Psi((x - x_0) + (y - y_0)\sqrt{d}) \\ &= \Psi(\beta)|(x - x_0)^2 - d(y - y_0)^2| \\ &\leq \Psi(\beta)[|x - x_0|^2 + |d||y - y_0|^2] \\ &\leq \Psi(\beta)\left(\frac{1}{4} + |d|\frac{1}{4}\right) \end{aligned}$$

showing that $\mathbb{Z}[\sqrt{d}]$ is indeed a Euclidean domain for $d = -2, -1, 2$.

Below is a summary of the ring hierarchy (recall that PID and UFD stand respectively for principal ideal domain and unique factorization domain):

integral domains $\supset$ UFD $\supset$ PID $\supset$ Euclidean domains

Note that though the Euclidean division may sound like an elementary concept, as soon as the ring we consider is fancier than $\mathbb{Z}$, it becomes quickly a difficult problem. We can see that from the fact that being Euclidean is stronger than being a principal ideal domain. All the inclusions are strict, since one may check that $\mathbb{Z}[\sqrt{-3}]$ is an integral domain but is not a UFD, $\mathbb{Z}[X]$ is a UFD which is not PID, while $\mathbb{Z}[(1 + i\sqrt{19})/2]$ is a PID which is not a Euclidean domain.

ring	ED	PID	UFD	ID
$\mathbb{Z}$	yes	yes	yes	yes
$F[X]$, F a field	yes	yes	yes	yes
$\mathbb{Z}[i]$	yes	yes	yes	yes
$\mathbb{Z}[\sqrt{\pm 2}]$	yes	yes	yes	yes
$\mathbb{Z}[\sqrt{3}]$	yes	yes	yes	yes
$\mathbb{Z}[(1+i\sqrt{19})/2]$	no	yes	yes	yes
$\mathbb{Z}[X]$	no	no	yes	yes
$\mathbb{Z}[\sqrt{-3}]$	no	no	no	yes

Table 3.1: Examples of rings we saw: that $\mathbb{Z}[\sqrt{3}]$ is a Euclidean domain is done in the exercises, that $\mathbb{Z}[X]$ is not a principal ideal domain is also shown in the exercises, it is enough to show that the ideal $\langle 2, X\rangle$ is not principal. Finally $\mathbb{Z}[\sqrt{-3}]$ is not a unique factorization domain because we saw that 2 is irreducible but not prime.

3.7 Irreducible polynomials

Recall the definition of irreducible that we have seen: a non-zero element a which is not a unit is said to be irreducible if $a = bc$ implies that either b or c is a unit. Let us focus on the case where the ring is a ring of polynomials $R[X]$ and R is an integral domain.

Definition 3.27. If R is an integral domain, then an irreducible element of $R[X]$ is called an irreducible polynomial.

In the case of a field F, then units of $F[X]$ are non-zero elements of F. Then we get the more familiar definition that an irreducible element of $F[X]$ is a polynomial of degree at least 1, that cannot be factored into two polynomials of lower degree.

Let us now consider the more general case where R is an integral domain (thus not necessarily a field, it may not even be a unique factorization domain). To study when polynomials over an integral domain R are irreducible, it is often more convenient to place oneselves in a suitable field that contains R, since division in R can be problematic. To do so, we will now introduce the field of fractions, also called quotient field, of R. Since there is not much more difficulty in treating the general case, that is, when R is a commutative ring, we present this construction.

Let S be a subset of R which is closed under multiplication, contains 1 and does not contain 0. This definition includes the set of all non-zero elements of an integral domain, or the set of all non-zero elements of a commutative ring that are not zero divisors. We define the following equivalence relation on $R \times S$:

$$(a, b) \sim (c, d) \iff s(ad - bc) = 0 \text{ for some } s \in S.$$

It is clearly reflexive and symmetric. Let us check the transitivity. Suppose that

$(a, b) \sim (c, d)$ and $(c, d) \sim (e, f)$. Then

$$s(ad - bc) = 0 \text{ and } t(cf - de) = 0$$

for some $s, t \in S$. We can now multiply the first equation by tf, the second by sb and add them

$$stf(ad - bc) + tsb(cf - de) = 0$$

to get

$$sdt(fa - be) = 0$$

which proves the transitivity.

What we are trying to do here is to mimic the way we deal with $\mathbb{Z}$. If we take non-zero $a, b, c, d \in \mathbb{Z}$, we can write down $a/b = c/d$, or equivalently $ad = bc$, which is also what $(a, b) \sim (c, d)$ satisfies by definition if we take R to be an integral domain. In a sense, (a, b) is some approximation of a/b.

Formally, if $a \in R$ and $b \in S$, we define the fraction a/b to be the equivalence class of the pair (a, b). The set of all equivalence classes is denoted by $S^{-1}R$. To make it into a ring, we define the following laws in a natural way:

- addition:
$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}.$$
- multiplication:
$$\frac{a}{b}\frac{c}{d} = \frac{ac}{bd}.$$
- additive identity:
$$\frac{0}{1} = \frac{0}{s},\ s \in S.$$
- additive inverse:
$$-\frac{a}{b} = \frac{-a}{b}.$$
- multiplicative identity:
$$\frac{1}{1} = \frac{s}{s},\ s \in S.$$

To prove that we really obtain a ring, we need to check that all these laws are well-defined.

Theorem 3.18. *With the above definitions, the set of equivalence classes $S^{-1}R$ is a commutative ring.*

1. *If R is an integral domain, so is $S^{-1}R$.*
2. *If R is an integral domain, and $S = R\backslash\{0\}$, then $S^{-1}R$ is a field.*

Proof. **Addition is well-defined.** If $a_1/b_1 = c_1/d_1$ and $a_2/b_2 = c_2/d_2$, then for some $s, t \in S$, we have

$$s(a_1d_1 - b_1c_1) = 0 \text{ and } t(a_2d_2 - b_2c_2) = 0.$$

We can now multiply the first equation by tb_2d_2 and the second by sb_1d_1 to get

$$tb_2d_2s(a_1d_1 - b_1c_1) = 0 \text{ and } sb_1d_1t(a_2d_2 - b_2c_2) = 0,$$

and adding them yields

$$st[d_2d_1(b_2a_1 + b_1a_2) - b_2b_1(d_2c_1 + d_1c_2)] = 0$$

that is

$$\frac{b_2a_1 + b_1a_2}{b_2b_1} = \frac{d_2c_1 + d_1c_2}{d_2d_1},$$

which can be rewritten as

$$\frac{a_1}{b_1} + \frac{a_2}{b_2} = \frac{c_1}{d_1} + \frac{c_2}{d_2}$$

and we conclude that addition does not depend on the choice of a representative in an equivalence class.

Multiplication is well-defined. We start as before. If $a_1/b_1 = c_1/d_1$ and $a_2/b_2 = c_2/d_2$, then for some $s, t \in S$, we have

$$s(a_1d_1 - b_1c_1) = 0 \text{ and } t(a_2d_2 - b_2c_2) = 0.$$

Now we multiply instead the first equation by ta_2d_2, the second by sc_1b_1 and we add them:

$$st[a_2d_2a_1d_1 - c_1b_1b_2c_2] = 0.$$

This implies, as desired, that

$$\frac{a_1a_2}{b_1b_2} = \frac{c_1c_2}{d_1d_2}.$$

To be complete, one should check that the properties of a ring are fulfilled, but this follows from the fact that addition and multiplication are carried the usual way.

1. We want to prove that $S^{-1}R$ is an integral domain. We assume that R is an integral domain, and we need to check the definition of an integral domain for $S^{-1}R$. Namely, suppose that $(a/b)(c/d) = 0$ in $S^{-1}R$, that is

$$\frac{a}{b}\frac{c}{d} = \frac{0}{1}.$$

This means that $(ac, bd) \sim (0, 1)$ and $acs = 0$ for some $s \in S$. Now $acs = 0$ is an equation in R, which is an integral domain, and $s \neq 0$, thus $ac = 0$, so either a or c is 0, and consequently either a/b or c/d is zero.

2. To conclude, we want to prove that $S^{-1}R$ is a field, assuming that R is an integral domain, and $S = R\backslash\{0\}$. We consider a/b a non-zero element of $S^{-1}R$, for which we need to find an inverse. Note that a and b are non-zero, thus they are both in S meaning that both a/b and b/a are in $S^{-1}R$ and b/a is the multiplicative inverse of a/b.

□

Definition 3.28. Let R be a commutative ring. Based on the above, the set of equivalence classes $S^{-1}R$ is a commutative ring, called the ring of fractions of R by S. If R is an integral domain, and $S = R\backslash\{0\}$, then $S^{-1}R$ is called the field of fractions or quotient field of R.

Now that we have defined a suitable field, we are left to prove that we can embed an integral domain R in its quotient field.

Proposition 3.19. *A commutative ring R can be embedded in its ring of fractions $S^{-1}R$, where S is the set of all its non-divisors of zero. In particular, an integral domain can be embedded in its quotient field, which is furthermore the smallest field containing R.*

Proof. Consider the following map:

$$f : R \to S^{-1}R,\ a \mapsto f(a) = a/1.$$

It is not hard to check that f is a ring homomorphism. If S has no zero divisor, we have that the kernel of f is given by the set of a such that $f(a) = a/1 = 0/1$, that is the set of a such that $sa = 0$ for some s. Since s is not a zero divisor, we have $a = 0$ and f is a monomorphism. □

Let us get back to the irreducible polynomials, and consider now the case where D is a unique factorization domain. It is not necessarily a field, but we now know how to embed it in a suitable field, namely its field of fractions, or quotient field. Take the polynomial $f(X) = a + abX$, $a \neq 0$ not a unit. Since we can factor it as

$$f(X) = a(1 + bX)$$

where a is not a unit by assumption, this polynomial is not irreducible. But we do not really have a factorization into two polynomials of lower degree. What happens here is that the constant polynomials are not necessarily units, unlike in the case of fields. To distinguish this case, we introduce the notion of primitive polynomial.

Definition 3.29. Let D be a unique factorization domain and let $f \in D[X]$. We call the greatest common divisor of all the coefficients of f the content of f, denoted by $c(f)$. A polynomial whose content is a unit is called a primitive polynomial.

Figure 3.2: Carl Friedrich Gauss (1777-1855)

We can now rule out the above example, and we will prove later that this allows us to say that a primitive polynomial is irreducible if and only if it cannot be factored into two polynomials of lower degree. Be careful however that "primitive polynomial" has a different meaning if it is defined over a field.

The next goal is to prove Gauss lemma, which in particular implies that the product of two primitive polynomials is a primitive polynomial.

We start with a lemma.

Lemma 3.20. *Let D be a unique factorization domain, and consider $f \neq 0, g, h \in D[X]$ such that $pf(X) = g(X)h(X)$ with p a prime. Then either p divides all the coefficients of g or p divides all the coefficients of h.*

Before starting the proof, let us notice that this lemma is somehow a generalization of the notion of prime. Instead of saying that $p|ab$ implies $p|a$ or $p|b$, we have $p|g(X)h(X)$ implies that $p|g(X)$ or $p|h(X)$ (dividing the whole polynomial means dividing all of its coefficients).

Proof. Denote

$$g(X) = g_0 + g_1X + \ldots + g_sX^s,\ h(X) = h_0 + h_1X + \ldots + h_tX^t.$$

Suppose by contradiction that p does not divide all coefficients of g and does not divide all coefficients of h either. Then let g_u and h_v be the coefficients of minimum index not divisible by p. Then the coefficient of X^{u+v} in $g(X)h(X)$ is

$$g_0h_{u+v} + g_1h_{u+v-1} + \ldots + g_uh_v + \ldots + g_{u+v-1}h_1 + g_{u+v}h_0.$$

By definition of u and v, p divides every term but g_uh_v, thus p cannot possibly divide the entire expression, and thus there exists a coefficient of $g(X)h(X)$ not divisible by p. This contradicts the fact that $p|g(X)h(X)$. □

Proposition 3.21. (Gauss Lemma). *Let f, g be non-constant polynomials in $D[X]$ where D is a unique factorization domain. The content of a product of polynomials is the product of the contents, namely*

$$c(fg) = c(f)c(g),$$

up to associates. In particular, the product of two primitive polynomials is primitive.

Proof. Let us start by noticing that by definition of content, we can rewrite

$$f(X) = c(f)f^*(X),\ g(X) = c(g)g^*(X),$$

where $f^*, g^* \in D[X]$ are primitive. Clearly

$$fg = c(f)c(g)f^*g^*.$$

Since $c(f)c(g)$ divides fg, it divides every coefficient of fg and thus their greatest common divisor:

$$c(f)c(g) \mid c(gf).$$

We now prove the converse, namely that $c(gf)| \ |c(g)c(g)$. To do that, we consider each prime p appearing in the factorization of $c(gf)$ and argue that $p \mid c(f)c(g)$. Let thus p be a prime factor of $c(gf)$. Since $fg = c(fg)(fg)^*$, we have that $c(fg)$ divides fg, that is

$$p \mid fg.$$

By the above lemma, either $p \mid f$ or $p \mid g$, say $p \mid f = c(f)f^*$, meaning that either $p \mid c(f)$ or $p \mid f^*$. Since f^* is primitive, p cannot possibly divide f^*, and thus

$$p \mid c(f) \Rightarrow p \mid c(f)c(g).$$

If p appears with multiplicity, we iterate the reasoning with the same p. □

We are now ready to connect irreducibility over a unique factorization domain and irreducibility over the corresponding quotient field or field of fractions.

Proposition 3.22. *Let D be a unique factorization domain with quotient field F. If f is a non-constant polynomial in $D[X]$, then f is irreducible over D if and only if f is primitive and f is irreducible over F.*

For example, this says that f is irreducible over $\mathbb{Z}$ if and only if f is primitive, and f is irreducible over $\mathbb{Q}$.

Proof. First assume that f is irreducible over D.

f **is primitive.** Indeed, if f were not primitive, then we could write

$$f = c(f)f^*,$$

where $c(f)$ denotes the content of f and f^* is primitive. Since we assume f is not primitive, its content cannot be a unit, which contradicts the irreducibility of f over D, and we conclude that f is primitive.

f is irreducible over F. Again assume by contradiction that f is not irreducible over F. Now F is a field, thus reducible means f can be factored into a product of two non-constant polynomials in $F[X]$ of smaller degree:

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

Since g, h are in $F[X]$, and F is the field of fractions of D, we can write

$$g(X) = \frac{a}{b}g^*(X), \ h(X) = \frac{c}{d}h^*(X), \ a, b, c, d \in D$$

and g^*, h^* primitive. Thus

$$f(X) = \frac{ac}{bd}g^*(X)h^*(X)$$

where g^*h^* is a primitive polynomial by Gauss Lemma. Since we have already proven (in the 1st part) that f is primitive, it must be that $ac/bd = u$ is a unit. But this would mean that

$$f(X) = ug^*(X)h^*(X)$$

which contradicts the fact that $f(X)$ is irreducible over $D[X]$ and we conclude that f is also irreducible over $F[X]$.

We are left to prove the converse. Let then f be a primitive and f be an irreducible polynomial over F. We do it by contraction, and assume that the primitive polynomial f is not irreducible over D:

$$f(X) = g(X)h(X).$$

Since f is primitive, $\deg g$ and $\deg h$ are at least 1. But then neither g not h can be a unit in $F[X]$ (these are units in F) and thus

$$f = gh$$

contradicts the irreducibility of f over F. □

In other words, we have proven that f irreducible over D is equivalent to f primitive and cannot be factored into two polynomials of lower degree in $F[X]$.

To conclude, we present a practical criterion to decide whether a polynomial in $D[X]$ is irreducible over F.

Proposition 3.23. (Eisenstein's criterion). *Let D be a unique factorization domain, with quotient field F and let*

$$f(X) = a_nX^n + \ldots + a_1X + a_0$$

be a polynomial in $D[X]$ with $n \geq 1$ and $a_n \neq 0$.

If p is a prime in D and p divides a_i, $0 \leq i < n$ but p does not divide a_n nor does p^2 divide a_0, then f is irreducible over F.

Figure 3.3: Ferdinand Eisenstein (1823-1852)

Proof. We first divide f by its content, to get a primitive polynomial. By the above proposition, it is enough to prove that this primitive polynomial is irreducible over D.

Let thus f be a primitive polynomial and assume by contradiction it is reducible, that is

$$f(X) = g(X)h(X)$$

with

$$g(X) = g_0 + \ldots + g_r X^r,\ h(X) = h_0 + \ldots + h_s X^s.$$

Notice that r cannot be zero, for if $r = 0$, then $g_0 = g$ would divide f and thus all a_i implying that g_0 divides the content of f and is thus a unit. But this would contradict the fact that f is reducible. We may from now on assume that

$$r \geq 1,\ s \geq 1.$$

Now by hypothesis, $p \mid a_0 = g_0 h_0$ but p^2 does not divide a_0, meaning that p cannot divide both g_0 and h_0. Let us say that

$$p \mid g_0$$

and p does not divide h_0 (and vice-versa).

By looking at the dominant coefficient $a_n = g_r h_s$, we deduce from the assumption that p does not divide a_n that p cannot possibly divide g_r. Let i be the smallest integer such that p does not divide g_i. Then

$$1 \leq i \leq r < n = r + s.$$

Let us look at the ith coefficient

$$a_i = g_0 h_i + g_1 h_{i-1} + \ldots + g_i h_0$$

and by choice of i, p must divide $g_0, \ldots, g_{i-1}$. Since p divides a_i by assumption, it thus must divide the last term $g_i h_0$, and either $p \mid g_i$ or $p \mid h_0$ by definition of prime. Both are impossible: we have chosen p dividing neither h_0 nor g_i. This concludes the proof. □

The main definitions and results of this chapter are

- **(2.1-2.2).** Definitions of: ring, zero divisor, unit, integral domain, division ring, subring, characteristic, ring homomorphism, ideal, quotient ring. Factor and 1st Isomorphism Theorem for rings.
- **(2.3-2.4).** Operations on ideals, Chinese Remainder Theorem, Correspondence Theorem for rings. Definitions of: principal ideal, maximal ideal, prime ideal, the characterization of the two latter in the commutative case.
- **(2.5).** Polynomial Euclidean division, number of roots of a polynomial.
- **(2.6).** Definitions of: associate, prime, irreducible, unique factorization domain, ascending chain condition, principal ideal domain, Euclidean domain. Connections between prime and irreducible. Hierarchy among UFD, PID and Euclidean domains.
- **(2.7).** Construction of ring of fractions. Definitions of: content of a polynomial, primitive polynomial. Gauss Lemma, Eisenstein's criterion.

Chapter 4

Exercises on Ring Theory

Exercises marked by (*) are considered difficult.

4.1 Rings, ideals and homomorphisms

Exercise 56. Let R be a ring and $x \in R$. Suppose there exists a positive integer n such that $x^n = 0$. Show that $1 + x$ is a unit, and so is $1 - x$.

Answer. The element $1 - x$ is a unit since

$$(1 - x)(1 + x + \ldots + x^{n-1}) = 1.$$

The element $1 + x$ is a unit since

$$(1 + x)(1 - x + x^2 - x^3 \ldots \pm x^{n-1}) = 1.$$

Exercise 57. Let R be a commutative ring, and I be an ideal of R. Show that

$$\sqrt{I} := \{x \in R \mid \text{ there exists } m \in \mathbb{N}^* \text{ such that } x^m \in I\}$$

is an ideal of R. **Answer.**

- Clearly, $0 \in \sqrt{I}$. If $a \in \sqrt{I}$, then $a^m \in I$ for some $m \geq 1$. Then $(-a)^m = (-1)^m a^m \in I$, so $-a \in \sqrt{I}$. Now let $a, b \in \sqrt{I}$, so $a^n \in I$ for some $n \geq 1$ and $b^m \in I$ for some $m \geq 1$. Now let us show that $(a + b)^{n+m} \in I$. We have $(a + b)^{n+m} = \sum_{j=0}^{n+m} \frac{n!}{j!(n+m-j)!} a^j b^{n+m-j}$ (because R is commutative). Now if $0 \leq j \leq n$, we have $n + m - j \geq m$, so $b^{n+m-j} \in I$ in this case (since $b^m \in I \Rightarrow b^i \in I$ for $i \geq m$). If $n + 1 \leq j \leq n + m$, we have $j \geq n + 1$, so $a^j \in I$ in this case (since $a^n \in I \Rightarrow a^i \in I$ for $i \geq n$). Therefore all the terms in the previous sum are in I and thus $(a+b)^{n+m} \in I$. Hence $a+b \in \sqrt{I}$. We just proved that $\sqrt{I}$ is an additive subgroup of R.

- Now we have to check the second property. Let $a \in \sqrt{I}$, and $r \in R$. We have $a^n \in I$ for some $n \geq 1$. Now $(ar)^n = a^n r^n$ because R is commutative, so $(ar)^n \in I$ and therefore $ar \in \sqrt{I}$. Therefore $\sqrt{I}$ is an ideal of R.

Exercise 58. Determine all rings of cardinality p and characteristic p.

Answer. Let R be a ring of characteristic p. Consider the ring homomorphism: $\varphi : \mathbb{Z} \to R$, the characteristic of R is the natural number p such that $p\mathbb{Z}$ is the kernel of φ. We can now factorize φ in an injective map $\mathbb{Z}/p\mathbb{Z} \to R$. If now we further assume that R has cardinality p, we have that $\mathbb{Z}/p\mathbb{Z}$ and R have same cardinality, and thus we have an isomorphism. This means that the only ring of cardinality and characteristic p is $\mathbb{Z}/p\mathbb{Z}$.

Exercise 59. Let R be a commutative ring. Let

$$Nil(R) = \{r \in R | \exists n \geq 1, r^n = 0\}.$$

1. Prove that $Nil(R)$ is an ideal of R.

2. Show that if $r \in Nil(R)$, then $1 - r$ is invertible in R.

3. Show, with a counter-example, that $Nil(R)$ is not necessarily an ideal anymore if R is not commutative.

1. - Clearly, $0 \in Nil(R)$. If $a \in Nil(R)$, then $a^m = 0$ for some $m \geq 1$. Then $(-a)^m = (-1)^m a^m = 0$, so $-a \in Nil(R)$. Now let $a, b \in Nil(R)$, so $a^n = 0$ for some $n \geq 1$ and $b^m = 0$ for some $m \geq 1$. Now let us show that $(a+b)^{n+m} = 0$. We have $(a+b)^{n+m} = \sum_{j=0}^{n+m} \frac{n!}{j!(n+m-j)!} a^j b^{n+m-j}$ (because R is commutative). Now if $0 \leq j \leq n$, we have $n+m-j \geq m$, so $b^{n+m-j} = 0$ in this case (since $b^m = 0 \Rightarrow b^i = 0$ for $i \geq m$). If $n+1 \leq j \leq n+m$, we have $j \geq n+1$, so $a^j = 0$ in this case (since $a^n = 0 \Rightarrow a^i = 0$ for $i \geq n$). Therefore all the terms in the previous sum are 0 and thus $(a+b)^{n+m} = 0$. Hence $a + b \in Nil(R)$. We just proved that $Nil(R)$ is an additive subgroup of R.
   - Now we have to check the second property. Let $a \in Nil(R)$, and $r \in R$. We have $a^n = 0$ for some $n \geq 1$. Now $(ar)^n = a^n r^n$ because R is commutative, so $(ar)^n = 0$ and therefore $ar \in Nil(R)$. Therefore $Nil(R)$ is an ideal of R.

2. If $r \in Nil(R)$, then $r^m = 0$ for some $m \geq 1$. Then $1 + r + r^2 + \cdots + r^{m-1}$ is the inverse of $1 - r$ since

$$(1-r)(1+r+r^2+\cdots+r^{m-1}) = 1+r+r^2+\cdots+r^{m-1}-r-r^2+\cdots+r^m = 1-r^m = 1.$$

3. If $R = M_2(\mathbb{C})$, let $a = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $b = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. Then $a^2 = b^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, so $a, b \in Nil(R)$, but $a + b$ does not lie in $Nil(R)$, since $(a + b)^2 = I_2$, and $I_2^n = I_2$ for all $n \geq 1$.

Exercise 60. Determine whether the following maps are ring homomorphisms:

1. $f_1 : \mathbb{Z} \longrightarrow \mathbb{Z}$ with $f_1(x) = x + 1$.
2. $f_2 : \mathbb{Z} \longrightarrow \mathbb{Z}$ with $f_2(x) = x^2$.
3. $f_3 : \mathbb{Z}/15\mathbb{Z} \longrightarrow \mathbb{Z}/15\mathbb{Z}$ with $f_3(x) = 4x$.
4. $f_4 : \mathbb{Z}/15\mathbb{Z} \longrightarrow \mathbb{Z}/15\mathbb{Z}$ with $f_4(x) = 6x$.

Answer.

1. Since $f_1(0) = 1$, f_1, f cannot be a ring homomorphism.
2. Since $f_2(x + y) = x^2 + y^2 + 2xy \neq x^2 + y^2 = f_2(x) + f_2(y)$, f_2 cannot be a ring homomorphism.
3. Since $f_3(xy) = 4xy \neq xy = f_3(x)f_3(y)$, f_3 cannot be a ring homomorphism.
4. Since $f_4(1) \neq 1$, f_4 cannot be a ring homomorphism!

Exercise 61. Let K be a division ring with center k.

1. Show that the center of the polynomial ring $K[X]$ is $k[X]$.
2. For any a in $K \backslash k$, show that the ideal generated by $X - a$ in $K[X]$ is in fact the whole ring $K[X]$.
3. Show that any ideal $I \subseteq K[X]$ has the form $K[X]h$ where $h \in k[X]$.

Answer.

1. Clearly $k[X]$ is in the center. Conversely, if $f = \sum a_i X^i$ is in the center, then $fa = af$ for all $a \in K$, showing that $a_i \in k$.
2. Fix $b \in K$ such that $ab \neq ba$. Then the ideal generated by $X - a$ contains
$$b(X - a) - (X - a)b = ab - ba \in K$$
since $ab \neq ba$ so $(X - a) = R$.
3. We may assume $I \neq 0$ and fix a monic polynomial of the least degree in I. By the usual Euclidean algorithm argument, we have that $I = K[X]h$. For any $a \in K$, we have $ha \in I = K[X]h$ so $ha = rh$ for some r in $K[X]$. By comparing the leading terms, we see that $r \in K$ and in fact $r = a$. Thus $ha = ah$ for any $a \in K$, which means that $h \in k[X]$.

Exercise 62. Consider the ring $\mathcal{M}_n(\mathbb{R})$ of real $n \times n$ matrices. Are the trace and the determinant ring homomorphisms?

Answer. The trace is not multiplicative, since

$$2 = \mathrm{Tr}\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) \neq \mathrm{Tr}\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) \cdot \mathrm{Tr}\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) = 4.$$

The determinant is not additive:

$$4 = \det\left(\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}\right) \neq \det\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) + \det\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) = 2.$$

Thus none of them are ring homomorphisms.

4.2 Quotient rings

Exercise 63. Compute the characteristic of the following rings R:

1. $R = \mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$,
2. $R = \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/10\mathbb{Z}$,
3. $R = \mathbb{Z}[j]/(2 - 5j)$, where j denotes a primitive 3rd root of unity ($j^3 = 1$ but $j^2 \neq 1$).

Answer. In this exercise, we use the notation $\overline{x}$ to denote an element in the quotient group involved.

1. For $1 \leq m \leq n-1$, we have $m \cdot \overline{1} = \overline{m} \neq 0$, since m is not a multiple of n. But $n \cdot \overline{1} = \overline{n} = \overline{0}$. So $\mathrm{char}(R) = n$ by definition of the characteristic.

2. If $m \in \mathbb{Z}$, we will denote by respectively by $\overline{m}, [m], \tilde{m}$ its class modulo $2, 4$ and 10. Assume that $m(\overline{1}, [1], \tilde{1}) = (\overline{0}, [0], \tilde{0})$. Then we have

$$(\overline{m}, [m], \tilde{m}) = (\overline{0}, [0], \tilde{0}),$$

   which implies that m is a multiple of $2, 4$ and 10. Hence m is a multiple of the lowest common multiple of $2, 4$ and 10, which is 20. Conversely, $20(\overline{1}, [1], \tilde{1}) = (\overline{20}, [20], \tilde{20}) = (\overline{0}, [0], \tilde{0})$. Therefore $\mathrm{char}(R) = 20$.

3. Here we have $(2-5j)(2-5j^2) = 4 - 10(j + j^2) + 25j^3 = 4 + 10 + 25 = 39$. Hence $39 \cdot \overline{1} = \overline{39} = \overline{(2-5j) \cdot (2-5j^2)} = \overline{0}$. Then the characteristic of R is finite and divides 39. Therefore the characteristic of R is $1, 3, 13$ or 39. Now let $c = \mathrm{char}(R) > 0$. Since $c \cdot 1_R$ lies in the ideal $(2 - 5j)$, then $c = (2 - 5j)(a + bj)$ for some $a, b, \in \mathbb{Z}$. Hence $|c|^2 = |2 - 5j|^2 |a + bj|^2$, so

$$c^2 = 39(a^2 + b^2 - ab)$$

   and therefore $39 | c^2$. The only value (among $1, 3, 13$ and 39) for which it is possible is $c = 39$. Thus $\mathrm{char}(R) = 39$.

Exercise 64. Prove the following isomorphisms:

1. $\mathbb{Z}[i]/(1+i) \simeq \mathbb{Z}/2\mathbb{Z}$.
2. $\mathbb{Z}[X]/(n, X) \simeq \mathbb{Z}/n\mathbb{Z}$, $n \geq 2$.
3. $\mathbb{Z}[X]/(n) \simeq (\mathbb{Z}/n\mathbb{Z})[X]$, $n \geq 2$.

Answer.

1. Consider $\varphi : m \in \mathbb{Z} \mapsto m \cdot 1_R = \overline{m} \in \mathbb{Z}[i]/(1+i)$. This is a ring homomorphism. It is surjective. Indeed, let $\overline{a+bi} \in \mathbb{Z}[i]/(1+i)$. We have $\overline{a+bi} = \overline{(b-a)+a(1+i)} = \overline{b-a}$, so $\overline{a+bi} = \varphi(b-a)$. Now $\ker(\varphi) = c \cdot \mathbb{Z}$, where $c = \mathrm{char}(R)$ by definition of the characteristic. By direct computation, we get $\mathrm{char}(R) = 2$ (since R is not the trivial ring and $(1+i)(1-i) = 2$). Therefore $\ker(\varphi) = 2\,\mathbb{Z}$. Now use the first isomorphism theorem.

2. Let us consider $\varphi : P \in \mathbb{Z}[X] \mapsto \overline{P(0)} \in \mathbb{Z}\,/n\,\mathbb{Z}$. This is the composition of the ring homomorphisms $P \in \mathbb{Z}[X] \mapsto P(0) \in \mathbb{Z}$ and $m \in \mathbb{Z} \mapsto \overline{m} \in \mathbb{Z}\,/n\,\mathbb{Z}$, so it is a ring homomorphism. It is surjective: for $\overline{m} \in \mathbb{Z}\,/n\,\mathbb{Z}$, we have $\varphi(m) = \overline{m}$, where $m \in \mathbb{Z} \subset \mathbb{Z}[X]$ is considered as a constant polynomial. Now we have $\ker(\varphi) = \{P \in \mathbb{Z}[X] | P(0) \text{ is divisible by } n\}$, which equals (n, X). Hence $\ker(\varphi) = (n, X)$; now applying the first isomorphism theorem, we get the result.

3. Consider the reduction modulo n, $\varphi : P \in \mathbb{Z}[X] \mapsto \overline{P} \in (\mathbb{Z}\,/n\,\mathbb{Z})[X]$. We have that φ is a ring homomorphism. It is surjective: let $f \in (\mathbb{Z}\,/n\,\mathbb{Z})[X]$, $f = \overline{a}_0 + \cdots + \overline{a}_m X^m, a_i \in \mathbb{Z}$. Then let $P = a_0 + \cdots + a_m X^m \in \mathbb{Z}[X]$. By definition of $\overline{P}$, we have $\varphi(P) = f$. Now let us compute the kernel of φ. Let $P = a_0 + \cdots + a_m X^m$. We have $\varphi(P) = 0 \iff \overline{a}_0 + \cdots + \overline{a}_m X^m = 0$. This is equivalent to say that $\overline{a}_i = \overline{0}$ for all i, which means that $n | a_i$ for all i. This is equivalent to say that $P = n \cdot Q$, for some $Q \in \mathbb{Z}[X]$. Hence $\ker(\varphi) = (n)$. Now apply the first isomorphism theorem.

Exercise 65. Let $A = \mathbb{C}[X; \sigma]$ be the ring of all skew polynomials $\sum a_i X^i$, $a_i \in \mathbb{C}$, where multiplication is defined by $Xa = \sigma(a)X$ for all $a \in \mathbb{C}$, and σ is the complex conjugation on $\mathbb{C}$.

- Show that the center $Z(A)$ of A is $Z(A) = \mathbb{R}[X^2]$.
- Show that $\bar{A} = A/A(X^2+1)$ is a ring.
- Show that $\bar{A}$ is isomorphic to $\mathbb{H}$, the division ring of Hamilton quaternions.

Answer.

- Note that $X^2 a = X\sigma(a)X = \sigma^2(a)X^2$ and more generally

$$(\sum a_i X^i)(\sum b_j X^j) = \sum_i \sum_j a_i \sigma^i(b_j) X^{i+j}.$$

Now if $\sum b_j X^j$ is in the center, then we must have

$$\sum_i \sum_j a_i \sigma^i(b_j) X^{i+j} = (\sum b_j X^j)(\sum a_i X^i)$$

thus X^j must be an even power of X so that when a_i anti-commute with X^j, $\sigma^j(a_i) = a_i$ since σ is of order 2. Furthermore, we must have that $\sigma^i(b_j) = b_j$ for any i, showing that b_j must be real, which shows that the center is $\mathbb{R}[X^2]$. (More formally, one can take a polynomial in the center, say $p(X)$, and compute $p(X)a = ap(X)$ for any $a \in \mathbb{C}$, which shows that $p(X) \in \mathbb{C}[X^2]$, then compute $p(X)X = Xp(X)$ which shows that $p(X) \in \mathbb{R}[X^2]$).

- For this quotient to be a ring, we need the ideal $A(X^2+1)$ to be two-sided. This is the case since $X^2 + 1$ belongs to the center by the point above.

- We can express the ring of Hamilton quaternions $\mathbb{H}$ in the form $\mathbb{H} = \mathbb{C} \oplus \mathbb{C}j$, and define

$$\varphi : A \to \mathbb{H},\ \varphi(X) = j,\ \varphi(a) = a,\ a \in \mathbb{C}.$$

Since $ja = \sigma(a)j$ in $\mathbb{H}$ for any $a \in \mathbb{C}$, φ gives a ring homomorphism from A to $\mathbb{H}$. This induces a ring homomorphism $\bar{\varphi} : \bar{A} \to \mathbb{H}$ since $\varphi(X^2 + 1) = j^2 + 1 = 0$. Since

$$\bar{\varphi}(\overline{a + bX}) = a + bj,$$

$\bar{\varphi}$ is an isomorphism. (This is the first isomorphism theorem for rings.)

4.3 The Chinese Remainder Theorem

Exercise 66. Show that the following rings are isomorphic:

$$\mathbb{Z}/72\mathbb{Z} \times \mathbb{Z}/84\mathbb{Z} \simeq \mathbb{Z}/36\mathbb{Z} \times 168\mathbb{Z}.$$

Answer. We have that $72 = 8 \cdot 9$ and $\gcd(8, 9) = 1$, thus $\mathbb{Z}_{72} \simeq \mathbb{Z}_8 \times \mathbb{Z}_9$. Similarly $\mathbb{Z}_{84} \simeq \mathbb{Z}_4 \times \mathbb{Z}_3 \times \mathbb{Z}_7$, $\mathbb{Z}_{36} \simeq \mathbb{Z}_4 \times \mathbb{Z}_9$ and $\mathbb{Z}_{168} \simeq \mathbb{Z}_8 \times \mathbb{Z}_3 \times \mathbb{Z}_7$. Thus

$$\begin{aligned} \mathbb{Z}_{72} \times \mathbb{Z}_{84} &\simeq \mathbb{Z}_8 \times \mathbb{Z}_9 \times \mathbb{Z}_4 \times \mathbb{Z}_3 \times \mathbb{Z}_7 \\ &\simeq \mathbb{Z}_8 \times \mathbb{Z}_{36} \times \mathbb{Z}_3 \times \mathbb{Z}_7 \\ &\simeq \mathbb{Z}_{36} \times \mathbb{Z}_{128}. \end{aligned}$$

Exercise 67. Show that $10^{99} + 1$ is a multiple of 247.

Answer. We have that

$$100 = 12 \cdot 8 + 4$$

thus

$$10^{100} = (10^{12})^8 \cdot 10^4 \equiv 10^4 \equiv (-3)^4 \equiv 3 \equiv -10 \mod 13$$

where the second equality uses that $a^{p-1} \equiv 1 \mod p$. Similarly ($100 = 18 \cdot 6 - 8$)

$$10^{100} \equiv 10^{-8} \equiv 2^8 \equiv 9 \equiv -10 \mod 19.$$

By the Chinese Theorem, we deduce that

$$10^{100} \equiv -10 \mod 247.$$

Since $\gcd(10, 247) = 1$, we can simplify by a factor of 10, and get

$$10^{99} \equiv -1 \mod 247$$

and thus $247 \mid 10^{99} + 1$.

Exercise 68. The battle of Hasting (October 14, 1066). "The men of Harold stood well, together, as their wont was, and formed thirteen squares, with a like number of men in every square thereof, and woe to the hardy Norman who ventured to enter thier redoubts; for a single blow of a saxon warhatched would break his lance and cut through his coat of mail... When Harold threw himself into the fray the Saxon were one mighty square of men, shouting the battle-cries 'Ut!', 'Olicross!', 'Godemite!'."

How many men were there in the army of Harald Hardrada? (This exercise is courtesy of C. Wuthrich).

Answer. The men of Harald formed thirteen squares, that is $13x^2$, when Harold threw himself into the battle $(+1)$, they were one mighty square of men (y^2). This gives the equation

$$y^2 = 13x^2 + 1.$$

We then have to look for the smallest integer solution. Using field theory instead, one can rewrite this equation as

$$1 = (y - \sqrt{13}x)(y + \sqrt{13}).$$

We are thus looking for an element $y + \sqrt{13}x$ of $K = \mathbb{Q}(\sqrt{13})$ which satisfies this equation. One can show that $\eta = \frac{3+\sqrt{13}}{2}$ satisfies this equation up to a sign -1, thus η with an even power satisfies it, and η and its powers are actually the only elements in K to satisfy it. We thus need to take an even power of η which will give us an element in the ring $\mathbb{Z}[\sqrt{13}]$. We find that $\eta^6 = 649 + 180\sqrt{13}$ is the first power to satisfy this condition. Finally, the smallest integer solution to the equation $y^2 = 13x^2 + 1$ is $x = 180$ and $y = 649$, that is, there were 421'200 men with Harald Hardrada. It is however known that his army was instead containing about 7'500 men.

4.4 Maximal and prime ideals

Exercise 69. Show that a non-zero principal ideal is prime if and only if it is generated by a prime element.

Answer. If p is prime then consider the principal ideal $pR = \{pr,\ r \in R\}$. To show that pR is prime, we have to show that if $ab \in pR$ then either a or b is in pR. If $ab \in pR$, then $ab = pr$ for some $r \in R$. Since p is prime, it has to divide either a or b, that is either $a = pa'$ or $b = pb'$. Conversely, take a principal ideal cR which is prime, thus if $ab \in cR$, either $a \in cR$, that is $a = ca'$, or $b \in cR$, that is $b = cb'$. We have thus shown that if $c|ab$, then $c|a$ or $c|b$.

Exercise 70. Are the ideals $(X, X+1)$, $(5, X^2+4)$ and $(X^2+1, X+2)$ prime/maximal in $\mathbb{Z}[X]$?

Answer.

- $I = (X, X+1) = \mathbb{Z}$ since $1 = (X+1) - X$, thus I is not a proper ideal and cannot be prime.
- Consider $\mathbb{Z}[X]/(5, X^2+4) \simeq \mathbb{Z}_5[X]/(X^2+4)$, and $(X^2+4) = (X-\bar{1})(X+\bar{1})$ is reducible modulo 5, thus this quotient is not an integral domain and thus the ideal is not prime.
- $I = (X^2+1, X+2) = (X+2, 5)$ since $(X+2)^2 - 4(X+2) + 5 = X^2+1$, then $\mathbb{Z}[X]/I \simeq \mathbb{Z}_5[X]/(X+\bar{2})$ where $X+\bar{2}$ is irreducible in $\mathbb{Z}_5[X]$ thus the quotient is a field and I is maximal.

Exercise 71.

1. Consider the ring $R = \mathbb{Z}[i]$ and the ideal $I = (1+i)$ in R. Is I prime? Is I maximal?
2. Consider the ring $R = \mathbb{Z}[j]$ and the ideal $I = (2 - rj)$ in R. Is I prime? Is I maximal? (j is a primitive 3rd root of unity.)
3. Consider the ring $R = \mathbb{Z}[X]$ and the ideal $I = (n)$ in R. Is I prime? Is I maximal?

Answer.

1. We have $\mathbb{Z}[i]/(1+i) \simeq \mathbb{Z}/2\mathbb{Z}$, which is a field, so $(1+i)$ is maximal (hence prime).
2. The characteristic of $\mathbb{Z}[j]/(2-5j)$ is 39 which is not a prime number (see Exercise 63), so $\mathbb{Z}[j]/(2-5j)$ is not an integral domain. Hence $(2-5j)$ is not prime and therefore not maximal.
3. We have $\mathbb{Z}[X]/(n) \simeq \mathbb{Z}/n\mathbb{Z}[X]$. We have that $\mathbb{Z}/n\mathbb{Z}[X]$ is an integral domain if and only if $\mathbb{Z}/n\mathbb{Z}$ is an integral domain. Hence (n) is a prime ideal if and only if n is a prime number. It is never maximal since $\mathbb{Z}/n\mathbb{Z}[X]$ is not a field for any n (X has no inverse).

Exercise 72. Consider the ring $R = K[X]$ and the ideal of R given by $I = (X-a)$, where K is a field, and $a \in K$. Is I maximal? Is I prime?

Answer. Let $\varphi : P \in K[X] \mapsto P(a) \in K$. This is a ring homomorphism, which is surjective: indeed, if $\lambda \in K$, then $\varphi(\lambda) = \lambda$, where $\lambda \in K \subset K[X]$

is viewed as a constant polynomial. We now determine the kernel of φ. Let $P \in K[X]$. We can write $P = Q(X).(X-a)+c$, for some $Q \in K[X]$ and $c \in K$. (Indeed, it suffices to proceed to the division of P by $X-a$. The remainder is either zero or has degree < 1, that is degree 0, which means that the remainder is a constant.) Then we have $P(a) = Q(a).(a-a)+c = c$. Therefore, $\varphi(P) = 0 \iff c = 0 \iff P$ is a multiple of $X-a$. Hence $\ker(\varphi) = (X-a)$ (the principal ideal generated by $X-a$). Using the first isomorphism theorem, we get that $K[X]/(X-a) \simeq K$. Since $K[X]/(X-a) \simeq K$, and K is a field, then $K[X]/(X-a)$ is a field as well and $(X-a)$ is maximal (hence prime).

Exercise 73. Let R be a commutative ring. Let

$$Nil(R) = \{r \in R | \exists n \geq 1, r^n = 0\}.$$

1. Show that $Nil(R)$ is contained in the intersection of all prime ideals of R.
2. Show that $Nil(R/Nil(R)) = 0$.

Answer.

1. Let $a \in Nil(R)$, so $a^n = 0$ for some $n \geq 1$. Assume that there is a prime ideal $\mathfrak{p}$ for which $a \notin \mathfrak{p}$. We have $a^n = 0 \in \mathfrak{p}$. Since $a^n = a^{n-1}.a$ and $\mathfrak{p}$ is a prime ideal, then $a^{n-1} \in \mathfrak{p}$ or $a \in \mathfrak{p}$. By assumption on a, we have $a \notin \mathfrak{p}$, so necessarily $a^{n-1} \in \mathfrak{p}$. But $a^{n-1} = a^{n-2}.a \in \mathfrak{p}$, so $a^{n-2} \in \mathfrak{p}$ for the same reasons, and by induction we get $a \in \mathfrak{p}$, a contradiction. Therefore a lies in all the prime ideals of R.
2. Let $\overline{a} \in Nil((R/Nil(R)))$, so $\overline{a}^n = \overline{0}$ for some $n \geq 1$. Then $\overline{a^n} = \overline{0}$, which means that $a^n \in Nil(R)$ by definition of the quotient ring. Therefore, there exists $m \geq 1$ such that $(a^n)^m = 0$, so $a^{nm} = 0$, which means that $a \in Nil(R)$. Hence $\overline{a} = \overline{0}$.

Exercise 74. Let $R = \mathbb{Z}[X]$, and let $n \geq 1$.

- Show that the ideal (n, X) is given by

$$(n, X) = \{p(X) \in \mathbb{Z}[X],\ p(0) \text{ is a multiple of } n\}.$$

- Show that (n, X) is a prime ideal if and only if n is a prime number.

Answer.

- Let $P \in (n, X)$, so $P = n.Q_1 + X.Q_2$ for some $Q_1, Q_2 \in \mathbb{Z}[X]$. Then $P(0) = n.Q_1(0) \in n\,\mathbb{Z}$ (we have $Q_1(0) \in \mathbb{Z}$ since $Q_1 \in \mathbb{Z}[X]$), that is $P(0)$ is a multiple of n. Conversely, assume that $P \in \mathbb{Z}[X]$ is such that $P(0)$ is a multiple of n, and write $P = a_nX^n + \cdots + a_1X + a_0$. Then $P(0) = a_0$, so by assumption $a_0 = n.m$ for some $m \in \mathbb{Z}$. Now we get $P = n.m + X.(a_nX^{n-1} + \cdots + a_2X + a_1)$, so $P \in (n, X)$.

- If n is not a prime number, then we can write $n = n_1.n_2, 1 < n_1, n_2 < n$. Now consider $P_1 = n_1, P_2 = n_2 \in \mathbb{Z}[X]$ (constant polynomials). We have $P_1.P_2 = n_1.n_2 = n \in (n, X)$, but P_1 and P_2 are not elements of (n, X). Indeed, $P_1(0) = n_1$ and $P_2(0) = n_2$, but n_1, n_2 are not multiples of n by definition. Hence (n, X) is not a prime ideal. Now assume that n is equal to a prime number p. First of all, $(p, X) \neq \mathbb{Z}[X]$, because $1 \notin (p, X)$ for example. Now let $P_1, P_2 \in \mathbb{Z}[X]$ such that $P_1.P_2 \in (p, X)$. Then $(P_1.P_2)(0)$ is a multiple of p by the previous point, that is $p|P_1(0).P_2(0)$. Since p is a prime number, it means that $p|P_1(0)$ or $p|P_2(0)$, that is $P_1 \in (p, X)$ or $P_2 \in (p, X)$. Hence (p, X) is a prime ideal.

4.5 Polynomial rings

Exercise 75. Set

$$E = \{p(X) \in \mathbb{Z}[X] \mid p(0) \text{ is even }\},\ F = \{q(X) \in \mathbb{Z}[X] \mid q(0) \equiv 0(\text{mod } 3)\}.$$

Check that E and F are ideals of $\mathbb{Z}[X]$ and compute the ideal $E + F$. Furthermore, check that $E \cdot F \subseteq \{p(X) \in \mathbb{Z}[X] | p(0) \equiv 0 \text{ (mod 6) }\}$.

Answer. If $p(X) = \sum_{k=0}^{n} p_k X^k$, then

$$E = \{p(X) \in \mathbb{Z}[X] \mid p_0 \in 2\mathbb{Z}\} \quad \text{and} \quad F = \{q(X) \in \mathbb{Z}[X] \mid q_0 \in 3\mathbb{Z}\}.$$

Thus E and F are ideals of $\mathbb{Z}[X]$ since $2\mathbb{Z}$ and $3\mathbb{Z}$ are ideals of $\mathbb{Z}$. If $\sum_k c_k X^k = \left(\sum_k p_k X^k\right) \cdot \left(\sum_k q_k X^k\right)$, then $c_0 = p_0 q_0$ and thus

$$E \cdot F \subseteq \{p(X) \in \mathbb{Z}[X] \mid p_0 \in 2\mathbb{Z} \cdot 3\mathbb{Z}\} = \{p(X) \in \mathbb{Z}[X] \mid p_0 \in 6\mathbb{Z}\}.$$

Similarly,

$$E + F = \{p(X) \in \mathbb{Z}[X] \mid p_0 \in 2\mathbb{Z} + 3\mathbb{Z}\} \underbrace{=}_{\text{Bezout}} \{p(X) \in \mathbb{Z}[X] \mid p_0 \in \mathbb{Z}\} = \mathbb{Z}[X].$$

Exercise 76. Show that if F is a field, the units in $F[X]$ are exactly the nonzero elements of F.

Answer. Let $f(X) \in F[X]$ of degree n, $f(X)$ is a unit if and only if there exists another polynomial $g(X) \in F[X]$ of degree m such that $f(X)g(X) = 1$. Because F is a field (thus in particular an integral domain), $f(X)g(X)$ is a polynomial of degree $n + m$, thus for the equality to hold, since 1 is a polynomial of degree 0, we need $n + m = 0$, thus both f and g are constant, satisfying $fg = 1$, that is they are units of F, that is nonzero elements since F is a field.

Exercise 77. There exists a polynomial of degree 2 over $\mathbb{Z}/4\mathbb{Z}$ which has 4 roots. True or false? Justify your answer.

Answer. Take the polynomial $2X(X - 1)$.

Exercise 78. Let R be a ring, and let $a \neq 0 \in R$ such that there exists an integer n with $a^n = 0$. Show that $R^* \subset (R[X])^*$ and $R^* \neq R[X]^*$, where R^* and $R[X]^*$ denote respectively the group of units of R and $R[X]$.

Answer. Clearly $R^* \subseteq R[X]^*$. We need to show that the inclusion is strict, that this, there exists an element in $R[X]^*$ which is not in R^*. Take $f(X) = 1 - aX$. We have

$$(1 - aX)(1 + aX + (aX)^2 + \ldots + (aX)^{n-1}) = 1,$$

and f does not belong to R^*.

Exercise 79. Let K be a field. Consider the ring $K[X, Y]$ of polynomials in indeterminates X and Y with coefficients in K.

1. Is $K[X, Y]$ an integral domain?
2. What are the units of $K[X, Y]$?
3. Consider the ideals $\mathcal{I}_1 = (X)$ and $\mathcal{I}_2 = (X, Y)$. Are they prime ideals of $K[X, Y]$?
4. Show that $\mathcal{J} = \{f \in K[X, Y],\ f(0, 0) = 0\}$ is an ideal.
5. Deduce using $\mathcal{J}$ that $K[X, Y]$ cannot be a principal ideal domain.

Answer.

1. Yes it is. It is a commutative ring (since K is a field). Furthermore, it has no zero divisor, since K as none.
2. So units of $K[X, Y]$ are polynomials $f \in K[X, Y]$ such that there exist $g \in K[X, Y]$ with $fg = 1$. Thus the degree of the polynomial fg is 0, and both f, g must be constant polynomials (since K is a field). Thus the units are those of K.
3. Both of them are for the same reason: $K[X, Y]/\mathcal{I}_1 \simeq K[Y]$ and $K[X, Y]/\mathcal{I}_2 \simeq K$, both of them are integral domains, thus both ideals are prime.
4. Take $f, g \in \mathcal{J}$, then $f - g$ belongs to $\mathcal{J}$, and if h is in $K[X, Y]$, we also have that $hf \in \mathcal{J}$.
5. Assume there exists $f \in K[X, Y]$ such that $(f) = \mathcal{J}$. Note that both X and Y belong to $\mathcal{J}$. Thus there must exist $g, h \in K[X, Y]$ such that $X = f(X, Y)g(X, Y)$ and $Y = f(X, Y)h(X, Y)$. Since X is of degree 1, and Y is of degree 1, we should have $f(X, Y) = aX + bY$. But now, if $a \neq 0$, $Y = h(X, Y)(aX + bY)$ is not possible, and if $b \neq 0$, $X = f(X, Y)g(X, Y)$ is not possible either.

4.6 Unique factorization and Euclidean division

Exercise 80.

Show that the ideal generated by 2 and X in the ring of polynomials $\mathbb{Z}[X]$ is not principal.

Answer. We have that

$$\langle 2, X\rangle = \{2r(X) + Xs(X),\ r(X), s(X) \in \mathbb{Z}[X]\},$$

and assume there exists $f(X) \in \mathbb{Z}[X]$ such that $\langle 2, X\rangle = (f(X))$. Since $2 \in (f(X))$, then $f(X) = \pm 2$. Since $X \in (f(X))$, we should have $X = \pm 2g(X)$, a contradiction.

Exercise 81. Let R be an integral domain in which every decreasing chain of ideals is finite. Show that R is a field.

Answer. Let $x \in R$, $x \neq 0$. Then $(x) \supset (x^2) \supset (x^3) \supset \ldots$ is a decreasing chain of ideals. It thus stabilizes at some point by assumption, that is, there is a k in $\mathbb{N}$ such that $(x^k) = (x^{k+1})$. In particular, there is an element $a \in R$ such that $ax^{k+1} = x^k$. Since R is an integral domain, we have $ax = 1$, and thus x is invertible, showing that R without the 0 element is a field.

Exercise 82. Show that if R is a unique factorization domain, then $R[X]$ is also a unique factorization domain.

Answer. Let us write $f(X) = a_0 + a_1X + a_2X^2 + \cdots + a_nX^n$, $a_j \in R$. Recall that $c(f)$ is the content of f defined as $\gcd(a_0, \ldots, a_n)$. We need to check that a factorization exists, and that it is unique.

- Existence: if $p \in R$ is irreducible, then p is also irreducible in $R[X]$. If $f(X) \in R[X]$, we can write $f(X) = d\tilde{f}(X)$ be factoring the content d, so that $c(\tilde{f}) = 1$. We can factor d into a product of irreducible in R. Now either $\tilde{f}$ is irreducible in $R[X]$, or it factors properly into a product of lower degree polynomials ($c(\tilde{f}) = 1$). All the factors will also have content 1, and we can only lower degree of factors finitely often, so we get a factorization of $\tilde{f}$, and thus one for f as product of irreducibles in $R[X]$.
- Uniqueness: by Theor 2.15, it suffices to prove that each irreducible element is prime, which we can do by proving that each irreducible element generates a prime ideal in $R[X]/$ If $p \in R$ is irreducible, then $R[X]/pR[X] = (R/p)[X]$ which is an integral domain.

Exercise 83. Let F be a field, let $f(X), g(X) \in F[X]$, and let $d(X)$ be a greatest common divisor of $f(X)$ and $g(X)$. Show that there are polynomials $u(X), v(X) \in F[X]$ such that

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

When does Bezout identity hold more generally?

Answer. Bezout identity works for general PID as follows (and thus in particular for $F[X]$). Take $a, b \in R$, where R is a PID. Consider the corresponding principal ideals aR and bR, we have that

$$aR + bR = cR$$

simply because R is a PID. Since $aR \subset cR$, $c|a$ and for the same reason $c|b$. Now consider $d = \gcd(a, b)$, then $d|a$ and $d|b$, and thus conversely dR contains aR and bR and thus cR, showing that $d|c$. But c must also divide d, showing that $c = d$, that is

$$aR + bR = \gcd(a, b)R,$$

in words, $\gcd(a, b)$ is some linear combination of a and b using coefficients in R. This does not work for arbitrary UFDs. For example, in $\mathbb{Z}[X]$, the polynomials X and 2 are coprime, but no linear combination of 2 and X gives 1. For more generalization of this notion, check the definition of Bezout domain.

Exercise 84. Show that $\mathbb{Z}[\sqrt{3}]$ is a Euclidean domain. (Hint: use the same technique as the one seen for $\mathbb{Z}[\sqrt{2}]$.)

Answer. Consider the ring

$$\mathbb{Z}[\sqrt{3}] = \{a + b\sqrt{3},\ a, b \in \mathbb{Z}\}$$

with

$$\Psi(a + b\sqrt{3}) = |a^2 - 3b^2|.$$

Take $\alpha, \beta \neq 0$ in $\mathbb{Z}[\sqrt{3}]$, and compute the division in $\mathbb{Q}(\sqrt{3})$:

$$\alpha/\beta = q',$$

with $q' = x + \sqrt{3}y$ with x, y rational. Let us now approximate x, y by integers x_0, y_0, namely take x_0, y_0 such that

$$|x - x_0| \leq 1/2,\ |y - y_0| \leq 1/2.$$

Take

$$q = x_0 + y_0\sqrt{3},\ r = \beta((x - x_0) + (y - y_0)\sqrt{3}),$$

where clearly $q \in \mathbb{Z}[\sqrt{3}]$, then

$$\begin{aligned} \beta q + r &= \beta(x_0 + y_0\sqrt{3}) + \beta((x - x_0) + (y - y_0)\sqrt{3}) \\ &= \beta(x + y\sqrt{3}) = \beta q' = \alpha, \end{aligned}$$

which at the same time shows that $r \in \mathbb{Z}[\sqrt{3}]$. So far this is exactly what we did in the lecture. We are also left to show that $\Psi(r) < \Psi(\beta)$. We have

$$\begin{aligned} \Psi(r) &= \Psi(\beta)\Psi((x - x_0) + (y - y_0)\sqrt{d}) \\ &= \Psi(\beta)|(x - x_0)^2 - d(y - y_0)^2| \\ &\leq \Psi(\beta)[|x - x_0|^2 + |d||y - y_0|^2] \\ &\leq \Psi(\beta)\left(\frac{1}{4} + |3|\frac{1}{4}\right) \end{aligned}$$

though here we notice that we get $\frac{1}{4}+|3|\frac{1}{4}=1$. So this is not good enough! But let us see what this means to get 1: this happens only if $|x-x_0|^2=|y-y_0|^2=1/4$, otherwise we do get something smaller than 1. Now if $|x-x_0|^2=|y-y_0|^2=1/4$, we have from the second equation that

$$\Psi=\Psi(\beta)|(x-x_0)^2-d(y-y_0)^2|=\Psi(\beta)|\frac{1}{4}-\frac{3}{4}|<1$$

and we are done.

Exercise 85. The goal of this exercise is to show that a principal ideal domain is a unique factorization domain in which every prime ideal is maximal. (Hint: To show that every prime is maximal, take a prime ideal $\mathcal{I}$ and a maximal ideal $\mathcal{M}$, and see what it means for $\mathcal{I}$ to be included in $\mathcal{M}$ in a PID). Note that the converse is true.

Answer. If we have a PID, it is a UFD (this is far from obvious, this was shown in the notes). We have to show that every prime ideal is maximal. Take $\mathcal{I}$ a prime ideal, and $\mathcal{M}$ a maximal ideal. Thus $\mathcal{I}\subseteq\mathcal{M}$ by maximality of $\mathcal{M}$. Now since we have a PID, we can write $\mathcal{I}=(a)$, $\mathcal{M}=(m)$ and $(a)\subseteq(m)$ showing that $m|a$. Thus $a=md$ for some d. But now a is prime (this follows from (a) being prime, see Exercise 69) thus it is irreducible (in a UFD, irreducible and prime are equivalent). Since a is irreducible, either m or d is a unit, and m cannot be (otherwise $\mathcal{M}$ would be R, which is impossible by definition of maximal ideal), thus d is a unit. Then a and m are associate, so they generate the same principal ideal, and $\mathcal{I}=\mathcal{M}$.

4.7 Irreducible polynomials

Exercise 86. Prove whether the following polynomials are reducible/irreducible over F.

1. t^2-2, $F=\mathbb{Q}$.
2. $\frac{2}{9}t^5+\frac{5}{3}t^4+t^3+\frac{1}{3}$, $F=\mathbb{Q}$.
3. t^4+15t^3+7, $F=\mathbb{Z}$, *hint: think of modulo.*
4. $t^{16}+t^{15}+t^{14}+\ldots+t^3+t^2+t+1$, $F=\mathbb{Q}$, *hint: this needs a trick.*

Answer.

1. Use Eisenstein's criterion with $p=2$.
2. This polynomial is irreducible if and only if
$$9f(t)=2t^5+15t^4+9t^3+3$$
is irreducible over $\mathbb{Q}$. Here Eisenstein's criterion can be applied with $p=3$, showing that f is irreducible.

3. Modulo 5, $f(t) \equiv t^4 + 2$. If this is reducible, then either it has a a factor of degree 1 (not possible, it is easy to try the 5 values), or it is a product of two factors of degree 2. The latter can be checked explicitly: if

$$t^4 + 2 = (t^2 + at + b)(t^2 + ct + d)$$

then $a + c = 0$, $ac + b + d = 0$, $bd = 2$. One can check all possible values and see that this is not possible either. Hence $t^4 + 2$ is irreducible modulo 5, and therefore the original polynomial was irreducible over $\mathbb{Z}$.

4. Notice that $f(t)$ is irreducible if and only if $f(t + 1)$ is. By expanding $f(t + 1)$, one can use Eisenstein's criterion with $q = 17$.

Exercise 87. True/False.

Q1. Let R be a ring, and let r be an element of R. If r is not a zero divisor of R, then r is a unit.

Q2. A principal ideal domain is a euclidean domain.

Q3. Hamilton's quaternions form a skew field.

Q4. The quotient ring $\mathbb{Z}[i]/(1 + i)\mathbb{Z}[i]$ is a field.

Q5. A field is a unique factorization domain.

Q6. The ideal $(5, i)$ in $\mathbb{Z}[i]$ is principal.

Q7. The polynomial $3x^4 + 15X^2 + 10$ is irreducible over $\mathbb{Q}$.

Q8. Let R be a ring, and M be a maximal ideal, then R/M is an integral domain.

Answer.

Q1. This cannot be true in general! Take $\mathbb{Z}$ for example. It has no zero divisor, but apart 1 and -1, no other element is a unit! Actually, in an integral domain, there is no zero divisor, which does not mean it is an field.

Q2. A euclidean domain is a principal ideal domain. The converse is not true. Take for example $\mathbb{Z}[(1 + i\sqrt{19})/2]$. It is a principal ideal domain, but it is not a euclidean domain.

Q3. A skew field is non-commutative field. Hamilton's quaternions are non-commutative, and we have seen that every non-zero quaternion is invertible (the inverse of q is its conjugate divided by its norm).

Q4. It is actually a field. You can actually compute the quotient ring explicitly, this shows that $\mathbb{Z}[i]/(1 + i)\mathbb{Z}[i]$ is isomorphic to the field of 2 elements $\{0, 1\}$.This can be done using the first isomorphism for rings.

Q5. It is true since every non-zero element is a unit by definition.

Q6. It is true! With no computation, we know it from the theory: We know that $\mathbb{Z}[i]$ is a euclidean domain, and thus it is a principal domain, so all ideals including this one are principal.

Q7. It is true! Use for example Eisenstein's criterion with $p = 5$.

Q8. Who said the ring R is commutative? The statement seen in the class is about commutative rings. It is not true for non-commutative rings. Here is an example: take $R = \mathbb{Z} + \mathbb{Z}i + \mathbb{Z}j + \mathbb{Z}k$ (ring of quaternions with integer coefficients), pR is a maximal ideal of R (p odd prime) but R/pR is actually isomorphic to $M_2(\mathbb{Z}/p\mathbb{Z})$ and thus is not an integral domain.

Chapter 5

Field Theory

Abstract field theory emerged from three theories, which we would now call Galois theory, algebraic number theory and algebraic geometry.

Field theoretic notions appeared, even though still implicitly, in the modern theory of solvability of polynomial equations, as introduced by Abel and Galois in the early nineteenth century. Galois had a good insight into fields obtained by adjoining roots of polynomials, and he proved what we call now the Primitive Element Theorem.

Independently, Dedekind and Kronecker came up with the notion of algebraic number fields, arising from three major number -theoretic problems: Fermat's Last Theorem, reciprocity laws and representation of integers by binary quadratic forms.

Algebraic geometry is the study of algebraic curves and their generalizations to higher dimensions, namely, algebraic varieties. Dedekind and Weber carried over to algebraic functions the ideas which Dedekind had earlier introduced for algebraic numbers, that is, define an algebraic function field as a finite extension of the field of rational functions.

At the end of the nineteenth century, abstraction and axiomatics started to take place. Cantor (1883) defined the real numbers as equivalence classes of Cauchy sequences,von Dyck (1882) gave an abstract definition of group (about thirty years after Cayley had defined a finite group). Weber's definition of a field appeared in 1893, for which he gave number fields and function fields as examples. In 1899, Hensel initiated a study of p-adic numbers, taking as starting point the analogy between function fields and number fields. It is the work of Steinitz in 1910 that initiated the abstract study of fields as an independent subject. A few examples of his results are: classification of fields into those of characteristic zero and those of characteristic p, development of the theory of transcendental extensions, recognition that it is precisely the finite, normal, separable extensions to which Galois theory applies, proof of the existence of the algebraic closure of any field.

Major developments in field theory and related areas that followed Steinitz's work include valuation theory, class field theory, infinite Galois theory and finite fields.

5.1 Field extension and minimal polynomial

Definition 5.1. If F and E are fields, and $F \subseteq E$, we say that E is an extension of F, and we write either $F \leq E$ or E/F.

Examples 5.1. Here are some classical examples:

1. $\mathbb{C} = \{a + bi,\ a, b \in \mathbb{R}\}$ is a field extension of $\mathbb{R}$.
2. $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2},\ a, b \in \mathbb{Q}\}$ is a field extension of $\mathbb{Q}$.
3. $\mathbb{Q}(i) = \{a + bi,\ a, b \in \mathbb{Q}\}$ is a field extension of $\mathbb{Q}$.

If E is an extension of F, then in particular E is an abelian group under addition, and we may multiply $x \in E$ by $\lambda \in F$. We can see that this endows E with a structure of F-vector space (the elements of E are seen as vectors, those of F as scalars). It then makes sense to speak of the dimension of E over F.

Definition 5.2. Let E/F be a field extension. The dimension of E as F-vector space is called the degree of the extension, written $[E : F]$. If $[E : F] < \infty$, we say that E is a finite extension of F, or that the extension E/F is finite.

Let us get back to our examples:

Examples 5.2.

1. Consider the field extension $\mathbb{C}/\mathbb{R}$. We have that $\mathbb{C}$ is a vector space of dimension 2 over $\mathbb{R}$. It is thus an extension of degree 2 (with basis $\{1, i\}$).
2. The field extension $\mathbb{Q}(\sqrt{(2)})/\mathbb{Q}$ is of degree 2, it is called a quadratic extension of $\mathbb{Q}$.
3. The field extension $\mathbb{Q}(i)/\mathbb{Q}$ is a also a quadratic field extension of $\mathbb{Q}$.
4. Both $\mathbb{Q}(\sqrt{(2)})/\mathbb{Q}$ and $\mathbb{Q}(i)/\mathbb{Q}$ are finite field extensions of $\mathbb{Q}$. Finite extensions of $\mathbb{Q}$ are called number fields.

If we look at $\mathbb{C}$, we see it is obtained by adding i to $\mathbb{R}$, and i is a root of the polynomial $X^2 + 1$. Similarly, $\mathbb{Q}(\sqrt{2})/\mathbb{Q}$ is obtained by adding a root of the polynomial $X^2 - 2$. In what follows, we will make formal the connection between roots of polynomials and field extensions.

Before we start, recall that if we have two fields E, F and a field homomorphism between them (that is, a ring homomorphism between two fields), then f is a monomorphism. We have seen the argument in the previous chapter already: the kernel of a ring homomorphism is an ideal, and a field has only trivial ideals, namely $\{0\}$ and itself, and it cannot be that the whole field is the kernel.

Theorem 5.1. *Let f be a non-constant polynomial over a field F. Then there is an extension E/F and an element $\alpha \in E$ such that $f(\alpha) = 0$.*

Proof. Recall that $F[X]$ is a unique factorization domain, thus f can be factored into a product of irreducible polynomials, and we may assume without loss of generality that f is itself irreducible. Consider now the ideal

$$\mathcal{I} = (f(X))$$

in $F[X]$, the ring of polynomials with indeterminate X and coefficients in F. Again using that $F[X]$ is a unique factorization domain, we have that $f(X)$ is irreducible and equivalently prime, implying that $(f(X))$ is prime. Now $F[X]$ is furthermore a principal ideal domain. This means that $\mathcal{I} = (f(X))$ is contained in a principal maximal ideal $(q(X))$, so that $q(X)$ divides the prime $f(X)$. Since $f(X) = q(X)g(X)$ for some $g(X)$, and $q(X)$ cannot be a unit because $f(X)$ is irreducible, $f(X)$ and $q(X)$ are associates, and $(f(X)) = (q(X))$, proving that $(p(X)) = \mathcal{I}$ is maximal. Thus by the characterization of maximal ideals with respect to their quotient ring, we have that

$$E = F[X]/\mathcal{I}$$

is a field. We now place an isomorphic copy of F inside E via the monomorphism

$$h : F \to E,\ a \mapsto a + \mathcal{I}.$$

This thus gives a field extension E/F. Now let

$$\alpha = X + \mathcal{I} \in E.$$

We are left to prove that α is a root of $f(X)$. If $f(X) = a_0 + a_1X + \ldots + a_nX^n$, then

$$\begin{aligned} f(\alpha) &= (a_0 + \mathcal{I}) + a_1(X + \mathcal{I}) + \ldots + a_n(X + \mathcal{I})^n \\ &= a_0 + \mathcal{I} + a_1X + a_1\mathcal{I} + \ldots + a_nX^n + \ldots + a_n\mathcal{I}^n \\ &= (a_0 + a_1X + \ldots + a_nX^n) + \mathcal{I} \\ &= f(X) + \mathcal{I} \end{aligned}$$

which is zero in E. □

The extension E is sometimes said to be obtained from F by adjoining a root of f.

Remark. Note that in the above proof, we have shown that a prime ideal in a principal ideal domain is maximal.

Definition 5.3. If E is an extension of F, an element $\alpha \in E$ is said to be algebraic over F if there is a non-constant polynomial $f \in F[X]$ such that $f(\alpha) = 0$. If α is not algebraic over F, it is said to be transcendental over F. If every element of E is algebraic over F, then E is said to be an algebraic extension of F.

Suppose that $\alpha \in E$ is algebraic over F. Thus there exists by definition a polynomial $f \in F[X]$ with $f(\alpha) = 0$. It thus makes sense to consider the set $\mathcal{I}$ of all polynomials $g \in F[X]$ such that $g(\alpha) = 0$. Clearly

- if g_1, g_2 are in $\mathcal{I}$, so does $g_1 \pm g_2$,
- if $g \in \mathcal{I}$ and $h \in F[X]$, then $gh \in \mathcal{I}$.

This tells us that $\mathcal{I} = \{g \in F[X],\ g(\alpha) = 0\}$ is an ideal of $F[X]$.

Since $F[X]$ is a principal ideal domain, we have

$$\mathcal{I} = (m(X))$$

for some $m(X)$ in $F[X]$. Any two generators of $\mathcal{I}$ are thus multiple of each others, so they must be of same degree, and since $m(X)$ is monic, it has to be unique. This polynomial $m(X)$ has the following properties:

1. If $g \in F[X]$, then $g(\alpha) = 0$ if and only if $m(X)$ divides $g(X)$. This is clear from the definition of $\mathcal{I}$.
2. $m(X)$ is the monic polynomial of least degree such that $m(\alpha) = 0$, which follows from the above property.
3. $m(X)$ is the unique monic irreducible polynomial such that $m(\alpha) = 0$. Indeed, if $m(X) = h(X)k(X)$ with $\deg h < \deg m$, $\deg k < \deg m$, then either $h(\alpha) = 0$ or $k(\alpha) = 0$, so that either $h(X)$ or $k(X)$ is a multiple of $m(X)$ by the first property, which is impossible. Thus $m(X)$ is irreducible. We are left to prove the unicity of $m(X)$. This comes from the fact that since $m(X)$ is monic, then if there were two irreducible monic polynomials $m(X)$ and $m'(X)$ such that $m(\alpha) = m'(\alpha) = 0$, they have α as common root, and thus $m(X)$ and $m'(X)$ cannot be distinct (see the proposition below).

Definition 5.4. The polynomial $m(X)$ is called the minimal polynomial of α over F. It may be denoted by $\min(\alpha, F)$ or $\mu_{\alpha,F}$.

Example 5.3. The polynomial $X^2 + 1$ is the minimal polynomial of i over $\mathbb{Q}$. It also the minimal polynomial of i over $\mathbb{R}$.

Proposition 5.2. *1. Let f and g be polynomials over the field F. Then f and g are relatively prime if and only if f and g have no common root in any extension of F.*

2. If f and g are distinct monic irreducible polynomials over F, then f and g have no common roots in any extension of F.

Proof. 1. If f and g are relatively prime, their greatest common divisor is 1, so there are polynomials $a(X)$ and $b(X)$ over F such that

$$a(X)f(X) + b(X)g(X) = 1.$$

If there is a common root say α, then we get that $0 = 1$, a contradiction.

Conversely, let us assume that the greatest common divisor $d(X)$ of $f(X)$ and $g(X)$ is non-constant and show that then $f(X)$ and $g(X)$ have a common root. By the above proposition, there exists E an extension of F in which $d(X)$ has a root α. Since $d(X)$ divides both $f(X)$ and $g(X)$, α is a common root of f and g in E.

2. By the first part, it is enough to show that f and g are relatively prime. Assume to the contrary that h is a non-constant divisor of the polynomials f and g which are irreducible. Then $f = f'h$ and $g = g'h$ with f', g' non-zero constant, and $h = \frac{f}{f'} = \frac{g}{g'}$, that is, $f = \frac{f'}{g'}g$. It is impossible for f to be a constant multiple of g, because f and g are monic and distinct.

□

If E is an extension of F and $\alpha \in E$ is a root of a polynomial $f \in F[X]$, one may consider the field $F(\alpha)$ generated by F and α, which is the smallest subfield of E containing both F and α. Alternatively, $F(\alpha)$ can be described as the intersection of all subfields of E containing F and α, or the set of all rational functions

$$\frac{a_0 + a_1\alpha + \cdots + a_m\alpha^m}{b_0 + b_a 1\alpha + \ldots + b_n\alpha^n}$$

with $a_i, b_j \in F$, $m, n = 0, 1, \ldots$ and the denominator is different from 0.

Theorem 5.3. *Let $\alpha \in E$ be algebraic over F, with minimal polynomial $m(X)$ over F of degree n.*

1. *We have $F(\alpha) = F[\alpha] = F_{n-1}[\alpha]$ where $F_{n-1}[\alpha]$ denotes the set of all polynomials of degree at most $n-1$ with coefficients in F.*
2. *$\{1, \alpha, \ldots, \alpha^{n-1}\}$ forms a basis for the vector space $F(\alpha)$ over the field F. Consequently $[F(\alpha) : F] = n$.*

Proof. Let us first prove that $F_{n-1}[\alpha]$ is a field. Let $f(X)$ be any non-zero polynomial over F of degree at most $n-1$. Since $m(X)$ is irreducible with $\deg f < \deg m$, $f(X)$ and $m(X)$ are relatively prime, and there exist polynomials $a(X)$ and $b(X)$ over F such

$$a(X)f(X) + b(X)m(X) = 1.$$

Using that α is a root of m, we get

$$a(\alpha)f(\alpha) = 1$$

so that any non-zero element of $F_{n-1}[\alpha]$ has an inverse, and $F_{n-1}[\alpha]$ is a field.

1. Any field containing F and α must contain all polynomials in α, and in particular all those of degree at most $n-1$. Thus

$$F_{n-1}[\alpha] \subset F[\alpha] \subset F(\alpha).$$

But $F(\alpha)$ is the smallest field containing F and α, so

$$F(\alpha) \subset F_{n-1}[\alpha]$$

and we conclude that

$$F(\alpha) = F[\alpha] = F_{n-1}[\alpha].$$

2. Now $1, \alpha, \ldots, \alpha^{n-1}$ certainly span $F_{n-1}[\alpha]$, and they are linearly independent because if a non-trivial linear combination of them were zero, this would yield a non-zero polynomial of degree less than that of $m(X)$ with α as a root, a contradiction.

□

Example 5.4. Let ζ_5 denote a primitive 5th root of unity (that is, $\zeta_5^5 = 1$ and $\zeta_5^k \neq 1$ for $1 \leq k \leq 4$). We have that $\zeta_5 \in \mathbb{Q}(\zeta_5)$ is algebraic over $\mathbb{Q}$, with minimal polynomial $X^4 + X^3 + X^2 + X + 1 = 0$ of degree 4 over $\mathbb{Q}$. A $\mathbb{Q}$-basis is given by $\{1, \zeta_5, \zeta_5^2, \zeta_5^3\}$ and $[\mathbb{Q}(\zeta_5) : \mathbb{Q}] = 4$.

Once we have a field extension K/F, we can take again K as base field and get another field extension E/K, yielding a tower of extensions $E/K/F$.

Proposition 5.4. *Consider the field extensions $E/K/F$.*

1. *If α_i, $i \in I$, form a basis for E over K, and β_j, $j \in J$ form a basis for K over F, then $\alpha_i\beta_j$, $i \in I$, $j \in J$, form a basis for E over F.*

2. *The degree is multiplicative, namely*

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

In particular, $[E : F]$ is finite if and only if $[E : K]$ and $[K : F]$ are finite.

Proof. 1. Take $\gamma \in E$. Then

$$\begin{aligned} \gamma &= \sum_{i \in I} a_i \alpha_i, \ a_i \in K \\ &= \sum_{i \in I} (\sum_{j \in J} b_{ij} \beta_j) \alpha_i, \ b_{ij} \in F. \end{aligned}$$

Thus $\alpha_i\beta_j$ span E over F. We now check the linear independence.

$$\sum_{i,j} \lambda_{ij} \alpha_i \beta_j = 0 \Rightarrow \sum_i \lambda_{ij} \alpha_i = 0$$

for all j and consequently $\lambda_{ij} = 0$ for all i, j which concludes the proof.

2. It is enough to use the first part, with

$$[E : K] = |I|, \ [K : F] = |J|, \ [E : F] = |I||J|.$$

□

Example 5.5. Consider the field extension $\mathbb{Q}(\zeta_8)/\mathbb{Q}$ where ζ_8 is a primitive 8th root of unity. We have that

$$\zeta_8 = \frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}$$

and $\mathbb{Q}(\zeta_8)/\mathbb{Q}$ is the same field extension as $\mathbb{Q}(i, \sqrt{2})/\mathbb{Q}$. We have

$$[\mathbb{Q}(i, \sqrt{2}) : \mathbb{Q}] = [\mathbb{Q}(i, \sqrt{2}) : \mathbb{Q}(\sqrt{2})][\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2 \cdot 2 = 4.$$

Recall that an algebraic extension is a field extension where every element is algebraic. The result below describes families of algebraic extensions.

Theorem 5.5. *If E is a finite extension of F, then E is an algebraic extension of F.*

Proof. Let $\alpha \in E$ with degree $[E : F] = n$. Then $1, \alpha, \ldots, \alpha^n$ are $n+1$ elements while the dimension is n, so they must be linearly dependent, say

$$a_0 + a_1\alpha + \ldots + a_n\alpha^n = 0, \ a_i \in F.$$

Take $p(X) = a_0 + a_1X + \ldots + a_nX^n \in F[X]$, α is a root of $p(X)$ and by definition α is algebraic over F. □

Examples 5.6.

1. By definition, a number field is a finite extension of $\mathbb{Q}$. Thus a number field is an algebraic extension of $\mathbb{Q}$.
2. The converse is not true. There are infinite algebraic extensions, for example, the field of all algebraic numbers over the rationals is algebraic and of infinite degree.

5.2 Splitting fields and algebraic closures

For $\alpha \in E$, an extension of F, we have introduced above $F(\alpha)$ as the intersection of all the subfields of E containing F and α. This can be of course generalized if we pick $\alpha_1, \ldots, \alpha_k \in E$, and $F(\alpha_1, \ldots, \alpha_k)$ is the intersection of all the subfields of E containing F and $\alpha_1, \ldots, \alpha_k$.

Definition 5.5. If E is an extension of F and $f \in F[X]$, we say that f splits over E if f can be written as $\lambda(X - \alpha_1) \cdots (X - \alpha_k)$ for some $\alpha_1, \ldots, \alpha_k \in E$ and $\lambda \in F$.

Definition 5.6. If K is an extension of F and $f \in F[X]$, we say that K is a splitting field for f over F is f splits over K but not over any proper subfield of K containing F.

Example 5.7. Consider the polynomial $f(X) = X^3 - 2$ over $\mathbb{Q}$. Its roots are

$$\sqrt[3]{2}, \ \sqrt[3]{2}\left(-\frac{1}{2} + i\frac{1}{2}\sqrt{3}\right), \ \sqrt[3]{2}\left(-\frac{1}{2} - i\frac{1}{2}\sqrt{3}\right).$$

Alternatively, if ζ_3 denotes a primitive 3rd root of unity, we can write the roots as

$$\sqrt[3]{2}, \zeta_3\sqrt[3]{2}, \zeta_3^2\sqrt[3]{2}.$$

The polynomial f is irreducible (for example using Eisenstein's criterion). Since it is also monic, it is the minimal polynomial of $\sqrt[3]{2}$, and

$$[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3.$$

Now since $\sqrt[3]{2}$ and $i\sqrt{3}$ (or ζ_3) generate all the roots of f, the splitting field of f is

$$K = \mathbb{Q}(\sqrt[3]{2}, i\sqrt{3}) = \mathbb{Q}(\sqrt[3]{2}, \zeta_3).$$

We finish by computing the degree of K over $\mathbb{Q}$. Clearly $i\sqrt{3}$ cannot belong to $\mathbb{Q}(\sqrt[3]{2})$ which is a subfield of $\mathbb{R}$, thus $[\mathbb{Q}(\sqrt[3]{2}, i\sqrt{3}) : \mathbb{Q}(\sqrt[3]{2})]$ is at least 2. Since $i\sqrt{3}$ is a root of $X^2+3 \in \mathbb{Q}(\sqrt[3]{2})[X]$, this degree is exactly 2. By multiplicativity of the degrees, we get that

$$[K : \mathbb{Q}] = 6.$$

Using that ζ_3 is a root of $X^2 + X + 1$ stays irreducible over $\mathbb{Q}(\sqrt{2})$ gives the same result.

Equivalently, K is a splitting field for f over F if f splits over K and K is generated over F by the roots $\alpha_1, \ldots, \alpha_k$ of f, that is $K = F(\alpha_1, \ldots, \alpha_k)$.

If $f \in F[X]$ and f splits over the extension E of F, then E contains a unique splitting field for f, namely $F(\alpha_1, \ldots, \alpha_k)$.

Here is a result on the degree of splitting fields. Note that the above example shows that this bound is tight.

Proposition 5.6. *If $f \in F[X]$ and $\deg f = n$, then f has a splitting field K over F with $[K : F] \leq n!$.*

Proof. First we may assume that $n \geq 1$, for if $n = 0$, then f is constant, and we take $K = F$ with $[K : F] = 1$.

Thus f has at least one root α_1, and by Theorem 5.1, there is an extension E_1 of F containing α_1. Since $f(\alpha_1) = 0$, the minimal polynomial $m_1(X)$ of α_1 divides $f(X)$, that is $f(X) = m_1(X)f'(X)$ for some $f'(X)$, and since $\deg f = n$, $\deg m_1(X) \leq n$, implying that $F(\alpha_1)/F$ has degree at most n.

We may then further write $f(X) = (X - \alpha_1)^{r_1} g(X)$ where $g(\alpha_1) \neq 0$ and $\deg g \leq n - 1$. If g is constant, then $f(X)$ has no other root than α_1, and its splitting field is $F(\alpha_1)/F$ whose degree is at most n which is indeed smaller than $n!$.

Now if g is non-constant, we can iterate on g the reasoning we did on f. Namely, we have that g has degree at least 1, and thus it has at least one root α_2. Invoking again Theorem 5.1, there is an extension of $F(\alpha_1)$ containing α_2 and the extension $F(\alpha_1, \alpha_2)$ has degree at most $n-1$ over $F(\alpha_1)$ (corresponding to the case where $r_1 = 1$). Thus we have

$$\begin{aligned} [F(\alpha_1, \alpha_2) : F] &= [F(\alpha_1, \alpha_2) : F(\alpha_1)][F(\alpha_1) : F] \\ &\leq (n-1)n. \end{aligned}$$

We can now continue inductively to reach that if $\alpha_1, \ldots, \alpha_n$ are all the roots of f, then

$$[F(\alpha_1, \alpha_2, \ldots, \alpha_n) : F] \leq n!.$$

□

If $f \in F[X]$ and f splits over E, then we may take any root α of f and adjoin it to F to get the extension $F(\alpha)$. More precisely:

Theorem 5.7. *If α and β are roots of the irreducible polynomial $f \in F[X]$ in an extension E of F, then $F(\alpha)$ is isomorphic to $F(\beta)$.*

Proof. If f is not monic, start by dividing f by its leading coefficient, so that we can assume that f is monic. Since f is monic, irreducible and $f(\alpha) = f(\beta) = 0$, f is the minimal polynomial of α and β, say of degree n. Now if $a \in F(\alpha)$, then a can be uniquely written as

$$a = a_0 + a_1\alpha + \ldots + a_{n-1}\alpha^{n-1}.$$

The map

$$a_0 + a_1\alpha + \ldots + a_{n-1}\alpha^{n-1} \mapsto a_0 + a_1\beta + \ldots + a_{n-1}\beta^{n-1}$$

defines a field isomorphism between $F(\alpha)$ and $F(\beta)$. □

When discussing field isomorphisms, one may want to emphasize the base field.

Definition 5.7. If E and E' are extensions of F, and $\iota : E \to E'$ is an isomorphism, we say that ι is an F-isomorphism if ι fixes F, that is, if

$$\iota(a) = a, \ a \in F.$$

Given a polynomial $f \in F[X]$, we have discussed its splitting field, namely the smallest field over which f splits. If F is $\mathbb{Q}$, $\mathbb{R}$ or more generally $\mathbb{C}$, not only we can find a splitting field for each polynomial, but we know that there is a field C with the property that any polynomial in $\mathbb{C}[X]$ splits over C, namely $C = \mathbb{C}$ itself.

We now would like to express this property in general, without having to assume that F is $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$. Namely, for a general field F, we want an extension C of F such that any polynomial in $C[X]$ splits over C. We will later on add the requirement that this extension is algebraic.

Proposition 5.8. *If C is a field, the following conditions are equivalent.*

1. *Every non-constant polynomial $f \in C[X]$ has at least one root in C.*
2. *Every non-constant polynomial $f \in C[X]$ splits over C.*
3. *Every irreducible polynomial $f \in C[X]$ is linear.*

4. C has no proper algebraic extension.

Proof. We prove 1. $\Rightarrow$ 2. $\Rightarrow$ 3. $\Rightarrow$ 4. $\Rightarrow$ 1.

1. $\Rightarrow$ 2. Take $f \in C[X]$ a non-constant polynomial. Since f has at least one root, we write $f = (X - \alpha_1)g$ for g some polynomial in $C[X]$. If g is constant, we are done since f splits. If g is non-constant, then again by assumption it has one root and $g = (X - \alpha_2)h$ for some h. We conclude by repeating inductively.

2. $\Rightarrow$ 3. Take $f \in C[X]$ which is irreducible, thus non-constant. By assumption it is a product of linear factors. But f is irreducible, so there can be only one such factor.

3. $\Rightarrow$ 4. Let E be an algebraic extension of C. Take $\alpha \in E$ with minimal polynomial f over C. Then f is irreducible and of the form $X - \alpha \in C[X]$ by assumption. Thus $\alpha \in C$ and $E = C$.

4. $\Rightarrow$ 1. Let f be a non-constant polynomial in $C[X]$, with root α. We can adjoin α to C to obtain $C(\alpha)$. But by assumption, there is no proper algebraic extension of C, so $C(\alpha) = C$ and $\alpha \in C$. Thus f has at least one root in C and we are done.

□

Definition 5.8. A field C as described in the above equivalent properties is said to be algebraically closed.

Examples 5.8.

1. The field $\mathbb{R}$ is not algebraically closed, since $X^2 + 1 = 0$ has not root in $\mathbb{R}$.
2. No finite field $\mathbb{F}$ is algebraically closed, since if $a_1, \ldots, a_n$ are all the elements of F, then the polynomial $(X - a_1) \ldots (X - a_n) + 1$ has no zero in $\mathbb{F}$.
3. The field $\mathbb{C}$ is algebraically closed, this is the fundamental theorem of algebra.
4. The field of all algebraic numbers is algebraically closed. (We will not prove this here, but for a proof that algebraic numbers in a field extension indeed form a field, see Corollary 5.11 below.)

We can embed an arbitrary field F in an algebraically closed field as follows.

Definition 5.9. An extension C of F is called an algebraic closure if C is algebraic over F and C is algebraically closed.

Examples 5.9. To get examples of algebraic closures, we thus need to start with known algebraically closed fields.

1. The field $\mathbb{C}$ is the algebraic closure of $\mathbb{R}$.

2. The field of all algebraic numbers is the algebraic closure of $\mathbb{Q}$.

Note that C is minimal among algebraically closed extensions of F. Indeed, let us assume that there is an algebraically closed field K such that $C/K/F$. Let $\alpha \in C$ but $\alpha \notin K$ (it exists if we assume that $C \neq K$). Then α is algebraic over F, and consequently algebraic over K. But since $\alpha \notin K$, the minimal polynomial of α over K cannot contain the factor $X - \alpha$, which contradicts that K is an algebraically closed field.

We can prove the following theorems (we will omit the proof).

Theorem 5.9. *1. Every field F has an algebraic closure.*

2. Any two algebraic closures C and C' of F are F-isomorphic.

3. If E is an algebraic extension of F, C is an algebraic closure of F, and ι is an embedding of F into C. Then ι can be extended to an embedding of E into C.

Let us now prove the first transitivity property of field extensions. Several will follow later on in this chapter.

Proposition 5.10. *1. If E is generated over F by finitely many elements $\alpha_1, \ldots, \alpha_n$ algebraic over F, then E is a finite extension of F.*

2. **(Transitivity of algebraic extensions).** *If E is algebraic over K, and K is algebraic over F, then E is algebraic over F.*

Proof. 1. Set $E_0 = F$, $E_k = F(\alpha_1, \ldots, \alpha_k)$, $1 \leq k \leq n$, in particular $E_n = F(\alpha_1, \ldots, \alpha_n) = E$ by definition of E. Then $E_k = E_{k-1}(\alpha_k)$, where α_k is algebraic over F, and hence over E_{k-1}. Now $[E_k : E_{k-1}]$ is the degree of the minimal polynomial of α_k over E_{k-1}, which is finite. By multiplicativity of the degrees, we conclude that

$$[E : F] = \prod_{k=1}^{n} [E_k : E_{k-1}] < \infty.$$

2. Let $\alpha \in E$ with minimal polynomial

$$m(X) = b_0 + b_1 X + \ldots + b_{n-1} X^{n-1} + X^n$$

over K since by assumption α is algebraic over K. The coefficients b_i are in K and thus are algebraic over F. Set $L = F(b_0, b_1, \ldots, b_{n-1})$, by the first part, L is a finite extension of F. Therefore $m(X) \in L[X]$, α is algebraic over L, and $L(\alpha)$ is a finite extension of L. This gives us the following tower of field extensions:

$$L(\alpha)/L = F(b_0, b_1, \ldots, b_{n-1})/F.$$

By transitivity of the degrees, since $[L : F] < \infty$ and $[L(\alpha) : L] < \infty$, we get that $[L(\alpha) : F] < \infty$. We conclude since we know that all finite extensions are algebraic, and thus α is algebraic over F.

□

Corollary 5.11. *If E is an extension of F and A is the set of all elements in E that are algebraic over F, then A is a subfield of E.*

Proof. If $\alpha, \beta \in A$, then the sum, difference, product and quotient (if $\beta \neq 0$) of α and β belong to $F(\alpha, \beta)$, which is a finite extension of F by the first part of the above proposition. This is thus an algebraic extension since all finite extensions are, and thus $\alpha + \beta$, $\alpha - \beta$, $\alpha\beta$ and α/β are in A, proving that A is a field. □

5.3 Separability

If f is a polynomial in $F[X]$, we have seen above that we can construct a splitting field K for f over F, and K is such that all roots of f lie in it. We can thus study the multiplicity of the roots of f in K.

Definition 5.10. An irreducible polynomial $f \in F[X]$ is separable if f has no repeated roots in a splitting field. It is called inseparable otherwise. Note that if f is not necessarily irreducible, then we call f separable if each of its irreducible factors is separable.

For example $f(X) = (X-1)^2(X-2) \in \mathbb{Q}$ is separable, since its irreducible factors $X-1$ and $X-2$ are separable.

We start by computing a criterion to test if a polynomial has multiple roots.

Proposition 5.12. *Consider*

$$f(X) = a_0 + a_1 X + \cdots + a_n X^n \in F[X]$$

and its formal derivative

$$f'(X) = a_1 + 2a_2 X + \cdots + n a_n X^{n-1}.$$

Then f has a repeated root in a splitting field if and only if the degree of the greatest common divisor of f and f' is at least 1.

Proof. Let us assume that f has a repeated root in its splitting field, say α. Then we can write

$$f(X) = (X - \alpha)^r h(X)$$

where $r \geq 2$ since we consider a repeated root. Now we compute the derivative of f:

$$f'(X) = r(X-\alpha)^{r-1} h(X) + (X-\alpha)^r h'(X)$$

and since $r - 1 \geq 1$, we have that $(X - \alpha)$ is a factor of both f and f'.

Conversely, let us assume that the greatest common divisor g of f and f' has degree at least 1, and let α be a root of g (in a splitting field). By definition of g, $X - \alpha$ is then a factor of both f and f'. We are left to prove that α is a repeated root of f. Indeed, if it were not the case, then $f(X)$ would be of the form $f(X) = (X - \alpha)h(X)$ where $h(\alpha) \neq 0$ and by computing the derivative, we would get (put $r = 1$ in the above expression for f') $f'(\alpha) = h(\alpha) \neq 0$ which contradicts the fact that $X - \alpha$ is a factor of f'. □

As a corollary of this result, we can exhibit two classes of separable polynomials.

Corollary 5.13. *1. Over a field of characteristic zero, every polynomial is separable.*

2. Over a field F of prime characteristic p, an irreducible polynomial f is inseparable if and only if f' is the zero polynomial (equivalently f is in $F[X^p]$).

Proof. 1. Without loss of generality, consider f an irreducible polynomial in $F[X]$, where F is of characteristic zero. If f is a polynomial of degree n, then its derivative f' is of degree less than n, and it cannot possibly be the zero polynomial. Since f is irreducible, the greatest common divisor of f and f' is either 1 or f, but it cannot be f since f' is of smaller degree. Thus it is 1, and f is separable by the above proposition.

2. We now consider the case where F is of characteristic p. As above, we take f an irreducible polynomial of degree n in $F[X]$ and compute its derivative f'. If f' is non-zero, we can use the same argument. But f' could also be zero, in which case the greatest common divisor of f and f' is actually f, and by the above proposition, f has a multiple root and is then not separable. That $f' = 0$ means that $f \in F[X^p]$ since we work in characteristic p.

□

Example 5.10. Polynomials over $\mathbb{R}[X]$ and $\mathbb{Q}[X]$ are separable.

Another class of separable polynomials are polynomials over finite fields, but this asks a little bit more work.

Lemma 5.14. *Let F be a finite field of characteristic p. Consider the map*

$$f : F \to F, \ f(\alpha) = \alpha^p.$$

Then f is an automorphism (called the Frobenius Automorphism). In particular, we have for all $\alpha \in F$ that

$$\alpha = \beta^p$$

for some $\beta \in F$.

Proof. We have that f is a ring automorphism since

$$\begin{aligned} f(1) &= 1 \\ f(\alpha+\beta) &= (\alpha+\beta)^p = \alpha^p + \beta^p = f(\alpha) + f(\beta) \\ f(\alpha\beta) &= (\alpha\beta)^p = \alpha^p\beta^p = f(\alpha)f(\beta). \end{aligned}$$

The second set of equalities uses the binomial expansion modulo p. Now f is a monomorphism since F is a field, and an injective map from a finite set to itself is necessarily surjective. □

Proposition 5.15. *Every polynomial is separable over a finite field F (of prime characteristic).*

Proof. Suppose that f is an irreducible polynomial which, by contradiction, has multiple roots in a splitting field. Using the criterion of the corollary, $f(X)$ must be in $F[X^p]$, namely

$$f(X) = a_0 + a_1 X^p + \cdots + a_n X^{np},\ a_i \in F.$$

Using the bijectivity of the Frobenius automorphism, we can write $a_i = b_i^p$, yielding

$$(b_0 + b_1 X + \cdots + b_n X^n)^p = b_0^p + b_1^p X^p + \cdots + b_n^p X^{np} = f(X)$$

which contradicts the irreducibility of f. □

Definition 5.11. If E is an extension of F and $\alpha \in E$, then α is said to be separable over F if α is algebraic over F and its minimal polynomial $\mu_{\alpha,F}$ is a separable polynomial. If every element of E is separable over F, we say that E is a separable extension of F or that E/F is separable.

Examples 5.11.

1. Typical examples of separable extensions are finite fields and number fields.

2. If F is a field with algebraic closure C, then C contains a smallest field containing all finite separable extensions of F, called the separable closure of F. It is a separable extension of F.

Here is a first result on how separability behaves in a tower of extensions.

Lemma 5.16. *If $E/K/F$ and E is separable over F, then K is separable over F and E is separable over K.*

Proof. **K/F is separable.** Since K is a subfield of E, every element $\beta \in K$ belongs to E, and every element of E is separable over F by assumption.

E/K is separable. Take $\alpha \in E$. Since E is separable over F, it is in particular algebraic over F and we may consider the minimal polynomial $\mu_{\alpha,F}$ of α over F. Denote by $\mu_{\alpha,K}$ the minimal polynomial of α over K, we have

$$\mu_{\alpha,K} \mid \mu_{\alpha,F}.$$

Since $\mu_{\alpha,F}$ has no repeated root, neither has $\mu_{\alpha,K}$, and E/K is separable. □

The converse is also true, and gives the transitivity of separable extensions: If K/F and E/K are separable, then E/F is separable.

It is less easy to construct inseparable extensions, but here is a classical example.

Example 5.12. Let $\mathbb{F}_p$ denote the finite field of integers modulo p. Consider the field $F = \mathbb{F}_p(t)$ of rational functions in t with coefficients in the finite field with p elements $\mathbb{F}_p$. We get a field extension of E/F by adjoining to F a root of the polynomial $X^p - t$ (one has to check that $X^p - t$ is irreducible over $\mathbb{F}_p[t]$). The extension E/F is inseparable since

$$X^p - t = X^p - (\sqrt[p]{t})^p = (X - \sqrt[p]{t})^p,$$

which has multiple roots.

Let E/F be a separable extension of F and let C be an algebraic closure of E. We next count the number of embeddings of E in C that fix F, that is, the number of F-monomorphisms of E into C. We start with a lemma.

Lemma 5.17. *Let $\sigma : E \to E$ be an F-monomorphism and assume that $f \in F[X]$ splits over E. Then σ permutes the roots of f, namely, if α is a root of f in E then so is $\sigma(\alpha)$.*

Proof. Write $f(X)$ as

$$f(X) = b_0 + b_1 X + \cdots + b_n X^n,\ b_i \in F.$$

If α is a root of f in E, then

$$f(\alpha) = b_0 + b_1\alpha + \cdots + b_n\alpha^n = 0.$$

Apply σ to the above equation, and use that σ is a field homomorphism that fixes F to get

$$b_0 + b_1\sigma(\alpha) + \cdots + b_n\sigma(\alpha)^n = 0,$$

showing that $\sigma(\alpha)$ is a root. □

Theorem 5.18. *Let E/F be a finite separable extension of degree n, and let σ be an embedding of F into an algebraic closure C. Then σ extends to exactly n embeddings of E in C. Namely, there are exactly n embeddings τ of E into C, such that the restriction $\tau|_F$ of τ to F coincides with σ. In particular, taking σ to be the identity on F, there are exactly n F-monomorphisms of E into C.*

Proof. We do a proof by induction. If $n = 1$, then $E = F$ and σ extends to exactly 1 embedding, namely itself.

We now assume that $n > 1$ and choose $\alpha \in E$, $\alpha \notin F$. Let $f = \mu_{\alpha,F}$ be the minimal polynomial of α over F of degree say r. It is irreducible and separable (E/F is separable by assumption). In order to use the induction hypothesis, we need to split the field extension E/F, which we do by considering the field extension $F(\alpha)$, which satisfies

$$E/F(\alpha)/F,\ [E : F(\alpha)] = n/r,\ [F(\alpha) : F] = r.$$

We first take care of the extension $F(\alpha)/F$. Let σ be an embedding of F into C, and define the polynomial $g = \sigma(f)$, where σ is applied on all the coefficients

of f. The polynomial g inherits the property of being irreducible and separable from f. Let β denotes a root of g. We can thus define a unique isomorphism

$$F(\alpha) \to (\sigma(F))(\beta),\ b_0 + b_1\alpha + \ldots + b_r\alpha^r \mapsto \sigma(b_0) + \sigma(b_1)\beta + \ldots + \sigma(b_r)\beta^r$$

and restricted to F it indeed coincides with σ. This isomorphism is defined by the choice of β, and there are exactly r choices for it, corresponding to the r roots of g (note that this is here that the separability of g is crucial). For each of these r isomorphisms, using the induction hypothesis on $[E : F(\alpha)] = n/r < n$, we can extend them to exactly n/r embeddings of E into C. This gives us a total of $n/r \cdot r$ distinct embeddings of E into C extending σ. We conclude by noting that we cannot have more than n such embeddings. □

We conclude by giving a nice description of finite separable field extensions.

Theorem 5.19. (Theorem of the Primitive Element). *If E/F is a finite separable extension, then*

$$E = F(\gamma)$$

for some $\gamma \in E$. We say that γ is a primitive element of E over F.

Proof. Since we have not studied finite fields yet, let us assume that F is an infinite field. (If you have already studied finite fields, then you know we can take γ to be any generator of the cyclic group $E^\times$).

We proceed by induction on the degree n of the extension E/F. If $n = 1$, then $E = F$ and we can take any element for α.

Let us thus assume $n > 1$, the assumption true up to $n - 1$, and say the degree of E/F is n. Choose $\alpha \in E$ but not in F. We now look at the field extension $E/F(\alpha)$. By induction hypothesis, there is a primitive element β such that

$$E = F(\alpha, \beta).$$

We are now going to prove that there exists a $c \in F$ such that

$$E = F(\alpha + c\beta),$$

that is

$$\gamma = \alpha + c\beta$$

will be the primitive element. We will show that it is enough to take $c \notin S$, where S is a finite subset of F defined as follows: let f be the minimal polynomial of α over F, and let g be the minimal polynomial of β over F, the exceptional set S consists of all $c \in F$ such that

$$c = \frac{\alpha' - \alpha}{\beta - \beta'}$$

for α' a root of f and β' a conjugate of β (we extend $F(\alpha, \beta)$ to a field L in which f and g both split to be able to speak of all their roots).

To show that γ is primitive for $c \notin S$, it is enough to prove that $F(\alpha + c\beta)$ contains β and $\alpha = \gamma - c\beta$ (clearly the reverse inclusion holds: $F(\alpha + c\beta) \subseteq F(\alpha, \beta)$). To this end, it is enough to show that the minimal polynomial of β over $F(\gamma)$ cannot have degree greater or equal to 2, implying that β is in $F(\gamma)$.

Note first that if we take the polynomial $h(X)$ defined by

$$h(X) = f(\gamma - cX) \in F(\gamma)[X]$$

and evaluate it in β, we get

$$h(\beta) = f(\gamma - c\beta) = f(\alpha + c\beta - c\beta) = 0.$$

Thus β is a root of h and the minimal polynomial of β over $F(\gamma)$ divides both g and h, so we are done if we show that the greatest common divisor of g and h in $F(\gamma)[X]$ cannot have degree greater or equal to 2.

Suppose the greatest common divisor does have degree≥ 2. Then g and h have as common root in L not only β, but also $\beta' \neq \beta$ in L. This is where we use the separability of g, since otherwise β could be a root with multiplicity 2. Then

$$f(\gamma - c\beta') = 0 \Rightarrow \gamma - c\beta' = \alpha'$$

for some root α' of f, which can be rewritten as

$$\alpha + c\beta - c\beta' = 0 \Rightarrow c = \frac{\alpha' - \alpha}{\beta - \beta'}$$

which is exactly what was ruled out by choosing $c \notin S$. □

Definition 5.12. A simple extension is a field extension which is generated by the adjunction of a single element.

Thus the primitive element Theorem above provides a characterization of the finite extensions which are simple.

Example 5.13. Number fields are simple extensions.

5.4 Normality

So far, we have considered two properties of field extensions (both of them being transitive): being algebraic and separable. We now introduce a third property, which is not transitive, the one of being normal.

Definition 5.13. An algebraic extension E/F is normal if every irreducible polynomial over F that has at least one root in E splits over E. If we call the other roots of this polynomial the conjugates of α, we can rephrase the definition by saying that if $\alpha \in E$, then all conjugates of α over F are in E.

Note that this definition assumes that we start with an algebraic extension.

Example 5.14. Consider the field extension $E = \mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$. The roots of the irreducible polynomial $f(X) = X^3 - 2$ are

$$\sqrt[3]{2}, \zeta_3\sqrt[3]{2}, \zeta_3^2\sqrt[3]{2},$$

where ζ_3 is a primitive 3rd root of unity (for example $\zeta_3 = e^{2\pi i/3}$). Thus E is not a normal extension.

We can give another characterization in terms of monomorphisms of E.

Theorem 5.20. *The finite extension E/F is normal if and only if every F-monomorphism of E into an algebraic closure C is actually an F-automorphism of E. (Finite could be replaced by algebraic, which we will not prove).*

Proof. If E/F is normal, then an F-monomorphism of E into C must map each element of E to one of its conjugates (as is the case in the proof of Lemma 5.17). Thus $\tau(E) \subseteq E$, but $\tau(E)$ is an isomorphic copy of E and thus has the same degree as E and $E = \tau(E)$, showing that τ is indeed an F-automorphism of E.

Conversely, consider $\alpha \in E$ and let β be a conjugate of α over F. There exists an F-monomorphism of E into C that carries α to β (the construction is given in the proof of Theorem 5.18). If all such embeddings are F-automorphisms of E, that means β must be in E, and we conclude that E/F is normal. □

Here is another characterization of normal extensions in terms of splitting fields.

Theorem 5.21. *The finite extension E/F is normal if and only if E is a splitting field for some polynomial f in $F[X]$.*

Proof. Let E/F be a finite normal extension of degree n, and let $\alpha_1, \ldots, \alpha_n$ be a basis for E over F. Consider for each α_i its minimal polynomial f_i over F. By definition of normal extension, since f_i has a root in E, then f_i splits over E, and so does the polynomial

$$f = f_1 \cdots f_n.$$

To prove that E is a splitting field, we are left to prove it is the smallest field over which f splits. This is here that we understand why we take such an f. If f were to split over a subfield K, that is K such that

$$F \subset K \subset E$$

then each $\alpha_i \in K$, and $K = E$ (this is a conclusion we cannot reach if we take for f only one f_i or a subset of them). This proves that E is a splitting field for f over F.

Conversely, let E be a splitting field for some f over F, whose roots are denoted by $\alpha_1, \ldots, \alpha_n$. Let τ be an F-monomorphism of E into an algebraic closure, that is τ takes each α_i into another root of f.

Since E is a splitting field for f, we have

$$F(\alpha_1, \ldots, \alpha_n) = E$$

and $\tau(E) \subset E$. Thus since E and $\tau(E)$ have same dimension, we get that

$$\tau(E) = E$$

and τ is actually an automorphism of E, and by the above theorem, we conclude the E/F is normal. □

As a corollary, we see how a subextension inherits the property of normality.

Corollary 5.22. *Let $E/K/F$ be a finite extension ($[E : F] < \infty$). If E/F is normal, so is E/K.*

Proof. Since E/F is normal, E is a splitting field for some polynomial $f \in F[X]$, that is E is generated over F by the roots of f. Since $f \in F[X] \subset K[X]$, f can also be seen as a polynomial in $K[X]$ and E is generated over K by the roots of f, and again by the above theorem, E/K is normal. □

There is no reason for an arbitrary field extension E/F to be normal. However, if E/F is finite (or more generally algebraic) one can always embed it in normal extension.

Definition 5.14. Let E/F be an algebraic extension. The normal closure of E/F is an extension field N of E such that N/E is normal and N is minimal with this property.

If E/F is finite, we can see it as follows: E is finitely generated over F, so it can be written as $E = F(\alpha_1, \ldots, \alpha_n)$. Let now K be a normal extension of F that contains E:

$$K/E/F.$$

Since K is normal, it must contain not only all the α_i but also all their conjugates. Let f_i be the minimal polynomial of α_i, $i = 1, \ldots, n$. Then we can rephrase the last statement and say that K must contain all the roots of f_i, $i = 1, \ldots, n$. Consider the polynomial

$$f = f_1 \cdots f_n.$$

Then K must contain the splitting field N for f over F. But N/F is normal, so N must be the smallest normal extension of F that contains E. Thus N is a normal closure of E over F.

The main definitions and results of this chapter are

- **(3.1).** Definitions of: field extension, minimal polynomial, degree of a field extension, field homomorphism, algebraic, transcendental. That the degree is multiplicative.
- **(3.2).** Definitions of: to split, splitting field, algebraically closed, algebraic closure. Transitivity of algebraic extensions.
- **(3.3).** Definition of separability, typical separable extensions, separability in extension towers, number of embeddings into an algebraic closure, primitive element Theorem.
- **(3.4).** Definition of normality, two equivalent characterizations of normal extensions.

Chapter 6

Exercises for Field Theory

Exercises marked by (*) are considered difficult.

6.1 Field extension and minimal polynomial

Exercise 88. 1. For which of the following $p(X)$ do there exist extensions $K(\alpha)$ of K for which α has minimal polynomial $p(X)$?

- $p(X) = X^2 - 4$, $K = \mathbb{R}$.
- $p(X) = X^2 + 1$, $K = \mathbb{Z}_5$ (integers modulo 5).
- $p(X) = X^3 + 2$, $K = \mathbb{Q}$.

In the case where you obtain a field extension, what is the degree of the extension?

2. Find an irreducible polynomial of degree 2 over the integers modulo 2. Use it to construct a field with 4 elements. Describe the obtained field.

Answer.

1. $p(X) = X^2 - 4 = (X - 2)(X + 2)$, it is not irreducible so it cannot be a minimal polynomial. Then $p(X) = X^2 + 1 = (X - 2)(X + 2)$ modulo 5, so it is not irreducible, and cannot be a minimal polynomial. Finally $X^3 + 2$ is irreducible, monic, we obtain the field extension $\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$, it is of degree 3.

2. Take the polynomial $X^2 + X + 1$, it has no root modulo 2 and is thus irreducible. We can construct a field using the generic construction that we know. The field $\mathbb{Z}_2[X]/(X^2+X+1)$ contains a root α of the polynomial, it is a field containing 4 elements. Indeed, it is of degree 2 (degree of the minimal polynomial), and a basis is given by $\{1, \alpha\}$, thus every element

can be written as $a + b\alpha$, $a, b \in \mathbb{Z}_2$. That makes 4 possible elements, and the field is described by

$$\mathbb{Z}_2[X]/(X^2 + X + 1) \simeq \{a + b\alpha, a, b \in \mathbb{Z}_2\}.$$

Exercise 89. 1. Show that $\mathbb{C}/\mathbb{R}$ is an algebraic extension.

2. Compute the degree of the following extensions: $\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$, $\mathbb{Q}(\sqrt{3}+\sqrt{2})/\mathbb{Q}$.

3. Let $E = \mathbb{Q}(\sqrt{2})$ and $F = \mathbb{Q}(i\sqrt{2})$. Show that -1 is a sum of 2 squares in F. Deduce that E and F are not isomorphic.

Answer.

1. $\mathbb{C}/\mathbb{R}$ is an extension of degree 2 (a $\mathbb{R}$-basis is $\{1, i\}$), it is thus finite, thus algebraic.

2. $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$ (a $\mathbb{Q}$-basis is $\{1, \sqrt[3]{2}, (\sqrt[3]{2})^2\}$), $[\mathbb{Q}(\sqrt{3} + \sqrt{2}) : \mathbb{Q}] = 4$ (a $\mathbb{Q}$-basis is $\{1, \sqrt{2}, \sqrt{3}, \sqrt{6}\}$, because $\mathbb{Q}(\sqrt{3} + \sqrt{2}) = \mathbb{Q}(\sqrt{3}, \sqrt{2})$).

3. In F, we have that $(i\sqrt{2})^2 + 1^2 = -1$. Since both fields have the same degree and knowing that a field homomorphism is always injective, we try to build a ring homomorphism f from F to E. Thus

$$f((i\sqrt{2})^2 + 1^2) = f(-1) \Rightarrow f((i\sqrt{2})^2) + f(1) = -f(1)$$

since f is a ring homomorphism, furthermore, it must send $f(1)$ to 1, thus we must have

$$f((i\sqrt{2}))^2 = -2$$

that is there must be an element of E whose square is negative which is not possible.

Exercise 90. Consider the extension $\mathbb{C}/\mathbb{R}$. What are all the $\mathbb{R}$-automorphisms of $\mathbb{C}$? Justify your answer.

Answer. Write an element $x \in \mathbb{C}$ as $x = a + ib$, $a, b \in \mathbb{R}$, and let σ be an $\mathbb{R}$-automorphisms. Thus

$$\sigma(x) = \sigma(a) + \sigma(i)\sigma(b) = a + \sigma(i)b$$

using for the first equality the property of ring homomorphism, and for the second one that σ fixes $\mathbb{R}$. Thus $\sigma(x)$ is determined by $\sigma(i)$. Since $i^2 = -1$, we have that $\sigma(i^2) = \sigma(-1)$, that is

$$\sigma(i)^2 + 1 = 0.$$

Thus either $\sigma(i) = i$ or $\sigma(i) = -i$, which are the only two possible $\mathbb{R}$-automorphims of $\mathbb{C}$.

Exercise 91. Prove that if $[K(u) : K]$ is odd, then $K(u) = K(u^2)$.

Answer. We first notice that $K(u^2) \subset K(u)$, thus

$$[K(u) : K] = [K(u) : K(u^2)][K(u^2) : K].$$

Since u is a root of the polynomial $X^2 - u^2$ in $K(u^2)[X]$, we have that $[K(u) : K(u^2)] \leq 2$, and it cannot be 2 because $[K(u) : K]$ is odd, thus $[K(u) : K(u^2)] = 1$ and the conclusion follows.

6.2 Splitting fields and algebraic closures

Exercise 92. What is the splitting field of the following polynomials?

1. $f(x) = (x^2 - 3)(x^3 + 1) \in \mathbb{Q}(x)$.
2. $f(x) = x^2 + x + 1 \in \mathbb{F}_2[x]$.

Answer.

1. We have that $f(X) = (x-\sqrt{3})(x+\sqrt{3})(x-1)(x^2+x+1)$, thus the splitting field of f must contain $\sqrt{3}$ and ζ_3, the primitive third root of unity. This then must be $\mathbb{Q}(i, \sqrt{3})$.
2. We have that $x^2 + x + 1$ is irreducible over $\mathbb{F}_2$, we can construct $\mathbb{F}_4$ as $\mathbb{F}_2[x]/(f(x))$, that is $\mathbb{F}_4 \simeq \mathbb{F}_2(w)$ where $w^2 + w + 1 = 0$. Thus the splitting field of f is $\mathbb{F}_4$.

6.3 Separability

6.4 Normality

Exercise 93. Show that $\mathbb{Q}(\sqrt[3]{5})/\mathbb{Q}$ is not normal.

Answer. The roots of $x^3 - 5$ are $\sqrt[3]{5}, \zeta_3\sqrt[3]{5}, \zeta_3^2\sqrt[3]{5}$, where ζ_3 denote a primite 3rd root of unity. Since $\mathbb{Q}(\sqrt[3]{5})/\mathbb{Q}$ is totally real, it cannot contain the complex roots.

Exercise 94. Are the following claims true or false? Justify your answer.

1. Every polynomial splits over some field.
2. The polynomial $x^3 + 5$ is separable over $\mathbb{F}_7$.
3. Every finite extension is normal.
4. Every separable extension is normal.
5. Every finite normal extension is a splitting field for some polynomial.

6. A reducible polynomial cannot be separable.

Answer.

1. This is true, for every root of the polynomial, there is a field that will contain this root, so that we can build a field extension containing all the roots (if the polynomial has coefficients in $\mathbb{R}$, then one can use $\mathbb{C}$, but $\mathbb{C}$ will not work if the polynomial has coefficients in a finite field).
2. True since $\mathbb{F}_7$ is a finite field.
3. False, $\mathbb{Q}(\sqrt[3]{5})/\mathbb{Q}$ is finite but not normal.
4. False, $\mathbb{Q}(\sqrt[3]{5})/\mathbb{Q}$ is separable (because $\mathbb{Q}$ is of characteristic zero) but not normal.
5. True, we proved this.
6. False, when a polynomial is reducible, the definition of separability applies on its irreducible factors, which may or may not be separable.

Exercise 95. True/False.

Q1. Every field has non-trivial extensions.

Q2. Every field has non-trivial algebraic extensions.

Q3. Extensions of the same degree are isomorphic.

Q4. Every algebraic extension is finite.

Q5. Every algebraic extension of $\mathbb{Q}$ is finite.

Q6. Every extension of a finite field is finite.

Q7. The polynomial $X^3 + 5$ is separable over Z_7 (= integers modulo 7).

Q8. Every finite extension is normal.

Q9. Every separable extension is normal.

Q10. Every K-monomorphism is a K-automorphism.

Q11. Every extension of a field of characteristic 0 is normal.

Answer.

Q1. That's true! We are not speaking of algebraic extensions necessarily. Even if you take $\mathbb{C}$, you can for example get function fields over $\mathbb{C}$ by adding an indeterminate.

Q2. We know that one of the characterizations of algebraically closed fields is that they have no non-trivial algebraic extensions! So that is one counter example.

Q3. False! It's the other way round: if two extensions are isomorphic, then they have the same degree.

Q4. False, it is the other way round! If an extension is finite, it is algebraic. If it is algebraic it does not have to be finite (take an algebraic closure).

Q5. Still false. Taking $\mathbb{Q}$ as the base field does not change anything to the problem. The same counter example as in the previous question holds: you can take an algebraic closure of $\mathbb{Q}$, it is algebraic and infinite.

Q6. This is still false! You can build a function field as a counter example.

Q7. It is true. We have proved this result in general for fields of characteristic zero and finite fields.

Q8. It's false! There is no connection between both concepts. For example, we know that $\mathbb{Q}(\alpha)$ with $\alpha^3 = 2$ is finite and not normal.

Q9. It is false! There is no connection, you can take as above $\mathbb{Q}(\alpha)$ with $\alpha^3 = 2$, it is separable and not normal.

Q10. This is false! For a counter example, take any extension which is not normal. You'll find a K-monomorphism which is not a K-automorphism.

Q11. This is wrong! Imagine this were true, then all number fields would be normal, this is surely not the case!!

Chapter 7

Galois Theory

Galois theory is named after the French mathematician Evariste Galois.

Galois was born in 1811, and had what could be called the life of a misunderstood genius. At the age of 15, he was already reading material written for professional mathematicians. He took the examination to the "Ecole Polytechnique" to study mathematics but failed and entered the "Ecole Normale" in 1828. He wrote his first paper at the age of 18. He tried to advertise his work, and sent his discoveries in the theory of polynomial equations to the Academy of Sciences, where Cauchy rejected his memoir. He did not get discouraged, and in 1830, he wrote again his researches to the Academy of Sciences, where this time Fourier got his manuscript. However, Fourier died before reading it.

A year later, he made a third attempt, and sent to the Academy of Sciences a memoir called "On the conditions of solvability of equations by radicals". Poisson was a referee, and he answered several months later, declaring the paper incomprehensible.

In 1832, he got involved in a love affair, but got rejected due to a rival, who challenged him to a duel. The night before the duel, he wrote a letter to his friend, where he reported his mathematical discoveries. He died during the duel with pistols in 1832.

It is after his death that his friend insisted to have his letter published, which was finally done by the mathematician Chevalier.

7.1 Galois group and fixed fields

Definition 7.1. If E/F is normal and separable, it is said to be a Galois extension, or alternatively, we say that E is Galois over F.

Take E/F a Galois extension of degree n. Since it is separable of degree n, we know that there are exactly n F-monomorphisms of E into an algebraic closure C. But E/F being also normal, every F-automorphism into C is actually and

Figure 7.1: Evariste Galois (1811-1832)

F-automorphism of E. Thus there are exactly $n = [E : F]$ F-automorphisms of E.

We can define the notion of a Galois group for an arbitrary field extension.

Definition 7.2. If E/F is a field extension, the Galois group of E/F, denoted by $\mathrm{Gal}(E/F)$, is the set of F-automorphisms of E. It forms a group under the composition of functions.

Example 7.1. If $E = \mathbb{Q}(\sqrt[3]{2})$, then $\mathrm{Gal}(E/\mathbb{Q}) = \{1\}$, that is the identity on E.

The above example illustrates the fact that though one can always define a Galois group, we need the extension to be actually Galois to say that the order of the Galois group is actually the degree of the field extension.

Definition 7.3. Let $G = \mathrm{Gal}(E/F)$ be the Galois group of the extension E/F. If H is a subgroup of G, the fixed field of H is the set of elements fixed by every automorphism in H, that is

$$\mathcal{F}(H) = \{x \in E,\ \sigma(x) = x \text{ for all } \sigma \in H\}.$$

Vice-versa, if K is an intermediate field, define

$$\mathcal{G}(K) = \mathrm{Gal}(E/K) = \{\sigma \in G,\ \sigma(x) = x \text{ for all } x \in K\}.$$

It is the group fixing K.

Galois theory has much to do with studying the relations between fixed fields and fixing groups.

Proposition 7.1. *Let E/F be a finite Galois extension with Galois group $G = \mathrm{Gal}(E/F)$. Then*

1. *The fixed field of G is F.*

2. If H is a proper subgroup of G, then the fixed field $\mathcal{F}(H)$ of H properly contains F.

Proof. 1. Let F_0 be the fixed field of G (and we have the field extensions $E/F_0/F$). We want to prove that $F_0 = F$.

We first note that if σ is an F-automorphism of E (that is σ is in G), then by definition of F_0, σ fixes everything in F_0, meaning that σ is an F_0-automorphism. Thus the F-automorphisms in the group G coincide with the F_0-automorphisms in the group G.

Now we further have that E/F_0 is Galois: indeed, we have $E/F_0/F$ with E/F Galois thus normal and separable, and E/F_0 inherits both properties.

We now look at the degrees of the extensions considered:

$$|\mathrm{Gal}(E/F_0)| = [E : F_0],\ |\mathrm{Gal}(E/F)| = [E : F],$$

since both are Galois. Furthermore by the first remark, the number of $F-$ and F_0- automorphisms in G coincide:

$$|\mathrm{Gal}(E/F_0)| = |\mathrm{Gal}(E/F)|$$

showing that

$$[E : F_0] = [E : F]$$

and by multiplicativity of the degrees

$$[E : F] = [E : F_0][F_0 : F] \Rightarrow [F_0 : F] = 1$$

and $F = F_0$.

2. In order to prove that $F \subsetneq \mathcal{F}(H)$, let us assume by contradiction that $F = \mathcal{F}(H)$.

Since we consider a finite Galois extension, we can invoke the Theorem of the Primitive Element and claim that

$$E = F(\alpha),\ \alpha \in E. \tag{7.1}$$

Consider the polynomial

$$f(X) = \prod_{\sigma \in H} (X - \sigma(\alpha)) \in E[X].$$

It is a priori in $E[X]$, but we will prove now that it is actually in $F[X]$. Since by contradiction we are assuming that $F = \mathcal{F}(H)$, it is enough to proof that $f(X)$ is fixed by H. Indeed, take $\tau \in H$, then

$$\prod_{\sigma \in H} (X - \tau\sigma(\alpha)) = \prod_{\sigma \in H} (X - \sigma(\alpha))$$

since $\tau\sigma$ ranges over all H as does σ.

Thus $f(X) \in F[X]$ and $f(\alpha) = 0$ (σ must be the identity once while ranging through H). Now on the one hand, we have

$$\deg f = |H| < |G| = [E : F]$$

since we assume that H is proper and E/F is Galois. On the other hand,

$$\deg f \geq [F(\alpha) : F] = [E : F]$$

since f is a multiple of the minimal polynomial of α over F (equality holds if f is the minimal polynomial of α over F), and $E = F(\alpha)$ by (7.1). We cannot possibly have $\deg f < [E : F]$ and $\deg f \geq [E : F]$ at the same time, which is a contradiction and concludes the proof.

□

7.2 The fundamental Theorem of Galois theory

The most significant discovery of Galois is that (surely not in these terms!) under some hypotheses, there is a one-to-one correspondence between

1. subgroups of the Galois group $\mathrm{Gal}(E/F)$
2. subfields M of E such that $F \subseteq M$.

The correspondence goes as follows:

- To each intermediate subfield M, associate the group $\mathrm{Gal}(E/M)$ of all M-automorphisms of E:

$$\begin{array}{rcl} \mathcal{G} = \mathrm{Gal} : \{\text{intermediate fields}\} & \to & \{\text{subgroups of } \mathrm{Gal}(E/F)\} \\ M & \mapsto & \mathcal{G}(M) = \mathrm{Gal}(E/M). \end{array}$$

- To each subgroup H of $\mathrm{Gal}(E/F)$, associate the fixed subfield $\mathcal{F}(H)$:

$$\begin{array}{rcl} \mathcal{F} : \{\text{subgroups of } \mathrm{Gal}(E/F)\} & \to & \{\text{intermediate fields}\} \\ H & \mapsto & \mathcal{F}(H). \end{array}$$

We will prove below that, under the right hypotheses, we actually have a bijection (namely $\mathcal{G}$ is the inverse of $\mathcal{F}$). Let us start with an example.

Example 7.2. Consider the field extension $E = \mathbb{Q}(i, \sqrt{5})/\mathbb{Q}$. It has four $\mathbb{Q}$-automorphisms, given by (it is enough to describe their actions on i and $\sqrt{5}$):

$$\begin{array}{lll} \sigma_1 : & i \mapsto i, & \sqrt{5} \mapsto \sqrt{5} \\ \sigma_2 : & i \mapsto -i, & \sqrt{5} \mapsto \sqrt{5} \\ \sigma_3 : & i \mapsto i, & \sqrt{5} \mapsto -\sqrt{5} \\ \sigma_4 : & i \mapsto -i, & \sqrt{5} \mapsto -\sqrt{5} \end{array}$$

thus

$$\mathrm{Gal}(E/\mathbb{Q}) = \{\sigma_1, \sigma_2, \sigma_3, \sigma_4\}.$$

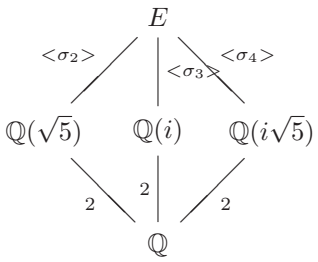
The proper subgroups of $\mathrm{Gal}(E/\mathbb{Q})$ are

$$\{\sigma_1\}, \{\sigma_1, \sigma_2\}, \{\sigma_1, \sigma_3\}, \{\sigma_1, \sigma_4\}$$

and their corresponding subfields are

$$E,\ \mathbb{Q}(\sqrt{5}),\ \mathbb{Q}(i),\ \mathbb{Q}(i\sqrt{5}).$$

We thus get the following diagram:



Theorem 7.2. *Let E/F be a finite Galois extension with Galois group G.*

1. *The map $\mathcal{F}$ is a bijection from subgroups to intermediate fields, with inverse $\mathcal{G}$.*
2. *Consider the intermediate field $K = \mathcal{F}(H)$ which is fixed by H, and $\sigma \in G$. Then the intermediate field*

$$\sigma K = \{\sigma(x),\ x \in K\}$$

is fixed by $\sigma H \sigma^{-1}$, namely $\sigma K = \mathcal{F}(\sigma H \sigma^{-1})$.

Proof. 1. We first consider the composition of maps

$$H \to \mathcal{F}(H) \to \mathcal{G}\mathcal{F}(H).$$

We need to prove that $\mathcal{G}\mathcal{F}(H) = H$. Take σ in H, then σ fixes $\mathcal{F}(H)$ by definition and $\sigma \in \mathrm{Gal}(E/\mathcal{F}(H)) = \mathcal{G}(\mathcal{F}(H))$, showing that

$$H \subseteq \mathcal{G}\mathcal{F}(H).$$

To prove equality, we need to rule out the strict inclusion. If H were a proper subgroup of $\mathcal{G}(\mathcal{F}(H))$, by the above proposition the fixed field $\mathcal{F}(H)$ of H should properly contain the fixed field of $\mathcal{G}\mathcal{F}(H)$ which is $\mathcal{F}(H)$ itself, a contradiction, showing that

$$H = \mathcal{G}\mathcal{F}(H).$$

Now consider the reverse composition of maps

$$K \to \mathcal{G}(K) \to \mathcal{F}\mathcal{G}(K).$$

This time we need to prove that $K = \mathcal{F}\mathcal{G}(K)$. But

$$\mathcal{F}\mathcal{G}(K) = \text{ fixed field by } \mathrm{Gal}(E/K)$$

which is exactly K by the above proposition (its first point).

2. It is enough to compute $\mathcal{F}(\sigma H \sigma^{-1})$ and show that it is actually equal to $\sigma K = \sigma\mathcal{F}(H)$.

$$\begin{aligned}
\mathcal{F}(\sigma H \sigma^{-1}) &= \{x \in E,\ \sigma\tau\sigma^{-1}(x) = x \text{ for all } \tau \in H\} \\
&= \{x \in E,\ \tau\sigma^{-1}(x) = \sigma^{-1}(x) \text{ for all } \tau \in H\} \\
&= \{x \in E,\ \sigma^{-1}(x) \in \mathcal{F}(H)\} \\
&= \{x \in E,\ x \in \sigma(\mathcal{F}(H))\} = \sigma(\mathcal{F}(H)).
\end{aligned}$$

□

We now look at subextensions of the finite Galois extension E/F and ask about their respective Galois group.

Theorem 7.3. *Let E/F be a finite Galois extension with Galois group G. Let K be an intermediate subfield, fixed by the subgroup H.*

1. *The extension E/K is Galois.*
2. *The extension K/F is normal if and only if H is a normal subgroup of G.*
3. *If H is a normal subgroup of G, then*

$$\mathrm{Gal}(K/F) \simeq G/H = \mathrm{Gal}(E/F)/\mathrm{Gal}(E/K).$$

4. *Whether K/F is normal or not, we have*

$$[K : F] = [G : H].$$

Proof. 1. That E/K is Galois is immediate from the fact that a subextension $E/K/F$ inherits normality and separability from E/F.

2. First note that σ is an F-monomorphism of K into E if and only if σ is the restriction to K of an element of G: if σ is an F-monomorphism of K into E, it can be extended to an F-monomorphism of E into itself thanks to the normality of E. Conversely, if τ is an F-automorphism of E, then $\sigma = \tau|_K$ is surely a F-monomorphism of K into E.

   Now, this time by a characterization of a normal extension, we have

$$K/F \text{ normal} \iff \sigma(K) = K \text{ for all } \sigma \in G.$$

Since $K = \mathcal{F}(H)$, we just rewrite

$$K/F \text{ normal} \iff \sigma(\mathcal{F}(H)) = \mathcal{F}(H) \text{ for all } \sigma \in G.$$

Now by the above theorem, we know that $\sigma(\mathcal{F}(H)) = \mathcal{F}(\sigma H \sigma^{-1})$, and we have

$$K/F \text{ normal} \iff \mathcal{F}(\sigma H \sigma^{-1}) = \mathcal{F}(H) \text{ for all } \sigma \in G.$$

We are almost there, we now use again the above theorem that tells us that $\mathcal{F}$ is invertible, with inverse $\mathcal{G}$, to get the conclusion:

$$K/F \text{ normal} \iff \sigma H \sigma^{-1} = H \text{ for all } \sigma \in G.$$

3. To prove this isomorphism, we will use the 1st isomorphism Theorem for groups. Consider the group homomorphism

$$\mathrm{Gal}(E/F) \to \mathrm{Gal}(K/F),\ \sigma \mapsto \sigma|_K.$$

This map is surjective (we showed it above, when we mentioned that we can extend $\sigma|_K$ to σ. Its kernel is given by

$$\mathrm{Ker} = \{\sigma,\ \sigma|_K = 1\} = H = \mathrm{Gal}(E/K).$$

Applying the 1st isomorphism Theorem for groups, we get

$$\mathrm{Gal}(K/F) \simeq \mathrm{Gal}(E/F)/\mathrm{Gal}(E/K).$$

4. Finally, by multiplicativity of the degrees:

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

Since E/F and E/K are Galois, we can rewrite

$$|G| = |H|[K : F].$$

We conclude by Lagrange Theorem:

$$[G : H] = |G|/|H| = [K : F].$$

□

7.3 Finite fields

We will provide a precise classification of finite fields.

Theorem 7.4. *Let E be a finite field of characteristic p.*

1. *The cardinality of E is*

$$|E| = p^n,$$

for some $n \geq 1$. It is denoted $E = \mathbb{F}_{p^n}$.

2. Furthermore, E is the splitting field for the separable polynomial

$$f(X) = X^{p^n} - X$$

over $\mathbb{F}_p$, so that any finite field with p^n elements is isomorphic to E. In fact, E coincides with the set of roots of f.

Proof. 1. Let $\mathbb{F}_p$ be the finite field with p elements, given by the integers modulo p. Since E has characteristic p, it contains a copy of $\mathbb{F}_p$. Thus E is a field extension of $\mathbb{F}_p$, and we may see E as a vector space over $\mathbb{F}_p$. If the dimension is n, then let $\alpha_1, \ldots, \alpha_n$ be a basis. Every x in E can be written as

$$x = x_1\alpha_1 + \cdots + x_n\alpha_n$$

and there are p choices for each x_i, thus a total of p^n different elements in E.

2. Let $E^\times$ be the multiplicative group of non-zero elements of E. If $\alpha \in E^\times$, then

$$\alpha^{p^n-1} = 1$$

by Lagrange's Theorem, so that

$$\alpha^{p^n} = \alpha$$

for all α in E (including $\alpha = 0$). Thus each element of E is a root of f, and f is separable.

Now f has at most p^n distinct roots, and we have already identified the p^n elements of E as roots of f.

□

Corollary 7.5. *If E is a finite field of characteristic p, then $E/\mathbb{F}_p$ is a Galois extension, with cyclic Galois group, generated by the Frobenius automorphism*

$$\sigma : x \mapsto \sigma(x) = x^p, \ x \in E.$$

Proof. By the above proposition, we know that E is a splitting field for a separable polynomial over $\mathbb{F}_p$, thus $E/\mathbb{F}_p$ is Galois.

Since $x^p = x$ for all x in $\mathbb{F}_p$, we have that

$$\mathbb{F}_p \subset \mathcal{F}(\langle\sigma\rangle)$$

that is $\mathbb{F}_p$ is contained in the fixed field of the cyclic subgroup generated by the Frobenius automorphism σ. But conversely, each element fixed by σ is a root of $X^p - X$ so $\mathcal{F}(\langle\sigma\rangle)$ has at most p elements. Consequently

$$\mathbb{F}_p = \mathcal{F}(\langle\sigma\rangle)$$

and

$$\mathrm{Gal}(E/\mathbb{F}_p) = \langle\sigma\rangle.$$

□

This can be generalized when the base field is larger than $\mathbb{F}_p$.

Corollary 7.6. *Let E/F be a finite field extension with $|E| = p^n$ and $|F| = p^m$. Then E/F is a Galois extension and $m|n$. Furthermore, the Galois group is cyclic, generated by the automorphism*

$$\tau : x \mapsto \tau(x) = x^{p^m}, \ x \in E.$$

Proof. If the degree $[E : F] = d$, then every x in E can be written as

$$x = x_1\alpha_1 + \cdots + x_d\alpha_d$$

and there are p^m choices for each x_i, thus a total of

$$(p^m)^d = p^n$$

different elements in E, so that

$$d = m/n \text{ and } m|n.$$

The same proof as for the above corollary holds for the rest. □

Thus a way to construct a finite field E is, given p and n, to construct $E = \mathbb{F}_{p^n}$ as a splitting field for $X^{p^n} - X$ over $\mathbb{F}_p$.

Theorem 7.7. *If G is a finite subgroup of the multiplicative group of an arbitrary field, then G is cyclic. Thus in particular, the multiplicative group $E^\times$ of a finite field E is cyclic.*

Proof. The proof relies on the following fact: if G is a finite abelian group, it contains an element g whose order r is the exponent of G, that is, the least common multiple of the orders of all elements of G.

Assuming this fact, we proceed as follows: if $x \in G$, then its order divides r and thus

$$x^r = 1.$$

Therefore each element of G is a root of $X^r - 1$ and

$$|G| \leq r.$$

Conversely, $|G|$ is a multiple of the order of every element, so $|G|$ is at least as big as their least common multiple, that is

$$|G| \geq r$$

and

$$|G| = r.$$

Since the order of $|G|$ is r, and it coincides with the order of the element g whose order is the exponent, we have that G is generated by g, that is $G = \langle g \rangle$ is cyclic. □

Since $E^\times$ is cyclic, it is generated by a single element, say α:

$$E = \mathbb{F}_p(\alpha)$$

and α is called a primitive element of E. The minimal polynomial of α is called a primitive polynomial.

Example 7.3. Consider the following irreducible polynomial

$$g(X) = X^4 + X + 1$$

over $\mathbb{F}_2$. Let α be a root of $g(X)$. A direct computation shows that α is primitive:

$$\alpha^0 = 1, \ldots, \ \alpha^4 = \alpha + 1, \ldots, \ \alpha^7 = \alpha^3 + \alpha + 1, \ldots, \ \alpha^{14} = 1 + \alpha^3.$$

7.4 Cyclotomic fields

Definition 7.4. A cyclotomic extension of a field F is a splitting field E for the polynomial

$$f(X) = X^n - 1$$

over F. The roots of f are called nth roots of unity.

The nth roots of unity form a multiplicative subgroup of the group $E^\times$ of non-zero elements of E, and thus must be cyclic. A primitive nth root of unity is an nth root of unity whose order in $E^\times$ is n. It is denoted ζ_n.

From now on, we will assume that we work in a characteristic $\mathrm{char}(F)$ such that $\mathrm{char}(F)$ does not divide n. (Otherwise, we have $n = m\mathrm{char}(F)$ and $0 = \zeta_n^n - 1 = (\zeta^m - 1)^{\mathrm{char}(F)}$ and the order of ζ_n is less than n.)

Example 7.4. The field $\mathbb{Q}(\zeta_p)$ where p is a prime and ζ_p is a primitive pth root of unity is a cyclotomic field over $\mathbb{Q}$.

Let us look at the Galois group $\mathrm{Gal}(E/F)$ of the cyclotomic extension E/F. Then $\sigma \in \mathrm{Gal}(E/F)$ must map a primitive nth root of unity ζ_n to another primitive nth root of unity ζ_n^r, with $(r, n) = 1$. We can then identify σ with r, and this shows that

$$\mathrm{Gal}(E/F) \simeq U_n$$

where U_n denotes the group of units modulo n. This shows that the Galois group is abelian.

Example 7.5. Consider the field extension $\mathbb{Q}(\zeta_3)/\mathbb{Q}$. We have

$$X^3 - 1 = (X - 1)(X^2 + X + 1).$$

The Galois group is given by:

$$\begin{aligned} \sigma : \zeta_3 &\mapsto \zeta_3^2 \\ \sigma^2 : \zeta_3 &\mapsto \zeta_3 \end{aligned}$$

and the group U_3 of units modulo 3 is $U_3 = \{1, 2\}$. Thus

$$\mathrm{Gal}(\mathbb{Q}(\zeta_3)/\mathbb{Q}) = \{\sigma, 1\} \simeq \{2, 1\} = (\mathbb{Z}/3\mathbb{Z})^\times.$$

Finally, since E/F is Galois (under the above assumption)

$$[E : F] = |\mathrm{Gal}(E/F)| = \varphi(n)$$

where $\varphi(n)$ is the Euler totient function.

From now on, we fix the base field $F = \mathbb{Q}$. This means that a primitive nth root of unity ζ_n is given by

$$\zeta_n = e^{i2\pi r/n}, \ (r, n) = 1.$$

Definition 7.5. The nth cyclotomic polynomial is defined by

$$\Psi_n(X) = \prod_{(i,n)=1} (X - \zeta_n^i),$$

where the product is taken over all primitive nth roots of unity in $\mathbb{C}$.

The degree of $\Psi_n(X)$ is thus

$$\deg(\Psi_n) = \varphi(n).$$

Example 7.6. For $n = 1, 2$, we have

$$\Psi_1(X) = X - 1, \ \Psi_2(X) = X - (-1) = X + 1.$$

Computing a cyclotomic polynomial is not that easy. Here is a formula that can help.

Proposition 7.8. *We have*

$$X^n - 1 = \prod_{d|n} \Psi_d(X).$$

In particular, if $n = p$ a prime, then d is either 1 or p and

$$X^p - 1 = \Psi_1(X)\Psi_p(X) = (X - 1)\Psi_p(X)$$

from which we get

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

Proof. We prove equality by comparing the roots of both monic polynomials.

If ζ is a nth root of unity, then by definition

$$\zeta_n^n = 0$$

and its order d divides n. Thus ζ is actually a primitive dth root of unity, and a root of $\Psi_d(X)$.

Conversely, if $d|n$, then any root of $\Psi_d(X)$ is a dth root hence a nth root of unity. □

Examples 7.7. For $n = 3$ and 5, we have a prime and thus we can use the above formula:

$$\begin{aligned}\Psi_3(X) &= X^2 + X + 1\\ \Psi_5(X) &= X^4 + X^3 + X^2 + X + 1.\end{aligned}$$

For $n = 4$ the primitive 4rth roots of unity are $\pm i$, and by definition

$$\Psi_4(X) = (X - i)(X + i) = X^2 + 1.$$

Finally for $n = 6$, the possible values for d are 1,2,3 and 6. Thus

$$\Psi_6(X) = \frac{X^6 - 1}{(X - 1)(X + 1)(X^2 + X + 1)} = X^2 - X + 1.$$

From the above examples, it is tempting to say that in general $\Psi_n(X)$ has integer coefficients. It happens to be true.

Proposition 7.9. *The nth cyclotomic polynomial $\Psi_n(X)$ satisfies*

$$\Psi_n(X) \in \mathbb{Z}[X].$$

Proof. We proceed by induction on n. It is true for $n = 1$ since $X - 1 \in \mathbb{Z}[X]$. Let us suppose it is true for $\Psi_k(X)$ where k is up to $n - 1$, and prove it is also true for n.

Using the above proposition, we know that

$$\begin{aligned}X^n - 1 &= \prod_{d|n} \Psi_d(X)\\ &= \Psi_n(X) \prod_{d|n, d<n} \Psi_d(X).\end{aligned}$$

The aim is to prove that $\Psi_n(X) \in \mathbb{Z}[X]$:

$$\Psi_n(X) = \frac{X^n - 1}{\prod_{d|n, d<n} \Psi_d(X)}.$$

First note that $\Psi_n(X)$ has to be monic (by definition), and both $X^n - 1$ and $\Psi_d(X)$ (by induction hypothesis) are in $\mathbb{Z}[X]$. We can thus conclude invoking the division algorithm for polynomials in $\mathbb{Z}[X]$. □

We conclude by proving the irreducibility of the cyclotomic polynomials.

Theorem 7.10. *The cyclotomic polynomial $\Psi_n(X)$ is irreducible over $\mathbb{Q}$.*

Proof. Let $f(X)$ be the minimal polynomial of ζ_n, a primitive nth root of unity over $\mathbb{Q}(X)$. We first note that by definition $f(X)$ is monic, and thus since $f(X)|X^n - 1$, we have

$$X^n - 1 = f(X)g(X) \tag{7.2}$$

and $f(X)$ and $g(X)$ must be in $\mathbb{Z}[X]$.

To prove that $\Psi_n(x)$ is irreducible, we will actually prove that

$$\Psi_n(X) = f(X).$$

To prove the equality, it is enough to show that every root of $\Psi_n(X)$ is a root of $f(X)$.

We need the following intermediate result: if p does not divide n, then

$$f(\zeta_n^p) = 0.$$

Let us prove this result. Suppose by contradiction that this is not the case, namely $f(\zeta_n^p) \neq 0$. By (7.2), we have

$$X^n - 1 = f(X)g(X),$$

which evaluated in $X = \zeta_n^p$ yields

$$(\zeta_n^p)^n - 1 = 0 = f(\zeta_n^p)g(\zeta_n^p)$$

implying by our assumption that $f(\zeta_n^p) \neq 0$ that

$$g(\zeta_n^p) = 0,$$

or in other words, ζ_n is a root of $g(X^p)$. But by definition of minimal polynomial, we have that $f(X)$ must then divide $g(X^p)$, that is

$$g(X^p) = f(X)h(X),\ h(X) \in \mathbb{Z}[X].$$

Since $g(X^p)$, $f(X)$ and $h(X)$ are in $\mathbb{Z}[X]$, we can look at their reduction modulo p, that is work in $\mathbb{F}_p[X]$. We will denote $\bar{p}(X)$ the polynomial obtained from $p(X)$ by taking all its coefficients modulo p: if $p(X) = \sum_{i=0}^n a_i X^i$, then $\bar{p}(X) = \sum_{i=0}^n (a_i \mod p) X^i$. Therefore

$$\bar{g}(X^p) = \bar{f}(X)\bar{h}(X) \in \mathbb{F}_p[X].$$

By working in $\mathbb{F}_p[X]$, we are now allowed to write that

$$\bar{g}(X^p) = \bar{g}(X)^p$$

and thus

$$\bar{g}(X)^p = \bar{f}(X)\bar{h}(X) \in \mathbb{F}_p[X].$$

This tells us that any irreducible factor of $\bar{f}(X)$ divides $\bar{g}(X)$ and consequently $\bar{f}$ and $\bar{g}$ have a common factor. Looking at (7.2) in $\mathbb{F}_p[X]$ gives

$$X^n - \bar{1} = \bar{f}(X)\bar{h}(X) \in \mathbb{F}_p[X].$$

Since $\bar{f}$ and $\bar{g}$ have a common factor, $X^n - \bar{1}$ has a multiple root, which cannot be since we have assumed that p does not divide n. This proves the claim.

To summarize, we have just proven that if p does not divide n, then $f(\zeta_n^p)$ is another root of f. Since all primitive nth roots of unity can be obtained from ζ_n by successive prime powers, we have that all primitive nth roots of unity are actually roots of $f(X)$, and we know that there are $\varphi(n)$ of them, which is also the degree of $\Psi_n(X)$. This concludes the proof, since

$$\deg f(X) \geq \varphi(n) = \deg(\Psi_n(X)) \Rightarrow f(X) = \Psi_n(X).$$

□

7.5 Solvability by radicals

The question of solvability by radicals is the one of solving polynomial equations under the restriction that we are only allowed to perform addition, subtraction, multiplication, division, and taking nth roots.

For example, we know (Fontana-Tartaglia, 1535) that for a cubic equation

$$X^3 + pX = q,$$

the solution is given by

$$X = \sqrt[3]{\frac{q}{2} + \sqrt{\frac{p^3}{27} + \frac{q^2}{4}}} + \sqrt[3]{\frac{q}{2} - \sqrt{\frac{p^3}{27} + \frac{q^2}{4}}}.$$

By the 16th century all polynomial equations of degree smaller or equal to 4 were solved. The natural question was then: what happens with quintic equations? Euler failed to give an answer, Lagrange (1770) proved that it depends on finding functions of the roots which are unchanged by certain permutations of the roots, and that this approach works up to degree 4 and fails for 5. Abel showed (1824) that quintics are insolvable by radicals. The next question thus became: decide whether or not a given equation can be solved by radicals. Liouville (1843) found the answer in Galois's papers.

The answer is to be found by connecting the problem with field theory as follows. We first need to define the notion of a radical extension. Informally, a radical extension is obtained by adjoining a sequence of nth roots. For example, to get a radical extension of $\mathbb{Q}$ containing

$$\sqrt[3]{11}\sqrt[5]{\frac{7+\sqrt{3}}{2}} + \sqrt[4]{1+\sqrt[3]{4}},$$

we must adjoin

$$\alpha = \sqrt[3]{11},\ \beta = \sqrt{3},\ \gamma = \sqrt[5]{\frac{7+\beta}{2}}, \delta = \sqrt[3]{4}, \epsilon = \sqrt[4]{1+\delta}.$$

This can be stated formally:

Definition 7.6. An extension E/F is radical if $E = F(\alpha_1, \ldots, \alpha_n)$ where for all $i = 1, \ldots, n$, there exists an integer $n(i)$ such that

$$\alpha_i^{n(i)} \in F(\alpha_1, \ldots, \alpha_{i-1}),\ i \geq 2.$$

The α_i's are said to form a radical sequence for E/F.

Example 7.8. The expression

$$\sqrt[3]{11}\sqrt[5]{\frac{7+\sqrt{3}}{2}} + \sqrt[4]{1+\sqrt[3]{4}}$$

is contained in $\mathbb{Q}(\alpha, \beta, \gamma, \delta, \epsilon)$, where

$$\alpha^3 = 11, \beta^2 = 3, \gamma^5 = \frac{7+\beta}{2}, \delta^3 = 4, \epsilon^4 = 1 + \delta.$$

Definition 7.7. Let f be a polynomial over a field F of characteristic zero (this is a simplifying assumption). We say that f is solvable (soluble) by radicals if there exists a field E containing a splitting field for f such that E/F is a radical extension.

We want to connect radical extensions and solvable groups. Here is the main theorem:

Theorem 7.11. *If F is a field of characteristic zero, and $F \subseteq E \subseteq M$ where M/F is a radical extension, then the Galois group of E/F is a solvable group.*

Thus a solvable (by radicals) polynomial has a solvable Galois group (of a splitting field over the base field).

Recall that a group G is solvable if G has a normal series

$$\{1\} = G_r \trianglelefteq G_{r-1} \trianglelefteq \ldots \trianglelefteq G_0 = G$$

with G_i/G_{i+1} abelian. The proof takes some fair amount of work, though the idea is simple. A radical extension is a series of extensions by nth roots. Such extensions have abelian Galois groups (to be proven though...), so the Galois group of a radical extension is made up by fitting together a sequence of abelian groups (unfortunately, the proof is not that simple...)

We can restate the above result in terms of polynomials.

Theorem 7.12. *Let f be a polynomial over a field E of characteristic zero. If f is solvable by radicals then its Galois group (that is the Galois group of its splitting field) over E is a solvable group.*

To find a polynomial which is not solvable by radicals, it suffices to find one whose Galois group is not solvable.

Lemma 7.13. *Let p be a prime, f an irreducible polynomial of degree p over $\mathbb{Q}$. Suppose that f has precisely two non-real zeros in $\mathbb{C}$. Then the Galois group of f over $\mathbb{Q}$ is the symmetric group S_p.*

Theorem 7.14. *The polynomial $X^5 - 6X + 3$ over $\mathbb{Q}$ is not solvable by radicals.*

The proof consists of showing that the polynomial is irreducible over $\mathbb{Q}$, by Eisenstein's criterion. Then f has exactly three real zeros with multiplicity 1 each, and the above lemma says that is Galois group is S_5. To conclude, we need to show that the symmetric group S_n is not solvable if $n \geq 5$.

7.6 Solvability by ruler and compasses

The ancient Greek philosopher Plato believed that the only perfect figures were the straight line and the circle, and this belief had a great impact in ancient Greek geometry: it restricted the instruments available for performing geometrical constructions to ruler and compasses.

Many constructions can be done just be using ruler and compasses, but three famous constructions could not be performed:

- duplication of the cube: find a cube twice the volume of a given cube.
- trisection of the angle: find an angle 1/3 the size of a given angle.
- quadrature of the circle: find a square of area equal to those of a given circle.

It is no wonder those problems remained unsolved (again, under these platonic constraints) since we will see, using our modern tools, that none of them are possible.

We start by formalizing the intuitive idea of a ruler and compass construction. Denote by P_0 the set of points in $\mathbb{R}^2$.

- operation 1 (ruler): through any 2 points of P_0, draw a straight line.
- operation 2 (compasses): draw a circle, whose center is a point of P_0 and whose radius is equal to the distance between some pairs of points in P_0.

Definition 7.8. The points of intersection of any two distinct lines or circles, drawn using operations 1 and 2 are said to be constructible from P_0 if there exists a sequence $r_1, \ldots, r_n$ of points of $\mathbb{R}^2$ such that for each $i = 1, \ldots, n$ the point r_i is constructible from the set $P_0 \cup \{r_1, \ldots, r_{i-1}\}$, $P_i = P_{i-1} \cup \{r_i\}$.

We can now bring field theory into play. With each stage, we associate the subfield of $\mathbb{R}$ generated by the coordinates of the points constructed. Denote by K_0 the subfield of $\mathbb{R}$ generated by the x- and y-coordinates of the points in P_0. If r_i has coordinates (x_i, y_i), then inductively we define

$$K_i = K_{i-1}(x_i, y_i)$$

to get

$$K_0 \subseteq K_1 \subseteq \ldots \subseteq K_n \subseteq \mathbb{R}.$$

Lemma 7.15. *With the above notation, x_i and y_i are zeros in K_i of quadratic polynomials over K_{i-1}.*

Proof. There are 3 cases to consider: line meets line, line meets circle and circle meets circle. We only give the proof of line meets circle.

Take 3 points $A = (p, q)$, $B = (r, s)$, $C = (t, u)$ in K_{i-1}, then draw a line between A and B, and a circle of center C with radius w. The equation of the line AB is

$$\frac{x-p}{r-p} = \frac{y-q}{s-q}$$

while the equation of the circle is

$$(x-t)^2 + (y-u)^2 = w^2.$$

Solving them yields

$$(x-t)^2 + \left(\frac{s-q}{r-p}(x-p) + q - u\right)^2 = w^2.$$

Now x, the first coordinate of the intersection point, is a zero of a quadratic polynomial over K_{i-1}. □

We note that fields obtained by adjoining the zeroes of a quadratic polynomial are extensions of degree 2.

Theorem 7.16. *If $r = (x, y)$ is constructible from a subset $P_0 \in \mathbb{R}^2$, and if K_0 is the subfield of $\mathbb{R}$ generated by the coordinates of the points of P_0, then the degrees $[K_0(x) : K_0]$ and $[K_0(y) : K_0]$ are powers of 2.*

Proof. We have that

$$[K_{i-1}(x_i) : K_{i-1}] = 1 \text{ or } 2,\ [K_{i-1}(y_i) : K_{i-1}] = 1 \text{ or } 2.$$

Using multiplication of degrees, we get

$$[K_{i-1}(x_i, y_i) : K_{i-1}] = [K_{i-1}(x_i, y_i) : K_{i-1}(x_i)][K_{i-1}(x_i) : K_{i-1}] = 1 \text{ or } 2 \text{ or } 4$$

with $K_i = K_{i-1}(x_i, y_i)$. Thus $[K_n : K_0]$ is a power of 2 implying that $[K_n : K_0(x)][K_0(x) : K_0]$ is a power of 2 from which we conclude that $[K_0(x) : K_0]$ is a power of 2, and similarly for y. □

We are now ready to discuss the impossibility proofs.

Theorem 7.17. *The cube cannot be duplicated using ruler and compass constructions.*

Proof. Take a cube whose side is the unit interval, that is of volume 1. We have $P_0 = \{(0, 0), (1, 0)\}$ and $K_0 = \mathbb{Q}$. If we could duplicate the cube, then we can construct a point $(\alpha, 0)$ such that the volume α^3 is equal to 2, that is

$$\alpha^3 = 2.$$

Now $[\mathbb{Q}(\alpha):\mathbb{Q}]$ is a power of 2, but α is a zero of t^3-2 which is irreducible (by Eisenstein) over $\mathbb{Q}$. This gives that

$$[\mathbb{Q}(\alpha):\mathbb{Q}]=3,$$

a contradiction to the fact that it should be a power of 2. □

Theorem 7.18. *The angle $\pi/3$ cannot be trisected using ruler and compass constructions.*

Proof. Constructing an angle trisecting $\pi/3$ is equal to constructing the point $(\alpha,0)$ given $(0,0)$ and $(1,0)$ where $\alpha=\cos(\pi/9)$. Knowing $\alpha=\cos(\pi/9)$, we can construct

$$\beta=2\cos(\pi/9).$$

Using that $\cos(3\theta)=4\cos^3(\theta)-3\cos(\theta)$ and $\cos(3\theta)=1/2$ when $\theta=\pi/9$, we have

$$1=8\cos^3(\theta)-6\cos(\theta)\Rightarrow\beta^3-3\beta-1=0.$$

Now $f(t)=t^3-3t-1$ is irreducible over $\mathbb{Q}$ (apply Eisenstein on $f(t+1)$) thus

$$[\mathbb{Q}(\beta):\mathbb{Q}]=3$$

contradicting the fact that it should be a power of 2. □

Theorem 7.19. *The circle cannot be squared using ruler and compass constructions.*

Proof. Without loss of generality, we assume that the circle is the unit circle centered at $(0,0)$. Constructing a square with area π is equivalent to constructing a point $(\sqrt{\pi},0)$. Since the smallest field with 0 and 1 is $\mathbb{Q}$, the field obtained from adjoining $(\sqrt{\pi},0)$ is $\mathbb{Q}(\sqrt{\pi})$. Thus $[\mathbb{Q}(\sqrt{\pi}):\mathbb{Q}]$ should be a power of 2, and in particular it should be algebraic, which is a contradiction (Lindeman's Theorem shows the transcendence of π, 1882). □

The main definitions and results of this chapter are

- **(4.1).** Definitions of: Galois extension, Galois group, fixed field.
- **(4.2).** The fundamental theorem of Galois theory, Galois groups of intermediate fields.
- **(4.3).** Characterization of finite fields, their Galois group, their multiplicative group.
- **(4.4).** Definition of cyclotomic field, primitive root of unity, cyclotomic polynomial. The Galois group of a cyclotomic field.

Chapter 8

Exercises on Galois Theory

Exercises marked by (*) are considered difficult.

8.1 Galois group and fixed fields

Exercise 96. Compute the Galois group of $X^4 - 2$ over $\mathbb{Q}$ and $\mathbb{F}_3$, the finite field with 3 elements.

Answer. Over $\mathbb{Q}$, we have

$$X^4 - 2 = (X^2 - \sqrt{2})(X^2 + \sqrt{2}) = (X - 2^{1/4})(X + 2^{1/4})(X - i2^{1/4})(X + i2^{1/4}),$$

while over $\mathbb{F}_3$, let w be a root of the irreducible polynomial $X^2 + X + 2 = 0$, then

$$w^2 = -w + 1,\ w^4 = -1,\ w^8 = 1$$

and

$$X^4 - 2 = X^4 + 1 = (X^2 - w^2)(X^2 + w^2) = (X - w)(X + w)(X - w^3)(X + w^3).$$

8.2 The fundamental Theorem of Galois theory

Exercise 97.

1. Compute the splitting field K of the polynomial $f(x) = x^4 - 2 \in \mathbb{Q}(x)$.
2. Show that K is a Galois extension.
3. Compute the degree of $K/\mathbb{Q}$.
4. Compute the $\mathbb{Q}$-automorphisms of K.
5. Do you recognize $\mathrm{Gal}(K/\mathbb{Q})$?

6. What are all the subgroups of $\mathrm{Gal}(K/\mathbb{Q})$?

7. What are all the intermediate subfields of $K/\mathbb{Q}$?

8. Among the intermediate subfields, which are normal?

Answer.

1. We have that $f(x) = (x^2-\sqrt{2})(x^2+\sqrt{2}) = (x-\sqrt[4]{2})(x+\sqrt[4]{2})(x+i\sqrt[4]{2})(x-i\sqrt[4]{2})$. Thus the splitting field of f is $\mathbb{Q}(i, \sqrt[4]{2})$.

2. It is a splitting field thus K is normal, it is separable because $\mathbb{Q}$ is of characteristic zero.

3. The degree is

$$[\mathbb{Q}(\sqrt[4]{2}, i) : \mathbb{Q}] = [\mathbb{Q}(\sqrt[4]{2}, i) : \mathbb{Q}(\sqrt[4]{2})][\mathbb{Q}(\sqrt[4]{2}) : \mathbb{Q}].$$

   The minimum polynomial of i over $\mathbb{Q}(\sqrt[4]{2})$ is x^2+1, so $[\mathbb{Q}(\sqrt[4]{2}, i) : \mathbb{Q}(\sqrt[4]{2})] = 2$. Since $f(x)$ is irreducible over $\mathbb{Q}$ (by Eisenstein), it is the minimal polynomial of $\sqrt[4]{2}$ over $\mathbb{Q}$, thus $[\mathbb{Q}(\sqrt[4]{2}) : \mathbb{Q}] = 4$ and finally the total degree is 8.

4. There are 8 of them. We have

$$\sigma(i) = i,\ \sigma(\sqrt[4]{2}) = i\sqrt[4]{2},$$

   and

$$\tau(i) = -i,\ \tau(\sqrt[4]{2}) = \sqrt[4]{2}$$

   and we can find the others by combining these two, namely:

$$\begin{aligned}
1 :&\quad \sqrt[4]{2} \mapsto \sqrt[4]{2}, &\quad i &\mapsto i\\
\sigma :&\quad \sqrt[4]{2} \mapsto i\sqrt[4]{2} &\quad i &\mapsto i\\
\sigma^2 :&\quad \sqrt[4]{2} \mapsto -\sqrt[4]{2} &\quad i &\mapsto i\\
\sigma^3 :&\quad \sqrt[4]{2} \mapsto -i\sqrt[4]{2} &\quad i &\mapsto i\\
\tau :&\quad \sqrt[4]{2} \mapsto \sqrt[4]{2} &\quad i &\mapsto -i\\
\sigma\tau :&\quad \sqrt[4]{2} \mapsto i\sqrt[4]{2} &\quad i &\mapsto -i\\
\sigma^2\tau :&\quad \sqrt[4]{2} \mapsto -\sqrt[4]{2} &\quad i &\mapsto -i\\
\sigma^3\tau :&\quad \sqrt[4]{2} \mapsto -i\sqrt[4]{2} &\quad i &\mapsto -i
\end{aligned}$$

5. This is the dihedral group of order 8.

6. - order 8: G, order 1: $\{1\}$.
   - order 4: there are 3 of them

     $S = \{1, \sigma, \sigma^2, \sigma^3\} \simeq C_4$, $T = \{1, \sigma^2, \tau, \sigma^2\tau\} \simeq C_2 \times C_2$, $U = \{1, \sigma^2, \sigma\tau, \sigma^3\tau\} \simeq C_2 \times C_2$.

- order 2, there are 5 of them, all isomorphic to C_2:

$$A = \{1, \sigma^2\},\ B = \{1, \tau\},\ C = \{1, \sigma\tau\},\ D = \{1, \sigma^2\tau\},\ E = \{1, \sigma^3\tau\}.$$

7. By Galois correspondence, we obtain the intermediate fiels as fixed fields of the subgroups. The subfields of degree 2 are the easiest to find:

$$\mathbb{Q}(i),\ \mathbb{Q}(\sqrt{2}),\ \mathbb{Q}(i\sqrt{2})$$

which are fixed by resp. S, T and U. By direct computation (that is, apply the automorphism on an element of the larger field, and solve the equation that describes that this element is fixed by this automorphism), we find that the others are:

$$\mathbb{Q}((1+i)\sqrt[4]{2}),\ \mathbb{Q}(i, \sqrt{2}),\ \mathbb{Q}(\sqrt[4]{2}).$$

fixed resp. by C, A and B.

8. The normal subgroups of G are G, S, T, U, A, I, thus their corresponding fixed fields are normal extensions of $\mathbb{Q}$.

Exercise 98. Let K be the subfield of $\mathbb{C}$ generated over $\mathbb{Q}$ by i and $\sqrt{2}$.

1. Show that $[K : \mathbb{Q}] = 4$.
2. Give a primitive element of K and its minimal polynomial.
3. Show that $\mathrm{Gal}(K/\mathbb{Q}) \simeq (\mathbb{Z}/2\mathbb{Z})^2$.
4. Give a list of all the subfields of K.

Answer.

1. Since $K = \mathbb{Q}(i, \sqrt{2})$, we can first build $\mathbb{Q}(\sqrt{2})/\mathbb{Q}$ which is of degree 2, because $x^2 - 2$ is irreducible, then we check that $x^2 + 1$ is irreducible over $\mathbb{Q}(\sqrt{2})$, so we obtain another extension of degree 2, by multiplicativity of the degrees, this gives an extension of degree 4.
2. For example, ζ_8, the primitive 8th root of unity, is a primitive element, with minimal polynomial $x^4 + 1$.
3. The Galois group is given by $\{1, \sigma, \tau, \sigma\tau\}$ where

$$\sigma : i \mapsto -i,\ \sqrt{2} \mapsto \sqrt{2},\ \tau : i \mapsto i,\ \sqrt{2} \mapsto -\sqrt{2}.$$

4. There is one for each subgroup of the Galois group. Since there are only subgroups of order 2 (but for the whole group and the trivial subgroup), we get 3 quadratic field extensions:

$$\mathbb{Q}(i),\ \mathbb{Q}(\sqrt{2}),\ \mathbb{Q}(i\sqrt{2}).$$

Exercise 99. 1. Show that $X^4 - 3 = 0$ is irreducible over $\mathbb{Q}$.

2. Compute the splitting field E of $X^4 - 3 = 0$.

3. Compute the Galois group of $E/\mathbb{Q}$.

4. Can you recognize this group?

5. Choose two proper, non-trivial subgroups of the Galois group above, and compute their corresponding fixed subfields.

Answer.

1. Use Eisenstein with $p = 3$.

2. The roots of $X^4 - 3$ are $i^j \sqrt[4]{3}$, $j = 0, 1, 2, 3$, thus the splitting field is $\mathbb{Q}(\sqrt[4]{3}, i)$.

3. As in previous exercise, with $\sqrt[4]{3}$ instead of $\sqrt[4]{2}$.

4. It is the dihedral group.

5. Again as in previous exercise.

Exercise 100. Consider the field extensions $M = \mathbb{Q}(\sqrt{2}, \sqrt{3})$ and $E = M(\alpha)$ where $\alpha = \sqrt{(2 + \sqrt{2})(3 + \sqrt{3})}$.

1. Show that M is a Galois extension of $\mathbb{Q}$ with Galois group $C_2 \times C_2$.

2. Denote by σ and τ the generators of the two cyclic groups of (1), so that the Galois group of M is written $\langle \tau \rangle \times \langle \sigma \rangle$.

   - Compute $\sigma(\alpha^2)/\alpha^2$ and deduce that $\alpha \notin M$. What is the degree of E over $\mathbb{Q}$?
   - Extend σ to an automorphism of E and show that this automorphism has order 4.
   - Similarly extend τ to an automorphism of E and compute its order. What is the Galois group of E over $\mathbb{Q}$?

Answer.

1. $M/\mathbb{Q}$ is clearly Galois because it is separable ($\mathbb{Q}$ is of characteristic 0) and normal.

2. - We have $\sigma(\alpha^2)/\alpha^2 = (\sqrt{2}-1)^2$ thus $\sigma(\alpha^2) = (\alpha(\sqrt{2}-1)^2$. If α were in M, then $\sigma(\alpha) = \pm\alpha(\sqrt{2}-1)$ and $\sigma^2(\alpha) = \alpha(\sqrt{2}-1)(-\sqrt{2}-1) = -\alpha$, a contradiction ($\sigma^2(\alpha) = \alpha$).
   - We have $\sigma^2(\alpha) = -\alpha$ thus $\sigma^4(\alpha) = \alpha$, $\sigma^4|M = 1$ and $\sigma^2 \neq 1$.

- We compute that $\tau(\alpha) = \frac{3-\sqrt{3}}{\sqrt{6}}\alpha$, we extend τ and its order is 4. We then compute that

$$\sigma\tau(\alpha) = \frac{3-\sqrt{3}}{-\sqrt{6}}(\sqrt{2}-1)\alpha,\ \tau(\sigma(\alpha) = (\sqrt{2}-1)\frac{3-\sqrt{3}}{\sqrt{6}}\alpha$$

  so σ and τ anticommute, and they are both of order 3, so the Galois group is the quaternion group.

Exercise 101. Let L/K be a Galois extension of degree 8. We further assume that there exists a subextension M/K of degree 4 which is not a Galois extension.

- Show that the Galois group G of L/K cannot be abelian.
- Determine the Galois group G of L/K.

Answer.

- A subextension M/K of degree 4 which is not Galois, means that there is a subgroup of order 2 which is not normal in G. Thus G cannot be abelian, since all subgroups of an abelian group are all normal.
- The only groups of order 8 which are not abelian are D_4 and Q_8. All the subgroups of Q_8 are normal, thus it must be D_4.

Exercise 102. Assume that the polynomial $X^4+aX^2+b \in \mathbb{Q}[X]$ is irreducible. Prove that its Galois group is:

1. the Klein group if $\sqrt{b} \in \mathbb{Q}$.
2. the cyclic group of order 4 if $\sqrt{a^2-4b}\sqrt{b} \in \mathbb{Q}$.

Answer.

1. Set $Y = X^2$, then

$$Y^2 + aY + b = (Y - y_1)(Y - y_2)$$

   with

$$y_1 = \frac{-a+\sqrt{a^2-4b}}{2}, y_2 = \frac{-a-\sqrt{a^2-4b}}{2}$$

   and $X = \pm\sqrt{Y}$ so that the four roots are $\pm\sqrt{y_1}, \pm\sqrt{y_2}$. Now $y_1y_2 = b$ thus

$$\sqrt{y_1}\sqrt{y_2} = \sqrt{b} \in \mathbb{Q}$$

   and if $\sigma(\sqrt{y_1}) = \sqrt{y_2}$, then we have that

$$\sigma(\sqrt{y_2}) = \sqrt{b}/\sigma(\sqrt{y_1}) = \sqrt{b}/\sqrt{y_2} = \sqrt{y_1}$$

   and all the elements of the Galois group have order 2, so that it must be the Klein group.

2. We have
$$y_1 - y_2 = \sqrt{a^2 - 4b},$$
thus
$$\sqrt{y_1}\sqrt{y_2}(y_1 - y_2) = \sqrt{b}\sqrt{a^2 - 4b} \in \mathbb{Q}.$$
Now take $\sigma(\sqrt{y_1}) = \sqrt{y_2}$ and if it were of order 2, then $\sigma(\sqrt{y_2}) = \sqrt{y_1}$ and
$$\sigma(\sqrt{y_1}\sqrt{y_2}(y_1 - y_2)) = \sqrt{y_1}\sqrt{y_2}(y_2 - y_1)$$
which contradicts that $\sqrt{y_1}\sqrt{y_2}(y_1 - y_2) \in \mathbb{Q}$ thus σ is of order 4 and the Galois group must be the cyclic group of order 4.

8.3 Finite fields

Exercise 103. Identify the finite fields $\mathbb{Z}[i]/(2+i)$ and $\mathbb{Z}[i]/(7)$.

Answer. $\mathbb{F}_5$ and $\mathbb{F}_{49}$

Exercise 104. Consider the following two polynomials $p(x) = x^2 - x - 1 \in \mathbb{F}_3[x]$ and $q(x) = x^2 + 1 \in \mathbb{F}_3[x]$. Consider the fields $\mathbb{F}_3[x]/(p(x)) \simeq \mathbb{F}_3(\alpha)$ where $p(\alpha) = 0$ and $\mathbb{F}_3[x]/(q(x)) \simeq \mathbb{F}_3(\beta)$ where $q(\beta) = 0$.

1. Compute $(\alpha + 1)^2$.
2. Deduce that the two fields $\mathbb{F}_3(\alpha)$ and $\mathbb{F}_3(\beta)$ are isomorphic.

Answer.

1. We have $(\alpha + 1)^2 = \alpha^2 - \alpha + 1 = (\alpha + 1) - \alpha + 1 = 2 = -1$.
2. We have that $\beta^2 = -1$ by definition of β and we have just shown above that $(\alpha+1)^2 = -1$, thus it is natural to map β to $\alpha+1$, that is $f : \mathbb{F}_3(\beta) \to \mathbb{F}_3(\alpha),\ a + b\beta \mapsto a + b(\alpha + 1)$. Check that f is a ring homomorphism. Then argue that a field homomorphism is always injective, and that both fields have same number of elements.

Exercise 105. Let $\mathbb{F}_2$ be the finite field with two elements.

1. Show that $\mathbb{F}_2(\beta) = \mathbb{F}_2[X]/(q(X))$ is a finite field, where $q(X) = X^2 + X + 1$ and $q(\beta) = 0$.
2. Consider the polynomial $r(Y) = Y^2 + Y + \beta \in \mathbb{F}_2(\beta)[Y]$, and set $L = \mathbb{F}_2(\beta)[Y]/(r(Y))$.
   - Is L a field? Justify your answer.
   - What is the cardinality of L? What is its characteristic? Justify your answers.

Answer.

1. It is enough to show that $q(X)$ is irreducible over $\mathbb{F}_2$, this generates the finite field $\mathbb{F}_4 \simeq \mathbb{F}_2(\beta)$.
2. - We have to see if $r(Y)$ is irreducible over $\mathbb{F}_4$. It is enough to evaluate it in β and $\beta + 1$ to see that it is not zero.
   - This creates an extension of degree 2 of $\mathbb{F}_4$, that is 16 elements. It has characteristic 2.

Exercise 106.
- Let $\mathbb{F}_p$ be a finite field, $p \geq 3$ a prime number. Show that the sum of all the elements of $\mathbb{F}_p$ is 0.
- Let $q = p^n$, p a prime. Show that if $q \neq 2$, then the sum of all elements of $\mathbb{F}_q$ is 0.
- Let $q = p^n$, p a prime. Show that the product of all the non-zero elements of a finite field $\mathbb{F}_q$ is -1.

Answer.

- There are many ways of doing that. Modulo p, one could simply notice that $1+2+\ldots+p-1$ is $p(p-1)/2$, if $p \geq 3$, p is an odd prime, thus $p-1$ is even, $(p-1)/2$ is an integer and thus mod p we do get 0.
- An element a in $\mathbb{F}_q$ satisfies that $a^{p^n} = a$, that is, it is a root of $X^{p^n} - X$. Now all the roots of this polynomial exactly coincide with the elements of $\mathbb{F}_q$, that is, we can write

$$X^{p^n} - X = \prod_{a \in \mathbb{F}_q} (X - a).$$

  If we develop the product, we get that the term in X^{p^n-1} has as coefficients exactly the sum of the elements of $\mathbb{F}_q$, which is thus 0.
- This follows from above. Now we just factor X from the polynomial $X^{p^n} - X$ to get

$$X^{p^n-1} - 1 = \prod_{a \in \mathbb{F}_q^*} (X - a).$$

  Now -1 corresponds to the constant term of the product, which is exactly the product over all non-zero elements of the finite field.

Exercise 107. Consider the finite fields $\mathbb{F}_2, \mathbb{F}_3$ and $\mathbb{F}_4$, and the polynomial $P(Y) = Y^3 + Y + 1$. Over which of these finite fields is $P(Y)$ irreducible? If possible, construct the corresponding field extension.

Answer. Since this polynomial is of degree 3, if it is reducible, that means there is at least one linear term, that is one root in the base field. It is thus irreducible over $\mathbb{F}_2$, however over $\mathbb{F}_3$, we have that $P(1) = 0$, and over $\mathbb{F}_4$, we have no root. Over $\mathbb{F}_2$, we get an extension of degree 3, that is $\mathbb{F}_8$, over $\mathbb{F}_4$, we get an extension of degree 3, that is $\mathbb{F}_{4^3}$.

4.4 Cyclotomic fields

Exercise 108. Let ζ be a primitive 20th root of unity in $\mathbb{C}$, and let $E = \mathbb{Q}(\zeta)$.

- Compute the Galois group $\mathrm{Gal}(E/\mathbb{Q})$.
- How many subfields of E are there which are quadratic extensions of $\mathbb{Q}$?
- Determine the irreducible polynomial of ζ over $\mathbb{Q}$.

Answer.

- We know that $\mathrm{Gal}(E/\mathbb{Q}) \simeq (\mathbb{Z}/20\mathbb{Z})^*$.
- There are 3 of them: $\mathbb{Q}(i\sqrt{5})$, $\mathbb{Q}(\sqrt{5})$ and $\mathbb{Q}(i)$.
- It is $X^8 - X^6 + X^4 - X^2 + 1$.

4.5 Solvability by radicals

4.6 Solvability by ruler and compasses

Exercise 109. True/False.

Q1. An extension having Galois group of order 1 is normal.

Answer.

Q1. It's false! If there is only one element, then it's the identity. Again $\mathbb{Q}(\alpha)$ with $\alpha^3 = 2$ has a Galois group with only the identity, and it is not normal!

Bibliography

[1] Robert Ash. *Abstract Algebra*. www.math.uiuc.edu/ r-ash.

[2] Israel Kleiner. The evolution of group theory: A brief survey. *http://www.jstor.org/stable/2690312*.

[3] Israel Kleiner. History of field theory. *A History of Abstract Algebra*.

[4] J. J. O'Connor and E F Robertson. History of ring theory. *http://www-history.mcs.st-andrews.ac.uk/HistTopics/Ring_theory.html*.

[5] Richard L. Roth. A history of lagrange's theorem on groups.

[6] I. Stewart. *Galois Theory*. Chapman and Hall.