ICTP

Diploma Course in Mathematics

NOTES OF THE ALGEBRA COURSE

Michela Brundu and Alessandro Logar

Dipartimento di Scienze Matematiche
P.le Europa 1, 34127 Trieste, Italy

Contents

Chapter I
Groups

1. Preliminaries

Definition 1.1. A set G together with an operation $\mu : G \times G \longrightarrow G$ is called a *group* if

i) μ is associative (i.e. $\mu(\mu(x,y),z) = \mu(x,\mu(y,z))$, for every $x,y,z \in G$);

ii) there exists an element e of G, called *neutral element*, such that $\mu(e,x) = \mu(x,e) = x$, for every $x \in G$;

iii) for every $x \in G$, there exists an element $x' \in G$, called *inverse* of x, such that $\mu(x,x') = \mu(x',x) = e$.

If, in addition, the following property holds:

iv) $\mu(x,y) = \mu(y,x)$, for every $x,y \in G$,

we say that G is a *commutative* or *abelian* group.

Notation. Usually we denote $\mu(x,y)$ by xy (or $x \cdot y$) or by $x+y$, and we say that G is, respectively, a *multiplicative* or an *additive* group. Note that the additive notation '+' is normally used for abelian groups. In a multiplicative group (respectively additive), the neutral element is usually denoted by 1_G or simply by 1 (resp. 0_G or 0) and the inverse of an element x by x^{-1} (resp. by $-x$).

Example 1.1.1. Here are some examples of groups (the notations here given, which are almost standard, will be used during all the notes).
$(\mathbb{Z},+)$, $(\mathbb{Q},+)$, $(\mathbb{R},+)$, $(\mathbb{C},+)$ are the groups of, respectively, *integer numbers, rational numbers, real numbers, complex numbers*, w.r.t. the addition;
$\mathbb{Q}^\times$, $\mathbb{R}^\times$, $\mathbb{C}^\times$, which are the groups of, respectively, not-zero rational, not-zero real and not-zero complex numbers w.r.t. the product;
the set $(\mathbb{R}^n,+)$ of n-uples of real numbers w.r.t. the addition;
the set $(M_{m,n}(\mathbb{R}),+)$ of the $m \times n$ real matrices w.r.t. the usual addition of matrices; the set $(GL_n(\mathbb{R}),\cdot)$ of invertible matrices of order n w.r.t. the product of matrices.
Some other examples of groups are:
the set $G := \{f : A \longrightarrow A \mid f \text{ is bijective}\}$ w.r.t. the composition of maps, where A is any set;
the set $\{f : \mathbb{R} \longrightarrow \mathbb{R}\}$, with the addition defined pointwise;
the set $S^1 := \{z \in \mathbb{C} \mid |z| = 1\}$, w.r.t. the product of $\mathbb{C}$; the set $\{z \in \mathbb{C} \mid z^n = 1\}$, again w.r.t. the product of $\mathbb{C}$ (n is any natural number).
Another class of groups is that given by the symmetries of some geometric figures. For instance, if X is a square, then the set of symmetries of X is a group of eight elements (denoted by D_8).

Definition 1.2. A *subgroup* of G is a subset H of G which itself forms a group with respect to the operation defining G; we will write $H \leq G$ (or $H < G$, if the subgroup H is proper, i.e. a proper subset of G).

Remark 1.3. It is easy to verify that a non-empty subset H of G is a subgroup if and only if $xy^{-1} \in H$ for every $x, y \in H$.

Example 1.3.1. The following are examples of subgroups:
$(\mathbb{Z}, +) \leq (\mathbb{Q}, +) \leq (\mathbb{R}, +) \leq (\mathbb{C}, +)$.
$\{z \in \mathbb{C} \mid z^n = 1\} \leq S^1 \leq \mathbb{C}^\times$.
If $n \in \mathbb{Z}$ is any element, we shall denote by (n) the set $\{mn \mid m \in \mathbb{Z}\}$; then (n) is a subgroup of $(\mathbb{Z}, +)$.

Proposition 1.4. *Given two subgroups H and J of G, then $H \cap J$ is a subgroup of G. More generally, if H_i, $i \in I$, is a family of subgroups of G, then $\bigcap_{i \in I} H_i$ is a subgroup of G.*

Proof. Let $x, y \in H \cap J$; then, since both H and J are subgroups, $xy^{-1} \in H$ and $xy^{-1} \in J$. So $xy^{-1} \in H \cap J$. Analogously for the case of any family of subgroups. □

Definition 1.5. The smallest subgroup of a group G containing two given subgroups H and J of G is called *product* of H and J and it is denoted by HJ. If for G we use the additive notation, hence if $(G, +)$ is an abelian group and H, J are subgroups of G, then the smallest subgroup of G which contains H and J is denoted by $H + J$.

Example 1.5.1. For example, take $(2), (3) \subseteq \mathbb{Z}$, then the subgroup $(2) + (3)$ is $\mathbb{Z}$, while the subgroup $(4) + (6)$ is (2).

Definition 1.6. Let G and G' be two groups w.r.t. the laws μ and μ', respectively. Then the set $G \times G'$ naturally becomes a group w.r.t. the law $\mu \times \mu'$ defined as follows: $(\mu \times \mu')((g_1, g_1'), (g_2, g_2')) = (\mu(g_1, g_2), \mu'(g_1', g_2'))$. Such group is called *product group* of G and G' and it is denoted by $G \times G'$. Analogously we define the product group $G_1 \times \cdots \times G_n$ of a finite family of groups.

Definition 1.7. Let G be a group, H be a subgroup and $g \in G$; we call, respectively, *left coset* and *right coset* of H with respect to g the two subsets:

$$gH = \{gh \mid h \in H\} \qquad Hg = \{hg \mid h \in H\}.$$

If the number of left (right) cosets of a subgroup H of G is finite, we say that H is of *finite index*. Such a number is called *index* of H in G and it is denoted by $[G : H]$.
If gH is a coset, g is called *representative* of gH.

Note that two representatives of the same coset are equal up to an element of H; i.e. $gH = fH$ if and only if there exists some $h \in H$ such that $g = fh$.

Definition 1.8. A subgroup H of G is a *normal subgroup* if $gH = Hg$ for every $g \in G$; we will write $H \triangleleft G$.

Example 1.8.1. Let $D_n := \{aI_n \mid a \in \mathbb{R}, a \neq 0\}$ (I_n is the $n \times n$ identity matrix). Then D_n is a normal subgroup of $GL_n(\mathbb{R})$.

Proposition 1.9. *Let G be a group and $H \leq G$. The following facts are equivalent:*
i) $H \triangleleft G$;
ii) $gHg^{-1} \subseteq H$ *for all* $g \in G$ *(where* $gHg^{-1} := \{ghg^{-1} \mid h \in H\}$*);*
iii) $gHg^{-1} = H$ *for all* $g \in G$.

Proof. i) $\Rightarrow$ ii). If $x \in gHg^{-1}$, then $x = ghg^{-1}$, for a suitable $h \in H$. But $gh = h'g$ by assumption, so $x = h' \in H$.
The other parts of the proof are analogous. □

If H is a normal subgroup of G, let us consider the set $\{gH \mid g \in G\}$. In this set we define a multiplication by: $g_1H \cdot g_2H := (g_1g_2)H$. This operation is well-defined, because H is normal. In fact if $g_1H = f_1H$, then $g_1 = f_1h$, for some $h \in H$. We want to show that $g_1H \cdot g_2H = f_1H \cdot g_2H$, i.e. that $(g_1g_2)H = (f_1g_2)H$ or, equivalently, that $g_1g_2 = f_1g_2k$, for some $k \in H$. By assumption $g_1g_2 = f_1hg_2$ and $hg_2 = g_2k$, for a suitable $k \in H$, since H is normal. So we are done.
The given set becomes in this way a group.

Definition 1.10. If H is a normal subgroup of G, the group $\{gH \mid g \in G\}$, endowed with the product defined above, is called *quotient group* of G by H and is denoted by G/H.

Remark 1.11. Note that a normal subgroup H of a group G defines the following relation: $x \sim y \Leftrightarrow xy^{-1} \in H$, which is easily seen to be an equivalence relation. Moreover the quotient set $G/\sim$ turns out to be G/H, since $[g] = \{x \in G \mid x \sim g\} = \{x \mid xg^{-1} \in H\} = Hg = gH$.

Example 1.11.1. If $n \in \mathbb{Z}$, then, since $\mathbb{Z}$ is abelian, (n) is a normal subgroup of it. The quotient $\mathbb{Z}/(n)$ is denoted by $\mathbb{Z}_n$ (or by $\mathbb{Z}_n^+$, if we want explicitly refer to the composition law considered). According to the definition given above, two elements $a, b \in \mathbb{Z}$ are equivalent (hence are the same element in $\mathbb{Z}_n$) if and only if $a-b$ is divisible by n. Therefore the set of elements equivalent to a (denoted by $[a]$) is $\{\ldots, -2n+a, -n+a, 0+a, n+a, 2n+a, \ldots\}$. In this way we see that $\mathbb{Z}_n$ is a finite group of n elements. We shall denote by $[0], [1], \ldots, [n-1]$ its elements (sometimes it can be convenient to denote them simply by $0, 1, \ldots, n-1$).

Example 1.11.2. Let G be the subgroup of $\mathbb{R}^\times$ given by all the positive real numbers. It is easy to verify that $\mathbb{C}^\times/G$ can be identified with the group S^1 defined above. In fact, two elements $a + \mathrm{i}b, c + \mathrm{i}d$ of $\mathbb{C}^\times$ are equivalent w.r.t. G if there exists $\lambda \in G$ s.t. $a + \mathrm{i}b = \lambda(c + \mathrm{i}d)$; in particular any element $a + \mathrm{i}b$ is equivalent to $(a + \mathrm{i}b)/\sqrt{a^2 + b^2} \in S^1$.

Definition 1.12. Let G and G' be two groups; a map $f : G \longrightarrow G'$ is a *group homomorphism* if $f(g_1g_2) = f(g_1)f(g_2)$, for every $g_1, g_2 \in G$. (Note that $f(1_G) = 1_{G'}$ and that $f(g^{-1}) = f(g)^{-1}$).
A group homomorphism which is injective, surjective or bijective is called, respectively, a *group monomorphism, epimorphism, isomorphism.* To mean that there exists an isomorphism between G and G' we shall write $G \cong G'$.
A group homomorphism from a group to itself is called an *endomorphism*; if it is also bijective, it is called an *automorphism.*

Examples 1.12.1. If H, G are groups, s.t. $H \leq G$, then the inclusion map $i : H \longrightarrow G$ is a group homomorphism.

Let us try to find all the group homomorphisms $f : \mathbb{Z} \longrightarrow \mathbb{Z}$. Set $m := f(1)$. If n is any positive integer, then $f(n) = f(1+\cdots+1) = f(1)+\cdots+f(1) = nm$, and if n is any negative integer (so $-n$ is positive), then $f(n) = -f(-n) = -f(1+\cdots+1) = -(-nm) = nm$. Therefore we proved that, if $f : \mathbb{Z} \longrightarrow \mathbb{Z}$ is any homomorphism, then there exists an $m \in \mathbb{Z}$ s.t. $f(n) = nm$ for any $n \in \mathbb{Z}$. Conversely, if $m \in \mathbb{Z}$, then it is easy to see that the map $f : \mathbb{Z} \longrightarrow \mathbb{Z}$ defined by $f(n) = nm$ is a group homomorphism.
Let $\phi : \mathbb{C}^\times \longrightarrow S^1$ be defined by: $\phi(a + \mathrm{i}b) := (a + \mathrm{i}b)/\sqrt{a^2 + b^2}$. It is easy to see that ϕ is a group homomorphism.
Let $\det : GL_n(\mathbb{R}) \longrightarrow \mathbb{R}^\times$ be the determinant map. It is well known that $\det(AB) = \det(A)\det(B)$ (Binet theorem), hence det is a homomorphism.
Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be defined by

$$f(a, b) := \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}.$$

The map f gives a rotation in the plane $\mathbb{R}^2$ of an angle ϕ and is a group homomorphism.

Example 1.12.2. Let G be any group. Then the set $\mathrm{Aut}\,(G)$ of all the automorphisms of G is a group w.r.t. the composition of maps. To each $g \in G$ there is associated a map $\tau_g : G \longrightarrow G$ defined by $\tau_g(x) := g^{-1}xg$, for all $x \in G$. The map τ_g, being a group homomorphism, is called *inner automorphism* of G induced by g. Let $\tau : G \longrightarrow \mathrm{Aut}\,(G)$ be defined by $\tau(g) := \tau_g$. It is a group homomorphism.

Definition 1.13. Let $f : G \longrightarrow G'$ be a group homomorphism; the set $\{g \in G \mid f(g) = 1_{G'}\}$ is called the *kernel* of f and it is denoted by $\ker(f)$. The image of f will be denoted by $\mathrm{Im}\,(f)$ (it is the set $\{f(g) \mid g \in G\}$).

Example 1.13.1. Let $f : \mathbb{Z} \longrightarrow \mathbb{Z}$ be the homomorphism $f(n) := mn$ for a fixed $m \in \mathbb{Z}$. Then $\mathrm{Im}\, f = (m)$ and $\ker f = (0)$, if $m \neq 0$.
If $\mathbb{C}^\times \longrightarrow S^1$ is defined as in 1.12.1, then $\ker(f) = \{a \in \mathbb{R} \mid a > 0\}$, i.e. it is the group G of 1.11.2.
If det is the map considered in 1.12.1, then ker(det) is denoted by $SL_n(\mathbb{R})$, and is called the *special linear group.*

Theorem 1.14. *If $f : G \longrightarrow G'$ is a group homomorphism, then:*
i) $\ker(f)$ *and* $\mathrm{Im}\,(f)$ *are subgroups of G and G', respectively;*
ii) $\ker(f)$ *is a normal subgroup of G;*
iii) f is injective if and only if $\ker(f) = 1$*;*
iv) if H is a subgroup of G, then $f(H) = \{\ f(h) \mid h \in H\}$ is a subgroup of $\mathrm{Im}\,(f)$.

Proof. i) Let $x, y \in \ker(f)$; then $f(xy^{-1}) = f(x)f(y^{-1}) = f(x)f(y)^{-1} = 1_G$; therefore $xy^{-1} \in \ker(f)$. Let now $z, t \in \mathrm{Im}\,(f)$; hence there exist $x, y \in G$ such that $z = f(x)$ and $t = f(y)$. Consider $xy^{-1} \in G$; $f(xy^{-1}) = f(x)f(y)^{-1} = zt^{-1}$; so $zt^{-1} \in \mathrm{Im}\,(f)$.
ii) Let us take $g \in G$ and $h \in \ker(f)$. Then $f(ghg^{-1}) = f(g)f(h)f(g)^{-1} = 1_{G'}$; hence $ghg^{-1} \in \ker(f)$, for all $g \in G$.
iii) Assume f is injective; then take $h \in \ker(f) : f(h) = 1_{G'} = f(1_G)$, so $h = 1_G$.
Conversely, let $\ker(f) = \{1_G\}$ and assume there exist $x, y \in G$ such that $f(x) = f(y)$.

Then $f(x)f(y)^{-1} = 1_{G'}$, so $f(xy^{-1}) = 1_{G'}$, i.e. $xy^{-1} \in \ker(f) = \{1_G\}$; therefore $x = y$.
iv) Let $z, t \in f(H)$; so $z = f(h), t = f(k)$ for some $h, k \in H$. Since H is a subgroup of G, $hk^{-1} \in H$, hence $zt^{-1} = f(h)f(k)^{-1} = f(hk^{-1}) \in f(H)$. □

Example 1.14.1. Let $f : \mathbb{Z} \longrightarrow \mathbb{Z}$ be the homomorphism given by $f(n) = mn$ (see example 1.12.1), then, if $H := (a)$ is a subgroup of $\mathbb{Z}$, its image $f(H)$ is the group (ma), which indeed is a subgroup of (m).

Let G be any group, $g \in G$ and $\tau_g : G \longrightarrow G$ be the inner automorphism induced by g. If $H \leq G$, then $\tau_g(H) = g^{-1}Hg$ is a subgroup of $\mathrm{Im}\,(\tau_g)$ (by 1.14.iv) and hence $g^{-1}Hg \leq G$.

Definition 1.15. If $H \leq G$ and $g \in G$, then the subgroup $g^{-1}Hg$ is said *conjugate* to H w.r.t. g.

Note that, by 1.9, that H has no conjugate subgroup (shortly: H is self-conjugate) if and only if it is normal.

Remark 1.16. Let $H \triangleleft G$; then the map $\pi : G \longrightarrow G/H$ defined by $\pi(g) = gH$ is a surjective homomorphism (called *canonical homomorphism* or *canonical projection* of G onto G/H) and $\ker(\pi) = H$.

Theorem 1.17. *(Fundamental theorem) Let $f : G \longrightarrow G'$ be a group homomorphism, $K = \ker(f)$ and $\pi : G \longrightarrow G/K$ be the canonical projection. Then there exists an injective homomorphism $h : G/K \longrightarrow G'$ such that $f = h \circ \pi$. In particular,* $\mathrm{Im}\,(f) \cong G/K$.

Proof. Let us define $h : G/K \longrightarrow G'$ by $h(gK) := f(g)$. We have to show that h is well defined, i.e. if $g_1K = g_2K$ then $f(g_1) = f(g_2)$. Since $K \triangleleft G$, from $g_1K = g_2K$ one has $g_2^{-1}g_1 \in K = \ker(f)$; so $f(g_2^{-1}g_1) = 1_{G'}$ i.e. $f(g_1) = f(g_2)$.
Moreover h is a homomorphism by definition, since f is a homomorphism.
Assume, finally, that $h(gK) = 1_{G'}$; this means that $f(g) = 1_{G'}$, so $g \in K$; this implies $gK = K = 1_{G/K}$. By definition $h(\pi(g)) = h(gK) = f(g)$. So, from $f = h \circ \pi$ and from the surjectivity of π, it follows that $\mathrm{Im}\,(f) = \mathrm{Im}\,(h) \cong G/K$, since h is injective. □

Proposition 1.18. *Let $H \triangleleft G$ and $\pi : G \longrightarrow G/H$ be the canonical projection. Then there is a one-to-one correspondence between the set $\mathcal{A} := \{K \mid H \leq K \leq G\}$ and $\mathcal{B} := \{K' \mid K' \leq G/H\}$ and this correspondence preserves the inclusions.*

Proof. Let $K \in \mathcal{A}$; the corresponding element in $\mathcal{B}$ is $K/H \leq G/H$. Conversely, if $K' \leq G/H$, then set $K := \pi^{-1}(K')$. Then K is a subgroup of G and $H \triangleleft K$. □

Definition 1.19. The *center* of a group G is defined to be:

$$Z(G) = \{x \in G \mid xg = gx \text{ for all } g \in G\}.$$

Example 1.19.1. It is easy to verify that the center of $GL_2(\mathbb{R})$ is given by the set of matrices aI, where $a \in \mathbb{R}^\times$ and I is the 2×2 identity matrix.
If we consider the automorphism τ defined in 1.12.2, we see that its kernel is $Z(G)$.

Definition 1.20. Let H be a subgroup of the group G; we define the *centralizer of H* to be:

$$C_G(H) = \{x \in G \mid xh = hx \text{ for all } h \in H\}.$$

Remark 1.21.
i) G is abelian iff $Z(G) = G$;
ii) for any H, $C_G(H) \supseteq Z(G)$;
iii) $C_G(G) = Z(G) = \bigcap_{x \in G} C_G(x)$.

Definition 1.22. Let H be a subgroup of the group G; we define the *normalizer of H* to be:

$$N_G(H) = \{x \in G \mid x^{-1}Hx = H\}.$$

Remark 1.23. Since $N_G(H)$ is the smallest subgroup of G in which H is normal, H is normal in G if and only if $N_G(H) = G$.

Definition 1.24. Let G be a group and $g_1, \ldots, g_n$ be elements of G. We call *subgroup generated by* $g_1, \ldots, g_n$, and we denote it by $\langle g_1, \ldots, g_n \rangle$, the smallest subgroup containing those elements, i.e. $\langle g_1, \ldots, g_n \rangle := \{x_1 \cdots x_p \mid x_i \in \{g_1, \ldots, g_n, g_1^{-1}, \ldots g_n^{-1}\}, p \in \mathbb{N}\}$.
In particular, $\langle g \rangle = \{g^p, p \in \mathbb{Z}\}$ is called *cyclic subgroup generated by g*.
If $G = \langle g \rangle$ for some $g \in G$, we say that G is a *cyclic group*.

Example 1.24.1. Let $a \in \mathbb{C}^\times$ s.t. $|a| = 1$. Using the trigonometric notation, let $a = \cos\phi + \mathrm{i}\sin\phi$. Then

$$\langle a \rangle \cong \begin{cases} \mathbb{Z}_n \text{ if } 2\pi/\phi = m/n \in \mathbb{Q} \ (m, n \text{ coprime}, n > 0) \\ \mathbb{Z} \quad \text{if } 2\pi/\phi \notin \mathbb{Q}. \end{cases}$$

Definition 1.25. The *order* of a group G is the number of its elements and it is denoted by $|G|$; G is *finite* if it has finite order. If $g \in G$, the *order* of g is $|\langle g \rangle|$, simply denoted by $|g|$.

Remark 1.26. Every subgroup of a cyclic group is cyclic; every quotient of a cyclic group is cyclic. In particular, if $G \leq \mathbb{Z}$ is a subgroup of $\mathbb{Z}$, then it is cyclic, hence it is of the form (n) for a suitable $n \in \mathbb{Z}$.
Moreover, if G is a cyclic group, then two possibilities can arise: either it is finite of order n and so isomorphic to $\mathbb{Z}_n$, or it is infinite and so it is isomorphic to $\mathbb{Z}$. To see this, suppose that $G =< g >$ and consider the following map:

$$f : \mathbb{Z} \longrightarrow G$$

defined by $a \mapsto g^a$. Clearly f is an epimorphism and

$$\ker(f) = \begin{cases} (0) \\ (n) \end{cases}$$

If $\ker(f) = (0)$, then $G \cong \mathbb{Z}$; if $\ker(f) = (n)$, then $G \cong \mathbb{Z}_n$ (see thm. 1.17).

Example 1.26.1 Let us recall that the set of the roots of $x^n - 1$ is an (abelian) subgroup of $\mathbb{C}^\times$ (see 1.3.1). It is well-known that these roots are distinct; for instance, using the trigonometric notation, those roots have the form

$$\varepsilon_k = \cos(2(k-1)\pi/n) + \mathrm{i}\sin(2(k-1)\pi/n)$$

for $k = 1, \ldots, n$.
So the (multiplicative) group $\{\varepsilon_1 = 1, \ldots, \varepsilon_n\}$ of nth roots of unity has order n; it is usually denoted by C_n.
Since $\varepsilon_k = \varepsilon_1^k$, for all k, then C_n is cyclic; therefore $C_n \cong \mathbb{Z}_n$. A generator of C_n is called a *primitive* nth root of unity.

Definition 1.27. If $|g| < \infty$, we say that g is a *torsion element.*
If a group G has no torsion elements, we say that G is *torsion free*; if every element of G is a torsion element, we say that G is *periodic* or a *torsion group.*

Remark 1.28. A finite group is obviously periodic; in particular a finite cyclic group is of the form $G = \{1, g, g^2, \ldots, g^{n-1}\}$, where $n = |G|$.
Clearly a periodic group is not necessarily finite: the orders of its elements can be even unbounded. For example, let us consider the additive group $\mathbb{Q}$ and its subgroup $\mathbb{Z}$. The quotient $\mathbb{Q}/\mathbb{Z}$, whose elements are the cosets $[a/b]$, is infinite and periodic; in fact $|[a/b]| = b$, if a and b are coprime and b positive.

2. Finite groups

Theorem 2.1. *(Lagrange) Let G be a finite group and $H \leq G$; then $|H|$ divides $|G|$.*

Proof. Just note that G can be partitioned as the union of a (finite) number of disjoint cosets gH, each containing $|H|$ elements. □

Corollary 2.2. *Let G be a finite group and $H \triangleleft G$; then $|G|/|H| = |G/H|$.*

Proof. The thesis follows from the proof of 2.1, noting that the number of the cosets gH is $|G/H|$ by definition. □

Example 2.2.1. As an example of Lagrange theorem, let us consider the group $G := \mathbb{Z}_{15}$ and its subgroup $H := \{0, 5, 10\}$. Then (note that here we use the additive notation): $0 + H = 5 + H = 10 + H = \{0, 5, 10\}$, $1 + H = 6 + H = 11 + H = \{1, 6, 11\}$, $2 + H = 7 + H = 12 + H = \{2, 7, 12\}$, $3 + H = 8 + H = 13 + H = \{3, 8, 13\}$, $4 + H = 9 + H = 14 + H = \{4, 9, 14\}$, hence the elements of G are divided into five classes of three elements each, as expected.

A natural question arises about the existence and the number of "substantially different" groups of a given order. More precisely, if $n \in \mathbb{N}$, we can consider the family of groups of order n. If this family is non empty, we can consider the following equivalence relation on it: two groups of order n, say G and H, are equivalent if there exists a group isomorphism from G to H. We call *type* of G the equivalence class of G.

Definition 2.3. For each positive integer n, let us denote by $\nu(n)$ the number of different types of groups of order n.

Remark 2.4. i) For each positive integer n, there exists at least one group of order n (i.e. $\nu(n) \geq 1$, for all n).
In fact it is enough to consider, for each n, the additive group $\mathbb{Z}_n$.
ii) For each positive integer n, there are only finitely many different types of groups of order n (i.e. $\nu(n) < \infty$, for all n).
In order to see this, let G be a finite set. A group structure on G is a map $\mu : G \times G \longrightarrow G$ satisfying some conditions. In particular, μ can be identified with a subset of $(G \times G) \times G$. Since the subsets of $G \times G \times G$ are finite, there are only a finite number of group structures on a finite set.

Example 2.4.1. Let us compute $\nu(4)$. First of all note that $\mathbb{Z}_4$ is a cyclic group, while $\mathbb{Z}_2 \times \mathbb{Z}_2$ is not cyclic. Hence $\nu(4) \geq 2$. Let G be any group of order 4. If it is cyclic, then $G \cong \mathbb{Z}_4$ by 1.26. Then assume G is not cyclic. By Lagrange theorem, any element in $G \setminus \{1\}$ has order 2. Take $a, b \in G$, $a \neq 1 \neq b, a \neq b$. Then it is easy to see that $G = \{1, a, b, ab\}$; otherwise $ab = 1$ would imply that $b = a^{-1}$, while $a^2 = 1$ gives $a^{-1} = a$. One can verify that the bijection $f : G \longrightarrow (a) \times (b)$ defined by:

$$f(1) := (1, 1); \quad f(a) := (a, 1); \quad f(b) := (1, b); \quad f(ab) := (a, b)$$

is a group homomorphism. Since (a) and (b) are cyclic groups of order 2, it follows that $G \cong \mathbb{Z}_2 \times \mathbb{Z}_2$.

Proposition 2.5. *For each prime number p, $\nu(p) = 1$.*

Proof. Let G be a group of order p prime. Take any element $g \in G$, $g \neq 1$. Then, as a consequence of Lagrange theorem, the subgroup $\langle g \rangle$ of G must be G, hence G is cyclic and we already observed that all the cyclic groups of fixed order are isomorphic (see 1.26). □

Corollary 2.6. *If $|G|$ is a prime number p, then $G \cong \mathbb{Z}_p$; in particular G is abelian since cyclic.* □

Definition 2.7. A group G is called a *p-group* if $|G| = p^n$, where p is a prime number and n is a positive integer.

Example 2.7.1. $\mathbb{Z}_9$ is a 3-group (since $9 = 3^2$), while $\mathbb{Z}_6$ is clearly not a p-group.

A fundamental result regarding finite groups is the following:

Theorem 2.8. *(Sylow) Let G be a finite group, with $|G| = p^m r$, where p, m, r are positive integers, p a prime number not dividing r. Then:*
a) G has a subgroup of order p^m (called a Sylow p-subgroup of G);
b) all the Sylow p-subgroups are conjugate; moreover denoting by n the number of distinct Sylow p-subgroups of G, then n divides r and $n \equiv 1 \pmod{p}$;
c) any subgroup of order p^h is contained in a Sylow p-subgroup;
d) there exists an element $x \in G$ of order p.

Proof. See [R], thm. 5.9; [J], 1.13; [L], Ch. I, 6.3. □

3. Series

Definition 3.1. Let G be a group. We call a *series of* G a finite sequence $G_0, G_1, \ldots, G_n$ of subgroups of G, such that

$$1 = G_0 \triangleleft G_1 \triangleleft \cdots \triangleleft G_n = G \qquad (a)$$

i.e. G_i is a normal subgroup of G_{i+1} for $i = 0, \ldots, n-1$.
The subgroups $G_0, G_1, \ldots, G_n$ are called *terms* of the series and the quotient groups G_{i+1}/G_i, for $i = 0, \ldots, n-1$, are called *factors* of the series.
We say that the series (a) is *proper* if $G_i \neq G_{i+1}$, for all i.

Definition 3.2. Let G be a group with a series (a). We say that another series

$$1 = H_0 \triangleleft H_1 \triangleleft \cdots \triangleleft H_m = G \qquad (b)$$

is a *refinement of* (a) if (a) can be obtained from (b) by deleting some terms. We say that (b) is a *proper refinement of* (a) if, furthermore, the two series are different.

Definition 3.3. A proper series which has no proper refinement is called a *composition series* and its factors are called *composition factors*.

Example 3.3.1. Infinite groups need not have a composition series: the additive group $\mathbb{Z}$ is abelian, so each subgroup is normal. But each subgroup of $\mathbb{Z}$ is isomorphic to $\mathbb{Z}$ itself; hence any series of $\mathbb{Z}$ has a proper refinement.

Remark 3.4. i) Every finite group G has a composition series. One can prove it by induction on $|G|$; in fact, let $G = G_n$ and let G_{n-1} be a normal proper subgroup of G with $|G_{n-1}|$ as large as possible. So $|G_{n-1}| < |G|$ and conclude by induction.
ii) By 1.18, the factors of a series have no proper normal subgroups if and only if they are composition factors.

Definition 3.5. Two series of G, say

$$1 = G_0 \triangleleft G_1 \triangleleft \cdots \triangleleft G_n = G \qquad (a)$$

and

$$1 = H_0 \triangleleft H_1 \triangleleft \cdots \triangleleft H_m = G \qquad (b)$$

are said to be *equivalent* if $m = n$ and there exists a one-to-one correspondence between the two sets of factors

$$\{G_n/G_{n-1}, G_{n-1}/G_{n-2}, \ldots, G_1/G_0\} \quad \text{and} \quad \{H_n/H_{n-1}, H_{n-1}/H_{n-2}, \ldots, H_1/H_0\}$$

such that corresponding factors are isomorphic groups.

Remark 3.6. This relationship defines an equivalence relation on the set of series of G.

The problem we want to solve is to understand which are the equivalence classes of such a relation, in particular regarding composition series. The answer will be given by the Jordan-Hölder theorem; for, we need some preliminary facts.

Lemma 3.7. *(Zassenhaus) Let G be a group and A, A^*, B, B^* be subgroups of G such that $A \lhd A^*, B \lhd B^*$. Then:*

i) $(A^* \cap B)A \lhd (A^* \cap B^*)A$;

ii) $(A \cap B^*)B \lhd (A^* \cap B^*)B$;

iii) $\dfrac{(A^* \cap B^*)A}{(A^* \cap B)A} \cong \dfrac{(A^* \cap B^*)B}{(A \cap B^*)B}$.

Proof. See [L], Ch. I, thm. 4.2. □

Theorem 3.8. *(Schreier) Any two series of G have equivalent refinements.*

Proof. Let us consider the two series of G:

$$1 = G_0 \lhd G_1 \lhd \cdots \lhd G_n = G \qquad (a)$$

$$1 = H_0 \lhd H_1 \lhd \cdots \lhd H_m = G. \qquad (b)$$

The main idea is to construct a refinement of (a) by inserting $m-1$ subgroups $G_{i1}, \ldots, G_{i,m-1}$ between G_{i-1} and G_i, for all i. Analogously, we will construct a refinement of (b) by inserting $n-1$ subgroups $H_{1j}, \ldots, H_{n-1,j}$ between H_{j-1} and H_j, for all j.
Set $G_{ij} = (G_i \cap H_j)G_{i-1}$ and $H_{ij} = (H_j \cap G_i)H_{j-1}$, for all $i = 1, \ldots, n$ and $j = 1, \ldots, m$.
Denote by (a') and (b'), respectively, these refinements of (a) and (b). Note that both (a') and (b') have nm terms. Let us set $G_{i0} = G_{i-1}$, $G_{im} = G_i$, $H_{0j} = H_{j-1}$, $H_{nj} = H_j$.
We want to show that (a') and (b') are equivalent; more precisely, that

$$\frac{G_{ij}}{G_{i,j-1}} \cong \frac{H_{ij}}{H_{i-1,j}}$$

for all i and j. But the two quotients

$$\frac{G_{ij}}{G_{i,j-1}} = \frac{(G_i \cap H_j)G_{i-1}}{(G_i \cap H_{j-1})G_{i-1}} \quad \text{and} \quad \frac{H_{ij}}{H_{i-1,j}} = \frac{(H_j \cap G_i)H_{j-1}}{(H_j \cap G_{i-1})H_{j-1}}$$

are isomorphic by 3.7 (part iii), applied to $A = G_{i-1}, A^* = G_i, B = H_{j-1}, B^* = H_j$. □

Corollary 3.9. *Let (a) and (b) be two series of G; if (a) and (b) are equivalent and (b) is a composition series, then (a) is a composition series.*

Proof. If (a) is not a composition series, it admits a refinement (a') equivalent to (b), by 3.8. But, by transitivity, (a) and (a') are equivalent. So (a) and (a') coincide; therefore (a) does not admit proper refinements, i.e. it is a composition series. □

Corollary 3.10. *Suppose that G has a composition series. Then every proper series of G has a refinement which is a composition series.*

Proof. Directly from 3.8. and 3.9. □

Corollary 3.11. *(Theorem of Jordan-Hölder) Any two composition series of G are equivalent.*

Proof. Directly from 3.8. □

An important example of the notions introduced above will be given in next section.

4. A fundamental example: symmetric groups

Definition 4.1. Let X_n be a set of n elements, where $n \in \mathbb{N}$; for simplicity $X_n = \{1, 2, \ldots, n\}$. The set $\{\sigma : X_n \longrightarrow X_n \mid \sigma \text{ is a bijection}\}$ has a natural group structure, given by the composition of maps, i.e. $\sigma \cdot \rho$ will denote the map $\sigma \circ \rho : X_n \longrightarrow X_n$ defined by $(\sigma \circ \rho)(i) = \sigma(\rho(i))$, for all $i \in X_n$.
Such a group is called the *group of permutations* of n objects or *symmetric group* of n objects or simply *symmetric group of order n* and it is denoted by S_n.

Notation. A permutation $\sigma \in S_n$ will be explicitly denoted by

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ a_1 & a_2 & \ldots & a_n \end{pmatrix}$$

meaning that $\sigma(i) = a_i$ for all i.
For example the identity of S_n is $\sigma_1 = \begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix}$.

Note that $|S_n| = n!$.

Example 4.1.1. Let us examine the first cases. For $n = 2$, $|S_2| = 2$, in fact $S_2 = \{\sigma_1, \sigma_2\}$, where

$$\sigma_1 = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \quad \text{and} \quad \sigma_2 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Obviously, $\sigma_1^2 = \sigma_2^2 = \sigma_1$.
For $n = 3$ we get S_3: $|S_3| = 6$, in fact $S_3 = \{\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5, \sigma_6\}$, where

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix},\ \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix},\ \sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix},$$

$$\sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix},\ \sigma_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix},\ \sigma_6 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

A direct computation leads to the following table of the group law (here the σ_i's written in the vertical column operate first on the set X_3):

	σ_1	σ_2	σ_3	σ_4	σ_5	σ_6
σ_1	σ_1	σ_2	σ_3	σ_4	σ_5	σ_6
σ_2	σ_2	σ_1	σ_4	σ_3	σ_6	σ_5
σ_3	σ_3	σ_5	σ_1	σ_6	σ_2	σ_4
σ_4	σ_4	σ_6	σ_2	σ_5	σ_1	σ_3
σ_5	σ_5	σ_3	σ_6	σ_1	σ_4	σ_2
σ_6	σ_6	σ_4	σ_5	σ_2	σ_3	σ_1

table 1

Now we are going to introduce a particular kind of permutations, those which involve only a subset of elements of X_n, moving the first to the second one, the second to the third one, and so on until the last one, sent to the first. More precisely:

Definition 4.2. Let r be an integer, $2 \leq r \leq n$. A *cycle of order* r or an *r-cycle* is an element $\sigma \in S_n$ such that there exists a subset $\{a_1, \ldots, a_r\} \subseteq \{1, \ldots, n\}$ such that $\sigma(a_1) = a_2, \ldots, \sigma(a_{r-1}) = a_r, \sigma(a_r) = a_1$, while $\sigma(i) = i$ if $i \notin \{a_1, \ldots, a_r\}$. We will denote such a cycle σ by $(a_1, \ldots, a_r)$.

Example 4.2.1. The permutation $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 4 & 2 & 5 \end{pmatrix}$ is a 3-cycle and precisely $\sigma = (2, 3, 4)$. Of course we can also express σ in other forms, like $(3, 4, 2)$ and $(4, 2, 3)$.

Remark 4.3. Note that, if σ is an r-cycle, then $\sigma^r = \sigma_1$, so the order of σ in the group S_n is really r in the sense of 1.25.
Note also that the composition of disjoint cycles (i.e. cycles involving disjoint sets of elements) commute, that is $(a_1, \ldots, a_r) \cdot (b_1, \ldots, b_s) = (b_1, \ldots, b_s) \cdot (a_1, \ldots, a_r)$, if $a_i \neq b_j$ for all i and j.

Clearly the elements of S_2 and S_3 are all cycles. In general this is no more true, but cycles are enough to generate any symmetric group.

Remark 4.4.
i) Any permutation is a composition of disjoint cycles;
ii) any r-cycle $(a_1, \ldots, a_r)$ can be immediately written as a composition of 2-cycles; in fact $(a_1, \ldots, a_r) = (a_1, a_r) \cdots (a_1, a_3)(a_1, a_2)$.
iii) any 2-cycle (mr) is the product $(1m)(1r)(1m)$.

Theorem 4.5. *For any $n \in \mathbb{N}$, we have:*
i) S_n is generated by the 2-cycles $(1, 2), (1, 3), \ldots, (1, n)$;
ii) S_n is generated by $(1, 2, \ldots, n)$ and $(1, 2)$.

Proof. *i)* follows from 4.4. *ii)* can be proved with similar arguments. □

The knowledge of cycles is therefore useful to understand the structure of the symmetric group. First of all let us compute their number.

Proposition 4.6. *The number of r-cycles of S_n is $\frac{n!}{(n-r)!r}$.*

Proof. A cycle $(a_1, \ldots, a_r)$ arises from $n(n-1)\cdots(n-r+1) = \frac{n!}{(n-r)!}$ choices between $\{1, \ldots, n\}$. Furthermore an r-cycle can be written in r different ways. □

4.7. Let $A_3 := \{\sigma_1, \sigma_4, \sigma_5\} \subseteq S_3$. From table 1, one easily checks that A_3 is a normal subgroup of S_3, by showing that $\sigma_i A_3 = A_3 \sigma_i$, for $i = 2, 3, 6$. Furthermore, since $|S_3| = 6$ and $|A_3| = 3$, then A_3 is a maximal subgroup of S_3 and it has no proper subgroups. So here a composition series is easily computable:

$$1 = \{\sigma_1\} \triangleleft A_3 \triangleleft S_3.$$

4.8. In S_4, $\operatorname{card}\{\text{2-cycles}\} = 6$, $\operatorname{card}\{\text{3-cycles}\} = 8$, $\operatorname{card}\{\text{4-cycles}\} = 6$ ($\operatorname{card}(A)$ denotes the number of elements of the set A).

Since $|S_4| = 24$, there are 4 other elements. One of them is the identity σ_1. The three elements left are $\sigma_2 = (1,2)(3,4), \sigma_3 = (1,3)(2,4), \sigma_4 = (1,4)(2,3)$, which are compositions of 2-cycles.

Let us compute the series of S_4.
Let $A_4 := \{3\text{-cycles}\} \cup \{\sigma_1, \sigma_2, \sigma_3, \sigma_4\}$. We will see later (in the general case) that A_4 is a subgroup of S_4. To see that A_4 is normal, we need to show that $\rho^{-1}\sigma\rho \in A_4$, for all $\sigma \in A_4$ and $\rho \in S_4$.
Since S_4 is generated by $\{(1,i)\ , i = 2,3,4\}$ by 4.5 and $(1,i)^{-1} = (1,i)$, we can write

$$\rho = (1,i_1)(1,i_2)\cdots(1,i_n),$$

then $\rho^{-1} = (1,i_n)^{-1}\cdots(1,i_2)^{-1}(1,i_1)^{-1} = (1,i_n)\cdots(1,i_2)(1,i_1)$. Hence

$$\rho^{-1}\sigma\rho = (1,i_n)\cdots(1,i_2)(1,i_1)\sigma(1,i_1)(1,i_2)\cdots(1,i_n).$$

To show the above claim, it is enough to show that $(1,i)\sigma(1,i) \in A_4$, for all $\sigma \in A_4$ and $i = 2,3,4$.
For instance, applying 4.3, one sees that

$$(1,2)\sigma_2(1,2) = (1,2)(1,2)(3,4)(1,2) = (1,2)(3,4) = \sigma_2 \in A_4$$

and, in the same way, that

$$(1,3)\sigma_2(1,3) = \sigma_4 \in A_4 \qquad \text{and} \qquad (1,4)\sigma_2(1,4) = \sigma_3 \in A_4.$$

Analogously for the other elements of A_4, so $A_4 \lhd S_4$.
Since $|A_4| = 12$, then $|S_4/A_4| = 2$, so $S_4/A_4 \cong \mathbb{Z}_2$.

Let now $V := \{\sigma_1, \sigma_2, \sigma_3, \sigma_4\}$. As we noted above, $\rho^{-1}\sigma_j\rho \in V$, for all $j = 1, \ldots, 4$ and $\rho \in S_4$; hence $V \lhd S_4$, in particular $V \lhd A_4$.
Since $|V| = 4$, then $|A_4/V| = 3$, so $A_4/V \cong \mathbb{Z}_3$.

Note also that $V \cong \mathbb{Z}_2 \times \mathbb{Z}_2$, since each element of V has order at most 2, hence it has a (normal) subgroup of order 2, e.g. $W = \{\sigma_1, \sigma_2\}$. So we have constructed a series

$$1 = \{\sigma_1\} \lhd W \lhd V \lhd A_4 \lhd S_4.$$

Since the orders of the quotients are prime numbers, this is a composition series (see 3.4 ii)). In particular the set of the composition factors of S_4 (uniquely determined by the Jordan Hölder theorem) is $\{\mathbb{Z}_2, \mathbb{Z}_3, \mathbb{Z}_2, \mathbb{Z}_2\}$.

We already stated that any permutation can be written as product of 2-cycles. This product doesn't need to be unique. Anyway the following fact holds:

Theorem 4.9. *If a permutation is a product of an even number of 2-cycles, then it cannot be written as a product of an odd number of 2-cycles.*

Proof. See [J], 1.6. □

So we can give the following definition:

Definition 4.10. A permutation $\sigma \in S_n$ is *even* if it can be given by a product of an even number of 2-cycles (*odd* otherwise).
The set of even permutations is denoted by A_n and it is called the *alternating group* of n objects. It is indeed a group, since it holds:

Proposition 4.11. *The set A_n of the even permutations is a subgroup of S_n.*

Proof. It immediately follows from 4.9. □

Remark 4.12. Note that the definition of A_3 and A_4 given above is consistent with the general definition given in 4.10. In 4.7, A_3 is the alternating group of 3 objects, in fact: $\sigma_2 = (2,3), \sigma_3 = (1,2), \sigma_6 = (1,3)$ and, by table 1, $\sigma_4 = \sigma_3\sigma_2, \sigma_5 = \sigma_2\sigma_3$. So the even permutations are $\sigma_1, \sigma_4, \sigma_5$.
In 4.8, A_4 is the alternating group of 4 objects (hence a subgroup of S_4); in fact $\sigma_2, \sigma_3, \sigma_4$ are even by definition and by 4.4 ii); moreover any 3-cycle can be written as product of two 2-cycles, e.g. $(1,2,3) = (1,3)(1,2)$.

Remark 4.13. Note that A_n is a normal subgroup of S_n since the product $\sigma_1\sigma_2$ of two permutation is even either if both σ_1 and σ_2 are even or if both σ_1 and σ_2 are odd. So $\sigma^{-1}\rho\sigma \in A_n$ for all $\rho \in A_n, \sigma \in S_n$. Observe also that $|A_n| = n!/2$, so $|S_n/A_n| = 2$ and S_n/A_n is $\mathbb{Z}_2$.

Remark 4.14. We observe that there exists the following alternative definition of A_n. Let $x_1, \ldots, x_n$ be variables. Consider the polynomial $\phi := \prod_{i>j}(x_i - x_j)$. If $\sigma \in S_n$, then call ϕ^σ the polynomial $\prod_{i>j}(x_{\sigma(i)} - x_{\sigma(j)})$. Then $\phi = \varepsilon\phi^\sigma$, where $\varepsilon = +1$ or -1. The permutation σ turns out to be even if $\varepsilon = +1$, odd if $\varepsilon = -1$. In this way we get the map $f : S_n \longrightarrow \{-1, 1\}$ which associates to any permutation its sign; f is a group homomorphism and A_n is its kernel. In particular A_n is a normal subgroup of S_n.

It is clear that any composition series of S_n must have A_n as biggest term. But it happens that, except in the case of S_4, one cannot go further in computing a chain of normal subgroups.

Definition 4.15. A group G is *simple* if it has no proper normal subgroups out of $\{1\}$.

The main result of this section is the following:

Theorem 4.16. *A_n is a simple group, for $n \neq 4$.*

Proof. See [R], thm. 5.28; [S], thm. 13.4. □

We will prove this result only in the case $n = 5$. First we need the following results:

Lemma 4.17. *If G is any group, $K \leq G$ and $N \triangleleft G$, then $N \cap K \triangleleft K$ and*

$$|K/(N \cap K)| \leq |G/N|.$$

Proof. Denoting by $i : K \longrightarrow G$ the inclusion map and by $\pi : G \longrightarrow G/N$ the canonical projection, let us consider the map $f := \pi \circ i : K \longrightarrow G/N$. Then $\ker(f) = N \cap K$, hence $N \cap K$ is normal in K.
Moreover, by 1.17, there exists a monomorphism $K/(N \cap K) \longrightarrow G/N$ and this proves the lemma. □

Lemma 4.18. *The elements of A_5 different from the identity σ_1 are of the following types: 3-cycles (20 of them), 5-cycles (24 of them) and 15 products of two disjoint 2-cycles. In particular, the possible order of an element of A_5 is 2,3 or 5.*

Proof. Just compute the number of 3-cycles and 5-cycles by 4.6. Note also that a 3-cycle is a product of two 2-cycles and a 5-cycle is a product of four 2-cycles; so both are even permutations (see 4.4 *ii*)).
Since $|A_5| = 60$, there are no more elements. □

Lemma 4.19. *The group A_5 has:*
i) 6 subgroups of order 5 (Sylow 5-subgroups);
ii) 10 subgroups of order 3 (Sylow 3-subgroups);
iii) 5 subgroups of order 4 (Sylow 2-subgroups).

Proof. *i*) Since $|A_5| = 60 = 2^2 \cdot 3 \cdot 5$, by Sylow theorem (2.8), we get that the number n_5 of distinct Sylow 5-subgroups divides 12 and $n_5 \equiv 1 \pmod 5$. Hence n_5 is either 1 or 6. Let K be one of these 5-subgroups; since $|K| = 5$, then K is cyclic, so its four elements different from σ_1 are of order 5. By lemma 4.18 there are 24 elements of order 5 in A_5; hence it must be $n_5 = 6$.
ii) In a similar way, one sees that the number n_3 of Sylow 3-subgroups divides 20 and $n_3 \equiv 1 \pmod 3$. So n_3 must be 1, 4 or 10. Again by 4.18, the number of elements of order 3 is 20; since in each Sylow 3-subgroup there are two such elements, we get $n_3 = 10$.
iii) Again, the number n_2 of Sylow 2-subgroups divides 15 and $n_2 \equiv 1 \pmod 2$. So n_2 must be 1, 3, 5 or 15. Let N be one of these subgroups; so either $N \cong \mathbb{Z}_4$ or $N \cong \mathbb{Z}_2 \times \mathbb{Z}_2$. Since in A_5 there is no element of order 4 (4.18), only the second possibility can occur. So N has exactly three elements of order 2. Since the total number of such elements in A_5 is 15, we get $n_2 = 5$. □

Theorem 4.20. *A_5 is a simple group.*

Proof. By 4.19, any subgroup of A_5 of order 5, 4 or 3 is not normal since the Sylow p-subgroups are conjugate (see 2.8).
Assume that there exists a proper normal subgroup N of A_5, $N \neq \{1\}$.
By Lagrange theorem, $|N| \in \{30, 20, 15, 12, 10, 6, 5, 4, 3, 2\}$.

I) $|N| < 30$.

Assume that $|N| = 30$. Let $A_5^{(1)}$ be the subgroup of A_5 of the permutations which fix 1. Since $A_5^{(1)} \cong A_4$, then $|A_5^{(1)}| = 12$. Furthermore $N \cap A_5^{(1)}$ is normal in $A_5^{(1)}$ and

$$\left| \frac{A_5^{(1)}}{N \cap A_5^{(1)}} \right| \leq \left| \frac{A_5}{N} \right| = 2$$

by 4.17. So $|N \cap A_5^{(1)}|$ is either 6 or 12.
In the first case $N \cap A_5^{(1)}$ should be a normal subgroup of $A_5^{(1)}$ of order 6, but A_4 has no normal subgroup of order 6 (see ex. 4.8).
In the second case, $A_5^{(1)}$ should be a subgroup of N, but 12 doesn't divide 30 and this gives *I*).

II) $|N|$ is not a multiple of 5.
If 5 divides $|N|$, by Sylow theorem (2.8) N has a subgroup of order 5, say P, which must be one of the six Sylow 5-subgroups of A_5, which are all conjugate, so of the form: $P_1 = P$, $P_2 = x_2^{-1}Px_2, \ldots, P_6 = x_6^{-1}Px_6$, for suitable $x_2, \ldots, x_6 \in A_5$.
Since N is normal, for each $x \in A_5$, we have $N = x^{-1}Nx \supset x^{-1}Px$. In particular N contains all subgroups $P_1, \ldots, P_6$. Moreover $P_i \cap P_j$ is a proper subgroup of P_i; so its order divides 5, hence $P_i \cap P_j = \{\sigma_1\}$, for all i and j.
This implies that $|N| \geq |\{\sigma_1\}| + 6 \times (|P_i| - 1) = 25$. But there are no allowed orders between 30 and 25.

III) $|N|$ is not a multiple of 3.
As in the previous case, we show that $|N| \geq 21$. But there are no allowed orders between 30 and 21.

IV) $|N| \neq 2$.
If $|N| = 2$, then N is of the form $N = \{\sigma_1, \sigma\}$. Since $x^{-1}Nx = N$ for every $x \in A_5$, it must be $x^{-1}\sigma x = \sigma$, for all $x \in A_5$. Since $\sigma \neq \sigma_1$, there exists i such that $\sigma(i) = h \neq i$. In A_5 there exists a permutation φ such that $\varphi(i) = i$ and $\varphi^{-1}(h) = k \neq h$. So $(\varphi^{-1}\sigma\varphi)(i) = k$ and $\sigma(i) = h$; therefore it cannot be $\varphi^{-1}\sigma\varphi = \sigma$. □

5. Solvable and simple groups

Definition 5.1. A group G is *solvable* if G admits a finite series of subgroups

$$1 = G_0 \triangleleft G_1 \triangleleft \cdots \triangleleft G_n = G \tag{a}$$

having abelian factors, i.e. such that G_{i+1}/G_i are abelian groups for $i = 0, \ldots, n-1$.

Example 5.1.1. Every abelian group G is solvable, with series $1 = G_0 \triangleleft G_1 = G$.

Example 5.1.2. The symmetric group S_3 is solvable. In fact in 4.7 we constructed a composition series

$$1 \triangleleft A_3 \triangleleft S_3.$$

Since $|A_3| = 3$ and $|S_3/A_3| = 2$, then both factors are abelian simple groups.

Example 5.1.3. The symmetric group S_4 is solvable. In fact in 4.8 we constructed a composition series

$$1 \triangleleft W \triangleleft V \triangleleft A_4 \triangleleft S_4$$

whose factors are: $\{\mathbb{Z}_2, \mathbb{Z}_3, \mathbb{Z}_2, \mathbb{Z}_2\}$; as we noted in 3.4 ii), all of them are abelian groups.

Let us present some fundamental properties of solvable groups. To do this, we first need a lemma.

Lemma 5.2. *Let G be a group, $H \triangleleft G$ and $A \leq G$.*
i) It holds: $H \cap A \triangleleft A$ and

$$\frac{A}{H \cap A} \cong \frac{HA}{H}.$$

ii) If $H \leq A$, $H \lhd G$, then $H \lhd A$, $A/H \lhd G/H$ and

$$\frac{G/H}{A/H} \cong \frac{G}{A}.$$

Proof. i) Consider the group homomorphism $f : A \longrightarrow HA/H$ defined by $f(a) = [a]$ (it is the composition of the inclusion $i : A \to HA$ and the canonical projection $\pi : HA \to HA/H$).
Note first that f is surjective. In fact: HA is generated by the elements of the form ha, $h \in H$, $a \in A$; so HA/H is generated by the elements $[ha]$. But $[ha] = [h][a] = [a] \in \mathrm{Im}\,(f)$. So $\mathrm{Im}\,(f) = HA/H$.
Moreover $\ker(f) = H \cap A$. Therefore $H \cap A \lhd A$ by 1.14 ii) and $A/H \cap A = A/\ker(f) \cong \mathrm{Im}\,(f) = HA/H$, by 1.17.
ii) Obviously $H \lhd A$. Consider the group homomorphism $f : G/H \longrightarrow G/A$ defined by $f([x]_H) = [x]_A$. Note that f is well-defined, since if $[x]_H = [y]_H$ then $xy^{-1} \in H \leq A$, hence $[x]_A = [y]_A$. Clearly f is surjective and $\ker(f) = A/H$. So, by 1.14 ii), $A/H \lhd G/H$ and, by 1.17:

$$\frac{G/H}{\ker(f)} = \frac{G/H}{A/H} \cong \mathrm{Im}\,(f) \cong G/A.$$

□

Theorem 5.3. *Let G be a group, H be a subgroup of G and N be a normal subgroup of G. Then:*
i) if G is solvable, then H is solvable;
ii) if G is solvable, then G/N is solvable;
iii) if N and G/N are solvable, then G is solvable.

Proof. i) Let

$$1 = G_0 \lhd G_1 \lhd \ldots \lhd G_n = G$$

be a series of G with abelian factors. Setting $H_i := G_i \cap H$, we get the series of H:

$$1 = H_0 \lhd H_1 \lhd \ldots \lhd H_n = H.$$

We want to show that the factors are abelian. Clearly

$$\frac{H_{i+1}}{H_i} = \frac{G_{i+1} \cap H}{G_i \cap H} = \frac{G_{i+1} \cap H}{G_i \cap (G_{i+1} \cap H)} \cong \frac{G_i(G_{i+1} \cap H)}{G_i}$$

the last isomorphism coming from 5.2 i). But the last quotient is a subgroup of G_{i+1}/G_i, which is abelian by assumption. Hence H_{i+1}/H_i is abelian, as required.
ii) Let

$$1 = G_0 \lhd G_1 \lhd \ldots \lhd G_n = G$$

be a series of G with abelian factors. Then G/N has a series

$$1 = N/N = G_0N/N \lhd G_1N/N \lhd \ldots \lhd G_nN/N = G/N.$$

We want to show that the factors

$$\frac{G_{i+1}N/N}{G_iN/N} \cong \frac{G_{i+1}N}{G_iN}$$

of such a series are abelian (this isomorphism coming from 5.2 ii)).
We have, applying again 5.2 i) and ii):

$$\frac{G_{i+1}N}{G_iN} = \frac{G_{i+1}(G_iN)}{G_iN} \cong \frac{G_{i+1}}{G_{i+1} \cap (G_iN)} \cong \frac{G_{i+1}/G_i}{(G_{i+1} \cap (G_iN))/G_i}$$

and this last group is abelian since it is a quotient of G_{i+1}/G_i, which is abelian by assumption.
iii) First note that any subgroup of G/N is of the form H/N, where $N \lhd H \leq G$, by 1.18. By assumption there exists a series of G/N with abelian factors; so this series must be of the form:

$$1 = N/N = G_0/N \lhd G_1/N \lhd \ldots \lhd G_s/N = G/N.$$

Also by assumption, there exists a series of N:

$$1 = N_0 \lhd N_1 \lhd \ldots \lhd N_r = N$$

with abelian factors. Let us consider the series of G:

$$1 = N_0 \lhd N_1 \lhd \ldots \lhd N_r = N = G_0 \lhd G_1 \lhd \ldots \lhd G_s = G.$$

Its quotients are either N_{i+1}/N_i or G_{i+1}/G_i which is isomorphic (by 5.2 ii)) to

$$\frac{G_{i+1}/N}{G_i/N}$$

and both are abelian by assumption. So G is solvable. □

We may reformulate the above theorem by saying that the class of solvable groups is closed under taking subgroups, quotients and extensions. Note that the class of abelian groups is closed under taking subgroups and quotients, but not extensions.

Let us recall a notion (given in 4.15) which is, in some sense, the "opposite" of that of solvability.

Definition 5.4. A group G is *simple* if its only normal subgroups are $\{1\}$ and G.

Remark 5.5. Every cyclic group of prime order is both simple and solvable. In fact, by Lagrange theorem, it has no subgroups out of $\{1\}$ and G; therefore it is simple. Moreover, a cyclic group is abelian and therefore solvable (see 5.1.1.).

The converse is also true; more precisely:

Theorem 5.6. *A group is both simple and solvable if and only if it is cyclic of prime order.*

Proof. Since G is solvable, it must have a proper series

$$1 = G_0 \triangleleft G_1 \triangleleft \cdots \triangleleft G_n = G$$

with abelian factors. But G is simple, so $G_{n-1} = 1$ and $G/G_{n-1} = G$ is abelian.
Since any element of an abelian group generates a cyclic (normal) subgroup and G is simple, then G itself must be cyclic with no proper subgroups. Hence G has prime order. □

The above theorem makes clear that simple groups and solvable groups are quite different classes.

In addition, combining this result with 4.15, we get an important consequence on symmetric groups.

Corollary 5.7. *The symmetric group S_n is not solvable if $n \geq 5$.*

Proof. If S_n were solvable, then by 5.3, also A_n would be solvable. But A_n is also simple, by 4.15; hence it must be cyclic of prime order by 5.6; while $|A_n| = n!/2$, which is not prime if $n \geq 5$. □

Finally we get the following interesting fact regarding finite solvable groups.

Remark 5.8. Let G be a finite solvable group and let

$$1 = G_0 \triangleleft G_1 \triangleleft \ldots \triangleleft G_n = G \qquad (a)$$

be a series with abelian factors. Then, by 3.4 i) and 3.10, (a) has a refinement

$$1 = H_0 \triangleleft H_1 \triangleleft \ldots \triangleleft H_m = G \qquad (b)$$

which is a composition series.
We claim that the factors H_{i+1}/H_i of (b) are cyclic of prime order.
In fact the factors of (b) are simple groups by 3.4 ii). Moreover $H_{i+1} \leq G$, hence it is solvable, by 5.3 i); finally observe that H_{i+1}/H_i is solvable, again by 5.3 ii).
Since H_{i+1}/H_i is both simple and solvable, we get the claim applying 5.6.

Chapter II
Commutative Rings

1. Preliminaries

Definition 1.1. A *ring* is a non-empty set R together with two binary operations $+$, $\cdot$ and two distinguished elements $0_R, 1_R$ (or simply $0, 1$) in R such that $(R, +)$ is an abelian group (0 is its neutral element), the product $\cdot$ is associative (i.e. $a \cdot (b \cdot c) = (a \cdot b) \cdot c$)), 1 is its neutral element, and the following distributive laws

$$a \cdot (b + c) = a \cdot b + a \cdot c$$
$$(b + c) \cdot a = b \cdot a + c \cdot a$$

hold for all $a, b, c \in R$.
A ring R is *commutative* if $ab = ba$ for all $a, b \in R$.

If $S \subseteq R$, then S is a *subring* of R if $1, 0 \in S$ and $+, \cdot$ induce a ring structure on S. Clearly the intersection of any set of subrings of R is a subring of R; hence if A is a subset of R, one can define the *subring generated* by A to be the intersection of all subrings of R which contain A. This is characterized by the properties: it is a subring, it contains A, and it is contained in every subring containing A.
In the sequel, we shall usually omit the symbol '$\cdot$' for the product.

Examples 1.1.1. 1) $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ are all rings. Moreover $\mathbb{Z}$ is a subring of $\mathbb{Q}$ which is a subring of $\mathbb{R}$, which is a subring of $\mathbb{C}$.
2) All the groups $\mathbb{Z}_n$ (defined in Ch.I, 1.11.1) are rings. The product is the product induced by $\mathbb{Z}$.
3) If A is any set, let $\Gamma := \{f : A \longrightarrow \mathbb{R}\}$. Then we can define in Γ the sum and the product pointwise, and with these operations Γ becomes a ring (0_Γ is the function which sends every element of A to $0 \in \mathbb{R}$, 1_Γ is the function which sends every element of A to $1 \in \mathbb{R}$).
4) If G is an abelian group, then on the set $\mathrm{End}\,(G)$ of group endomorphisms one can define a structure of abelian group pointwise (as in the previous example). Moreover, if we consider the product defined in Ch.I, 1.12.2 (i.e. the composition of maps), then $\mathrm{End}\,(G)$ becomes a ring (the *ring of endomorphisms* of G).
5) Let $\mathbb{Z}[\sqrt{2}]$ be the set of the real numbers of the form: $m + \sqrt{2}n$ $(m, n \in \mathbb{Z})$. Then $\mathbb{Z}[\sqrt{2}]$ is a subring of $\mathbb{R}$.
6) Let $R = M_{n,n}(\mathbb{R})$ be the set of all $n \times n$ matrices over $\mathbb{R}$. Then $(R, +)$ is an abelian group (see Ch.I, 1.1.1) and with the multiplication row by column it becomes a ring.
All the previous examples are commutative rings, except 4) and 6).

Let R be any ring. Here we list some properties, which are an easy consequence of the axioms of rings. For instance, it holds: $-(a \cdot b) = (-a) \cdot b = a \cdot (-b)$. If $a \in R$, then

$a \cdot 0 = 0 \cdot a = 0$. If $n \in \mathbb{Z}$, then by na we mean $a + a + \cdots + a$ (n times, if n is positive), or $-a - a - \cdots - a$ ($-n$ times, if n is negative). Then it holds:

$$\begin{aligned} n(a+b) &= na + nb; \\ (n+m)a &= na + ma; \\ (nm)a &= n(ma) \end{aligned}$$

for all $m, n \in \mathbb{Z}$ and all $a, b \in R$.

Definition 1.2. A ring is called a *domain* (or an *integral domain*) if the condition

$$a, b \in R, \quad ab = 0 \quad \Rightarrow \quad a = 0 \text{ or } b = 0,$$

holds.
Note that in a domain the cancellation law holds, i.e. if $ab = ac$ and $a \neq 0$, then $b = c$ (analogously $ba = ca,\ a \neq 0 \ \Rightarrow\ b = c$).

Examples 1.2.1. The rings in 1) and 5) in 1.1.1 are domains; while $\mathbb{Z}_n$ is a domain if and only if n is prime. The rings considered in 3), 4), 6) in general are not domains.

Definition 1.3. If R is a ring, an element $a \in R$ is a *left (right) zero-divisor* if there exists an element $b \in R$, $b \neq 0$ such that $ab = 0$ (resp. $ba = 0$).

For example, in $\mathbb{Z}_6$ the element [2] is a (left and right) zero-divisor, since $2 \cdot 3 = 6 \equiv 0$ in $\mathbb{Z}_6$.

Remark 1.4. A ring is a domain if and only if the only zero-divisor is the element 0.

Definition 1.5. An element $a \in R$ is called *invertible* or a *unit* if there exists $b \in R$ such that $ab = ba = 1_R$. It is obvious that the set of units of R is a multiplicative group w.r.t. the product defined in R, called *group of units* of R.

Example 1.5.1. The group of units of $\mathbb{Z}$ is $\{1, -1\}$. The group of units of $\mathbb{Z}_n$ is $\{[m] \,|\, m, n \text{ coprime}\}$. The group of units of $M_{n,n}(\mathbb{R})$ is $GL_n(\mathbb{R})$. The groups of units of $\mathrm{End}\,(G)$ is $\mathrm{Aut}\,(G)$.

Definition 1.6. A ring R is a *division ring* (also a *skew field*) if every non-zero element is a unit, i.e. if $(R \setminus \{0\}, \cdot)$ is the group of units. A commutative division ring is called a *field.*

Examples 1.6.1. $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ are fields; $\mathbb{Z}_n$ is a field if and only if n is prime (hence $\mathbb{Z}_n$ is a domain if and only if it is a field). More generally, a finite domain R is a field, in fact if $a \in R$, $a \neq 0$, then the set $\{a^n \mid n \in \mathbb{N}\}$ is finite, so there exist $m, n \in \mathbb{N}$, $m \neq n$, such that $a^n = a^m$, so there exists $r \in \mathbb{N}$ such that $a^r = 1$, therefore $a(a^{r-1}) = 1$.

Remark 1.7. A division ring is a domain; in fact if $ab = 0$ and $a \neq 0$, then there exists the (two-sided) inverse a^{-1} of a. So $a^{-1}ab = 0$ which gives $b = 0$. The converse, in general, is not true: $\mathbb{Z}$ is a domain, but not a field.

Definition 1.8. Let R be a ring and $a, b \in R$. We say that b is a *factor* or *divisor* of a if there exists $c \in R$ such that $a = bc$ or $a = cb$. In this case we shall write $b|a$ (b divides a) and a is called a *multiple* of b.

Remark 1.9. Units are factors of every element, since $a = u(u^{-1}a)$ for any $a \in R$, for any unit u.

Definition 1.10. Let $a, b \in R$; if $b|a$ and $a|b$, then a and b are *associates* and we shall write $a \sim b$ (in this case a and b differ by a unit factor).
If $b \,|\, a$ but $a \not| \, b$ (a is not a factor of b) then we say that b is a *proper factor* of a.

Definition 1.11. An element $a \in R$ is said to be *irreducible* if a is not a unit and a has no proper factors other than units (i.e. if $a = bc$ then either b or c is a unit).

Example 1.11.1. In $\mathbb{Z}$ an element is irreducible if and only if it is a prime number (different from 1 and -1). Therefore a factorization of an integer number n into prime factors is a factorization into irreducible factors. Moreover the expression

$$n = p_1 \cdots p_s$$

where the p_i's are prime numbers, is essentially unique, since two factorizations of n differ, at most, by a permutation of the prime factors and by the sign of each factor. E.g. $6 = 2 \cdot 3 = 3 \cdot 2 = (-2) \cdot (-3) = \dots$.

In general we have the following:

Definition 1.12. Let R be a ring and $a \in R$ be any element. The expression

$$a = p_1 \cdots p_s$$

is an *essentially unique factorization* of a *into irreducible elements* p_i's if for any other factorization

$$a = q_1 \cdots q_t$$

where the q_i's are irreducible elements, we have $t = s$ and $q_i \sim p_{i'}$ for a suitable permutation $i \mapsto i'$ of $\{1, 2, \dots, s\}$.

Definition 1.13. A domain R is *factorial* (also UFD: *Unique Factorization Domain*) if every non-unit element of $R \setminus \{0\}$ has an essentially unique factorization into irreducible elements.

Example 1.13.1. As observed before, $\mathbb{Z}$ is a UFD. Trivially, any field is a UFD. We will give in the sequel other examples of factorial rings (see section 3).

Definition 1.14. A *ring homomorphism* is a map $f : R \longrightarrow R'$ between two rings R and R' such that

$$f(1_R) = 1_{R'} \quad \text{and} \quad \begin{cases} f(a+b) = f(a) + f(b) \\ f(ab) = f(a)f(b) \end{cases} \quad \text{for every } a, b \in R.$$

A ring homomorphism is an *epimorphism*, a *monomorphism*, an *isomorphism* if it is, respectively, surjective, injective, bijective. If $R' = R$, then f is called an *endomorphism*; if, moreover, it is also bijective, it is an *automorphism*.

Example 1.14.1. If R is a ring and $a \in R$, the map $f_a : R \longrightarrow R$ defined by $f_a(x) := ax$ is a group homomorphism (monomorphism if and only if a is a non zero-divisor), but not a ring homomorphism, unless $a = 1$ i.e. $f = \mathrm{id}_R$.

Definition 1.15. If $f : R \longrightarrow R'$ is a ring homomorphism, its *kernel* is the set $\ker(f) := \{a \in R \mid f(a) = 0\}$. The *image* of f is the set $\mathrm{Im}(f) := \{f(a) \mid a \in R\}$.

Remark 1.16. Since a ring homomorphism is, in particular, a group homomorphism between the underlying additive group structures, $\ker(f)$ and $\mathrm{Im}(f)$ are subgroups of R and R', respectively (see Ch.I, 1.14). It is easy to see that $\mathrm{Im}(f)$ is also a subring of R', while $\ker(f)$ is not a subring of R, since $1 \notin \ker(f)$. However $\ker(f)$ has the important structure of ideal that will be introduced in next section.

Let us denote by $\mathrm{Aut}_{\mathcal{G}}(R)$ and by $\mathrm{Aut}_{\mathcal{R}}(R)$ the sets of (additive) group automorphisms and ring automorphisms of a ring R, respectively (these sets turn out to be groups with respect to the composition of maps). In general, as observed above, it holds $\mathrm{Aut}_{\mathcal{R}}(R) \subseteq \mathrm{Aut}_{\mathcal{G}}(R)$.

Example 1.16.1. If $R = \mathbb{Z}_n$, the strict inclusion holds. In fact, let $f : \mathbb{Z}_n \longrightarrow \mathbb{Z}_n$ be a group automorphism; if $f[1] = [a]$ for some $a \in \mathbb{Z}$, then $f[m] = [am]$, for every m. Therefore $\mathrm{Im}(f) = \langle [a] \rangle$; hence f is an automorphism if and only if a and n are coprime, i.e. $[a]$ is a unit in the ring $\mathbb{Z}_n$. Therefore $\mathrm{Aut}_{\mathcal{G}}(\mathbb{Z}_n)$ is isomorphic to the group of units of $\mathbb{Z}_n$. On the other hand, since $f : \mathbb{Z}_n \longrightarrow \mathbb{Z}_n$ defined by $f[m] = [am]$ is a ring homomorphism if and only if $[a] = [1]$, we have that $\mathrm{Aut}_{\mathcal{R}}(\mathbb{Z}_n)$ consists only of the identity map.

Warning. Since we are mostly interested in giving results for commutative rings, from now on, we shall assume that all the rings considered are commutative.

2. Ideals (part I)

Note that the kernel of a ring homomorphism $f : R \longrightarrow R'$ satisfies the following property: if $x \in \ker(f)$ and $r \in R$, then xr belongs to $\ker(f)$, since $f(xr) = f(x) \cdot f(r) = 0 \cdot f(r) = 0$.
This leads to the following definition:

Definition 2.1. If R is a ring and $I \subseteq R$, then I is an *ideal* if it is a subgroup of $(R, +)$ and if the following condition

$$\text{if } a \in R \text{ and } b \in I, \quad \text{then} \quad ab \in I$$

holds.
An ideal I is *proper* if it is properly contained in R.

Obviously the kernel of a ring homomorphism is an ideal.

If I is an ideal, then it is a normal subgroup of $(R,+)$, since R is abelian; so we can consider the quotient group R/I, whose elements are of the kind $[a] = a+I = \{a+b \mid b \in I\}$. It is clear that $(R/I,+)$ is an abelian group since $[a]+[b] = [a+b]$ (see Ch.I, 1.10). If we define in a similar way a product by $[a]\cdot[b] := [ab]$, it is immediate to verify that this product is well defined (this is a consequence of the condition defining an ideal given in 2.1).

Definition 2.2. The ring $(R/I,+,\cdot)$ constructed above is called the *quotient ring* of R w.r.t. I.

The canonical projection $\pi : R \longrightarrow R/I$ (defined in Ch.I, 1.16), whose kernel is I itself, turns out to be a ring homomorphism (epimorphism).

The following results are analogous to the results 1.14, 1.17 and 1.18 given in Ch.I. Also the proofs are quite similar to the proofs given in there.

Proposition 2.3. *If $f : R \longrightarrow R'$ is a ring homomorphism, then f is a monomorphism if and only if* $\ker(f) = (0)$. □

Theorem 2.4. *(Fundamental theorem of ring homomorphisms) Let $f : R \longrightarrow R'$ be a ring homomorphism, $I := \ker(f)$ and $\pi : R \longrightarrow R/I$ be the canonical projection. Then there exists an injective ring homomorphism $h : R/I \longrightarrow R'$ such that $f = h \circ \pi$. In particular,* $\mathrm{Im}\,(f) \cong R/I$. □

Theorem 2.5. *Let $I \subseteq R$ be an ideal and $\pi : R \longrightarrow R/I$ be the canonical projection. Then there is a one-to-one correspondence between the set $\mathcal{A} := \{J \subseteq R \mid I \subseteq J,\ J \text{ ideal of } R\}$ and $\mathcal{B} := \{J' \mid J' \text{ ideal of } R/I\}$ and this correspondence preserves the inclusions.* □

It is easy to see that if $\{I_s \mid s \in \Sigma\}$ is any family of ideals of a ring R, then $\cap_{s\in\Sigma} I_s$ is an ideal.

Definition 2.6. If $A \subseteq R$ is any subset, then the *ideal generated by A* is the intersection of all the ideals containing A and it is denoted by (A).

Note that (A) is the smallest ideal in R which contains A. It is easy to verify that (A) is the the following set:

$$\{a_1r_1 + \cdots + a_nr_n \mid a_i \in A,\ r_i \in R,\ n \in \mathbb{N}\}$$

in fact this set is an ideal, it contains A and if an ideal I contains A, then (A) is surely contained in I. Hence we have a representation of (A) in terms of its elements.

Definition 2.7. If $A = \{a\}$, then the ideal generated by A is denoted by (a) and it is said *principal* ideal generated by a. It is clear that $(a) = \{ar \mid r \in R\}$.
More generally, if $A = \{a_1,\ldots,a_m\}$ is a finite set of elements, then the ideal (A) is $\{\sum_{i=1}^m a_ir_i \mid r_i \in R\}$ and is denoted by $(a_1,\ldots,a_m)$, instead of $(\{a_1,\ldots a_m\})$. If $I = (a_1,\ldots,a_m)$, then I is said a *finitely generated ideal* and $\{a_1,\ldots,a_m\}$ is called a *system of generators* of I.

Examples 2.7.1. 1) Let us compute all the ideals of $\mathbb{Z}$. Assume $I \subseteq \mathbb{Z}$ is an ideal. Hence, in particular, I is a subgroup of $\mathbb{Z}$; therefore, as seen in Ch.I, 1.3.1 and 1.26, I is the set

$\{mn \mid m \in \mathbb{Z}\}$, so it is the principal ideal (n) generated by n. In this way we see that in $\mathbb{Z}$ all the ideals are principal.
2) We already defined the group quotient $\mathbb{Z}_n = \mathbb{Z}/(n)$. Since (n) is an ideal, this is the quotient ring of $\mathbb{Z}$ w.r.t. (n) (already considered in 1.1.1). Also in the ring $\mathbb{Z}_n$ all the ideals are principal. In fact, by 2.5, an ideal of $\mathbb{Z}_n$ is of the form $(m)/(n)$, i.e. of the form $([m])$.
3) If F is a field, then it has only two ideals, which are (0) and $(1) = F$.

Definition 2.8. A ring R whose ideals are all principal is said *principal ideals ring*, shortly PIR. If, furthermore, it is a domain, then it is said *principal ideals domain*, shortly PID.

As we noted before, $\mathbb{Z}$ is a PID and $\mathbb{Z}_n$ is a PIR (it is a PID if and only if n is a prime number). Furthermore a field F is a PID, since its only ideals (0) and $(1) = F$ are principal.

Let R be a ring and $a \in R$. Recall that if $n \in \mathbb{Z}$, then with the notation $n \cdot a$ we mean $a + \cdots + a$ (n times, if $n > 0$) or $-a - \cdots - a$ ($-n$ times, if $n < 0$). The map $\phi : \mathbb{Z} \longrightarrow R$ defined by $\phi(n) := n \cdot 1_R$ is a ring homomorphism. From 2.4 the subring $\mathrm{Im}\,(\phi)$ is isomorphic to $\mathbb{Z}/\ker(f)$.

Definition 2.9. Let R be a ring and f be as above. Two possibilities can arise:
either $\ker(f) = (0)$, i.e. f is a monomorphism, hence R contains (an isomorphic copy of) $\mathbb{Z}$; or $\ker(f) = (n)$, hence R contains (an isomorphic copy of) $\mathbb{Z}_n$.
In the first case we say that R is of *characteristic zero*; in the second one, that R is of *characteristic* n. The characteristic of the ring R will be denoted by $\mathrm{char}\,(R)$.

Proposition 2.10. *If R is an integral domain, then its characteristic is either 0 or a prime number.*

Proof. Assume $\mathrm{char}\,(R) = n$. If $n \neq 0$, then suppose $n = rs$ ($r \leq n,\ s \leq n$). We know that $n \cdot 1_R = 0$. Hence $(rs) \cdot 1_R = 0$, but $(rs) \cdot 1_R = (r \cdot 1_R) \cdot (s \cdot 1_R)$. Since R is a domain, then for example $r \cdot 1_R = 0$ and this implies $r = n$, so n is prime. □

3. Polynomial rings

Suppose that R is a (commutative, as usual) ring and assume it is a subring of a ring R'. Let $U \subseteq R'$ be any subset. Then we can consider the intersection of all the subrings of R', containing R and U. This is a ring and is denoted by $R[U]$. It is the smallest subring of R' containing R and U. If $U = \{u_1, \ldots, u_n\}$ is a finite set, then the ring $R[U]$ is denoted by $R[u_1, \ldots, u_n]$. If U, V are subsets of R', then it holds: $R[U][V] = R[U \cup V]$. In particular we see that $R[u_1][u_2] \cdots [u_n] = R[u_1, u_2, \ldots, u_n]$.
We can see in more details how are represented the elements of $R[u]$. It is clear that in $R[u]$ we have any element of the form $a_0 + a_1 u + \cdots + a_m u^m$, where $a_0, \ldots, a_m \in R$, which is called a *polynomial expression in u with coefficients in R*. It is easy to verify that the set of polynomial expressions in u is indeed a subring of R', and therefore we get that

$$R[u] = \{a_0 + a_1 u + \cdots + a_m u^m \mid a_0, \ldots, a_m \in R,\ m \in \mathbb{N}\}.$$

In considering the polynomial expressions in u one difficulty can arise: it may happen in fact that we have two different-looking expressions for the same element. For example, consider the rings $\mathbb{Q} \subseteq \mathbb{R}$, and take $u = \sqrt{2} \in \mathbb{R}$. Then in the ring $\mathbb{Q}[\sqrt{2}]$ an element can be given by many different polynomial expressions; for instance $1 + 3\sqrt{2} + 2(\sqrt{2})^2 = 5 + 3\sqrt{2} = -1 + \sqrt{2} - (\sqrt{2})^2 + (\sqrt{2})^3 + 2(\sqrt{2})^4 = \dots$.
Hence we want to construct a ring in which the following condition holds:

$$a_0 + a_1 x + \cdots + a_m x^m = b_0 + b_1 x + \cdots + b_n x^n \;\Rightarrow\; m = n,\ a_i = b_i \text{ , for all } i = 1, \dots, n.$$

To do this, it is necessary to define polynomial expressions in a symbol "x" having no algebraic relation with the elements of R. For instance, in the previous example the element $\sqrt{2}$ does have the following algebraic relation: $(\sqrt{2})^2 = 2 \in \mathbb{Q}$. The "good" notion will turn out to be that one of *polynomials*. Let us first start with the abstract definition of polynomials; then we will show that it corresponds to the usual idea introduced in the basic algebra courses.

Definition 3.1. Let R be any ring (here we do not assume anymore that it is a subring of a bigger ring). By a *polynomial* with coefficients in R we mean a sequence $(a_0, a_1, \dots, a_n, \dots)$ of elements of R where the a_i's are all zero but a finite number of them.

Definition 3.2. Let now R' be the set of all polynomials with coefficients in R, i.e. the set of the sequences $(a_0, a_1, \dots, a_n, \dots)$ such that $a_i \in R,\ a_i = 0$ for almost all i. In other words, $R' = \{f : \mathbb{N} \longrightarrow R \mid f(i) = 0 \text{ for almost all } i\}$. In R' we can define a sum pointwise:

$$(a_0, a_1, \dots, a_n, \dots) + (b_0, b_1, \dots, b_n, \dots) := (a_0 + b_0, a_1 + b_1, \dots, a_n + b_n, \dots).$$

In this way $(R', +)$ becomes an abelian group (the zero is the sequence $(0, 0, \dots, 0, \dots)$). Then in R' we can define a product as follows:

$$(a_0, a_1, \dots, a_n, \dots) \cdot (b_0, b_1, \dots, b_n, \dots) := (c_0, c_1, \dots, c_n, \dots),$$

where

$$c_i := \sum_{j=0}^{i} a_j b_{i-j} = \sum_{j+k=i} a_j b_k.$$

With this product R' becomes a commutative ring (with the unity $1 = (1, 0, 0, \dots)$), said *ring of polynomials over R*.

Notation. Usually we express a polynomial in a simpler way, by making some identifications. Firstly, since the map $R \longrightarrow R'$ defined by $a \mapsto (a, 0, 0, \dots)$ is a ring monomorphism, we may identify R with its image in R', so we consider R as a subring of R'. On the other hand, let $x := (0, 1, 0, \dots)$. It holds, using the product law defined above:

$$x^k = (0, 0, \dots, 0, 1, 0, \dots)$$

where 1 is placed in the $(k+1)$-th position. Note that $x^0 = (1, 0, \dots, 0) = 1_R = 1_{R'}$. Therefore we have:

$$\begin{aligned}(a_0, a_1, \dots, a_n, 0, 0, \dots) &= (a_0, 0, \dots)(1, 0, \dots) + (a_1, 0, \dots)(0, 1, 0, \dots) + \cdots \\ &\qquad + (a_n, 0, \dots)(0, \dots, 1, 0, \dots) \\ &= a_0 + a_1 x + \cdots + a_n x^n.\end{aligned}$$

It is natural, using the analogous of the notation of polynomial expressions, to denote the ring R' by $R[x]$. The addition and multiplication that we get from the above definitions are clearly the usual addition and multiplication of polynomials.

Definition 3.3. The ring $R[x]$ is called the *ring of polynomials over R in the indeterminate (or variable) x.*

We want now to study the polynomial ring $R[x]$ in one variable in more details. Let us first recall the following

Definition 3.4. If $f(x) = a_0 + a_1x + \cdots + a_nx^n$ is an element of $R[x]$, then the expression a_ix^i occurring in $f(x)$ is called *monomial* of degree i of $f(x)$; the element $a_i \in R$ is called *coefficient* of the monomial a_ix^i. The *degree* of $f(x)$, denoted by $\deg(f)$, is the greatest i such that $a_i \neq 0$ (i.e. if $f(x) = a_0 + a_1x + \cdots + a_nx^n$ and $a_n \neq 0$, then $\deg(f) = n$).

Remark 3.5. A polynomial has degree 0 if and only if it belongs to $R \setminus \{0\}$. We set $\deg(0) = -\infty$. Two polynomials are equal if and only if they have the same degree and have coefficients respectively equal, i.e. $a_0 + a_1x + \cdots + a_nx^n = b_0 + b_1x + \cdots + b_mx^m$ if and only if $n = m$ and $a_i = b_i$ for all $i = 1, \ldots, n$.

It is immediate to verify:

Proposition 3.6. *If R is an integral domain, $f, g \in R[x]$, then:*

$$\deg(f + g) \leq \max\{\deg(f), \deg(g)\}, \qquad \deg(fg) = \deg(f) + \deg(g). \qquad \square$$

Proposition 3.7. *If R is an integral domain, then $R[x]$ is an integral domain.*

Proof. If $f, g \in R[x]$, $f \neq 0 \neq g$, say $f = a_nx^n + \cdots + a_0$ and $g = b_mx^m + \cdots + b_0$, then $fg = a_nb_mx^{n+m} + \cdots + a_0b_0$, but $a_n \neq 0 \neq b_m$, so $a_nb_m \neq 0$ since R is a domain. $\square$

The following result (also known as *Gauss lemma*) shows that the UFD property, as well as other important properties, is preserved by polynomial extensions.

Theorem 3.8. *If R is a UFD, then the ring of polynomials $R[x]$ is a UFD.*

Proof. See [J], theorem 2.25 or [L], Ch.V, theorem 6.3. $\square$

Let us now give some interesting results regarding polynomial rings over a field.

Proposition 3.9. *Let $f, g \in K[x]$, where K is a field, and suppose $f \neq 0$. Then there exist unique polynomials $q, r \in K[x]$ such that:*

$$g = fq + r$$

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

Proof. By induction on $\deg(g)$. If $\deg(g) = -\infty$, then $g = 0$ and we take $q = r = 0$. If $\deg(g) = 0$, then $g \in K$, and it is easy to choose suitable q and r (namely, if $\deg(f) = 0$, choose $q = 1/f$ and $r = 0$; if $\deg(f) > 0$, choose $q = 0$ and $r = g$).

Hence we can assume $\deg(g) = n > 0$ and the theorem proved for polynomials of degree less than n. If $\deg(f) > \deg(g)$, we take $q = 0$, $r = g$. Otherwise, let

$$f = a_m x^m + \cdots + a_0, \qquad g = b_n x^n + \cdots + b_0$$

where $a_m \neq 0 \neq b_n$, $m \leq n$. Let

$$g_1 := g - b_n a_m^{-1} x^{n-m} f.$$

It is easy to see that $\deg(g_1) < \deg(g)$, hence, by induction, there exist q_1, r_1 such that $g_1 = fq_1 + r_1$ with $\deg(r_1) < \deg(f)$. Then set:

$$q := q_1 - b_n a_m^{-1} x^{n-m}, \qquad r := r_1.$$

In this way we get $g = fq + r$ and $\deg(r) < \deg(f)$, as required.
To see the uniqueness, suppose that

$$g = fq_1 + r_1 = fq_2 + r_2, \qquad \text{where } \deg(r_1), \deg(r_2) < \deg(f).$$

Then $f(q_1 - q_2) = r_2 - r_1$. From 3.6, by comparing the degrees of the polynomials involved, and by recalling that $f \neq 0$, we see that $q_1 = q_2$ and $r_1 = r_2$. □

Definition 3.10. Using the above notations, q is called the *quotient* and r the *remainder* (on dividing g by f).

We are going to mention two kind of consequences of the previous result: on one hand about the divisibility (and zeros) of a polynomial; on the other hand about the ideals in $K[x]$.
Recall that if $f, g \in K[x]$, then f *divides* g ($f|g$) if there exists a polynomial $h \in K[x]$ such that $g = fh$.

Proposition 3.11. *Let $f(x)$ be a non-zero polynomial in $K[x]$. Then $f(x)$ has a zero $a \in K$ (i.e. $f(a) = 0$) if and only if $x - a$ divides $f(x)$.*

Proof. From 3.9 we have that there exist $q, r \in K[x]$ such that $f = q(x - a) + r$ with $\deg(r) < \deg(x - a) = 1$, hence r is a constant. Since $f(a) = 0$, from $f = q(x - a) + r$ we get: $f(a) = 0 = r$, so $x - a$ divides f. The converse is trivial. □

As an easy consequence of this result we get:

Corollary 3.12. *If $f \in K[x]$ is a polynomial of degree n, then it has at most n zeros in K.* □

Theorem 3.13. *If K is a field, then the ring $K[x]$ is a PID.*

Proof. Note first that $K[x]$ is a domain by 3.7. Let now $I \subseteq K[x]$ be an ideal and assume $I \neq (0)$. Let $f \in I$ be a non-zero polynomial of minimum degree. If $g \in I$, then we can divide g by f and, according to 3.9, we get: $g = qf + r$, with $\deg(r) < \deg(f)$. From $r = g - qf$, we see that $r \in I$, and therefore $r = 0$, since its degree is less than the degree of f. This shows that $I = (f)$. □

A polynomial $d \in K[x]$ is a *greatest common divisor* (gcd) of f and g if d divides f and g and further, whenever e divides f and g, then e divides d.

Note that if d is a gcd of $f, g \in K[x]$, and if $a \in K$, $a \neq 0$, then ad is another gcd. Conversely, if d and e are two gcd's, then there exists an $a \in K$, $a \neq 0$ such that $e = ad$.

An immediate consequence of 3.13 is the following:

Corollary 3.14. *Let $f, g \in K[x]$. Then the ideal generated by f and g is the ideal (d), where d is a* gcd *of f and g.*

Proof. From 3.13 we know that $(f, g) = (d)$ for a suitable $d \in K[x]$. From $f \in (d)$ we deduce that $d|f$; analogously $d|g$. Since $d \in (f, g)$, there exist $a, b \in K[x]$ such that $d = af + bg$, hence if e divides f and g, then e divides d, and this shows that d is a gcd of f and g. □

The following method (known as the Euclidean Algorithm) shows that if $f, g \in K[x]$, then there exists a gcd of f and g and, moreover, it is possible to compute it.

Let $f, g \in K[x]$, $f, g \neq 0$. Set $r_{-1} := g$, $r_0 := f$. With the division algorithm we can construct polynomials q_i and r_i as follows:

$$r_j = q_{j+2}r_{j+1} + r_{j+2} \qquad j = -1, 0, 1, \ldots$$

where $\deg(r_0) > \deg(r_1) > \deg(r_2) > \cdots$. Since the degree of the r_j's decreases, we must eventually reach a point where the process stops; and this can only happen when some $r_{s+2} = 0$. Hence $r_s = q_{s+2}r_{s+1}$. □

Proposition 3.15. *With the above notation r_{s+1} is a* gcd *for f and g.*

Proof. First note that r_{s+1} divides f and g. In fact it divides $r_s (= q_{s+2}r_{s+1})$, hence it divides r_{s-1} since $r_{s-1} = q_{s+1}r_s + r_{s+1}$, and so on. Eventually we get r_{s+1} divides r_0, which is f and r_{-1}, which is g.
Suppose e divides f and g. Then it follows inductively that e divides r_i. Hence e divides r_{s+1}. Therefore r_{s+1} is a gcd. □

Remark 3.16. Let $f, g \in K[x]$ be non-zero polynomials, then in 3.14 we saw that any gcd d is the generator of the ideal (f, g). In particular there exist $a, b \in K[x]$ such that $d = af + bg$. It is easy to see that from the computation of the gcd with the Euclidean algorithm one can read off the polynomials a and b.

Let us first remark that, if R is a domain, then the units of $R[x]$ are precisely the elements of R which are units. In particular, if K is any field, the units of $K[x]$ are the non-zero elements of K.
From this observation, accordingly to 1.11, we get that in $K[x]$ a polynomial $f(x)$ is reducible (over K) if and only if it is the product of two polynomials of $K[x]$ of smaller degree. Any polynomial of degree one in $K[x]$ is clearly irreducible. The converse is not true, in general. For instance, $x^2 - 2 \in \mathbb{Q}[x]$ is irreducible; in fact if $x^2 - 2 = (ax+b)(cx+d)$, then $ac = 1$, $ad + bc = 0$, $bd = -2$, and there are no solutions (in $\mathbb{Q}$) to these equations.
Anyway $x^2 - 2$ may be reducible in a suitable polynomial ring. In fact $x^2 - 2 \in \mathbb{Q}[x] \subset \mathbb{R}[x]$ and in $\mathbb{R}[x]$ the equality $x^2 - 2 = (x + \sqrt{2})(x - \sqrt{2})$ holds. Therefore, if $R \subseteq R'$ are two factorial domains, it may happen that an element $a \in R$ is irreducible in R but reducible in R'.
Finally remark that, if $R \subseteq R'$ and R' is a UFD, then R is not necessarily a UFD. In fact $\mathbb{Z}[\sqrt{5}] \subset \mathbb{R}$, $\mathbb{R}$ is obviously a UFD, while $\mathbb{Z}[\sqrt{5}]$ is not a UFD. To see this, consider the factorizations $4 = 2 \cdot 2 = (\sqrt{5} - 1)(\sqrt{5} + 1)$; it is straightforward to verify that 2 is not associate in $\mathbb{Z}[\sqrt{5}]$ neither to $\sqrt{5} - 1$ nor to $\sqrt{5} + 1$.

Using the definition of polynomials in one variable, we can inductively define the ring of polynomials in n variables, by the equality:

$$R[x_1, \ldots, x_n] := R[x_1, \ldots, x_{n-1}][x_n].$$

Hence a polynomial in $R[x_1, \ldots, x_n]$ is an expression like

$$p(x_1, \ldots, x_n) = \sum a_{i_1 \ldots i_n} x_1^{i_1} \cdots x_n^{i_n}, \qquad \text{where } a_{i_1 \ldots i_n} \in R.$$

The *monomials* of $p(x_1, \ldots, x_n)$ are the expressions $a_{i_1 \ldots i_n} x_1^{i_1} \cdots x_n^{i_n}$. The *total degree* of such a monomial is $i_1 + \cdots + i_n$. The *total degree* of a polynomial p is the greatest total degree of its monomials. Of course p can also be considered as a polynomial in the variable x_i (i.e. as an element of $S[x_i]$, where $S = R[x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n]$). The *degree of p w.r.t. x_i* is the degree of p as a polynomial in $S[x_i]$.
An immediate consequence of 3.8 is:

Corollary 3.17. *If K is any field, then the ring of polynomials $K[x_1, \ldots, x_n]$ is a UFD.*

Proof. In fact K is clearly a UFD and to see the result we can use induction, recalling that $K[x_1, \ldots, x_n] = K[x_1, \ldots, x_{n-1}][x_n]$. □

The following result will be frequently applied in next chapter.

Theorem 3.18. *Let $f : R \longrightarrow S$ be a ring homomorphism and fix $\alpha_1, \ldots, \alpha_n \in S$. Then there exists a unique ring homomorphism*

$$\phi : R[x_1, \ldots, x_n] \longrightarrow S$$

such that $\phi|_R = f$ and $\phi(x_i) = \alpha_i$ $(i = 1, \ldots, n)$.

Proof. If $p(x_1, \ldots, x_n) = \sum a_{i_1 \ldots i_n} x_1^{i_1} \cdots x_n^{i_n}$, define

$$\phi(p) := \sum f(a_{i_1 \ldots i_n}) \alpha_1^{i_1} \cdots \alpha_n^{i_n}.$$

Clearly, ϕ is a ring homomorphism and satisfies the requirements. Uniqueness follows straightforward. □

4. Ideals (part II)

As we already remarked, the intersection of a finite number of ideals is still an ideal. We are going to define two other operations between ideals.

Definition 4.1. Let I, J be two ideals of a ring R. The smallest ideal containing them, i.e. $(I \cup J)$, is called the *sum* of I and J and it is denoted by $I + J$. More generally, we can define the sum of any family $\{I_t \mid t \in T\}$ of ideals by: $\sum_{t \in T} I_t := (\cup_{t \in T} I_t)$.
The *product* of two ideals I, J is defined as the ideal generated by the set $\{ab \,|\, a \in I,\ b \in J\}$ and is denoted by IJ.

Remark 4.2. It is clear that $I + J = \{a + b \,|\, a \in I, b \in J\}$.
Moreover a finitely generated ideal $I = (a_1, \ldots, a_n)$ is exactly $(a_1) + \cdots + (a_n)$. In general, $(a_1, \ldots, a_n) + (b_1, \ldots, b_m) = (a_1, \ldots, a_n, b_1, \ldots, b_m)$.
By definition, the elements of IJ are of the form $\sum_i a_i b_i$, where $a_i \in I$, $b_i \in J$.
Note that $IJ \subseteq I \cap J$ and, in general, the opposite inclusion doesn't hold.
As a particular case of the product, we can define the *power* of an ideal by setting $I^0 := R$, $I^1 := I$, $I^2 := II$ and, in general, $I^n := I(I^{n-1})$.

Example 4.2.1 Let us consider the case of the ring $\mathbb{Z}$. If $(m), (n) \subseteq \mathbb{Z}$ are two ideals, then $(m) + (n)$ is the ideal (r), where $r := \gcd(m, n)$, while $(m)(n)$ is the ideal (mn) generated by the product of m and n. Here $(m) \cap (n) = (s)$, where $s = \operatorname{lcm}(m, n)$ is the least common multiple of m and n. Hence $(m) \cap (n) = (m)(n)$ if and only if m and n are coprime.

Let us now introduce two very important notions in ideals theory, mainly relevant from the geometric point of view.
Let us first note that in the ring $\mathbb{Z}$ a prime number p satisfies the following property: if $p|nm$ then either $p|n$ or $p|m$; in terms of ideals, if $nm \in (p)$, then either $n \in (p)$ or $m \in (p)$. This naturally leads to the following

Definition 4.3. An ideal $\mathcal{P}$ of a ring R is called *prime* if it verifies the property:

$$ab \in \mathcal{P} \;\Rightarrow\; a \in \mathcal{P} \text{ or } b \in \mathcal{P}.$$

Proposition 4.4. *Let R be a domain and $a \in R$. If (a) is prime then a is irreducible. If R is a UFD then also the converse is true, i.e. if a is irreducible, then (a) is prime.*

Proof. If $a = 0$, then a is irreducible since R is a domain. So assume that $0 \neq a = bc$; then $bc \in (a)$. By assumption, either $b \in (a)$ or $c \in (a)$; hence either there exists $p \in R$ such that $b = ap$ (so $1 = pc$ i.e. c is a unit) or there exists $q \in R$ such that $c = aq$ (so $1 = bq$ i.e. b is a unit).
Suppose now that R is a UFD and that a is irreducible. If $bc \in (a)$, then $bc = da$ for a suitable $d \in R$. Hence a is one of the irreducible factors in bc; using the UFD hypothesis, a must be a factor of b or of c, so $b \in (a)$ or $c \in (a)$ and (a) is prime. □

Definition 4.5. A proper ideal $\mathcal{M}$ of a ring R is *maximal* if it verifies the property:

$$I \text{ is a proper ideal, } I \supseteq \mathcal{M} \;\Rightarrow\; I = \mathcal{M}.$$

Using Zorn's lemma, it is not difficult to prove that:

Theorem 4.6. *Every ring $R \neq 0$ has at least a maximal ideal.*

Proof. See [AMD], theorem 1.3. □

Corollary 4.7. *If I is a proper ideal of a ring R, then there exists a maximal ideal of R containing I.*

Proof. Apply 4.6 to the ring R/I and take into account 2.5. □

Theorem 4.8. *Let R be a ring. It holds:*
i) An ideal $\mathcal{P}$ is prime if and only if $R/\mathcal{P}$ is a domain.
ii) An ideal $\mathcal{M}$ is maximal if and only if $R/\mathcal{M}$ is a field.
iii) Any maximal ideal is prime.

Proof. i) Assume $\mathcal{P}$ prime and let $[x]$, $[y] \in R/\mathcal{P}$ such that $[x][y] = [0]$. This means $[xy] = [0]$, i.e. $xy \in \mathcal{P}$. By assumption, either $x \in \mathcal{P}$ or $y \in \mathcal{P}$, hence either $[x] = [0]$ or $[y] = [0]$. The converse is analogous.
ii) Assume $\mathcal{M}$ is maximal. Let $[x] \in R/\mathcal{M}$, $[x] \neq [0]$. Therefore $x \notin \mathcal{M}$, so $(x)+\mathcal{M}$ contains $\mathcal{M}$ properly. Then, since $\mathcal{M}$ is maximal, it must be $(x) + \mathcal{M} = R$, so $1 \in (x) + \mathcal{M}$, hence $1 = ax + m$ for some $a \in R$ and $m \in \mathcal{M}$. Taking the classes modulo $\mathcal{M}$, we get $[1] = [a][x]$, i.e. $[x]$ is a unit in $R/\mathcal{M}$. The converse is analogous.
iii) Follows from i) and ii) since any field is a domain. □

Example 4.8.1. 1) In $\mathbb{Z}$ a number p is prime if and only if (p) is prime if and only if (p) is maximal.
2) Let $R := K[x, y, z]$ be the ring of polynomials in x, y, z over K. Then the ideal (x, y) is prime. To see this, consider the map:

$$f : K[x, y, z] \longrightarrow K[z]$$

defined by $x \mapsto 0$, $y \mapsto 0$, $z \mapsto z$ (see 3.18). The map f is an epimorphism. Let us compute its kernel I. It is clear that $(x, y) \subseteq I$. Conversely suppose that $F \in I$. Write F as follows: $F(x, y, z) = xF_1(x, y, z) + yF_2(x, y, z) + F_3(z)$, where $F_3 \in K[z]$. Then, since $f(F) = 0$, we get $F_3 = 0$, so $F \in (x, y)$. Hence $I = (x, y)$. From 2.4 we get:

$$K[x, y, z]/I \cong K[z].$$

Since $K[z]$ is a domain also $K[x, y, z]/I$ is a domain and this shows that (x, y) is a prime ideal. Since $K[z]$ is not a field, (x, y) is not a maximal ideal. In fact $(x, y) \subset (x, y, z) \subset (1)$. Using the same procedure considered above to see that (x, y) is prime, one can show that (x, y, z) is a maximal ideal.

Corollary 4.9. *The ideal (0) of a ring R is prime if and only if R is a domain.* □

As observed in the above example, in general a prime ideal is not maximal. A class of rings in which the converse holds is given by the PID's, where the following result holds:

Theorem 4.10. *Let R be a PID and $a \in R$. Then a is irreducible if and only if (a) is prime if and only if (a) is maximal.*

Proof. By 4.4 and 4.8 iii), it is enough to show that, if R is a PID, then:
a irreducible $\Rightarrow$ (a) maximal.
Assume that (a) is not maximal; then there exists a proper principal ideal of R (being R a PID), say (b), such that $(b) \supset (a)$. So $a = bc$, for some $c \in R$. But a is irreducible, then either b is a unit (hence $(b) = R$) or c is a unit (hence $(a) = (b)$). In both cases we get a contradiction. □

An immediate consequence is:

Corollary 4.11. *If $p(x) \in K[x]$ is an irreducible polynomial over K of degree > 0, then $K[x]/(p)$ is a field.* □

As an application of the above results, we shall conclude this section with some examples. First of all we add the following:

Remark 4.12. If R is any ring, if $f, g \in R[x]$ and if the coefficient of the highest degree monomial of f is a unit (for instance if f is monic), then we can use the procedure explained in the proof of 3.9 to divide g by f and we get $g = fq + r$, where r is a polynomial whose degree is smaller than the degree of f.

Example 4.12.1. Let us verify that the ideal $I := (xy - 1) \subseteq K[x, y]$ is a prime ideal (where K is any field).
From 4.4, it follows that it is enough to check that $xy-1$ is irreducible. Assume $xy-1 = fg$, $f, g \in K[x, y]$, where f and g are not units. We can consider f and g as polynomials in y with coefficients in $K[x]$. From 3.6 it follows that $\deg_y(f) \leq 1$ and $\deg_y(g) \leq 1$. Analogously, $\deg_x(f) \leq 1$ and $\deg_x(g) \leq 1$, therefore we can assume that $f(x, y) = a_0 + a_1x + a_2y$ and $g(x, y) = b_0 + b_1x + b_2y$, hence $fg = xy - 1$ gives (using 3.5) $a_2b_2 = 0$ and $a_1b_1 = 0$. Therefore we can assume $a_2 = 0$. If $a_1 = 0$, then f is a unit, against the assumption. So $b_1 = 0$; from $(a_0 + a_1x)(b_0 + b_2y) = xy - 1$ we get $a_0b_2 = 0$, $a_1b_0 = 0$, $a_0b_0 = 1$, $a_1b_2 = 1$. The first equation gives $a_0 = 0$ (and this contradicts the third equation) or $b_2 = 0$ (and this contradicts the last equation). Hence $xy - 1$ is irreducible and the ideal is prime.

Example 4.12.2. Let us verify that the ideal $I := (y^2 - x^3) \subseteq K[x, y]$ is prime (K any field).
Let us consider the map $\phi : K[x, y] \longrightarrow K[t]$ defined by $\phi(a) := a$ for all $a \in K$, $\phi(x) := t^2$, $\phi(y) := t^3$ and then extend ϕ to a ring homomorphism using 3.18; therefore if $F(x, y) \in K[x, y]$ is any polynomial, then $\phi(F(x, y)) = F(t^2, t^3))$.

Claim: $\ker(\phi) = (y^2 - x^3)$.

Proof (of the claim). First of all it is clear that $(y^2 - x^3) \subseteq \ker(\phi)$. Hence we have to verify only the other inclusion. Take $F(x, y) \in \ker(\phi)$. We can consider F and $(y^2 - x^3)$ as polynomials in y with coefficients in $K[x]$. Since the coefficient of y^2 in $y^2 - x^3$ is 1, so invertible, we can apply 4.12 and divide F by $(y^2 - x^3)$. We get:

$$F(x, y) = (y^2 - x^3)q(x, y) + r(x, y), \tag{1}$$

where $r(x, y)$ is a polynomial of degree at most one in y. So $r(x, y) = r_0(x) + r_1(x)y$ for suitable $r_0 = r_{00} + r_{01}x + r_{02}x^2 + \cdots$, $r_1 = r_{10} + r_{11}x + r_{12}x^2 + \cdots \in K[x]$. If we apply ϕ to (1), we get $r(t^2, t^3) = 0$, so $r_0(t^2) + r_1(t^2)t^3 = 0$. By expanding we get:

$$r_{00} + r_{01}t^2 + r_{10}t^3 + r_{02}t^4 + r_{11}t^5 + r_{03}t^6 + r_{12}t^7 + \cdots = 0$$

and this forces $r_0 = r_1 = 0$ (see 3.5). Hence $F(x, y) \in (y^2 - x^3)$ so the claim is proved.
From 2.4 we get that $K[x, y]/\ker(\phi)$ is isomorphic to a subring of $K[t]$. Since $K[t]$ is a domain, $K[x, y]/(y^2 - x^3)$ is a domain, hence $(y^2 - x^3)$ is prime.

Example 4.12.3. Let us verify that $I := (y - x^2, z - x^3) \subseteq K[x, y, z]$ is a prime ideal (K is any field).
Using the same procedure seen in the previous example, we can consider the map

$$\phi : K[x, y, z] \longrightarrow K[t]$$

defined by $\phi(a) := a$ for every $a \in K$, $\phi(x) := t$ $\phi(y) := t^2$ and $\phi(z) := t^3$. Let us verify that $\ker(\phi) = (y - x^2, z - x^3)$. Take $F(x, y, z) \in \ker(\phi)$. Then we can divide, in $K[x, z][y]$, F by $y - x^2$ and we get:

$$F(x, y, z) = (y - x^2)q_1(x, y, z) + r_1(x, z).$$

Then we can divide, in $K[x][z]$, $r_1(x, z)$ by $(z - x^3)$ and we get $r_1(x, z) = (z - x^3)q_2(x, z) + r_2(x)$. Since $\phi(F) = 0$, we get $\phi(r_2(x)) = 0$, therefore $r_2 = 0$ and we get $F \in (y - x^2, z - x^3)$. Since the other inclusion is obvious, we see that $\ker(\phi) = (y - x^2, z - x^3)$. Therefore $K[x, y, z]/\ker(\phi) \cong K[t]$, so the given ideal is prime.

Example 4.12.4. If K is a field and $a_1, \ldots, a_n \in K$, then the ideal

$$(x_1 - a_1, \ldots, x_n - a_n) \subseteq K[x_1, \ldots, x_n]$$

is maximal.
Proof. Let $\phi : K[x_1, \ldots, x_n] \longrightarrow K$ be defined by $\phi(x_i) := a_i$ and $\phi(k) := k$ for every $k \in K$. Then ϕ is surjective and clearly $(x_1 - a_1, \ldots, x_n - a_n) \subseteq \ker(\phi)$. Suppose conversely that $F(x_1, \ldots, x_n) \in \ker(\phi)$. Since in the polynomial $x_n - a_n$ the coefficient of x_n is a unit, we can divide F by $x_n - a_n$ and we get:

$$F(x_1, \ldots, x_n) = (x_n - a_n)Q_1(x_1, \ldots, x_n) + F_1(x_1, \ldots, x_{n-1}).$$

(Note that the remainder is a polynomial in $x_1, \ldots, x_{n-1}$ only, since its degree in x_n must be zero).
Now we can divide F_1 by $x_{n-1} - a_{n-1}$ and we get: $F_1 = (x_{n-1} - a_{n-1})Q_2 + F_2$, where F_2 is a polynomial in $x_1, \ldots, x_{n-2}$ only. Then we can divide F_2 by $x_{n-2} - a_{n-2}$ and so on. After n divisions we get:

$$F = (x_n - a_n)Q_1 + (x_{n-1} - a_{n-1})Q_2 + \cdots + (x_1 - a_1)Q_n + R,$$

where R is an element of K. Applying ϕ, we get $R = 0$, so $F \in (x_1 - a_1, \ldots, x_n - a_n)$. Therefore

$$\ker(\phi) = (x_1 - a_1, \ldots, x_n - a_n)$$

and since $K[x_1, \ldots, x_n]/\ker(\phi) \cong K$, the ideal $(x_1 - a_1, \ldots, x_n - a_n)$ is maximal (see 4.8).

Example 4.12.5. The ideal $(x^2 + 2x - 1, y^2 - 2y - 1) \subseteq \mathbb{Q}[x, y]$ is not prime.
Proof. First of all note that

$$\mathbb{Q}[x]/(x^2 + 2x - 1) \cong \mathbb{Q}[\sqrt{2}].$$

To see this, consider the map $\phi : \mathbb{Q}[x] \longrightarrow \mathbb{Q}[\sqrt{2}]$ defined by $\phi(a) := a$ for all $a \in \mathbb{Q}$ and $\phi(x) := -1 + \sqrt{2}$. Then take $F(x) \in \ker(\phi)$; using 3.9 we get: $F(x) = (x^2 + 2x - 1)Q(x) + R(x)$, where $\deg(R) \leq 1$, i.e. $R(x) = a + bx$, $a, b \in \mathbb{Q}$. Applying ϕ we get $R(-1+\sqrt{2}) = 0 = a + (-1 + \sqrt{2})b$ which can be satisfied only if $a = b = 0$ since a and b are rational numbers. Hence $R = 0$. In this way we see that the ring $S := \mathbb{Q}[x]/(x^2 + 2x - 1)$ is isomorphic to $\mathbb{Q}[\sqrt{2}]$. It is not hard to see that $\mathbb{Q}[x, y]/(x^2 + 2x - 1, y^2 - 2y - 1) \cong S[y]/(y^2 - 2y - 1)$, hence

$$\mathbb{Q}[x, y]/(x^2 + 2x - 1, y^2 - 2y - 1) \cong \mathbb{Q}[\sqrt{2}][y]/(y^2 - 2y - 1)$$

and $y^2 - 2y - 1 = (y - 1 - \sqrt{2})(y - 1 + \sqrt{2})$ in $\mathbb{Q}[\sqrt{2}][y]$, so it is reducible. Therefore $\mathbb{Q}[\sqrt{2}][y]/(y^2 - 2y - 1)$ is not a domain and the ideal $(x^2 + 2x - 1, y^2 - 2y - 1)$ is not prime (see 4.8).

5. Noetherian Rings

We saw some examples of rings where all the ideals are principal (like $\mathbb{Z}$, the ring of polynomials $K[x]$ over a field K, the rings $\mathbb{Z}_n$). Note however that this property is quite strong, and in particular, is not preserved by polynomial extensions (for examples $K[x, y] = K[x][y]$ is not a PID anymore). A natural generalization of 'principal ideal' is that of 'finitely generated ideal'. In this section we shall study the class of rings where all the ideals are finitely generated and we shall see that this property, as well as other nice properties (like being a domain, or a UFD) is preserved under polynomial extensions.

Let R be a commutative ring. It holds:

Proposition 5.1. *The following conditions are equivalent:*

i) Every chain of ideals

$$I_1 \subseteq I_2 \subseteq I_3 \subseteq \cdots$$

is stationary (i.e. there exists $n \in \mathbb{N}$ such that $I_n = I_{n+1} = \cdots$).

ii) If Σ is any non-empty subset of ideals of R, then Σ has a maximal element (i.e. there exists $J \in \Sigma$ such that if $I \in \Sigma$, $I \supseteq J$, then $I = J$).

Proof. i) $\Rightarrow$ ii) Let Σ be any non-empty subset of ideals of R and let $I_1 \in \Sigma$. If I_1 is not maximal, then there exists an element $I_2 \in \Sigma$ such that $I_1 \subseteq I_2$. If I_2 is not maximal, then there exists $I_3 \in \Sigma$ such that $I_1 \subseteq I_2 \subseteq I_3$. From i) it follows that after a finite number of steps we pick up a maximal element of Σ.
ii) $\Rightarrow$ i) Let $I_1 \subseteq I_2 \subseteq I_3 \subseteq \cdots$ be an increasing sequence of ideals; set $\Sigma := \{I_1, I_2, I_3, \ldots\}$. Then, by hypothesis, there exists a maximal element I_n for Σ. It is clear that the sequence of the I_i's is stationary for $i > n$. □

Definition 5.2. A ring R is *noetherian* if it satisfies one (and hence both) of the conditions of the above proposition.

Examples 5.2.1. 1) If K is a field, then it is a noetherian ring (since it has only two ideals).

2) The ring of integers $\mathbb{Z}$ is noetherian. In fact we know that all the ideals of $\mathbb{Z}$ are principal. If $(a) \subseteq (b)$, then b divides a, and from this it is immediate to see that every chain of ideals is stationary.
3) The ring $K[x]$ of polynomials in one indeterminate over a field K is noetherian (the proof is analogous to the previous one given for $\mathbb{Z}$ since also $K[x]$ is a PID and a UFD). A more general result will be proved in 5.5.
4) Let $R := \{f : \mathbb{Z} \longrightarrow \mathbb{Z}\}$ (the addition and the multiplication in R are defined pointwise). Then R is not noetherian. To see this, let us consider the functions:

$$\delta_k(n) := \begin{cases} 0 \text{ if } n \neq k \\ 1 \text{ if } n = k. \end{cases} \qquad k = 1, 2, \dots$$

Let $I_k := (\delta_1, \dots, \delta_k)$. We have $\delta_{k+1} \notin I_k$ (if $\delta_{k+1} = \sum_{j=1}^{k} u_j \delta_j$ for suitable $u_j \in R$, then $\delta_{k+1}(k+1) = \sum_{j=1}^{k} u_j \delta_j(k+1) = 0$). Therefore the following is an infinite chain of ideals:

$$I_1 \subset I_2 \subset I_3 \subset \cdots$$

Recall (see 2.7) that an ideal I of a ring R is finitely generated if there exist elements $a_1, \dots, a_n \in I$ such that $I = (a_1, \dots, a_n)$. The following is the main characterization of noetherian rings:

Theorem 5.3. *A ring R is noetherian if and only if every ideal of R is finitely generated.*

Proof. Let $I \subseteq R$ be an ideal. Take $\Sigma := \{J \subseteq I \mid J \text{ is finitely generated}\}$. The set Σ is not empty (for instance (0) is in Σ), hence it has a maximal element N. If $N \neq I$, then take $a \in I \setminus N$. Then the ideal (N, a) should be in Σ but contains N properly. Hence $I = N$ and so I is finitely generated.
To see the converse, let $I_1 \subseteq I_2 \subseteq I_3 \subseteq \cdots$ be a chain of ideals. Set $I := \cup_{j=1}^{\infty} I_j$. Then I is finitely generated, say $I = (f_1, \dots, f_m)$. Hence there exists an n such that $f_1, \dots, f_m \in I_n$. Therefore $I_n = I_{n+1} = \cdots$. □

Proposition 5.4. *If R is a noetherian ring and $I \subseteq R$ is an ideal, then the ring R/I is noetherian.*

Proof. This result follows immediately from the inclusion-preserving bijection that there is between the ideals of R containing I and the ideals of R/I (see 2.5). □

A fundamental theorem regarding noetherian rings is the following:

Theorem 5.5. *(Hilbert Basis theorem) If R is a noetherian ring, then the ring of polynomials $R[x]$ is noetherian.*

Proof. Suppose $R[x]$ is not noetherian. Then there exists an ideal $I \subseteq R[x]$ which is not finitely generated. We can recursively define a sequence of polynomials $f_1, f_2, \dots$ as follows:
- take $f_1 \in I$ such that $\deg(f_1)$ is minimal;
- suppose $f_1, \dots, f_k$ are already defined; then define f_{k+1} as a polynomial of minimal degree among those in $I \setminus (f_1, \dots, f_k)$.

The sequence $f_1, f_2, \ldots$ can be found, since I is not finitely generated. Let $\deg(f_k) = n_k$ and let $f_k = a_k x^{n_k} + \textit{terms of lower degree}$. From the choice of the f_k's, we have $n_1 \leq n_2 \leq n_3 \leq \cdots$.
Let us consider the following chain of ideals of R:

$$(a_1) \subseteq (a_1, a_2) \subseteq (a_1, a_2, a_3) \subseteq \cdots$$

Since R is noetherian, this chain is stationary, hence there exists k such that $a_{k+1} = \sum_{i=1}^{k} u_i a_i$ $(u_i \in R)$. Let:

$$g := f_{k+1} - \sum_{i=1}^{k} u_i x^{n_{k+1}-n_i} f_i.$$

We have: $g \notin (f_1, \ldots, f_k)$ (since $f_{k+1} \notin (f_1, \ldots, f_k)$). But $g \in I$ and $\deg(g) < n_{k+1}$, hence we get a contradiction. □

Corollary 5.6. *If R is a noetherian ring, then the ring of polynomials $R[x_1, \ldots, x_n]$ is noetherian.*

Proof. It follows by induction, recalling that $R[x_1, \ldots, x_n] = R[x_1, \ldots, x_{n-1}][x_n]$. □

Remark 5.7. In particular, from the above corollary, we get that if K is any field, then $K[x_1, \ldots, x_n]$ is a noetherian ring. This result has an important geometric meaning: it claims that any set of zeros of a family of polynomials in the affine space $\mathbb{A}^n_K$ can also be described as the zeros of a *finite* set of polynomials.

Chapter III
Galois Theory

1. Preliminaries

If K is a field, recall that the *characteristic* of K is either a positive integer (in this case it is the smallest $n \in \mathbb{N}$ such that $n \cdot 1 = 0$) or it is 0 (if $n \cdot 1 \neq 0$ for all $n \in \mathbb{N}$). Since a field is a domain, its characteristic, if not zero, is a prime number (see Ch.II, 2.10). For instance $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ are fields of characteristic zero, while for each prime p, $\mathbb{Z}_p$ is a field of characteristic p. In general, if K is a field of characteristic zero, then it contains (an isomorphic copy of) $\mathbb{Q}$, since it contains an isomorphic copy of $\mathbb{Z}$ (see Ch.II, 2.9); if $\mathrm{char}\,(K) = p$, then it contains $\mathbb{Z}_p$.

We recall the following result:

Theorem 1.1. *(Fundamental theorem of algebra) If $f(x) = \sum_{i=0}^{n} a_i x^i \in \mathbb{C}[x]$ is a polynomial of degree $n > 0$, then there exists $\alpha \in \mathbb{C}$ such that $f(\alpha) = 0$.*

For the proof see, for instance, [A] (thm.1.3.4) or [L1] (Ch.II, thm.6.3 and Ch.V, section 1) or [M] (Ch.III, ex. 3.6). □

Corollary 1.2. *If $f(x) = \sum_{i=0}^{n} a_i x^i \in \mathbb{C}[x]$ is a polynomial of degree $n > 0$, then there exist $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ (not necessarily all distinct) such that $f(x) = a_n(x - \alpha_1) \cdots (x - \alpha_n)$.*

Proof. It is an immediate consequence of 1.1 and of Ch.II, 3.11. □

Remark 1.3. If K and L are fields, then a ring homomorphism $f : K \longrightarrow L$ is also a field homomorphism, in the sense that $f(k^{-1}) = f(k)^{-1}$ for all $k \in K \setminus \{0\}$. In the sequel we shall call such a homomorphism simply a *(field) morphism*. Furthermore, note that a morphism $f : K \longrightarrow L$ is always injective, hence we can consider K as a subfield of L.

Definition 1.4. If $i : K \longrightarrow L$ is a morphism, then L is said a *field extension* of K. If L is a field extension of K, with an abuse of notation, we shall usually write $K \subseteq L$.
If K is a field and X is any subset of K, the *field generated* by X is the intersection of all the subfields of K containing X. It is the smallest subfield of K containing X.

Example 1.4.1. Let $X := \{1, \mathrm{i}\} \subseteq \mathbb{C}$, (where $\mathrm{i}^2 = -1$), then it is easy to see that the field generated by X is the field $\{a + \mathrm{i}b \mid a, b \in \mathbb{Q}\}$.

Let us now introduce a kind of field extensions, which will be particularly important in this chapter.

Definition 1.5. Let $K \subseteq L$ be a field extension and let Y be any subset of L; the field generated by $X := K \cup Y$ is usually denoted by $K(Y)$ and it is called *subfield of L generated by Y over K*. Note that $K \subseteq K(Y) \subseteq L$.

For instance if $K = \mathbb{Q}$, $L = \mathbb{R}$ and $Y = \{\sqrt{2}\}$, it is easy to see that $\mathbb{Q}(\sqrt{2}) = \{a + \sqrt{2}b \mid a, b \in \mathbb{Q}\}$. If $K = \mathbb{R}$, $L = \mathbb{C}$ and $Y = \{\mathrm{i}\}$, then $\mathbb{R}(\mathrm{i}) = \mathbb{C}$.

Definition 1.6. A *simple extension* of a field K is an extension $K \subseteq L$ such that $L = K(\alpha)$, for a suitable $\alpha \in L$.

Let us recall that in Ch.II, section 3 we defined $K[\alpha]$ to be the smallest subring of L containing K and α; analogously here $K(\alpha)$ is the smallest subfield of L containing K and α. In general $K[\alpha] \subseteq K(\alpha)$; more precisely $K(\alpha)$ consists of the quotients $f(\alpha)/g(\alpha)$ of polynomial expressions $f(\alpha), g(\alpha) \in K[\alpha]$ $(g(\alpha) \neq 0)$. We shall characterize those α's for which $K[\alpha]$ is a field, hence it coincides with $K(\alpha)$.

Example 1.6.1. It is not difficult to verify that $\mathbb{Q}(\sqrt{2}, \sqrt{3}) = \mathbb{Q}(\sqrt{2}+\sqrt{3})$; here $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ and $\mathbb{Q}(\sqrt{2}+\sqrt{3})$ are considered subfields of $\mathbb{R}$.
In fact $\sqrt{2}+\sqrt{3} \in \mathbb{Q}(\sqrt{2}, \sqrt{3})$, so $\mathbb{Q}(\sqrt{2}+\sqrt{3}) \subseteq \mathbb{Q}(\sqrt{2}, \sqrt{3})$. Conversely, we want to see that $\sqrt{2}, \sqrt{3} \in \mathbb{Q}(\sqrt{2}+\sqrt{3})$. First note that $\sqrt{6} = 1/2(\sqrt{2}+\sqrt{3})^2 - 5/2 \in \mathbb{Q}(\sqrt{2}+\sqrt{3})$, hence $(\sqrt{2}+\sqrt{3})\sqrt{6} \in \mathbb{Q}(\sqrt{2}+\sqrt{3})$. So $3\sqrt{2}+2\sqrt{3} \in \mathbb{Q}(\sqrt{2}+\sqrt{3})$, therefore $\sqrt{2} = (3\sqrt{2}+2\sqrt{3}) - 2(\sqrt{2}+\sqrt{3}) \in \mathbb{Q}(\sqrt{2}+\sqrt{3})$ and $\sqrt{3} = 3(\sqrt{2}+\sqrt{3}) - (3\sqrt{2}+2\sqrt{3}) \in \mathbb{Q}(\sqrt{2}+\sqrt{3})$.
In particular, this example shows that $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ is a simple extension of $\mathbb{Q}$, although it is not given with the notation of a simple extension.

Definition 1.7. Let $K \subseteq L$ be an extension. An element $\alpha \in L$ is *algebraic* over K if there exists a polynomial $f(x) \in K[x]$ such that $f(\alpha) = 0$. Otherwise α is *transcendental* over K.

For instance in the extension $\mathbb{Q} \subseteq \mathbb{R}$, $\sqrt{3} \in \mathbb{R}$ is algebraic over $\mathbb{Q}$ since the polynomial $x^2 - 3 \in \mathbb{Q}[x]$ has $\sqrt{3}$ as a zero. It is possible to prove that the real numbers e and π are transcendental over $\mathbb{Q}$ (see, for instance, [S] Ch.6).

Remark 1.8. One can show that the set $\{\alpha \in \mathbb{R} \mid \alpha \text{ is algebraic over } \mathbb{Q}\}$ is numerable. Hence in particular, its complement in $\mathbb{R}$, which is the set of transcendental elements over $\mathbb{Q}$, has the same cardinality of $\mathbb{R}$ itself.

An algebraic element is strictly related to a particular polynomial, as follows. Let $K \subseteq L$ and let $\alpha \in L$ be algebraic over K. Then, among the polynomials $f(x) = \sum_{i=0}^{n} a_i x^i$ admitting α as a zero, we can choose a polynomial of minimum degree and monic (i.e. with $a_n = 1$). From the proposition below, it follows that this polynomial is unique.

Proposition - Definition 1.9. *Let $K \subseteq L$, $\alpha \in L$ be algebraic over K and let $p(x) \in K[x]$ be a monic polynomial such that $p(\alpha) = 0$ and $\deg(p)$ is minimum. Then:*
i) $p(x)$ is unique; it is called the minimum polynomial of α over K;
ii) $p(x)$ is irreducible and divides every polynomial in $K[x]$ which has α as a zero;
iii) in particular, if a polynomial is monic, irreducible and has α as a zero, then it is the minimum polynomial of α over K.

Proof. i) If $p(x)$ and $q(x)$ are both monic, of minimum degree and with α as a zero, then the polynomial $h(x) := p(x) - q(x)$, is of degree lower than $\deg(p) = \deg(q)$ and has α as a zero, hence is the zero polynomial. Therefore $p(x) = q(x)$.
ii) If $f(x)$ is such that $f(\alpha) = 0$, then we can divide f by p and we get: $f = ps + r$ with $\deg(r) < \deg(p)$ (see Ch.II, 3.9). Since $f(\alpha) = 0$, also $r(\alpha) = 0$, hence $r(x) = 0$. Therefore p divides any polynomial having α as zero. It follows immediately that p is irreducible.
iii) Let $q(x)$ be a monic, irreducible polynomial having α as a zero; then, by ii), $p(x)$ divides

$q(x)$. But $q(x)$ is irreducible, so necessarily $q(x) = a \cdot p(x)$, where a is a unit, i.e. $a \in K$. Since both $q(x)$ and $p(x)$ are monic, we get $a = 1$. □

Example 1.9.1. Let $\alpha = \sqrt{1+\sqrt{5}} \in \mathbb{R}$. Let us verify that α is algebraic over $\mathbb{Q}$, and let us find its minimum polynomial. Since $\alpha^2 = 1 + \sqrt{5}$, we get $(\alpha^2 - 1)^2 = 5$, i.e. $\alpha^4 - 2\alpha^2 - 4 = 0$. Let $p(x) := x^4 - 2x^2 - 4 \in \mathbb{Q}[x]$, then α is a root of $p(x)$, so it is algebraic over $\mathbb{Q}$. Recalling that a polynomial is irreducible if and only if it cannot be expressed as product of two polynomials of lower degree (see Ch.II, section 3), we can directly verify that $p(x)$ is irreducible over $\mathbb{Q}$, then from 1.9 we have that p is the minimum polynomial of α. Of course we could get other polynomials over $\mathbb{Q}$ which admit α as a root. For instance, from $\alpha = \sqrt{1+\sqrt{5}}$, we get: $\alpha^4 = (1+\sqrt{5})^2$, i.e. $(\alpha^4 - 6) = 2\sqrt{5}$, so $(\alpha^4 - 6)^2 = 20$. Then if we define $f(x) := x^8 - 12x^4 + 16$, α is a root of $f(x)$. It is immediate to verify that, according to 1.9, p divides f.

Let $K \subseteq L$ be an extension and let $K[x]$ be the ring of polynomials over K. If $\alpha \in L$, we can define a (ring) homomorphism

$$\phi_\alpha : K[x] \longrightarrow L$$

by $\phi_\alpha(k) := k$, for every $k \in K$, $\phi_\alpha(x) := \alpha$ (see Ch.II, 3.18). The image of ϕ_α is $K[\alpha]$, the smallest subring of L containing K and α, defined in Ch.II, section 3. It is clear that α is transcendental over K iff $\ker(\phi_\alpha) = (0)$. If α is algebraic then $\ker(\phi_\alpha)$ is a principal (non-zero) ideal of $K[x]$ (see Ch.II, 3.13); moreover $p(x)$ (the minimum polynomial of α over K) divides every element of $\ker(\phi_\alpha)$, by 1.9. Hence $\ker(\phi_\alpha) = (p(x))$.

Example 1.9.2. Let us consider $\mathbb{Q}[\sqrt{3}] \subseteq \mathbb{R}$. Take the map $\phi : \mathbb{Q}[x] \longrightarrow \mathbb{Q}[\sqrt{3}]$ defined by $\phi(k) := k$ for every $k \in \mathbb{Q}$, $\phi(x) := \sqrt{3}$, hence

$$\phi(a_0 + a_1 x + \cdots + a_m x^m) = a_0 + a_1\sqrt{3} + \cdots + a_m(\sqrt{3})^m.$$

Obviously ϕ is surjective; moreover $\ker(\phi) = (x^2 - 3)$, hence $\mathbb{Q}[x]/(x^2-3)$ is a field, and so $\mathbb{Q}[\sqrt{3}]$ is a field. It is also easy to directly verify that $\mathbb{Q}[\sqrt{3}]$ is a field: first of all observe that $\mathbb{Q}[\sqrt{3}] = \{a + \sqrt{3}b \mid a, b \in \mathbb{Q}\}$; if $a + \sqrt{3}b \in \mathbb{Q}[\sqrt{3}]$ and $a + \sqrt{3}b \neq 0$ (i.e. $(a, b) \neq (0, 0)$), then it is invertible in $\mathbb{Q}[\sqrt{3}]$, since $(a + \sqrt{3}b)^{-1} = a/(a^2 - 3b^2) - b/(a^2 - 3b^2)\sqrt{3}$.

Using the above observation and recalling that in a PID an irreducible element generates a maximal ideal (see Ch.II, 4.10), we get the following result:

Proposition 1.10. *Let $K \subseteq L$ be an extension and $\alpha \in L$. Then:*

i) if α is transcendental over K, then $K[\alpha] \cong K[x]$;

ii) if α is algebraic over K, then $K[\alpha]$ is a field; in particular $K[\alpha] = K(\alpha)$. Conversely, if $K[\alpha]$ is a field, then α is algebraic over K.

Proof. Since $K[\alpha] = \mathrm{Im}(\phi_\alpha) \cong K[x]/\ker(\phi_\alpha)$, we get immediately i).
If α is algebraic and p is the minimum polynomial of α, then it is irreducible, hence $K[\alpha] \cong K[x]/(p)$ is a field. Since $K(\alpha)$ is the smallest subfield of L containing K and α, then $K[\alpha] = K(\alpha)$. If conversely, $K[\alpha]$ is a field and α were not algebraic (i.e. transcendental), then by i) $K[\alpha]$ would be isomorphic to the ring of polynomials $K[x]$, which is not a field (for instance x is not invertible). □

Recall that if $K \subseteq L$ is an extension and $\alpha_1, \ldots, \alpha_n \in L$, then $K[\alpha_1, \ldots, \alpha_n]$ is the smallest subring of L which contains K and $\alpha_1, \ldots, \alpha_n$ (see Ch.II, section 3). The result in 1.10 can be generalized for any set $\{\alpha_1, \ldots, \alpha_n\}$ of algebraic elements:

Proposition 1.11. *If $K \subseteq L$ is a field extension and if $\alpha_1, \ldots, \alpha_n \in L$ are algebraic over K, then $K[\alpha_1, \ldots, \alpha_n] = K(\alpha_1, \ldots, \alpha_n)$.*

Proof. We can use induction on n. If $n = 1$, the result is given by 1.10. If $n > 1$, then, by induction, $K[\alpha_1, \ldots, \alpha_{n-1}] = K(\alpha_1, \ldots, \alpha_{n-1})$, hence we have:

$$K[\alpha_1, \ldots, \alpha_n] = K[\alpha_1, \ldots, \alpha_{n-1}][\alpha_n] = K(\alpha_1, \ldots, \alpha_{n-1})[\alpha_n] = K(\alpha_1, \ldots, \alpha_n).$$

The last equality follows again from 1.10, since α_n is algebraic over $K(\alpha_1, \ldots, \alpha_{n-1})$. □

Notation 1.12. If $K \subseteq L$ is a field extension, then L can be considered as a vector space over K (if $v \in L$, $\lambda \in K$, then λv is simply the product of the two elements as elements of L). The dimension of L as a K vector space is denoted by $[L : K]$.

Examples 1.12.1. $[\mathbb{C} : \mathbb{R}] = 2$ (in fact 1 and i, where $\mathrm{i}^2 = -1$, is a basis); $[\mathbb{Q}(\pi) : \mathbb{Q}] = \infty$; in fact 1, π, $\pi^2, \ldots \pi^n, \ldots$ are linearly independent, since π is transcendental over $\mathbb{Q}$.

Theorem 1.13. *If $K \subseteq L \subseteq M$ are fields, then*

$$[M : K] = [M : L] \cdot [L : K].$$

Proof. (We consider only the case $[M : K] < \infty$, although the proof for the general case is also straightforward). If $x_1, \ldots, x_n$ is a basis of L over K and if $y_1, \ldots, y_m$ is a basis of M over L, it is easy to verify that the mn elements $x_i y_j$, $i = 1, \ldots, n$, $j = 1, \ldots, m$ are a basis for M over K. □

Theorem 1.14. *Let $K \subseteq K(\alpha)$ be a simple extension. Then α is transcendental over K iff $[K(\alpha) : K] = \infty$. If α is algebraic, then $[K(\alpha) : K] = \deg(p)$, where p is the minimum polynomial of α over K.*

Proof. If α is transcendental, then $1, \alpha, \alpha^2, \ldots, \alpha^n, \ldots$ are linearly independent (otherwise a linear dependency relation among $1, \alpha, \alpha^2, \ldots, \alpha^n$ would give a polynomial in $K[x]$ vanishing on α). So the dimension of $K(\alpha)$ is infinite. If α is algebraic, let $p(x)$ be the minimum polynomial of α over K and let $n = \deg(p)$. Then $1,\ \alpha, \alpha^2, \ldots, \alpha^{n-1}$ are linearly independent over K. Take now any element f of $K[\alpha]$ $(= K(\alpha)$ by 1.10), $f = u_0 + \cdots + u_m \alpha^m$. Let $\tilde{f}(x) := u_0 + \cdots + u_m x^m \in K[x]$. Then, dividing $\tilde{f}$ by p we get $\tilde{f}(x) = p(x)q(x) + r(x)$, with $\deg(r(x)) < \deg(p(x))$. Then $\tilde{f}(\alpha) = f = r(\alpha)$ and $r(\alpha)$ is a linear combination of $1,\ \alpha, \ldots, \alpha^{n-1}$. So $1,\ \alpha, \ldots, \alpha^{n-1}$ is a basis of $K(\alpha)$ over K. □

Example 1.14.1. Let us consider $\mathbb{Q}[\sqrt[3]{2}] \subseteq \mathbb{R}$. Then $\sqrt[3]{2}$ is algebraic over $\mathbb{Q}$ and its minimum polynomial is $x^3 - 2$. From the proof of 1.14 we have that 1, $\sqrt[3]{2}$, $\sqrt[3]{4}$ are a basis of $\mathbb{Q}[\sqrt[3]{2}]$ over $\mathbb{Q}$, hence $\mathbb{Q}[\sqrt[3]{2}] = \{a + \sqrt[3]{2}b + \sqrt[3]{4}c \mid a, b, c \in \mathbb{Q}\}$. In the same way, we reobtain the representation for the elements of $\mathbb{Q}[\sqrt{3}]$ considered in 1.9.2.

Definition 1.15. The extension $K \subseteq L$ is a *finite extension* if $[L : K] < \infty$. In this case, if $[L : K] = n$, then n is the *degree* of the extension.

Clearly, from the above theorem, if α is algebraic over K, then $K(\alpha)$ is a finite extension. A more general result holds:

Theorem 1.16. *If $\alpha_1, \ldots, \alpha_n$ are algebraic over K, then $K[\alpha_1, \ldots, \alpha_n]$ is a finite extension of K.*

Proof. First note that $K[\alpha_1, \ldots \alpha_i]$ is a field for all $i = 1, \ldots, n$ by 1.11. Since α_{i+1} is algebraic over K, then it is algebraic over $K[\alpha_1, \ldots, \alpha_i]$; hence $[K[\alpha_1, \ldots, \alpha_{i+1}] : K[\alpha_1, \ldots, \alpha_i]]$ is finite by 1.14. Therefore

$$[K[\alpha_1, \ldots, \alpha_n] : K] = [K[\alpha_1, \ldots, \alpha_n] : K[\alpha_1, \ldots, \alpha_{n-1}]] \cdots [K[\alpha_1] : K]$$

by 1.13, so it is finite. □

Definition 1.17. An extension $K \subseteq L$ is *algebraic* if every element of L is algebraic over K.

Example 1.17.1. It is easy to verify that $\mathbb{Q} \subseteq \mathbb{Q}(\sqrt{3})$ is an algebraic extension of $\mathbb{Q}$, with a direct computation. In fact from 1.9.2 it follows that $\mathbb{Q}(\sqrt{3}) = \mathbb{Q}[\sqrt{3}]$ therefore any element of $\mathbb{Q}(\sqrt{3})$ is of the kind $a + \sqrt{3}b$ $(a, b \in \mathbb{Q})$. Hence fix an element $a + \sqrt{3}b$ in $\mathbb{Q}(\sqrt{3})$. Then let $f(x) := x^2 - 2ax + a^2 - 3b^2 \in \mathbb{Q}[x]$. Since $f(a + \sqrt{3}b) = 0$, any element of $\mathbb{Q}(\sqrt{3})$ is algebraic over $\mathbb{Q}$, and the extension is algebraic.
This argument essentially comes from the fact that $\sqrt{3}$ is algebraic over $\mathbb{Q}$.
More generally we have:

Theorem 1.18. *If $K \subseteq L$ is a finite extension, then it is an algebraic extension. In particular, if $\alpha_1, \ldots, \alpha_n$ are algebraic over K, then $K[\alpha_1, \ldots, \alpha_n]$ is an algebraic extension of K.*

Proof. If $\alpha \in L$ and if $[L : K] = n$, then $1, \alpha, \ldots, \alpha^n$ are $n + 1$ elements in L, so must be linearly dependent over K. Hence there exist $\lambda_0, \ldots, \lambda_n \in K$ such that the polynomial $f(x) := \lambda_0 + \cdots + \lambda_n x^n \in K[x]$ has α as a zero. This shows that α is algebraic over K. In particular, since $K[\alpha_1, \ldots, \alpha_n]$ is a finite extension of K by 1.16, then it is an algebraic extension of K. □

Definition 1.19. Let $A := \{\alpha \in \mathbb{C} \mid \alpha \text{ is algebraic over } \mathbb{Q}\}$. The elements of A are called *algebraic numbers*.

Using the above results, we can show that A is a field. In fact if $\alpha, \beta \in A$, then $\mathbb{Q}[\alpha, \beta]$ is a finite algebraic extension by 1.18. Therefore, $\mathbb{Q}[\alpha, \beta]$ is a field and $\mathbb{Q}[\alpha, \beta] \subseteq A$. So $\alpha - \beta$ and $\alpha\beta^{-1} \in \mathbb{Q}[\alpha, \beta] \subseteq A$.

2. Extensions of Morphisms

Warning. In this section, as well as in the next ones, we shall restrict the attention to fields which are subfields of $\mathbb{C}$. This is done in order to simplify several proofs. Note however that most of the results established here still hold in the general case.

Let $f(x) \in K[x]$, $f(x) := a_0 + a_1 + \cdots + a_n x^n$ be a polynomial over a field K. The *formal derivative* of f is the polynomial

$$Df(x) := a_1 + 2a_2 x + \cdots + n a_n x^{n-1}.$$

It is a straightforward verification to see that if f, g are polynomials over K, then $D(f+g) = Df + Dg$, $D(fg) = D(f)g + fD(g)$; if $\lambda \in K$, then $D(\lambda) = 0$, and $D(\lambda f) = \lambda D(f)$.

Proposition 2.1. *If $f(x) \in K[x]$ is an irreducible polynomial over $K \subseteq \mathbb{C}$ of degree n, then it has n distinct roots in $\mathbb{C}$.*

Proof. Let $f(x)$ be irreducible over $K[x]$ and assume that it has a multiple root in $\mathbb{C}$, this means that $f(x) = (x-\alpha)^2 g(x)$, for a suitable $\alpha \in \mathbb{C}$ (clearly this factorization of $f(x)$ occurs in $\mathbb{C}[x]$, while f is irreducible in $K[x]$!). So α is a root also for $Df(x) \in K[x]$. Hence f and Df have $x-\alpha$ as common factor in $\mathbb{C}[x]$. Let $g := \gcd(f, Df)$; then $\deg(g) \geq 1$. One can compute g from f and Df with the division algorithm (see Ch.II, 3.15). Let us remark that all the coefficients of the polynomials produced by the division algorithm are in the field K containing the coefficients of f and Df; therefore also g is surely in $K[x]$. We have: $\deg(g) \geq 1$, g divides f and f is irreducible; then we necessarily get $f = g$ up to a unit. This, *in a field of characteristic zero* is a contradiction, since f should divide Df, and $\deg(Df) < \deg(f)$. □

Theorem 2.2. *Let $K \subseteq L$ ($L \subseteq \mathbb{C}$, as specified above); assume $[L : K] = n$ (n finite) and let $\phi : K \longrightarrow \mathbb{C}$ be a morphism. Then there exist exactly n morphisms $\psi_1, \ldots, \psi_n : L \longrightarrow \mathbb{C}$ extending ϕ (i.e. such that $\psi_i\big|_K = \phi$ for all $i = 1, \ldots, n$).*

Proof. Assume first that L is a simple extension: $L = K(\alpha)$ for a suitable $\alpha \in L$. The minimum polynomial of α is a polynomial $p(x) = \sum_{i=0}^{n} a_i x^i$, ($a_n = 1$) of degree n (see 1.14). Then $1, \alpha, \ldots, \alpha^{n-1}$ is a basis of L as a K-vector space. Let $\psi : L \longrightarrow \mathbb{C}$ be an extension of ϕ. Then, since ψ is a field morphism,

$$\psi(\lambda_0 + \lambda_1\alpha + \cdots + \lambda_{n-1}\alpha^{n-1}) = \phi(\lambda_0) + \phi(\lambda_1)\psi(\alpha) + \cdots + \phi(\lambda_n)\psi(\alpha)^{n-1}$$

for every $\lambda_0, \ldots, \lambda_{n-1} \in K$. Hence ψ is uniquely determined by $\psi(\alpha)$.
We want to show that $\psi(\alpha)$ can assume n values, which are precisely the n distinct roots of a suitable polynomial $q(x) \in \mathbb{C}[x]$, and conversely that for each root β of $q(x)$ we can construct a morphism $\psi_\beta : L \longrightarrow \mathbb{C}$, extending ϕ.
Let $\Phi : K[x] \longrightarrow \mathbb{C}[x]$ be defined by:

$$\Phi(b_0 + \cdots + b_m x^m) := \phi(b_0) + \cdots + \phi(b_m)x^m,$$

(here $K[x]$ and $\mathbb{C}[x]$ are rings of polynomials and $b_0 + \cdots + b_m x^m$ is any polynomial of $K[x]$).
It is a straightforward verification to see that Φ is a ring homomorphism. It is clear that, for any polynomial $f(x) \in K[x]$, it holds

$$\psi(f(\alpha)) = \Phi(f(x))_{|x=\psi(\alpha)}$$

where the last expression means the value of the polynomial $\Phi(f(x))$ computed in $x = \psi(\alpha)$. Setting

$$q(x) := \Phi(p(x)) = \sum_{i=0}^{n} \phi(a_i)x^i$$

we have

$$0 = \psi(0) = \psi(p(\alpha)) = \Phi(p(x))_{|x=\psi(\alpha)} = q(\psi(\alpha))$$

so $\psi(\alpha)$ is a zero of $q(x)$.
Conversely, let β be a root of $q(x)$. Setting

$$\psi_\beta(\lambda_0 + \lambda_1\alpha + \cdots + \lambda_{n-1}\alpha^{n-1}) := \Phi(\lambda_0 + \lambda_1 x + \cdots + \lambda_{n-1}x^{n-1})(\beta)$$

we clearly have $\psi_\beta(\alpha) = \beta$.
Claim. The map $\psi_\beta : K(\alpha) \longrightarrow \mathbb{C}$ is a field morphism.
Since ψ_β is additive, i.e. $\psi_\beta(u+v) = \psi_\beta(u) + \psi_\beta(v)$, for every $u, v \in K(\alpha)$, then we have only to show that $\psi_\beta(uv) = \psi_\beta(u)\psi_\beta(v)$, for every $u, v \in K(\alpha)$. If $u = u_0 + \cdots + u_{n-1}\alpha^{n-1}$, let $\tilde{u}$ be the polynomial $u_0 + \cdots + u_{n-1}x^{n-1} \in K[x]$ and define analogously $\tilde{v}$. If we divide $\tilde{u} \cdot \tilde{v}$ by p, we get: $\tilde{u} \cdot \tilde{v} = fp + r$, where r is the remainder of the division. Hence $uv = (\tilde{u} \cdot \tilde{v})(\alpha) = r(\alpha)$, so $\psi_\beta(uv) = \psi_\beta(r(\alpha)) = \Phi(r)(\beta)$. Moreover

$$\begin{aligned}\psi_\beta(u)\psi_\beta(v) &= \psi_\beta(\tilde{u}(\alpha))\psi_\beta(\tilde{v}(\alpha)) = \Phi(\tilde{u})(\beta)\Phi(\tilde{v})(\beta) \\ &= (\Phi(\tilde{u})\Phi(\tilde{v}))(\beta) = \Phi(\tilde{u}\tilde{v})(\beta) = \Phi(fp + r)(\beta) = \Phi(r)(\beta).\end{aligned}$$

The last equality follows from $\Phi(fp)(\beta) = \Phi(f)(\beta)\Phi(p)(\beta) = 0$, since $\Phi(p)(\beta) = q(\beta) = 0$.
Hence we have $\psi_\beta(uv) = \psi_\beta(u)\psi_\beta(v)$, and this shows the claim.
To conclude the proof in the case of simple extensions we just need to observe that $p(x)$ has n distinct roots in $\mathbb{C}$ by 2.1. Moreover a root of p is a root of q and $\deg(q) = \deg(p)$. So p and q have the same roots.
Let now L be any finite extension, and let $\alpha_1, \ldots, \alpha_n$ be a basis of L over K. Then we have the following chain of fields:

$$K \subseteq K(\alpha_1) \subseteq K(\alpha_1)(\alpha_2) \subseteq \cdots \subseteq K(\alpha_1, \ldots, \alpha_{n-1})(\alpha_n) = L$$

and we see that every intermediate field is a simple extension of the previous one. Let $\phi : K \longrightarrow \mathbb{C}$ be a morphism and set $n_1 := [K(\alpha_1) : K]$. Then, from the first part of the proof, we have that there exist exactly n_1 morphisms $\sigma_1, \ldots, \sigma_{n_1} : K(\alpha_1) \longrightarrow \mathbb{C}$ extending ϕ. For the same reason, each $\sigma_i : K(\alpha_1) \longrightarrow \mathbb{C}$ can be extended in exactly n_2 morphisms $\sigma_{ij} : K(\alpha_1)(\alpha_2) \longrightarrow \mathbb{C}$, $j = 1, \ldots, n_2$ (where $n_2 := [K(\alpha_1)(\alpha_2) : K(\alpha_1)]$), and so on. Recalling that, by 1.13,

$$[L : K] = [K(\alpha_1, \ldots, \alpha_{n-1})(\alpha_n) : K(\alpha_1, \ldots, \alpha_{n-1})] \cdots [K(\alpha_1) : K],$$

we get the result. □

We shall prove (see theorem 2.8) that if L is a finite algebraic extension of K, then L is a simple extension of K, and one can give a proof of this result which do not depend on 2.2. Using this fact, one can deduce 2.2 only from the first part of the proof given here.

Remark 2.3. Let $f(x) \in K[x]$ be an irreducible polynomial and suppose that $\alpha, \beta \in \mathbb{C}$ are two zeros of f. Note that the first part of the above proof (the definition of the map ψ_β) can be used to prove the following relevant fact:
there exists a unique morphism $\psi : K(\alpha) \longrightarrow K(\beta)$ which is the identity map on the elements of K and such that $\psi(\alpha) = \beta$.

Example 2.3.1. Let us compute all the morphisms $\psi : \mathbb{Q}[\sqrt{2}] \longrightarrow \mathbb{C}$ such that $\psi|_{\mathbb{Q}} = \mathrm{id}_{\mathbb{Q}}$. The minimum polynomial of $\sqrt{2}$ over $\mathbb{Q}$ is $x^2 - 2$. Following the procedure (and the notations) of 2.2, we have: $\psi(a + \sqrt{2}b) = a + \psi(\sqrt{2})b$, so $\psi(\sqrt{2})$ must be a root of $\Phi(x^2-2) = x^2-2$, hence $\psi(\sqrt{2}) = \sqrt{2}$ or $\psi(\sqrt{2}) = -\sqrt{2}$. Therefore we find two morphisms:

$$\begin{aligned} &\psi_1 : \mathbb{Q}[\sqrt{2}] \longrightarrow \mathbb{C} \quad \text{such that} \quad \psi_1(a + \sqrt{2}b) = a + \sqrt{2}b \\ &\psi_2 : \mathbb{Q}[\sqrt{2}] \longrightarrow \mathbb{C} \quad \text{such that} \quad \psi_1(a + \sqrt{2}b) = a - \sqrt{2}b \end{aligned}$$

Note that, according to 2.2, $\left[\mathbb{Q}[\sqrt{2}] : \mathbb{Q}\right] = 2$.

Example 2.3.2. Let us now compute all the morphisms $\psi : \mathbb{Q}[\sqrt{2}, \sqrt{3}] \longrightarrow \mathbb{C}$ such that $\psi|_{\mathbb{Q}} = \mathrm{id}_{\mathbb{Q}}$. First of all recall that $\mathbb{Q}[\sqrt{2}, \sqrt{3}] = \{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} \mid a, b, c, d \in \mathbb{Q}\}$. Let's consider the morphism $\phi := \psi|_{\mathbb{Q}[\sqrt{2}]}$. Then, from example 2.3.1 above, we have two possibilities for ϕ: either $\phi(a + b\sqrt{2}) = a + b\sqrt{2}$ or $\phi(a + b\sqrt{2}) = a - b\sqrt{2}$. We have $\mathbb{Q}[\sqrt{2}, \sqrt{3}] = \mathbb{Q}[\sqrt{2}][\sqrt{3}]$, and $x^2 - 3$ is the minimum polynomial of $\sqrt{3}$ over $\mathbb{Q}[\sqrt{2}]$. Therefore $\psi(\sqrt{3})$ must be $\sqrt{3}$ or $-\sqrt{3}$. In this way we get the following four morphisms $\psi_1, \ldots, \psi_4 : \mathbb{Q}[\sqrt{2}, \sqrt{3}] \longrightarrow \mathbb{C}$:

$$\begin{aligned} \psi_1(a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}) &= a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}, \\ \psi_2(a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}) &= a + b\sqrt{2} - c\sqrt{3} - d\sqrt{6}, \\ \psi_3(a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}) &= a - b\sqrt{2} + c\sqrt{3} - d\sqrt{6}, \\ \psi_4(a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}) &= a - b\sqrt{2} - c\sqrt{3} + d\sqrt{6}. \end{aligned}$$

We remark that in this example the image of any ψ_i $(i = 1, \ldots, 4)$ is the field $\mathbb{Q}[\sqrt{2}, \sqrt{3}]$, hence $\psi_1, \ldots, \psi_4$ are automorphisms of $\mathbb{Q}[\sqrt{2}, \sqrt{3}]$.

Definition 2.4. Let $K \subseteq L$ be an extension. A morphism $\sigma : L \longrightarrow \mathbb{C}$ such that $\sigma|_K$ is the identity will be called a *K-morphism of L*. We denote by $\mathcal{I}(L, K)$ the set of K-morphisms of L.

From 2.2, $\mathcal{I}(L, K)$ has n elements, where $n = [L : K]$ (which is assumed finite). It is clear that if $K \subseteq H \subseteq L$, then $\mathcal{I}(L, H)$ is a subset of $\mathcal{I}(L, K)$. Hence $\mathcal{I}(L, \cdot)$ is a (contravariant) map that associates to every subfield of L containing K a subset of $\mathcal{I}(L, K)$.
More precisely, let $\mathcal{H} := \{H \mid K \subseteq H \subseteq L\}$ and consider the map

$$\mathcal{I}(L, \cdot) : \mathcal{H} \longrightarrow \{\text{subsets of } \mathcal{I}(L, K)\}$$

defined by

$$H \mapsto \mathcal{I}(L, H).$$

If $H_1 \subseteq H_2$ are elements of $\mathcal{H}$, then $\mathcal{I}(L, H_1) \supseteq \mathcal{I}(L, H_2)$.

Conversely, if $T = \{t_1, \ldots, t_s\} \subseteq \mathcal{I}(L, K)$, then we can construct a field defined by:

$$L^T := \{a \in L \mid t(a) = a \text{ for every } t \in T\}.$$

It is easy to see that L^T is a field contained in L and containing K.
Equivalently, there is a map

$$L^{(\cdot)} : \{\text{subsets of } \mathcal{I}(L,K)\} \longrightarrow \mathcal{H}$$

defined by

$$T \mapsto L^T.$$

Proposition 2.5. *Let $T \subseteq \mathcal{I}(L,K)$ be any subset and let H be any field $K \subseteq H \subseteq L$. Then it holds:*
i) $T \subseteq \mathcal{I}(L, L^T)$;
ii) $H = L^{\mathcal{I}(L,H)}$.

Proof. i) is immediate, since for any $t \in T$, for any $a \in L^T$, we have $t(a) = a$; so $t|_{L^T} = \mathrm{id}_{L^T}$, i.e. $t \in \mathcal{I}(L, L^T)$.
ii) It is clear that $H \subseteq L^{\mathcal{I}(L,H)}$, since for any $h \in H$, $t \in \mathcal{I}(L,H)$, we have $t(h) = h$; hence $h \in L^{\mathcal{I}(L,H)}$. To see the other inclusion, first recall (see 2.2) that $\sharp\mathcal{I}(L,H) = [L : H]$. From 1.13 we have that $[L : H] = [L : L^{\mathcal{I}(L,H)}][L^{\mathcal{I}(L,H)} : H]$. If $\sigma \in \mathcal{I}(L,H)$, then σ is a morphism from L to $\mathbb{C}$ extending the identity on $L^{\mathcal{I}(L,H)}$; in fact: take $x \in L^{\mathcal{I}(L,H)}$, so $\sigma(x) = x$, hence $\sigma \in \mathcal{I}(L, L^{\mathcal{I}(L,H)})$. Therefore $\mathcal{I}(L,H) \subseteq \mathcal{I}(L, L^{\mathcal{I}(L,H)})$, so $[L : L^{\mathcal{I}(L,H)}] \geq \sharp(\mathcal{I}(L,H))$. So, using the above equalities, we get:

$$\sharp\mathcal{I}(L,H) = [L : H] \geq \sharp\big([\mathcal{I}(L,H)\big)[L^{\mathcal{I}(L,H)} : H].$$

Therefore $[L^{\mathcal{I}(L,H)} : H] \leq 1$, and this gives $L^{\mathcal{I}(L,H)} = H$. □

As a consequence, we get:

Corollary 2.6. *The map*

$$\mathcal{I}(L,\cdot) : \mathcal{H} \longrightarrow \{\textit{subsets of } \mathcal{I}(L,K)\}$$

defined by $H \mapsto \mathcal{I}(L,H)$, is injective.

Proof. It is an immediate consequence of 2.5 ii). □

Corollary 2.7. *Let, as usual, $[L : K] < \infty$. Then the number of fields H such that $K \subseteq H \subseteq L$, i.e. $\sharp(\mathcal{H})$, is finite.*

Proof. It immediately follows from the fact that $\mathcal{I}(L,K)$ is finite and from 2.6. □

Theorem 2.8. *(Abel's theorem) If $K \subseteq L$ is a finite extension, then it is simple.*

Proof. If $\alpha_1, \ldots, \alpha_n$ is a K-basis of L, then $L = K(\alpha_1, \ldots, \alpha_n)$ and α_i are algebraic over K, from 1.18. It is enough to prove the result for $n = 2$. Let $L = K(\alpha, \beta)$, α and β algebraic over K. Let $h \in \mathbb{N}$, and consider the field $L_h := K(\alpha + h\beta)$. Since $K \subseteq L_h \subseteq L$, from 2.7 it follows that there exist $h, k \in \mathbb{N}$, $h \neq k$, such that $L_h = L_k$. In particular $\alpha + k\beta \in K(\alpha + h\beta)$, hence

$$\beta(k - h) = (\alpha + h\beta) - (\alpha + k\beta) \in K(\alpha + h\beta).$$

Since $k - h \in K(\alpha + h\beta)$, we get that $\beta = \beta(k-h)(k-h)^{-1} \in K(\alpha + h\beta)$. But if $\beta \in K(\alpha + h\beta)$, then $\alpha \in K(\alpha + h\beta)$ since $\alpha = (\alpha + h\beta) - h\beta$, therefore $K(\alpha, \beta) \subseteq K(\alpha + h\beta)$ and since the other inclusion is obvious, we get the proof. □

3. Galois Correspondence

Once more, let us recall that we always assume that all the fields considered are subfields of $\mathbb{C}$.

A natural question arising from 2.6 is the following: given an extension $K \subseteq L$, *which subsets of* $\mathcal{I}(L,K)$ *are of the type* $\mathcal{I}(L,H)$*?* Or, equivalently, *what is the image of the map*

$$\mathcal{I}(L,\cdot) : \mathcal{H} \longrightarrow \{\textit{subsets of } \mathcal{I}(L,K)\}?$$

It is possible to give a satisfactory answer to this question if we consider a particular kind of finite extensions $K \subseteq L$, as we shall see in this section (see 3.13).

Definition 3.1. A finite extension $K \subseteq L$ is a *normal* (or a *Galois*) *extension* if $\phi(L) \subseteq L$ for every $\phi \in \mathcal{I}(L,K)$.

Remark 3.2. If we have that $\phi(L) \subseteq L$, then $\phi(L) = L$. In fact if $\alpha_1, \ldots, \alpha_n$ is a basis of L over K, then $\phi(\alpha_1), \ldots, \phi(\alpha_n)$ is a basis of $\phi(L)$ over $\phi(K) = K$. Therefore $\phi(L)$ is a subvector space of L of the same dimension of L over K, therefore $\phi(L) = L$.

The following is a characterization of normal extensions:

Proposition 3.3. *The following are equivalent:*

i) $K \subseteq L$ *is a normal extension;*

ii) If $f(x) \in K[x]$ *is an irreducible polynomial which has a zero in* L*, then all its zeros are in* L*.*

Proof. i) $\Rightarrow$ ii) Let $f(x) \in K[x]$ be an irreducible polynomial and suppose $\alpha \in L$ is a root of f. Let $\beta \in \mathbb{C}$ be another zero of f and consider the K-morphism $\sigma : K(\alpha) \longrightarrow \mathbb{C}$ defined by $\sigma(\alpha) := \beta$ (see 2.3). From 2.2 we can extend σ to a map $\phi : L \longrightarrow \mathbb{C}$. By i), $\phi(L) \subseteq L$, so $\beta = \phi(\alpha) \in L$.
ii) $\Rightarrow$ i) Let $\phi : L \longrightarrow \mathbb{C}$ be any K-morphism and let $\alpha \in L$. We want to see that $\phi(\alpha) \in L$. Assume that $f(x) \in K[x]$ is the minimum polynomial of α over K. Then $0 = \phi(0) = \phi(f(\alpha)) = f(\phi(\alpha))$, so $\phi(\alpha)$ is a zero of f and, therefore, by ii), it is in L. $\square$

Definition 3.4. If $K \subseteq L$ is an extension, a K-*automorphism* of L is a K-morphism (see 2.4) whose image is contained in L, i.e. it is a morphism $\phi : L \longrightarrow L \subseteq \mathbb{C}$ such that $\phi|_K = \mathrm{id}_K$.
Let $\mathcal{G}(L,K)$ be the set of K-automorphisms of L. It is clearly a group under composition and it is called the *Galois group* of L over K.

Since the K-automorphisms are particular K-morphisms, we have that $\mathcal{G}(L,K) \subseteq \mathcal{I}(L,K)$; clearly $K \subseteq L$ is a normal extension iff equality holds.

Definition 3.5. Let $f(x) \in K[x]$ and let $L := K(\alpha_1, \ldots, \alpha_n)$, where $\alpha_1, \ldots, \alpha_n$ are all the roots of f in $\mathbb{C}$; the field L is called the *splitting field* of $f(x)$ over K. By the *Galois group of* f we mean the group $\mathcal{G}(L,K)$.

Lemma 3.6. *Let $f(x) \in K[x]$ be a polynomial and let $\alpha_1, \ldots, \alpha_n$ be the roots of f in $\mathbb{C}$. Set $L := K(\alpha_1, \ldots, \alpha_n)$ the splitting field of f over K. If $\phi : L \longrightarrow \mathbb{C}$ is any K-morphism, then for any $i \in \{1, \ldots, n\}$ there exists a $j \in \{1, \ldots, n\}$ such that $\phi(\alpha_i) = \alpha_j$ (i.e. ϕ sends roots of f into roots of f).*
Moreover, any K-morphism $\phi : L \longrightarrow \mathbb{C}$ is determined by the values of $\phi(\alpha_1), \ldots, \phi(\alpha_n)$, since

$$\phi(F(\alpha_1, \ldots, \alpha_n)) = F(\phi(\alpha_1), \ldots, \phi(\alpha_n)),$$

for every $F \in K[\alpha_1, \ldots, \alpha_n] = L$. In particular $\phi(L) \subseteq L$.

Proof. If $\phi : L \longrightarrow \mathbb{C}$ is any K-morphism, then $0 = \phi(f(\alpha_i)) = f(\phi(\alpha_i))$, hence $\phi(\alpha_i) \in \{\alpha_1, \ldots, \alpha_n\}$. By 1.11, $L = K[\alpha_1, \ldots, \alpha_n]$. Since ϕ is a ring homomorphism, we get $\phi(F(\alpha_1, \ldots, \alpha_n)) = F(\phi(\alpha_1), \ldots, \phi(\alpha_n))$, as requested. This completes the proof. □

Theorem 3.7. *Let $L := K(\alpha_1, \ldots, \alpha_n)$ be the splitting field of f over K (where $\alpha_1, \ldots, \alpha_n$ are the roots of f in $\mathbb{C}$). It holds:*
i) L is a normal extension of K.
ii) $\mathcal{G}(L, K)$ can be embedded in S_n (the symmetric group of n objects).

Proof. If $\phi : L \longrightarrow \mathbb{C}$ is any K-morphism, then $\phi(L) \subseteq L$, by 3.6; so i) is proved.
Let $\phi \in \mathcal{G}(L, K)$, then, from 3.6, it follows that for each $i \in \{1, \ldots, n\}$, there exists $\sigma(i) \in \{1, \ldots, n\}$ such that $\phi(\alpha_i) = \alpha_{\sigma(i)}$. In this way we get a permutation $\sigma \in S_n$, so we define a map $\Psi : \mathcal{G}(L, K) \longrightarrow S_n$ and it is easy to verify that it is a group homomorphism. If $\phi \in \ker \Psi$, then $\phi(\alpha_i) = \alpha_i$, $i = 1, \ldots, n$, so by 3.6, $\phi = \mathrm{id}_L$, hence Ψ is injective. □

Recall that if G is a group and X is a set, an *action of G on X* is a map $G \times X \longrightarrow X$ ($(g, x) \mapsto g \cdot x$) such that $1_G \cdot x = x$, for all $x \in X$, and $g_1 \cdot (g_2 \cdot x) = (g_1 g_2) \cdot x$, for all $g_1, g_2 \in G$, for all $x \in X$.
If $x \in X$, then the *orbit* of x is the set $\{g \cdot x \mid g \in G\}$. The set of all orbits gives a partition of X. The action is *transitive* if for all $x, y \in X$ there exists a $g \in G$ such that $g \cdot x = y$. Hence an action is transitive on each orbit; in particular it is transitive on X iff the orbit of any $x \in X$ is the whole set X.

If $L = K(\alpha_1, \ldots, \alpha_n)$ is the splitting field of f and if $G := \mathcal{G}(L, K)$ is the Galois group of f, then we can define an action of G on the set $\{\alpha_1, \ldots, \alpha_n\}$ of the roots of f as follows: $\phi \cdot \alpha_i := \phi(\alpha_i)$, for all $\phi \in G$, $i = 1, \ldots, n$. This means that we realize G as a subgroup of S_n (see 3.7), which acts on $\{\alpha_1, \ldots, \alpha_n\}$.

Theorem 3.8. *If f is irreducible over $K[x]$, then the above action is transitive; in general, if p is an irreducible factor of f, then the set of its roots is an orbit of the action of G on $\{\alpha_1, \ldots, \alpha_n\}$. Conversely, if $\{\beta_1, \ldots, \beta_k\} \subseteq \{\alpha_1, \ldots, \alpha_n\}$ is an orbit, then the polynomial $p := \prod_{i=1}^{k}(x - \beta_i)$ is an irreducible factor of f.*

Proof. Let f be irreducible in $K[x]$ and let α_i, α_j be two roots of f in $\mathbb{C}$. Then f is, up to a unit, the minimum polynomial of α_i (see 1.9), hence from 2.3 there exists a K-morphism $\phi_1 : K(\alpha_i) \longrightarrow K(\alpha_j)$ such that $\phi(\alpha_i) = \alpha_j$. From 2.2 we can extend ϕ_1 to a morphism $\phi : L \longrightarrow \mathbb{C}$. Since L is normal (by 3.7), then $\phi \in G$. Since $\phi(\alpha_i) = \phi_1(\alpha_i) = \alpha_j$, we see that the action is transitive.

Take now a prime factor p of f and let $\{\beta_1, \ldots, \beta_k\} \subseteq \{\alpha_1, \ldots, \alpha_n\}$ be the roots of p. Since p is irreducible, we can use the first part of the proof to find a K-morphism

$$\tau_i : K(\beta_1, \ldots, \beta_k) \longrightarrow \mathbb{C}$$

such that $\tau_i(\beta_1) = \beta_i$ $(i = 1, \ldots, k)$. Using 2.2 we can extend τ_i to a K-morphism $\psi_i : L \longrightarrow L$. This shows that $\{\beta_1, \ldots, \beta_k\}$ are all in the same orbit, since $\psi_i \in G$ and $\psi_i(\beta_1) = \beta_i$.
Conversely, if α is an element of the orbit of β_1, then $\alpha = \psi(\beta_1)$ for some $\psi \in G$. Therefore $0 = \psi(0) = \psi(p(\beta_1)) = p(\psi(\beta_1)) = p(\alpha)$, hence α is a root of p. So $\{\beta_1, \ldots, \beta_k\}$ is an orbit.
Finally, if α is a root of f, we can find an irreducible factor p of f such that $p(\alpha) = 0$ and the roots of p give the orbit of α.
Hence we see that there is a bijection between the prime factors of f and the orbits of the action. □

Example 3.8.1. Let's consider again the extension $\mathbb{Q} \subseteq \mathbb{Q}[\sqrt{2}, \sqrt{3}]$. In example 2.3.2 we found that there are exactly four morphisms $\psi_i : \mathbb{Q}[\sqrt{2}, \sqrt{3}] \longrightarrow \mathbb{C}$, $i = 1, \ldots, 4$, and their image is always $\mathbb{Q}[\sqrt{2}, \sqrt{3}]$, hence the extension $\mathbb{Q} \subseteq \mathbb{Q}[\sqrt{2}, \sqrt{3}]$ is normal. Observe that $\mathbb{Q}[\sqrt{2}, \sqrt{3}]$ is the splitting field of the polynomial $f(x) = (x^2 - 2)(x^2 - 3)$, so it turns out to be a normal extension also from i) of 3.7. The Galois group of $\mathbb{Q}[\sqrt{2}, \sqrt{3}]$ over $\mathbb{Q}$ is then the set $G := \{\psi_1, \psi_2, \psi_3, \psi_4\}$. We can embed G in S_4 as follows (see the proof of 3.8): let's label with $1, 2, 3, 4$ the elements $\sqrt{2}, -\sqrt{2}, \sqrt{3}, -\sqrt{3}$ respectively, which are the four roots of $f(x)$. Then

$$\psi_1 \mapsto \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \qquad \psi_2 \mapsto \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 3 \end{pmatrix},$$
$$\psi_3 \mapsto \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix}, \qquad \psi_4 \mapsto \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix},$$

It is easy to verify that the group G is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$. The action of G on $\{1, 2, 3, 4\}$ (i.e. on the set $\{\sqrt{2}, -\sqrt{2}, \sqrt{3}, -\sqrt{3}\}$), has the following two orbits: $\{1, 2\} = \{\psi \cdot 1 \mid \psi \in G\}$ and $\{3, 4\} = \{\psi \cdot 3 \mid \psi \in G\}$, according to the fact that $f(x) = (x^2 - 2)(x^2 - 3)$ is the product of two irreducible polynomials over $\mathbb{Q}[x]$.

Proposition 3.9. *Let $K \subseteq L$ be a normal extension of degree n. Then there exists an irreducible polynomial $f(x) \in K[x]$ with roots $\alpha_1, \ldots, \alpha_n$ such that $L = K(\alpha_1, \ldots, \alpha_n)$, i.e. any normal extension is the splitting field of a suitable irreducible polynomial.*

Proof. From 2.8, there exists $\alpha \in L$ such that $L = K(\alpha)$. Let $f(x)$ be the minimum polynomial of α over K. Then $\deg(f) = n$ (see 1.14). Set $\alpha_1 := \alpha, \alpha_2, \ldots, \alpha_n$ the roots of f. Then, from 3.3, since L is a normal extension of K and f is irreducible, $\alpha_i \in L$ for every i; therefore $L \supseteq K(\alpha_1, \ldots, \alpha_n)$.
Conversely, $L = K(\alpha_1) \subseteq K(\alpha_1, \ldots, \alpha_n)$; so we get the requested equality. □

Remark 3.10. From 3.7. and 3.9, we have that $K \subseteq L$ is a normal extension iff L is a splitting field of an irreducible polynomial in $K[x]$.

We saw in the example 3.8.1, that $\mathbb{Q}[\sqrt{2},\sqrt{3}]$ is a normal extension of $\mathbb{Q}$. Hence it should be the splitting field of an *irreducible* polynomial. From example 1.6.1, we have that $\mathbb{Q}[\sqrt{2},\sqrt{3}] = \mathbb{Q}[\sqrt{2}+\sqrt{3}]$. It is easy to verify that $p(x) := x^4 - 10x^2 + 1$ is the minimum polynomial of $\sqrt{2}+\sqrt{3}$ over $\mathbb{Q}$ and its splitting field is $\mathbb{Q}[\sqrt{2}+\sqrt{3}]$.

Definition 3.11. Let $K \subseteq L$ be an algebraic extension. A *normal closure* of L over K is an extension M of L such that
1) $K \subseteq M$ is normal;
2) if $L \subseteq H \subseteq M$, and $K \subseteq H$ is normal, then $H = M$
(i.e. M is the smallest extension of L which is normal over K).

Theorem 3.12. *Let $K \subseteq L$ be a finite extension. Then we can find a field N which is a finite extension of K and is a normal closure of K.*

Proof. We know from 2.8 that $L = K(\alpha)$. Let f be the minimum polynomial of α over K and let N be the splitting field of f over K. Then $N \supseteq L$, and $K \subseteq N$ is a normal extension (by 3.7). If $L \subseteq H \subseteq N$ and $K \subseteq H$ is normal, since $\alpha \in L$ it follows $\alpha \in H$, hence H contains a zero of f. Since $K \subseteq H$ is normal, it contains all the zeros of f (see 3.3) so $H \supseteq N$. □

Example 3.12.1. The extension $\mathbb{Q} \subseteq \mathbb{Q}(\sqrt[3]{2})$ is not normal (in fact $\mathbb{Q}(\sqrt[3]{2})$ contains only one of the three roots of $x^3 - 2$). As in the above proof, we get that the normal closure of the considered extension is

$$\mathbb{Q}\left(\sqrt[3]{2}, -1/2 + \mathrm{i}\sqrt{3}/2, -1/2 - \mathrm{i}\sqrt{3}/2\right).$$

Lemma 3.13. *Let $K \subseteq H \subseteq L$ and assume that $K \subseteq L$ is a normal extension. Then the following facts hold:*
i) $H \subseteq L$ is a normal extension.
Equivalently, if $\mathcal{I}(L,K) = \mathcal{G}(L,K)$, then $\mathcal{I}(L,H) = \mathcal{G}(L,H)$. In this case, the map considered in 2.6 becomes

$$\mathcal{G}(L,\cdot) : \mathcal{H} \longrightarrow \mathcal{M}$$

defined by $H \mapsto \mathcal{G}(L,H) = \mathcal{I}(L,H)$, where

$$\mathcal{H} := \{H \mid K \subseteq H \subseteq L\} \quad \text{and} \quad \mathcal{M} := \{M \mid M \leq \mathcal{G}(L,K)\};$$

ii) $K \subseteq H$ is normal iff $\phi(H) = H$ for every $\phi \in \mathcal{G}(L,K)$.

Proof. *i)* If $\phi : L \longrightarrow \mathbb{C}$ is any H-morphism, then it is a K-morphism, so $\phi(L) = L$; this means that ϕ is a H-automorphism of L.
Note also that the map $\mathcal{G}(L,\cdot)$ is well-defined, since $\mathcal{G}(L,H)$ is not only a subset, but a subgroup of $\mathcal{G}(L,K)$.
ii) Assume that $K \subseteq H$ is normal and let $\phi \in \mathcal{G}(L,K)$. Clearly $\phi_{|H}$ is a K-morphism of H; but $K \subseteq H$ is normal, hence $\phi_{|H}$ is a K-automorphism, i.e. $\phi(H) = H$.
Conversely, let ψ be a K-morphism of H; by 2.2 we can extend ψ to $\phi \in \mathcal{I}(L,K) = \mathcal{G}(L,K)$. Then $\psi(H) = \phi(H) = H$, hence ψ is a K-automorphism of H; so $K \subseteq H$ is normal. □

The following theorem gives an answer to the question considered at the very beginning of this section (for the case of normal extensions).

Theorem 3.14. *(Fundamental theorem of Galois) Let $K \subseteq L$ be a normal extension. Then:*

i) the map $\mathcal{G}(L,\cdot) : \mathcal{H} \longrightarrow \mathcal{M}$ is a bijection, whose inverse is given by the map

$$M \mapsto L^M = \{a \in L \mid t(a) = a \text{ for all } t \in M\};$$

ii) if $H_1, H_2 \in \mathcal{H}$, it holds: $H_1 \subseteq H_2$ iff $\mathcal{G}(L,H_1) \supseteq \mathcal{G}(L,H_2)$;

iii) let $H \in \mathcal{H}$; H is a normal extension of K iff $\mathcal{G}(L,H)$ is a normal subgroup of $\mathcal{G}(L,K)$;

iv) if $K \subseteq H$ is normal, then any K-automorphism of L is a K-automorphism of H; in this way we get an epimorphism of groups:

$$\mathcal{G}(L,K) \longrightarrow \mathcal{G}(H,K)$$

whose kernel is $\mathcal{G}(L,H)$. Therefore $\mathcal{G}(H,K) \cong \mathcal{G}(L,K)/\mathcal{G}(L,H)$.

Proof. *i)* If $M \in \mathcal{M}$ and $H \in \mathcal{H}$, we already know that $M \subseteq \mathcal{G}(L,L^M)$ and that $H = L^{\mathcal{G}(L,H)}$ (see 2.5). So we have only to see that $M \supseteq \mathcal{G}(L,L^M)$. Let m be the order of the group M. If we can prove that $\sharp(\mathcal{G}(L,L^M)) \leq m$, we are done.
Since $\sharp(\mathcal{G}(L,L^M)) = [L : L^M]$ (see 2.2), it is enough to see that $[L : L^M] \leq m$. Let $\alpha \in L$ be such that $L = K(\alpha)$ (see 2.8). Since $K \subseteq L^M \subseteq L$, so $L = L^M(\alpha)$ holds. Denote by $\sigma_1 = 1, \ldots, \sigma_m$ the elements of M. The following polynomial:

$$f(x) := (x - \sigma_1(\alpha))(x - \sigma_2(\alpha)) \cdots (x - \sigma_m(\alpha))$$

is a polynomial of degree m, and since $\sigma_1(\alpha) = \alpha$, then $f(\alpha) = 0$. If we can prove that its coefficients are in L^M, then *i)* is proved; in fact $[L : L^M] = \deg(p)$, where p is the minimum polynomial of α over L^M (see 1.14) and p divides f (see 1.9). So $\deg(p) \leq \deg(f)$, i.e. $[L : L^M] \leq m$.
Let $\tau \in M$, then the polynomial $(x-\tau\sigma_1(\alpha)) \cdots (x-\tau\sigma_m(\alpha))$ is again $f(x)$, since $\{\tau\sigma_i \mid i = 1, \ldots, m\} = M$. From this, we have that the coefficients of f are fixed by any $\tau \in M$, so they are in L^M.
ii) It is immediate to verify.
iii) Let H be such that $K \subseteq H \subseteq L$. We want to see that $K \subseteq H$ is normal if and only if

$$\phi\mathcal{G}(L,H)\phi^{-1} = \mathcal{G}(L,H) \quad \text{for every } \phi \in \mathcal{G}(L,K).$$

Step 1. If $\phi \in \mathcal{G}(L,K)$, then $\mathcal{G}(L,\phi(H)) = \phi\mathcal{G}(L,H)\phi^{-1}$. In fact: $\psi \in \mathcal{G}(L,\phi(H)) \Leftrightarrow \psi\phi(h) = \phi(h)$ for every $h \in H \Leftrightarrow \phi^{-1}\psi\phi(h) = h$ for every $h \in H \Leftrightarrow \phi^{-1}\psi\phi \in \mathcal{G}(L,H) \Leftrightarrow \psi \in \phi\mathcal{G}(L,H)\phi^{-1}$.
Step 2. $K \subseteq H$ is normal iff $\phi(H) = H$ for every $\phi \in \mathcal{G}(L,K)$ (from 3.13 *ii)*).
Step 3. From *i)* the bijection $\mathcal{G}(L,\cdot) : \mathcal{H} \longrightarrow \mathcal{M}$ gives that

$$\phi(H) = H \Leftrightarrow \mathcal{G}(L,\phi(H)) = \mathcal{G}(L,H).$$

Step 4. We can conclude that $K \subseteq H$ is normal $\Leftrightarrow \mathcal{G}(L,\phi(H)) = \mathcal{G}(L,H)$ for every ϕ $\Leftrightarrow \phi\mathcal{G}(L,H)\phi^{-1} = \mathcal{G}(L,H)$ for every $\phi \Leftrightarrow \mathcal{G}(L,H)$ is normal in $\mathcal{G}(L,K)$; this gives *iii)*.
iv) If $\phi \in \mathcal{G}(L,K)$, then $\phi(H) = H$ since $K \subseteq H$ is normal (see 3.13), so $\phi_{|H} \in \mathcal{G}(H,K)$. In this way we get a map $\mathcal{G}(L,K) \longrightarrow \mathcal{G}(H,K)$, defined by $\phi \mapsto \phi_{|H}$, which is surjective (by 2.2). This map is clearly a group homomorphism. Its kernel is the group

$$\{\phi \in \mathcal{G}(L,K) \mid \phi(h) = h \text{ for every } h \in H\} = \mathcal{G}(L,H).$$

□

4. Solvability by Radicals

Now we want to show how to use the above results in order to see when a polynomial equation is solvable by radicals. First of all we should explain the meaning of "solvable by radicals". By this we mean that the solutions are given by some expressions like: $\sqrt{7}\sqrt[4]{2-\sqrt[5]{3}}-\sqrt[3]{17}$ or $\sqrt[3]{\sqrt[5]{1-\sqrt{7}}+\sqrt[7]{3-\sqrt{8}}}\ldots$.
For instance, the solutions of $x^2+2x-4=0$ are: $x_1=1+\sqrt{5}$ and $x_2=1-\sqrt{5}$. The solutions of $x^6+2x^3-5=0$ are: $\sqrt[3]{1+\sqrt{6}}$ and $\sqrt[3]{1-\sqrt{6}}$ (where $\sqrt[3]{\cdot}$ can have three values). Take this last example. The solutions lie in $\mathbb{Q}(a)(b)=\mathbb{Q}(a,b)$, where $a^2=6$ and b is such that $b^3=1+a\in\mathbb{Q}(a)$, hence we can generalize this situation as follows:

Definition 4.1. An extension $K\subseteq L$ is *radical* if $L=K(\alpha_1,\ldots,\alpha_r)$, where for each $i=1,\ldots,r$ there exists an $n(i)\in\mathbb{N}$ such that

$$\alpha_i^{n(i)}\in K(\alpha_1,\ldots,\alpha_{i-1}).$$

In this case the chain of subfields we get:

$$K\subseteq K(\alpha_1)\subseteq K(\alpha_1,\alpha_2)\subseteq\cdots\subseteq K(\alpha_1,\ldots,\alpha_r)$$

is called a *tower* of subfields.

Clearly a radical extension is a finite algebraic extension.

Definition 4.2. Let $f\in K[x]$ be a polynomial, and let N be the splitting field of f over K. Then f is *solvable by radicals* if there exists a field L such that $N\subseteq L$ and L is a radical extension of K.

The aim of this section is to prove the following theorem:

Theorem 4.3. *A polynomial f is solvable by radicals if and only if its Galois group is a solvable group.*

The complete proof of this theorem will be given in this section. Unfortunately it requires some technical lemmas that we are going to present here.

Lemma 4.4. *If $K\subseteq L$ is a radical extension and M is a normal closure of L over K, then $K\subseteq M$ is radical.*

Proof. Let $L=K(\alpha_1,\ldots,\alpha_r)$, where $\alpha_i^{n(i)}\in K(\alpha_1,\ldots,\alpha_{i-1})$ and let f_i be the minimum polynomial of α_i over K. If $f:=\prod f_i$, then all the zeros of f must lie in M (see 3.3). If N is the splitting field of f over K, then $N\subseteq M$ and N is a normal extension of K (see 3.7), hence $N=M$, so M is the splitting field of f over K, i.e. $M=K(\cup\beta_{ij})$, where β_{ij} $(j=1,\ldots,\deg(f_i))$ are all the zeros of f_i (clearly $\cup\beta_{ij}$ denotes the set of all β_{ij}'s). Fix any β_{ij} and define

$$\phi:K(\alpha_1,\ldots,\alpha_{i-1})(\alpha_i)\longrightarrow\mathbb{C}$$

by: $\phi(u):=u$ for all $u\in K(\alpha_1,\ldots,\alpha_{i-1})$, $\phi(\alpha_i):=\beta_{ij}$. We can extend ϕ to $\psi:\ N\longrightarrow\mathbb{C}$ (see 2.2). Since N is a normal extension of K and $\psi(k)=k$ for every $k\in K$, $\psi(N)\subseteq N$, so

ψ is an automorphism of N such that $\psi(u) = u$ for every $u \in K(\alpha_1, \ldots, \alpha_{i-1})$, i.e. ψ is not only a K-automorphism, but also a $K(\alpha_1, \ldots, \alpha_{i-1})$-automorphism of N. By hypothesis, there exists an $a_i \in K(\alpha_1, \ldots, \alpha_{i-1})$ such that $\alpha_i^{n(i)} - a_i = 0$. Hence $\psi(\alpha_i^{n(i)} - a_i) = \beta_{ij}^{n(i)} - a_i = 0$. So β_{ij} is radical over $K(\alpha_1, \ldots, \alpha_{i-1})$. Therefore the following extensions of fields show that N is a radical extension of K:

$$K \subseteq K(\alpha_1, \ldots, \alpha_r) \subseteq K(\alpha_1, \ldots, \alpha_r)(\beta_{11}) \subseteq K(\alpha_1, \ldots, \alpha_r, \beta_{11})(\beta_{12}) \subseteq \cdots$$
$$\cdots \subseteq K(\alpha_1, \ldots, \alpha_r)(\cup \beta_{ij}) = N.$$

□

Let us recall (see Ch.I, 1.26.1) that C_n denotes the (cyclic) group of the n^{th} roots of unity. Let us set $C_n := \{\varepsilon_1 = 1, \ldots, \varepsilon_n\}$.

Lemma 4.5. *Let $K(\subseteq \mathbb{C})$ be a field, and let L be the splitting field of the polynomial $x^n - 1$ over K ($n \in \mathbb{N}$). Then $\mathcal{G}(L, K)$ is abelian.*

Proof. Let us consider the group monomorphism

$$\Psi : \mathcal{G}(L, K) \longrightarrow S_n$$

defined in the proof of 3.7, i.e.

$$\Psi : \phi \mapsto \begin{pmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_n \\ \phi(\varepsilon_1) & \phi(\varepsilon_2) & \cdots & \phi(\varepsilon_n) \end{pmatrix}$$

(see also 3.6). Note that $\mathrm{Aut}\,(C_n)$, the group of automorphisms of C_n, is a subgroup of S_n and the permutation $\begin{pmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_n \\ \phi(\varepsilon_1) & \phi(\varepsilon_2) & \cdots & \phi(\varepsilon_n) \end{pmatrix}$ is an automorphism of C_n, since ϕ preserves products. Therefore $\mathrm{Im}(\Psi) \subseteq \mathrm{Aut}\,(C_n)$. But $\mathrm{Aut}\,(C_n) \cong \mathrm{Aut}\,(\mathbb{Z}_n)$ and this is isomorphic to the group of units of $\mathbb{Z}_n$ (see Ch.I, 1.17.1); then $\mathrm{Aut}\,(C_n)$ is abelian and this concludes the proof. □

Lemma 4.6. *Let $K(\subseteq \mathbb{C})$ be a field where $x^n - 1$ splits. Let $a \in K$ and let L be a splitting field for $x^n - a$ over K. Then $\mathcal{G}(L, K)$ is abelian.*

Proof. Let α be a zero of $x^n - a$. All the zeros of $x^n - a$ are $\varepsilon\alpha$, where $\varepsilon \in C_n \subset K$. In particular, $L = K(\alpha)$. If $\phi : L \longrightarrow L$ is a K-automorphism of L, since L is a splitting field of $x^n - a$ over K, then by 3.6 $\phi(\alpha) = \varepsilon\alpha$ for a suitable $\varepsilon \in C_n$. Analogously, if ψ is another K-automorphism of L, $\psi(\alpha) = \eta\alpha$, for a suitable $\eta \in C_n$. Hence $\phi(\psi(\alpha)) = \phi(\eta\alpha) = \eta\phi(\alpha) = \eta\varepsilon\alpha = \psi(\phi(\alpha))$. Therefore $\mathcal{G}(L, K)$ is abelian. □

Lemma 4.7. *Let p be a prime number and assume that the field K contains all the p^{th} roots of unity. Let $K \subseteq L$ be a normal extension of degree p. Then there exists an element $d \in L$ such that $d^p \in K$ and $L = K(d)$.*

Proof. Let $c \in L \setminus K$. Then $L = K(c)$; in fact $K \subseteq K(c) \subseteq L$ and

$$p = [L : K] = [L : K(c)][K(c) : K]$$

so necessarily $[L : K(c)] = 1$. Moreover $[L : K] = \sharp\mathcal{I}(L, K)$, by 2.2; on the other hand $K \subseteq L$ is normal, then $\mathcal{I}(L, K) = \mathcal{G}(L, K)$; hence $\mathcal{G}(L, K)$ is cyclic of order p. Let $\phi \in \mathcal{G}(L, K)$ be a generator. Put $c_i := \phi^{i-1}(c)$, $(i = 1, \ldots, p)$. Then $c_1 = c$, and $\phi(c_i) = c_{i+1}$ $(i = 1, \ldots, p-1)$, $\phi(c_p) = \phi^p(c) = c_1$. Let $C_p = \{\varepsilon_1, \ldots, \varepsilon_p\} \subseteq K$ be the set of the p^{th} roots of 1 and set

$$d_i := c_1 + c_2\varepsilon_i + c_3\varepsilon_i^2 + \cdots + c_p\varepsilon_i^{p-1}, \quad i = 1, \ldots, p \tag{1}$$

(the elements d_i's so defined are called the *Lagrange resolvents*). Since $\varepsilon_i \in K$, $\phi(\varepsilon_i) = \varepsilon_i$, then it is easy to see that $\phi(d_i) = \varepsilon_i^{-1} d_i$ and so $\phi(d_i^p) = d_i^p$. From this it follows that d_i^p is fixed by every element of $\mathcal{G}(L, K)$, hence $d_i^p \in L^{\mathcal{G}(L,K)} = K$ (by 3.14). Consider now (1) as a linear system in the variables $c_1, \ldots, c_p$. It is a $p \times p$ linear system whose determinant is the Vandermonde determinant. Its value is known to be $\prod_{i>j}(\varepsilon_i - \varepsilon_j)$, hence it is in K and is not zero. Therefore (1) is a system with exactly one solution. We can find it using for instance the Cramer's rule, and we see that every c_i is a polynomial in $d_1, \ldots, d_p$ with coefficients in K. In particular $c \in K(d_1, \ldots, d_p)$, so $L = K(c) = K(d_1, \ldots, d_p)$. Hence there exists an index i such that $d_i \notin K$, and therefore $L = K(d_i)$. Since $d_i^p \in K$, the lemma is proved. □

Lemma 4.8. *Let $f(x) \in K[x]$ and let $K \subseteq K'$ be an extension. Call N the splitting field of f over K and N' the splitting field of f over K'. Then the Galois group $\mathcal{G}(N', K')$ of f over K' is (isomorphic to) a subgroup of the Galois group $\mathcal{G}(N, K)$ of f over K.*

Proof. Let $\alpha_1, \ldots, \alpha_n$ be the roots of f in $\mathbb{C}$, hence $N = K(\alpha_1, \ldots, \alpha_n)$ and $N' = K'(\alpha_1, \ldots, \alpha_n)$. Let $\phi \in \mathcal{G}(N', K')$, then $\phi(\alpha_i) \in \{\alpha_1, \ldots, \alpha_n\}$ and $\phi(k') = k'$ for every $k' \in K'$. Therefore $\phi(k) = k$ for every $k \in K$ and $\phi(N) \subseteq N$, so $\phi|_N \in \mathcal{G}(N, K)$. This defines a map

$$\mathcal{G}(N', K') \longrightarrow \mathcal{G}(N, K)$$

given by $\phi \mapsto \phi|_N$. Moreover, if $\psi \in \mathcal{G}(N', K')$ and $\psi|_N = \phi|_N$, then $\psi(\alpha_i) = \phi(\alpha_i)$, so $\psi = \phi$ by 3.6. Therefore the previous map is injective. It is clear that is also a group homomorphism. □

Now we have all the necessary lemmas to give the proof of 4.3.

Proof (of theorem 4.3). Suppose first that $f(x) = 0$ is solvable by radicals over K. Let N be the splitting field of f over K. We want to show that $\mathcal{G}(N, K)$ is solvable. Let us divide the proof in some steps.
I) N is contained in a normal radical extension M of K.
In fact N is contained in a radical extension L of K by assumption (see 4.2). Let M be the normal closure of $K \subseteq L$. Then, by 4.4, $K \subseteq M$ is a radical extension, and $K \subseteq M$ is normal; clearly $N \subseteq M$.
II) Since $K \subseteq M$ is radical, there exists a chain of subfields

$$K \subseteq K(\alpha_1) \subseteq K(\alpha_1, \alpha_2) \subseteq \cdots \subseteq K(\alpha_1, \ldots, \alpha_r) = M$$

where $\alpha_i^{n(i)} \in K(\alpha_1, \ldots, \alpha_{i-1})$.
Let n be the l.c.m. of the $n(i)$'s and let ε be a primitive n^{th}-root of the unity.

III) We want to show that $K \subseteq M(\varepsilon)$ is a normal radical extension.
In fact, $K \subseteq M$ is normal, then M is the splitting field of a polynomial $g \in K[x]$ (see 3.9); therefore $M(\varepsilon)$ is the splitting field of $g(x)(x^n - 1)$, since it contains all its roots, being ε primitive. Hence $K \subseteq M(\varepsilon)$ is normal. Moreover the tower of subfields

$$K = K_1 \subseteq K_2 = K_1(\varepsilon) \subseteq K_3 = K_2(\alpha_1) \subseteq \cdots \subseteq K_{r+2} = K_{r+1}(\alpha_r) = M(\varepsilon) \qquad (2)$$

gives $M(\varepsilon)$ as a radical extension of K, since $K_i \subseteq K_{i+1}$ is clearly radical and $\varepsilon^n = 1 \in K$, so $K \subseteq K(\varepsilon)$ is radical.
IV) The extensions $K_i \subseteq K_{i+1}$ appearing in the tower (2) are splitting fields for equations of the kind $x^{n(i-1)} - a_{i-1} = 0$.
In fact, since ε is a primitive n^{th}-root of unity, then $K_1(\varepsilon)$ is the splitting field over K of $x^n - 1$. Let now $i > 1$; then $K_{i+1} = K_i(\alpha_{i-1})$ where $\alpha_{i-1}^{n(i-1)} = a_{i-1}$, for a suitable $a_{i-1} \in K_i$. Hence K_{i+1} is the splitting field of $x^{n(i-1)} - a_{i-1}$ over K_i. In fact, if β is a root of $x^{n(i-1)} - a_{i-1}$, then $\beta^{n(i-1)} = a_{i-1}$, so $(\beta/\alpha_{i-1})^{n(i-1)} = 1$, and since $n(i-1) \mid n$, also $(\beta/\alpha_{i-1})^n = 1$, so $\beta/\alpha_{i-1} \in K_1(\varepsilon)$, hence $\beta \in K_1(\varepsilon, \alpha_{i-1}) \subset K_i(\alpha_{i-1})$.
V) The extensions $K_i \subseteq K_{i+1}$ in (2) are normal and the Galois groups $\mathcal{G}(K_{i+1}, K_i)$ are abelian.
The above extensions are clearly normal, since splitting fields, by IV). From 4.5, we get that $\mathcal{G}(K_1(\varepsilon), K)$ is abelian. Since, for $i > 1$, K_{i+1} is the splitting field of $x^{n(i-1)} - a_{i-1}$ over K_i, we can apply 4.6 and we get that $\mathcal{G}(K_{i+1}, K_i)$ is abelian.
VI) Final step.
Let $H := \mathcal{G}(M(\varepsilon), K)$. From (2) and 3.14 *ii*) we get the chain of subgroups

$$H = \mathcal{G}(M(\varepsilon), K_1) \geq \mathcal{G}(M(\varepsilon), K_2) \geq \cdots \geq \mathcal{G}(M(\varepsilon), M(\varepsilon)).$$

Set $H_i := \mathcal{G}(M(\varepsilon), K_i)$. Since K_{i+1} is a normal extension of K_i, then H_{i+1} is a normal subgroup of H_i (see 3.14, iii)) and $H_i/H_{i+1} \cong \mathcal{G}(K_{i+1}, K_i)$ (by 3.14, iv)), so H_i/H_{i+1} is abelian. This shows that $\mathcal{G}(M(\varepsilon), K)$ is a solvable group.
Since $K \subseteq N \subseteq M(\varepsilon)$ are both normal extensions of K, then $\mathcal{G}(N, K)$ is a quotient of $\mathcal{G}(M(\varepsilon), K)$ (see 3.14 iv)), hence it is solvable (see Ch.I, 5.3).

Conversely, assume that the Galois group $G = \mathcal{G}(N, K)$ of f over K is solvable, where N is the splitting field of f over K. Let $n := |G| = [N : K]$. Let $K' = K(\varepsilon)$, where ε is a primitive n^{th}-root of unity, and let N' be the splitting field of f over K'. By 4.8, $\mathcal{G}(N', K')$ is isomorphic to a subgroup H of G. Hence H is solvable (Ch.I, 5.3) and has a composition series:

$$1 = H_{r+1} \lhd \cdots \lhd H_2 \lhd H_1 = H$$

whose composition factors H_i/H_{i+1} are cyclic of prime order p_i $(1 \leq i \leq r)$ (see Ch.I, 5.8).
Setting $K_i := (N')^{H_i}$ we have the corresponding increasing chain of subfields

$$K(\varepsilon) = K' = K_1 \subseteq K_2 \subseteq \cdots \subseteq K_{r+1} = N'. \qquad (3)$$

Since $H_i = \mathcal{G}(N', K_i)$ for all i (see the bijection in 3.14), then $\mathcal{G}(N', K_{i+1}) \lhd \mathcal{G}(N', K_i)$. So, by the Galois correspondence (see 3.14, iii) and iv)), $K_i \subseteq K_{i+1}$ is a normal extension

and $\mathcal{G}(K_{i+1}, K_i) \cong \mathcal{G}(N', K_i)/\mathcal{G}(N', K_{i+1}) = H_i/H_{i+1}$, which is cyclic of order p_i. Since $H_{i+1} \leq H_i \leq H \leq G$ and $|H_i| = p_i|H_{i+1}|$, then by Lagrange theorem, $p_i \,|\, n \ (= |G|)$; then $C_{p_i} \leq C_n$; moreover $K_i \supseteq K(\varepsilon)$, then it contains a primitive n^{th}-root of unity; so K_i contains C_n and hence it contains also the ${p_i}^{\text{th}}$ roots of 1.
We can now apply 4.7 and we get that $K_{i+1} = K_i(\alpha_i)$, where $\alpha_i^{p_i} \in K_i$. From this and from the fact that $K \subseteq K(\varepsilon)$ is a radical extension ($\varepsilon^n = 1 \in K$), we get that N' contains a tower of radical extensions of K

$$K \subseteq K(\varepsilon) = K_1 \subseteq K_2 \subseteq \cdots \subseteq K_{r+1} = N'$$

which extends (3), i.e. N' is a radical extension of K, and since N' contains the splitting field N, then $f(x) = 0$ is solvable by radicals over K. □

These theorems allow us to show that there are equations which are not solvable by radicals. It is enough in fact to give a polynomial whose Galois group is not solvable (for instance S_m with a suitable $m \geq 5$: see Ch.I, 5.7).

Theorem 4.9. *Let p be a prime number, $f \in \mathbb{Q}[x]$ irreducible over $\mathbb{Q}$ of degree p. Suppose that f has exactly two non-real zeros in $\mathbb{C}$. Then the Galois group of f over $\mathbb{Q}$ is the whole symmetric group S_p.*

Proof. First note that f is, up to a constant, a minimum polynomial of α over $\mathbb{Q}$, where α is any root of f in $\mathbb{C}$; hence $p = [\mathbb{Q}(\alpha) : \mathbb{Q}]$ by 1.14.
Let L be the splitting field of f, then the Galois group of f over $\mathbb{Q}$ is $G := \mathcal{G}(L, \mathbb{Q})$. Using the identification considered in theorem 3.7, we can assume that $G \leq S_p$. We have: $\mathbb{Q} \subseteq \mathbb{Q}(\alpha) \subseteq L$, hence $p = [\mathbb{Q}(\alpha) : \mathbb{Q}]$ divides $[L : \mathbb{Q}] = |\mathcal{G}(L, \mathbb{Q})|$ (see 1.13 and 2.2).
By Sylow theorem (Ch.I, 2.8 *d*)), G has an element of order p. The only elements of S_p of order p are the p-cycles, so G contains a p-cycle. Complex conjugation is a $\mathbb{Q}$-automorphism of $\mathbb{C}$ which induces a $\mathbb{Q}$-automorphism of L, since L is normal. This fixes the $p-2$ real roots of f and exchanges the two non-real roots. Hence G, as subgroup of S_p, contains a 2-cycle. But if a subgroup of S_p contains a 2-cycle and a p-cycle, it is necessarily S_p (see Ch.I, 4.5). Therefore $G = S_p$. □

It is easy to construct polynomials with the characteristics considered in the above theorem. For instance the polynomial $x^5 - 6x + 3 \in \mathbb{Q}[x]$ is not solvable by radicals, since it has exactly 2 non real roots and we can apply 4.9 and 4.3 (see also Ch.I, 5.7).

As an example of applications of the previous results, we want to sketch how to find a formula for the solution of cubic equations. First of all note:

Remark 4.10. If $f(x) = x^3 - s_1x^2 + s_2x - s_3 \in K[x]$, then the Galois group of f is a (not necessarily proper) subgroup of S_3 (see 3.7). Since S_3, and hence any subgroup of it, is solvable (see Ch.I, 5.1.2 and 5.3), any cubic equation is solvable by radicals.

For technical reasons (which will be clear later) we assume that the field K contains the cubic roots of 1, i.e. $1, -1/2 + \sqrt{3}/2\,\mathrm{i}, -1/2 - \sqrt{3}/2\,\mathrm{i}$. For instance

$$K = \mathbb{Q}(-1/2 + \sqrt{3}/2\,\mathrm{i}, -1/2 - \sqrt{3}/2\,\mathrm{i}).$$

Suppose further that f is irreducible and the Galois group of f is S_3; for instance, f could be a polynomial with exactly one real root (see 4.9).
If $\alpha, \beta, \gamma \in \mathbb{C}$ are the roots of $f(x) = 0$, then $L := K(\alpha, \beta, \gamma) = K[\alpha, \beta, \gamma]$ is the splitting field of f. The Galois group G of f is assumed, as said, to be the whole S_3. The elements of G are then:

$$\sigma_1 \cdots \begin{cases} \alpha \mapsto \alpha \\ \beta \mapsto \beta \\ \gamma \mapsto \gamma \end{cases} \leftrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \sigma_2 \cdots \begin{cases} \alpha \mapsto \alpha \\ \beta \mapsto \gamma \\ \gamma \mapsto \beta \end{cases} \leftrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix},$$

$$\sigma_3 \cdots \begin{cases} \alpha \mapsto \beta \\ \beta \mapsto \alpha \\ \gamma \mapsto \gamma \end{cases} \leftrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad \sigma_4 \cdots \begin{cases} \alpha \mapsto \beta \\ \beta \mapsto \gamma \\ \gamma \mapsto \alpha \end{cases} \leftrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix},$$

$$\sigma_5 \cdots \begin{cases} \alpha \mapsto \gamma \\ \beta \mapsto \beta \\ \gamma \mapsto \alpha \end{cases} \leftrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \quad \sigma_6 \cdots \begin{cases} \alpha \mapsto \gamma \\ \beta \mapsto \alpha \\ \gamma \mapsto \beta \end{cases} \leftrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

The composition series of S_3 is:

$$\{\sigma_1\} \triangleleft A_3 \triangleleft S_3, \tag{4}$$

where

$$A_3 := \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \right\}$$

is abelian since cyclic of order 3. So

$$S_3/A_3 = \left\{ \left[\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \right], \left[\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \right] \right\}$$

is abelian, since cyclic of order 2 (see Ch.I, 4.7, 4.12, 5.1.2). From the Galois correspondence (theorem 3.14), it is clear that $L^{\{\sigma_1\}} = L$ and $L^{S_3} = K$; so we have the tower

$$L^{S_3} = K \subseteq L^{A_3} \subseteq L = L^{\{\sigma_1\}}.$$

It remains to compute $H := L^{A_3}$. It is useful in the next computations, to use the following relations:

$$\begin{cases} \alpha + \beta + \gamma = s_1 \\ \alpha\beta + \alpha\gamma + \beta\gamma = s_2 \\ \alpha\beta\gamma = s_3 \end{cases} \tag{5}$$

coming from the equality $f(x) = (x - \alpha)(x - \beta)(x - \gamma)$.

We want now to sketch the proof of the fact that $H = K[\phi, \psi]$, where

$$\phi := \alpha^2\beta + \beta^2\gamma + \gamma^2\alpha \quad \text{and} \quad \psi := \alpha\beta^2 + \beta\gamma^2 + \gamma\alpha^2.$$

In order to see this, consider first of all the elements of the form

$$F(u,v) := \alpha^u\beta^v + \beta^u\gamma^v + \gamma^u\alpha^v.$$

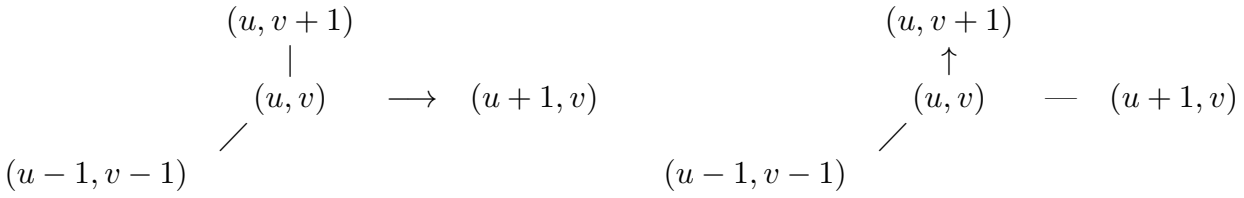
They are clearly in H since they are invariant under $A_3 = \{\sigma_1, \sigma_4, \sigma_5\}$. It is easy to see that

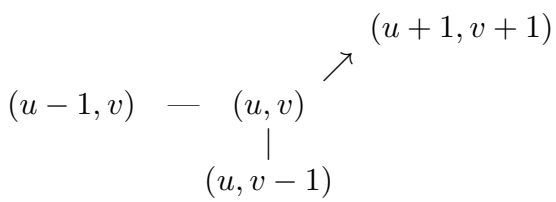
$$\begin{aligned} F(u+1,v) &= s_1F(u,v) - F(u,v+1) - s_3F(u-1,v-1) \\ F(u,v+1) &= s_1F(u,v) - F(u+1,v) - s_3F(u-1,v-1) \\ F(u+1,v+1) &= s_2F(u,v) - s_3F(u,v-1) - s_3F(u-1,v) \end{aligned} \tag{6}$$

Recall now that any symmetric polynomial in α, β, γ can be given as a polynomial expression in s_1, s_2, s_3, hence in particular, it is in K (see [S], Ch.2, Thm. 2.9).
Therefore the symmetric polynomials $F(u,0)$, $F(0,v)$ and $F(u,u)$ are in $K \subset K[\phi,\psi]$, for every u and v.
Moreover we also have $F(2,1) = \phi \in K[\phi,\psi]$ and $F(1,2) = \psi \in K[\phi,\psi]$. Therefore, applying (6), we get for instance:

$$F(3,1) = s_1F(2,1) - F(2,2) - s_3F(1,0) \in K[\phi,\psi].$$

In general we get that $F(u,v) \in K[\phi,\psi]$ for every u and v, using a sort of induction, based on the following schemes (obtained from (6)):





Let now $f(\alpha,\beta,\gamma)$ be any polynomial in H. Taking into account that f must be invariant under A_3, if $c\alpha^i\beta^j\gamma^h$ is a monomial in f, then also $\sigma_4(c\alpha^i\beta^j\gamma^h) = c\alpha^j\beta^h\gamma^i$ and $\sigma_5(c\alpha^i\beta^j\gamma^h) = c\alpha^h\beta^i\gamma^j$ must be monomials in f, hence

$$f = c\alpha^i\beta^j\gamma^h + c\alpha^j\beta^h\gamma^i + c\alpha^h\beta^i\gamma^j + \text{other monomials.}$$

Let us assume that $i = \min\{i,j,h\}$. Hence $c\alpha^i\beta^j\gamma^h + c\alpha^j\beta^h\gamma^i + c\alpha^h\beta^i\gamma^j = cs_3^iF(j-i,h-i)$ and in this way we see that $f \in K[\phi,\psi]$, therefore $H \subseteq K[\phi,\psi]$; the other inclusion follows from the fact that $\phi, \psi \in H$.

We know from 3.14 that $K \subseteq H$ has S_3/A_3 as the Galois group and $H \subseteq L$ has A_3 as Galois group. Note that $[H:K] =\mid S_3/A_3 \mid= 2$ and $[L:H] = |A_3| = 3$.

To better understand the following considerations, we suggest to keep in mind lemma 4.7 and its proof (which will be applied here to the two extensions $K \subseteq H$ and $H \subseteq L$). Observe that $\phi+\psi = s_1s_2-3s_3 \in K$, so $\psi \in K[\phi]$; therefore, if $\phi \notin K$, then $K[\phi,\psi] = K[\phi]$. Setting

$$\eta := \left[\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}\right]$$

the generator of S_3/A_3, then $\eta(\phi) = \psi$. Following 4.7 we see that the two Lagrange resolvents are: $\phi - \psi$ and $\phi + \psi$. Necessarily $\phi - \psi \notin K$, since $\phi + \psi \in K$, as remarked above. Moreover $(\phi - \psi)^2$ must be in K, accordingly to 4.7. In fact, using the relations (5) one verifies that

$$(\phi - \psi)^2 = s_1^2s_2^2 + 18s_1s_2s_3 - 27s_3^2 - 4s_1^3s_3 - 4s_2^3 \in K.$$

Hence $H = K[d]$, where $d = \phi - \psi$ and

$$d^2 = s_1^2s_2^2 + 18s_1s_2s_3 - 27s_3^2 - 4s_1^3s_3 - 4s_2^3. \tag{7}$$

To know the value of ϕ and ψ w.r.t. d, one has to solve the linear system:

$$\begin{cases} \phi + \psi = s_1s_2 - 3s_3 \\ \phi - \psi = d \end{cases} \tag{8}$$

Now we consider the extension $H \subseteq L$. Its Galois group is A_3. It is clear that at least one of the three roots α, β, γ is not in H. Assume $\alpha \notin H$. Then, following again the proof of 4.7, $L = H(\alpha)$. A generator of A_3 is $\theta := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$. Note that $\theta(\alpha) = \beta$, $\theta(\beta) = \gamma$ and $\theta(\gamma) = \alpha$. Let, as usual, $C_3 = \{1, \varepsilon, \varepsilon^2\}$ ($\varepsilon = -1/2 + \sqrt{3}/2\mathrm{i}$) be the set of cubic roots of 1. We assume, as said, they are in K (hence in H). The Lagrange resolvents are:

$$\begin{aligned} d_1 &= \alpha + \beta + \gamma \\ d_2 &= \alpha + \beta\varepsilon + \gamma\varepsilon^2 \\ d_3 &= \alpha + \beta\varepsilon^2 + \gamma\varepsilon \end{aligned} \tag{9}$$

At least one of the d_i's is not in H. Since $d_1 = s_3 \in K$, then at least one of d_2, d_3, say δ, is not in H. Then $\delta^3 \in H = K[\phi,\psi] = K[d]$. In fact it is easy to verify that

$$d_2^3 = s_1^3 - 6s_3 - 3(s_1s_2 - 3s_3) + 3\varepsilon\phi + 3\varepsilon^2\psi + 6s_1 \in H \tag{10}$$

and, analogously

$$d_3^3 = s_1^3 - 6s_3 - 3(s_1s_2 - 3s_3) + 3\varepsilon\psi + 3\varepsilon^2\phi + 6s_1 \in H. \tag{11}$$

Hence

$$K \subseteq K[d] \subseteq K[d][\delta].$$

so any element of $L = K[d][\delta]$ (in particular α, β, γ) can be expressed as a polynomial expression in d and δ. To find the value of α, β and γ we can consider (9) as a linear system in the variables α, β, γ. To simplify the notations, we shall assume that $\alpha + \beta + \gamma = 0$ (i.e. $s_1 = d_1 = 0$). This is not restrictive, since we can always have this condition setting $y = x - s_1/3$ in the polynomial $f(x)$. Solving the system (9) and using all the relations $(7), (8), (10), (11)$ defining ϕ, ψ, d_2, d_3 found so far, we find:

$$\alpha = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}$$

where $p := s_2$ and $q := s_3$.
This is the well known formula for the solution of cubic equations. Of course it could also be find without an explicit use of Galois theory. Let us conclude this section by recalling that there exists a similar formula also for the equations of degree four, according to the fact that S_4 is solvable.

References

[A] Ash R. *Complex variables*, Academic press, New York (1971)

[AMD] Atiyah M.F., Macdonald I.G. *Introduction to Commutative Algebra*, Addison-Wesley, Reading, Mass. (1969)

[J] Jacobson N. *Basic Algebra I*, Freeman and co., San Francisco (1974)

[L] Lang S. *Algebra*, 2nd edition, Addison-Wesley, Reading, Mass. (1974)

[L1] Lang S. *Complex analysis*, 2nd edition, Springer Verlag, New York (1985)

[M] Massey W.S. *Singular Homology Theory*, Springer Verlag, New York (1980)

[R] Rose J. *A course on Group Theory*, Cambridge Univ. Press, Cambridge (1978)

[S] Stewart I. *Galois Theory*, Chapman and Hall, London (1989)

Useful books regarding Galois theory are also:

Edwards H. *Galois Theory*, Springer Verlag, New York (1984)

Jacobson N. *Lectures in Abstract Algebra III*, D. Van Nostrand Co. (1964)

Procesi C. *Elementi di Teoria di Galois*, Decibel Editrice, I 77