ALGEBRA 2

E.N.S. 2014/15

Jan Nekovář

Introduction

This is a second course in algebra. It treats elementary properties of commutative rings (and modules over them), field extensions and Galois theory, as well as basic commutative algebra from a geometric perspective.

Throughout the course we are going to emphasise close links between abstract algebraic theory and concrete questions arising in arithmetic and geometry (cf. [Re, 9.9]), even though the course will not quite follow the historical development of the subject.

The purely abstract point of view, according to which a group is a set equipped with a binary operation satisfying suitable axioms, completely misses the point that most groups of interest naturally occur as symmetry groups. A typical example is provided by matrix groups – including orthogonal, symplectic and unitary groups – studied in the course Algebra 1.

The groups appearing in this course – Galois groups – are symmetry groups of polynomial equations in one variable. They can be realised concretely as permutation groups (subgroups of the symmetric group S_n), but many of them are miraculously related to matrix groups, as predicted by the Langlands programme (in fact, as we shall see, examples of Galois representations predated Galois theory itself).

Similarly, a commutative ring should not be considered merely as an abstract object, but as a ring of functions on a suitable "space". Depending on the context one can consider continuous functions (in topology), smooth functions (in differential geometry), holomorphic functions (in complex geometry) or polynomial functions (in algebraic geometry). We are going to be interested in purely algebraic objects (but see IV.12 below), a typical example of which is the polynomial ring $\mathbf{C}[x_1, \dots, x_n]$ (the ring of functions on the n-dimensional affine space over $\mathbf{C}$ with coordinates $x_1, \dots, x_n$) and its quotients $\mathbf{C}[x_1, \dots, x_n]/(f_1, \dots, f_m)$, which are rings of functions on subsets of the affine space defined by systems of polynomial equations $f_1(x_1, \dots, x_n) = \dots = f_m(x_1, \dots, x_n) = 0$.

There is a natural duality between points and functions. At its origin is the equation

$$f(x) = 0 \qquad (f = \text{ function}, x = \text{ point}),$$

considered traditionally for fixed f and variable x, but ever since a pioneering work of Dedekind and Weber in the 1880's also for fixed x and variable f (see 2.6 below). This duality leads to a dictionary between intuitively obvious – but much less obvious to define – geometric concepts (point, map, dimension, reducibility, singularity, multiplicity, localisation to a neighbourhood of a point) and their algebraic counterparts.

This geometric point of view applies equally well to rings naturally occurring in arithmetic, as anticipated by Kronecker. For example, it is well-known that the rings $\mathbf{Z}$ and $K[X]$ (for any field K) have very similar algebraic properties. Going one step further, one can adjoin to $\mathbf{Z}$ (resp. to $K[X]$) a square root i of -1 (resp. a square root Y of X) and obtain rings

$$\mathbf{Z}[i] = \{a + bi \mid a, b \in \mathbf{Z}\}, \qquad K[X,Y]/(Y^2 - X) = \{a + bY \mid a, b \in K[X]\} \quad (Y^2 = X).$$

The ring $K[X]$ (resp. $K[X,Y]/(Y^2 - X)$) is the ring of functions on a line L over the field K (resp. on the plane curve C over K defined by the equation $C : Y^2 - X = 0$). The inclusion of the rings $K[X] \subset K[X,Y]/(Y^2 - X)$ corresponds to the projection map $\alpha : C \longrightarrow L$, $\alpha(X,Y) = X$.

For any $t \in K$ the fibre $\alpha^{-1}(t)$ of α at t is given by the equation $Y^2 - t = 0$. For $t \neq 0$ (resp. $t = 0$) this represents two points – possibly defined over a bigger field than K – with multiplicity one (resp. a point with multiplicity two). In other words, the map α is unramified outside $t = 0$ and ramified – with ramification index equal to two – at $t = 0$.

The inclusion of rings $\mathbf{Z} \subset \mathbf{Z}[i]$ admits a similar geometric interpretation (see ?? below). In particular, it is ramified at the prime number $p = 2$, which plays the rôle analogous to that of the point $t = 0$.

A slightly more complicated example is provided by the inclusion $\mathbf{Z} \subset \mathbf{Z}[2i] = \{a + b \cdot 2i \mid a, b \in \mathbf{Z}\}$ and its geometric counterpart $K[X] \subset K[X, Y]/(Y^2 - X^3)$, which corresponds to the curve $C' : Y^2 - X^3 = 0$ with a singular point $(X, Y) = (0, 0)$ (and its projection to the line L).

To sum up, the rings $\mathbf{Z}$ and $\mathbf{Z}[i]$ are, from a geometric perspective, one-dimensional non-singular objects, while $\mathbf{Z}[2i]$ is a one-dimensional object with a singular point above the prime number $p = 2$ (see ?? below).

Arithmetic properties of the ring $\mathbf{Z}[i]$ of gaussian integers are closely intertwined with arithmetic of the quadratic form $x^2 + y^2 = (x + iy)(x - iy)$. Gauss showed that $\mathbf{Z}[i]$ is a unique factorisation domain (UFD): any non-zero element can be written as a product of "prime elements" in an essentially unique way. Together with an explicit description of such prime elements this yields a conceptual proof of a celebrated result of Fermat, according to which

$$\exists\, x, y \in \mathbf{Z} \quad q = x^2 + y^2 \quad \Longleftrightarrow \quad q \equiv 1 \pmod 4,$$

for any prime number $q \neq 2$.

A study of representations of prime numbers by quadratic forms

$$x^2 + xy + cy^2 = \left(x + y\,\frac{1 + \sqrt{1 - 4c}}{2}\right)\left(x + y\,\frac{1 - \sqrt{1 - 4c}}{2}\right) \qquad (c \in \mathbf{Z})$$

naturally leads to a consideration of the following rings

$$\mathbf{Z}\left[\frac{1 + \sqrt{1 - 4c}}{2}\right] = \left\{u + v\,\frac{1 + \sqrt{1 - 4c}}{2} \mid u, v \in \mathbf{Z}\right\}.$$

If $|1 - 4c|$ is equal to a prime number $p \leq 19$ ($\Longleftrightarrow$ $c = \pm 1, 2, \pm 3, -4, 5$) – and for certain other values of p as well – the corresponding ring is a UFD and an explicit description of its prime elements implies that

$$\exists\, x, y \in \mathbf{Z} \quad q = x^2 + xy + cy^2 \quad \Longleftrightarrow \quad q \equiv \text{ square } (\mathrm{mod}\, p),$$

for any prime number q.

For $c = 6$ and $p = |1 - 4c| = 23$ the ring $\mathbf{Z}[(1 + i\sqrt{23})/2]$ is no longer a UFD and representability of prime numbers by the quadratic form $x^2 + xy + 6y^2$ is no longer given by a simple congruence condition.

Closely related rings $\mathbf{Z}[\zeta_p] = \{a_0 + a_1\zeta_p + \cdots + a_{p-2}\zeta_p^{p-2} \mid a_j \in \mathbf{Z}\}$ (where $p \neq 2$ is a prime number and $\zeta_p = e^{2\pi i/p}$) naturally appeared in attempts at proving "Fermat's Last Theorem", namely, that the equation

$$x^p + y^p = z^p, \qquad x^p + y^p = \prod_{j=0}^{p-1}(x + \zeta_p^j y)$$

has no solutions $x, y, z \in \mathbf{Z} \setminus \{0\}$. If the ring $\mathbf{Z}[\zeta_p]$ is a UFD (which was tacitly assumed by some authors), then Fermat's method of infinite descent applies and Fermat's Last Theorem can be proved by this method. However, Kummer discovered that $\mathbf{Z}[\zeta_{23}]$ is not a UFD (in fact, $\mathbf{Z}[\zeta_p]$ is a UFD $\Longleftrightarrow$ $p \leq 19$), and therefore the standard method does not work. By an amazing leap of imagination, he managed to come up with a remedy: in each of the rings $\mathbf{Z}[\zeta_p]$ unique factorisation holds for "ideal numbers". Moreover, if the "difference" between the ideal and usual numbers in $\mathbf{Z}[\zeta_p]$ – the ideal class group – has size "prime to p", then a suitably modified method of infinite descent applies and Fermat's Last Theorem holds for exponent p. The exact nature of Kummer's ideal numbers was not quite apparent (even though he described "prime ideal numbers"). It was Dedekind who transformed them into ideals as we now know them. Rings such as $\mathbf{Z}[\zeta_p]$ in which there is unique factorisation for ideals are called Dedekind rings (see IV.15). They are non-singular and one-dimensional (and are, essentially, characterised by these properties).

Being a UFD is a global property of the ring $\mathbf{Z}[\zeta_p]$, analogous to the existence of functions on curves with arbitrarily prescribed zeroes and poles. A geometric counterpart of the ideal class group is an obstruction to the existence of such functions. In the complex analytic situation, this class group first appeared in the context of Abel's addition theorem for elliptic functions and their generalisations, and was studied by Abel, Jacobi and Riemann, among others. However, these topics are beyond the scope of the present course.

There will be many concrete examples scattered throughout these notes. The reader may begin by contemplating the following question: the standard picture of the unit circle $C : x^2 + y^2 - 1 = 0$ shows its real points $C(\mathbf{R})$. What are its complex points $C(\mathbf{C})$, rational points $C(\mathbf{Q})$ or points $C(\mathbf{Z}/n\mathbf{Z})$ with coordinates in $\mathbf{Z}/n\mathbf{Z}$?

Notation and conventions. The symbols $\subset$ and $\subseteq$ will be used interchangeably. A strict inclusion will be denoted by $\subsetneq$. The symmetric and the alternating group on n elements will be denoted, respectively, by S_n and A_n. The dihedral group with $2n$ elements will be denoted by D_{2n} (**not** D_n); thus $C_n \subset D_{2n} \subset S_n$ (where C_n is the cyclic group of order n).

I. Commutative rings (elementary properties)

1. Basic concepts

(1.1) Recall that a ring is a set A equipped with two binary operations, addition and multiplication, satisfying the usual properties: $(A,+)$ is an abelian group with neutral element 0, multiplication is associative and the distributive law $x(y+z)=xy+xz$, $(x+y)z=xz+yz$ holds for all $x,y,z\in A$. All rings in this course will be **unital** (there is a (unique) unit $1=1_A\in A$ such that $1x=x1=1$ for all $x\in A$) and **commutative** ($xy=yx$ for all $x,y\in A$). Note that $1\neq 0 \iff A\neq\{0\}=0$.

For a ring A we denote by $A[x_1,\dots,x_n]$ the ring of polynomials in the variables $x_1,\dots,x_n$ with coefficients in A.

(1.2) An element $x\in A$ is **invertible** if there is $y\in A$ (necessarily unique) such that $xy=1$. The set of invertible elements A^* is a group with respect to multiplication.

A ring $A\neq 0$ is a **domain** (= **integral domain**) if $x,y\in A\setminus\{0\}$ implies that $xy\in A\setminus\{0\}$. It is a **field** if $A^*=A\setminus\{0\}$. For example, $\mathbf{Z}/n\mathbf{Z}$ is a domain $\iff$ $\mathbf{Z}/n\mathbf{Z}$ is a field $\iff$ $n=p$ is a prime number.

Any domain A naturally embeds into its **fraction field** $\mathrm{Frac}(A)$, which consists of fractions $\frac{x}{y}$ ($x\in A$, $y\in A\setminus\{0\}$) subject to the relations

$$\frac{x}{y}=\frac{x'}{y'} \iff xy'=yx'.$$

An element $x\in A$ corresponds to $\frac{x}{1}\in\mathrm{Frac}(A)$ and the operations are given by the usual formulas

$$\frac{x}{y}+\frac{x'}{y'}=\frac{xy'+yx'}{yy'},\qquad \frac{x}{y}\,\frac{x'}{y'}=\frac{xx'}{yy'}.$$

We have $\mathrm{Frac}(\mathbf{Z})=\mathbf{Q}$ and, for any field K, $\mathrm{Frac}(K[x_1,\dots,x_n])=K(x_1,\dots,x_n)$ (the field of rational functions in n variables over K).

Most rings of interest to us will be related, in one way or another, to polynomial rings over $\mathbf{Z}$ or over a field.

(1.3) A ring A is a **subring** of a ring B if $A\subset B$ and if the operations "addition" and "multiplication" on A are induced by those on B (in particular, A and B have the same zero 0 and the same unit 1). For any subset $S\subset B$ we denote by $A[S]$ the intersection of all subrings of B containing A and S; it is a subring of B. If $S=\{b_1,\dots,b_n\}$ is finite, then $A[b_1,\dots,b_n]=\{f(b_1,\dots,b_n)\mid f\in A[x_1,\dots,x_n]\}$.

If, in addition, A and B are fields (in which case we say that A is a **subfield** of B), we denote by $A(S)$ the intersection of all subfields of B containing A and S; it is again a subfield of B and $A(b_1,\dots,b_n)=\{f(b_1,\dots,b_n)/g(b_1,\dots,b_n)\mid f,g\in A[x_1,\dots,x_n],\ g(b_1,\dots,b_n)\neq 0\}$.

(1.4) A **ring homomorphism** $f:A\longrightarrow B$ is a map between rings such that

$$\forall x,y\in A\quad f(x+y)=f(x)+f(y),\qquad \forall x,y\in A\quad f(xy)=f(x)f(y),\qquad f(1)=1$$

(the third condition is not always a consequence of the first two: consider $f:\mathbf{Z}/6\mathbf{Z}\longrightarrow\mathbf{Z}/6\mathbf{Z}$, $f(x)=3x$). We say that f is a **ring isomorphism** if there exists a ring homomorphism $g:B\longrightarrow A$ (necessarily unique) inverse to f. In this purely algebraic context an isomorphism is the same thing as a bijective homomorphism (which ceases to be the case if one considers algebraic objects equipped with a topology and continuous homomorphisms between them, since a continuous bijection does not necessarily have a continuous inverse).

The **image** of f

$$\mathrm{Im}(f)=\{f(x)\mid x\in A\}\subset B$$

is a subring of B. The **kernel** of f

$$\mathrm{Ker}(f)=\{x\in A\mid f(x)=0\}\subset A$$

is **not** a subring of A (unless $f=0$). It is an **ideal of** A (and all ideals arise in this way). The morphism f is injective $\iff$ $\mathrm{Ker}(f)=\{0\}=0$.

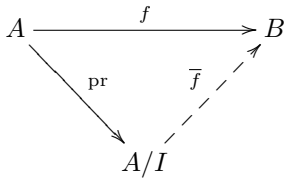
2. Ideals

(2.1) Why do we care about ideals? There are at least three different reasons: for an algebraist, ideals are kernels of ring homomorphisms; for a geometer, any system of (say, polynomial) equations is equivalent to a system given by an ideal; for a number theorist, ideals appear in the absence of unique factorisation.

(2.2) Definition. *An* **ideal** *of a ring A is an additive subgroup $I \subset (A,+)$ such that $AI \subset I$ (i.e., such that $ax \in I$ for all $a \in A$ and $x \in I$).*

(2.3) Examples of ideals. For any non-empty subset $S \subset A$ we denote by (S) the intersection of all ideals of A containing S. It is an ideal of A, equal to the set of all finite linear combinations $\sum_{s\in S} a_s s$ ($a_s \in A$, all but finitely many a_s are equal to 0). If $S = \{x\}$ consists of one element, then $(S) = Ax = \{ax \mid a \in A\}$ is the **principal ideal** generated by x. For example, $(0) = \{0\} = 0$, $(1) = A$ (more generally, $(x) = A \iff x \in A^*$). An ideal I is **finitely generated** if $I = (S)$ for a finite set S.

(2.4) Proposition. *(i) The kernel of any ring homomorphism $f : A \longrightarrow B$ is an ideal of A.*
(ii) For any ideal I of A the quotient abelian group A/I has a unique ring structure for which the canonical projection $\mathrm{pr} : A \longrightarrow A/I$ *($\mathrm{pr}(x) = x + I$) is a ring homomorphism. The kernel of* pr *is equal to* $\mathrm{Ker}(\mathrm{pr}) = I$.
(iii) (Universal property) For any ring homomorphism $f : A \longrightarrow B$ and any ideal I of A such that $I \subset \mathrm{Ker}(f)$ there is a unique ring homomorphism $\overline{f} : A/I \longrightarrow B$ satisfying $\overline{f} \circ \mathrm{pr} = f$:



(iv) (Isomorphism theorem) For any ring homomorphism $f : A \longrightarrow B$, the homomorphism $\overline{f}$ from (iii) for $I = \mathrm{Ker}(f)$ induces a ring isomorphism $\overline{f} : A/\mathrm{Ker}(f) \xrightarrow{\sim} \mathrm{Im}(f)$.

Proof. (i) Immediate. (ii) One needs to check that the product $\mathrm{pr}(x)\mathrm{pr}(y) := \mathrm{pr}(xy)$ is well-defined. If $\mathrm{pr}(x) = \mathrm{pr}(x')$ and $\mathrm{pr}(y) = \mathrm{pr}(y')$, then $x - x', y - y' \in I$, hence $xy - x'y' = x(y - y') + y(x - x') \in AI + AI \subset I + I = I$, which means that $\mathrm{pr}(xy) = \mathrm{pr}(x'y')$. (iii) The map $\overline{f}(\mathrm{pr}(x)) := f(x)$ is well-defined and has the required properties. (iv) The ring homomorphism $\overline{f} : A/\mathrm{Ker}(f) \longrightarrow \mathrm{Im}(f)$ is bijective, hence is an isomorphism.

(2.5) The ring A/I consists of the equivalence classes for the relation $x \equiv y \pmod I \iff x - y \in I$. The notation for the equivalence class containing $x \in A$ will vary according to the author's mood; it can be written as $\mathrm{pr}(x) = x + I = \overline{x} = x \pmod I$.

If A is a ring and $f = X^d + a_1X^{d-1} + \cdots + a_d \in A[X]$ a monic polynomial of degree $d \geq 1$, then the quotient ring $A[X]/(f)$ is equal to $\{g \pmod{(f)} \mid g \in A[X], \deg(g) < d\}$.

In particular, if $f = X - a$ ($a \in A$) has degree $\deg(f) = 1$, then the ideal $(X - a)$ coincides with the kernel of the evaluation map

$$\mathrm{ev}_a : A[X] \longrightarrow A, \qquad g(X) \mapsto g(a)$$

and we deduce from Proposition 2.4(iv) an isomorphism $\overline{\mathrm{ev}}_a : A[X]/(X - a) \xrightarrow{\sim} A$.

(2.6) Ideals in geometry. What is a geometric object? According to H. Cartan, such an object consists, at least locally, of a space X and a ring of functions $O(X)$ on X. The author of these notes likes to think of X as being the body, and $O(X)$ the soul.

For any subset $Y \subset X$ the set of functions vanishing on Y

$$I = \{f \in O(X) \mid \forall y \in Y \quad f(y) = 0\} \subset O(X)$$

is an ideal of $O(X)$. If Y is a sufficiently nice subspace of X, then it makes sense to consider the ring of functions $O(Y)$ on Y and the restriction map $\mathrm{res} : O(X) \longrightarrow O(Y)$. In this case $I = \mathrm{Ker}(\mathrm{res})$ and

$\overline{\text{res}} : O(X)/I \xrightarrow{\sim} \text{Im}(\text{res}) \subset O(Y)$. In particular, if res is surjective (which will be the case in affine algebraic geometry), then $O(Y) \xrightarrow{\sim} O(X)/I$ can be recovered from I.

If $Y' \supset Y$, then the corresponding ideal satisfies $I' \subset I$. In particular, "small" subspaces Y (= points) should correspond to "big" ideals $I \subset A$ (and vice versa).

Conversely, for any non-empty subset $S \subset O(X)$ its zero locus

$$Y = \{x \in X \mid \forall f \in S \quad f(x) = 0\} = \{x \in X \mid \forall f \in I \quad f(x) = 0\}$$

coincides with the zero locus of the ideal $I = (S)$ generated by S. More precisely, the two systems of equations $\forall f \in S \quad f(x) = 0$ (resp. $\forall f \in I \quad f(x) = 0$) are equivalent.

In the last part of this course we shall investigate the correspondence between subspaces and ideals in the algebraic situation, when X is an n-dimensional affine space over an algebraically closed field K and $O(X)$ is the polynomial ring in n variables over K.

(2.7) Ideals in arithmetic. The ring of integers $\mathbf{Z}$ has the unique factorisation property: any non-zero element can be written in a unique way (up to a sign) as a product of prime numbers. This is no longer the case for more general rings occurring in number theory. For example, in the ring $A = \mathbf{Z}[i\sqrt{5}] = \{a + bi\sqrt{5} \mid a, b \in \mathbf{Z}\}$ the two factorisations

$$21 = 3 \cdot 7 = (4 + i\sqrt{5})(4 - i\sqrt{5})$$

cannot be refined any further – all elements $3, 7, 4 \pm i\sqrt{5}$ are "irreducible" in A. One can try to follow Kummer's ideas and imagine that there are additional factorisations involving new objects, "divisors":

$$3 = P_1P_2, \qquad 7 = Q_1Q_2, \qquad 4 + i\sqrt{5} = P_1Q_1, \qquad 4 - i\sqrt{5} = P_2Q_2.$$

If true, then one can identify each divisor $D \in \{P_1, P_2, Q_1, Q_2\}$ with the set of elements of A divisible by D, which will be an ideal of A. The above factorisations then become equalities between ideals (for a suitable notion of product, defined in 2.8 below):

$$(3) = P_1P_2, \qquad (7) = Q_1Q_2, \qquad (4 + i\sqrt{5}) = P_1Q_1, \qquad (4 - i\sqrt{5}) = P_2Q_2.$$

This is, indeed, true (see 6.9 below). Moreover, A is a Dedekind ring (see IV.15 below) which means that unique factorisation holds for non-zero ideals of A.

(2.8) Operations on ideals. If I and J are ideals of A, so is their intersection $I \cap J$, sum $I + J = \{x + y \mid x \in I, y \in J\}$, product $IJ = (\{xy \mid x \in I, y \in J\}) = \{\sum_{\alpha=1}^{r} x_\alpha y_\alpha \mid x_\alpha \in I, y_\alpha \in J, r \geq 0\}$ and the **radical of** I: $\sqrt{I} = \{x \in A \mid \exists n \geq 1 \quad x^n \in I\}$. More generally, the intersection of any collection of ideals of A is an ideal. If $I = (x_1, \ldots, x_m)$ and $J = (y_1, \ldots, y_n)$, then $I + J = (x_1, \ldots, x_m, y_1, \ldots, y_n)$ and $IJ = (x_1y_1, \ldots, x_1y_n, \ldots, x_my_n)$.

For any ring homomorphism $f : A \longrightarrow B$, the inverse image of any ideal J of B is an ideal of A, since $I = f^{-1}(J) = \text{Ker}(\text{pr} \circ f : A \longrightarrow B \longrightarrow B/J)$.

For any ideal I of A there is a natural bijection between the set of ideals $\overline{J}$ of A/I and the set of ideals J of A containing I: $\overline{J} = J/I$ and $J = \text{pr}^{-1}(\overline{J})$. Moreover, $(A/I)/\overline{J} = A/J$.

(2.9) Example. If $A = \mathbf{Z}$ and $I = (m)$, $J = (n)$ for integers $m, n \geq 1$, then $(m) + (n) = (m, n) = (\gcd(m, n))$, $(m) \cap (n) = (\text{lcm}(m, n))$, $(m)(n) = (mn)$ and $\sqrt{(m)} = (m_0)$, where m_0 is the product of distinct prime numbers dividing m.

3. Product of rings, idempotents

(3.1) This section is a variation on the Chinese remainder theorem, which states that $\mathbf{Z}/mn\mathbf{Z} \xrightarrow{\sim} \mathbf{Z}/m\mathbf{Z} \times \mathbf{Z}/n\mathbf{Z}$ whenever $(m, n) = (1)$.

(3.2) Let $A_1, \ldots, A_n$ $(n \geq 2)$ be rings. The cartesian product $A = A_1 \times \cdots \times A_n$ has a natural ring structure

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n), \qquad (x_1, \ldots, x_n)(y_1, \ldots, y_n) = (x_1y_1, \ldots, x_ny_n),$$

with $1_A = (1, \ldots, 1) = (1_{A_1}, \ldots, 1_{A_n})$. The projections

$$p_k : A \longrightarrow A_k, \qquad p_k(x_1, \ldots, x_n) = x_k$$

are ring homomorphisms, but the inclusions

$$i_k : A_k \longrightarrow A, \qquad i_k(x_k) = (0, \ldots, x_k, \ldots, 0)$$

are not, since $e_k := i_k(1_{A_k}) = (0, \ldots, 1, \ldots, 0) \neq 1_A$. The elements $e_1, \ldots, e_n \in A$ have the following properties.

(3.2.1) $e_k^2 = e_k$ (each e_k is an idempotent);
(3.2.2) $e_j e_k = 0$ if $j \neq k$ (orthogonality);
(3.2.3) $e_1 + \cdots + e_n = 1$.

(3.3) Proposition. *If A is a ring and $e_1, \ldots, e_n \in A$ are elements satisfying (3.2.1)–(3.2.3), then, for each $k = 1, \ldots, n$, the subset $A_k := e_k A = \{e_k a \mid a \in A\} \subset A$ equipped with the operations addition and multiplication coming from A is a ring with unit $1_{A_k} = e_k$. The maps $p_k : A \longrightarrow A_k$, $p_k(a) = e_k a$ are ring homomorphisms and they induce a ring isomorphism*

$$p = (p_1, \ldots, p_n) : A \xrightarrow{\sim} A_1 \times \cdots \times A_n, \qquad p(a) = (e_1 a, \ldots, e_n a),$$

whose inverse is given by $p^{-1}(a_1, \ldots, a_n) = e_1 a_1 + \cdots + e_n a_n$. For any ideal I in A the image $I_k = p_k(I) = e_k I$ is an ideal in A_k and $p(I) = I_1 \times \cdots \times I_n$.

Proof. Easy exercise.

(3.4) Proposition. *If A is a ring and I, J are ideals such that $I + J = (1)$, then $I \cap J = IJ$ and the map*

$$A/(I \cap J) \xrightarrow{\sim} A/I \times A/J, \qquad x + (I \cap J) \mapsto (x + I,\, x + J)$$

is a ring isomorphism.

Proof. The map $\alpha : A \longrightarrow A/I \times A/J$, $\alpha(x) = (x + I, x + J)$ is a ring homomorphism with kernel $\mathrm{Ker}(\alpha) = I \cap J$, inducing an isomorphism $\overline{\alpha} : A/(I \cap J) \xrightarrow{\sim} \mathrm{Im}(\alpha)$. We must show that α is surjective: by assumption there exist $i \in I$ and $j \in J$ such that $i + j = 1$; then $\alpha(i) = (0, 1)$, $\alpha(j) = (1, 0)$ and $\alpha(xj + yi) = (x + I, y + J)$ for all $x, y \in A$. The inclusion $IJ \subset I \cap J$ always holds. Conversely, if $a \in I \cap J$, then $a = ai + aj \in JI + IJ = IJ$.

(3.5) Corollary (Chinese remainder theorem). *If A is a ring and $I_1, \ldots, I_n$ are ideals of A such that $I_j + I_k = (1)$ whenever $j \neq k$, then $I_1 \cap \cdots \cap I_n = I_1 \cdots I_n$ and the map*

$$A/(I_1 \cap \cdots \cap I_n) \xrightarrow{\sim} A/I_1 \times \cdots \times A/I_n, \qquad x + (I_1 \cap \cdots \cap I_n) \mapsto (x + I_1, \ldots, x + I_n)$$

is a ring isomorphism.

(3.6) Example of idempotents. For each $n \geq 1$ we have the idempotents $e_n, 1 - e_n \in \mathbf{Z}/10^n\mathbf{Z}$ corresponding to $(1, 0), (0, 1)$ under the isomorphism

$$\mathbf{Z}/10^n\mathbf{Z} \xrightarrow{\sim} \mathbf{Z}/2^n\mathbf{Z} \times \mathbf{Z}/5^n\mathbf{Z}$$

(in other words, $e_n \equiv 1 \pmod{2^n}$ and $e_n \equiv 0 \pmod{5^n}$). Here is a table of the first few values of e_n:

n	$e_n \pmod{10^n}$	$1 - e_n \pmod{10^n}$
1	5	6
2	25	76
3	625	376
4	0625	9376

If we keep increasing n, we obtain in the limit two numbers $e = \cdots 0625$ and $1 - e = \cdots 9376$ with infinitely many decimal digits which satisfy $e^2 = e, (1 - e)^2 = 1 - e$.

Question: what are they? Where do e and $1 - e$ live?

Answer: e and $1 - e$ are 10-adic integers, living in the projective limit $\mathbf{Z}_{10} = \varprojlim_n \mathbf{Z}/10^n\mathbf{Z}$.

(3.7) Definition. *Let $A_1, A_2, \ldots$ be sets (resp. groups, resp. rings) and $A_1 \xleftarrow{f_1} A_2 \xleftarrow{f_2} A_3 \xleftarrow{f_3} \cdots$ maps (resp. group homomorphisms, resp. ring homomorphisms). The* **projective limit** *of this system consists of compatible sequences of elements:*

$$A = \varprojlim_n A_n = \{a = (a_n)_{n\geq 1} \mid a_n \in A_n,\ f_n(a_{n+1}) = a_n \text{ for all } n \geq 1\}.$$

It is a set (resp. a group, resp. a ring, with respect to operations computed termwise in each A_n). If each map f_n is a continuous map between topological spaces A_{n+1} and A_n, then $A \subset \prod_{n=1}^{\infty} A_n$ has a natural topology, induced by the product topology of the A_n.

(3.8) Example. For any integer $b > 1$, the ring of b**-adic integers** is defined as

$$\mathbf{Z}_b = \varprojlim_n \mathbf{Z}/b^n\mathbf{Z} = \{(a_n)_{n\geq 1} \mid a_n \in \mathbf{Z}/b^n\mathbf{Z},\ a_{n+1} \equiv a_n \pmod{b^n}\}$$

(with each finite ring $\mathbf{Z}/b^n\mathbf{Z}$ being considered with its discrete topology). The elements of $\mathbf{Z}_b$ can be written explicitly in terms of an infinite (to the left!) b-adic expansion

$$x = \cdots x_2x_1x_0 = \sum_{i=0}^{\infty} x_i b^i, \qquad x_i \in \{0, 1, \ldots, b-1\}.$$

For example, $\cdots 1111 = \sum_{i=0}^{\infty} b^i = 1/(1-b) \in \mathbf{Z}_b$. If we allow finitely many b-adic digits with negative powers of b, then we obtain b**-adic numbers** $\mathbf{Q}_b \supset \mathbf{Z}_b$. Elements of $\mathbf{Q}_b$ can be written as

$$x = \cdots x_2x_1x_0, x_{-1}\cdots x_{-m} = \sum_{i=-m}^{\infty} x_i b^i, \qquad x_i \in \{0, 1, \ldots, b-1\}$$

(with m depending on x). The sets $x + b^n\mathbf{Z}_b$ ($x \in \mathbf{Q}_b$, $n \in \mathbf{Z}$) form a basis of the natural topology of $\mathbf{Q}_b$.

If $b = p_1^{r_1} \cdots p_k^{r_k}$, where $r_i \geq 1$ and p_i are distinct prime numbers, then the Chinese remainder theorem

$$\mathbf{Z}/b^n\mathbf{Z} \xrightarrow{\sim} \mathbf{Z}/p_1^{nr_1}\mathbf{Z} \times \cdots \times \mathbf{Z}/p_k^{nr_k}\mathbf{Z}$$

implies that

$$\mathbf{Z}_b \xrightarrow{\sim} \mathbf{Z}_{p_1} \times \cdots \times \mathbf{Z}_{p_k}, \qquad \mathbf{Q}_b \xrightarrow{\sim} \mathbf{Q}_{p_1} \times \cdots \times \mathbf{Q}_{p_k}.$$

In particular, it is enough to consider p-adic numbers (and p-adic integers), for prime numbers p.

(3.9) Exercise. *(i) Why is $\mathbf{Z}_{100} = \mathbf{Z}_{10}$?*
(ii) For any prime number p the ring $\mathbf{Z}_p$ is a domain and $\mathbf{Q}_p$ is a field (equal to $\mathrm{Frac}(\mathbf{Z}_p)$).
(iii) Show that $\mathbf{Z}_b^* = \{a \in \mathbf{Z}_b \mid a \pmod b \in (\mathbf{Z}/b\mathbf{Z})^*\}$.
(iv) All ideals of $\mathbf{Z}_p$ (where p is a prime number) are of the form (p^n) ($n \geq 0$).
(v) There is a natural inclusion $\mathbf{Z} \subset \mathbf{Z}_b$ (resp. $\mathbf{Q} \subset \mathbf{Q}_b$) whose image is dense in the natural topology of $\mathbf{Q}_b$.

(3.10) Example. For any ring A, the projective limit $\varprojlim_n A[x]/(x^n)$ is naturally identified with the ring of formal power series

$$A[[x]] = \{\sum_{i=0}^{\infty} a_i x^i \mid a_i \in A\}.$$

Note that a power series lies in $A[[x]]^* \iff$ its constant term lies in A^*, in analogy with 3.9(iii).

If $A = K$ is a field, then the rings $K[[x]]$ and $\mathbf{Z}_p$ have very similar properties (both are complete discrete valuation rings – see IV.14 below). The only ideals of $K[[x]]$ are (x^n) ($n \geq 0$) and the fraction field of $K[[x]]$ is naturally identified with the field of Laurent series

$$K((x)) = \bigcup_{m\geq 0} \{\sum_{i=-m}^{\infty} a_i x^i \mid a_i \in K\}.$$

(3.11) Exercise (Lifting of idempotents). *Let I be an ideal of a ring A.*
(i) If $e \in A$ satisfies $e^2 \equiv e \pmod{I^n}$ $(n \geq 1)$, then $e' = 3e^2 - 2e^3 \in A$ satisfies $e'^2 \equiv e' \pmod{I^{n+1}}$ and $e' \equiv e \pmod{I^n}$.
(ii) If $A/I = A_1 \times A_2$ is a non-trivial product of two rings, so is $\widehat{A} = \varprojlim_n A/I^n$.
(iii) Where does the expression $3e^2 - 2e^3$ come from? [From an explicit version of Hensel's Lemma 4.12.]

4. Regular functions and ring-valued points in algebraic geometry

In algebraic geometry one studies solutions of systems of polynomial equations. The question posed in the Introduction suggests that it may be of interest to consider solutions with coordinates in arbitrary rings.

(4.1) Consider the circle $C : x^2 + y^2 - 1 = 0$ (over an arbitrary base ring A) as a subspace of the plane with coordinates x and y. The discussion in 2.6 makes it plausible that the quotient

$$O(C) = A[x, y]/(x^2 + y^2 - 1)$$

of the polynomial ring $A[x, y]$ (= the ring of regular functions on the plane over A) by the principal ideal $(x^2 + y^2 - 1)$ (= the ideal of functions which vanish on C) should be the ring of regular functions on C.

This is, indeed, a good object to consider – it determines the set of points of C with values in an arbitrary ring $B \supset A$ (more generally, in an arbitrary A-algebra), by Proposition 4.4 below.

Moreover, the ring $O(C)$ also determines the tangent space T_bC at each point $b = (u, v)$ of C, as explained in 4.6 below. Recall the standard formula for the tangent space in the case $A = \mathbf{R}$: it is the affine line

$$\frac{\partial(x^2 + y^2 - 1)}{\partial x}(u, v)\,(x - u) + \frac{\partial(x^2 + y^2 - 1)}{\partial y}(u, v)\,(y - v) = 2u(x - u) + 2v(y - v) = 0. \tag{4.1.1}$$

The terminology in algebraic geometry is somewhat awkward: one considers as the tangent space T_bC the line passing through the origin

$$\frac{\partial(x^2 + y^2 - 1)}{\partial x}(u, v)\,x + \frac{\partial(x^2 + y^2 - 1)}{\partial y}(u, v)\,y = 2ux + 2vy = 0, \tag{4.1.2}$$

which is the direction of the "physical tangent space" (4.1.1).

(4.2) Definition. *Let A be a ring. An* **A-algebra** *is a pair (B, i) consisting of a ring B and a ring homomorphism $i : A \longrightarrow B$ (which is often omitted from notation). A* **homomorphism of A-algebras** *$(B, i) \longrightarrow (B', i')$ is a ring homomorphism $f : B \longrightarrow B'$ such that $f \circ i = i'$. The set of such homomorphisms will be denoted by $\mathrm{Hom}_{A-\mathrm{Alg}}(B, B')$.*

(4.3) If A is a subring of B, then B is an A-algebra via the inclusion map.

Any ring B has a canonical structure of a $\mathbf{Z}$-algebra: $n \in \mathbf{Z}$ is mapped to $1_B + \cdots + 1_B$ (sum of n terms) if $n > 0$ (resp. to $-(1_B + \cdots + 1_B)$ (sum of $-n$ terms) if $n < 0$). Ring homomorphisms then coincide with homomorphisms of $\mathbf{Z}$-algebras.

(4.4) Proposition-Definition. *Let A be a ring.*
(i) For any A-algebra B the map

$$\mathrm{Hom}_{A-\mathrm{Alg}}(A[x_1, \ldots, x_n], B) \xrightarrow{\sim} B^n, \qquad \alpha \mapsto (\alpha(x_1), \ldots, \alpha(x_n))$$

is bijective. Its inverse is given by sending $b = (b_1, \ldots, b_n) \in B^n$ to the evaluation map $\mathrm{ev}_b : A[x_1, \ldots, x_n] \longrightarrow B$, which assigns to a polynomial $f \in A[x_1, \ldots, x_n]$ its value at b: $\mathrm{ev}_b(f) = f(b) = f(b_1, \ldots, b_n)$.
(ii) Let I be an ideal of $A[x_1, \ldots, x_n]$. We would like to define Z to be the space defined by the system of polynomial equations $f = 0$ for $f \in I$. In concrete terms, we define, for any A-algebra B,

$$Z(B) = \{b = (b_1, \ldots, b_n) \in B^n \mid \forall f \in I \quad f(b) = 0\}$$

to be the set of B-valued points of Z. We also define the ring of regular functions on Z to be $O(Z) = A[x_1, \dots, x_n]/I$.
(iii) For any A-algebra B the map

$$\mathrm{Hom}_{A-\mathrm{Alg}}(O(Z), B) \longrightarrow B^n, \qquad \beta \mapsto (\beta \circ \mathrm{pr}(x_1), \dots, \beta \circ \mathrm{pr}(x_n))$$

gives rise to a bijection $\mathrm{Hom}_{A-\mathrm{Alg}}(O(Z), B) \xrightarrow{\sim} Z(B) \subset B^n$.

Proof. (i) Immediate. (iii) A homomorphism of A-algebras $\alpha : A[x_1, \dots, x_n] \longrightarrow B$ factors as $\beta \circ \mathrm{pr} : A[x_1, \dots, x_n] \longrightarrow A[x_1, \dots, x_n]/I = O(Z) \longrightarrow B$ (for unique β) if and only if $\alpha(I) = 0 \iff \forall f \in I \quad 0 = \alpha(f) = f(\alpha(x_1), \dots, \alpha(x_n)) \iff (\alpha(x_1), \dots, \alpha(x_n)) \in Z(B)$.

(4.5) Why do we need B? Why is it not enough to consider only A-valued points? A simple reason is that $Z(A)$ may well be empty (for example, for the conic $Z : x^2 + y^2 + 1 = 0$ over $A = \mathbf{R}$), while $Z(B)$ can be large for bigger rings $B \supset A$ (such as $\mathbf{C} \supset \mathbf{R}$).

A more interesting example (for $A = \mathbf{C}$) is provided by $Z : y = 0$ (a horizontal line) and $Z' : y^2 = 0$ (the same line with multiplicity two). These are different objects, but their complex points coincide: $Z(\mathbf{C}) = Z'(\mathbf{C}) = \{(x, 0) \mid x \in \mathbf{C}\}$. Their respective rings of regular functions are equal to $O(Z) = \mathbf{C}[x, y]/(y) = \mathbf{C}[x]$ and $O(Z') = \mathbf{C}[x, y]/(y^2) = \mathbf{C}[x] + \mathbf{C}[x]\varepsilon$, where $\varepsilon = y \pmod{(y^2)} \neq 0 \in O(Z')$, but $\varepsilon^2 = 0$. This is a typical example of an infinitesimal element in algebraic geometry.

(4.6) Dual numbers and the tangent space. More generally, let $O(Z) = A[x_1, \dots, x_n]/I$ be as in Proposition 4.4(ii). For any A-algebra B we let the **dual numbers** over B be the B-algebra $B[\varepsilon] = B + B\varepsilon$, where $\varepsilon^2 = 0$. Intuitively, we should think of ε as a tangent vector attached to a point.

The Taylor expansion for polynomials implies that, for any $f \in A[x_1, \dots, x_n]$ and any $b + b'\varepsilon = (b_1 + b'_1\varepsilon, \dots, b_n + b'_n\varepsilon) \in B[\varepsilon]^n$, we have

$$f(b + b'\varepsilon) = f(b) + \varepsilon \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(b)\, b'_j.$$

In particular,

$$Z(B[\varepsilon]) = \{b + b'\varepsilon \in B[\varepsilon]^n \mid b \in Z(B),\, b' \in (T_b Z)(B)\},$$

where $T_b Z$ denotes the **tangent space** (as mentioned above, it is the direction of the physical tangent space, but we are not going to argue with the standard terminology) of Z at b, defined by the system of linear equations

$$T_b Z : \forall f \in I \qquad \frac{\partial f}{\partial x_1}(b)\, x_1 + \dots + \frac{\partial f}{\partial x_n}(b)\, x_n = 0$$

(it is enough to consider these equations only for a set of generators of I).

In particular, in the situation considered in 4.5 we have

$$Z(\mathbf{C}[\varepsilon]) = \{(x, 0) \mid x \in \mathbf{C}[\varepsilon]\} \neq Z'(\mathbf{C}[\varepsilon]) = \{(x, \varepsilon y') \mid x \in \mathbf{C}[\varepsilon],\, y' \in \mathbf{C}\}.$$

Indeed, tangent vectors to the double line Z' have enough space to move around, but tangent vectors to the usual line Z have to be horizontal.

(4.7) Points in $C(\mathbf{Z}/n\mathbf{Z})$, tangent spaces and Hensel's lemma. What can one say about the points of the circle $C : x^2 + y^2 - 1 = 0$ with values in $\mathbf{Z}/n\mathbf{Z}$, i.e., about the set of all solutions of the congruence $x^2 + y^2 - 1 \equiv 0 \pmod{n}$?

If $n = p_1^{r_1} \cdots p_k^{r_k}$, where p_i are distinct prime numbers, then the Chinese remainder theorem implies that $C(\mathbf{Z}/n\mathbf{Z}) = C(\mathbf{Z}/p_1^{r_1}\mathbf{Z}) \times \cdots \times C(\mathbf{Z}/p_k^{r_k}\mathbf{Z})$. We can assume, therefore, that $n = p^r$ $(r \geq 1)$ is a prime power. The next step is to relate $C(\mathbf{Z}/p^r\mathbf{Z})$ and $C(\mathbf{Z}/p^{r+1}\mathbf{Z})$: given a solution $(u, v) \in \mathbf{Z}^2$ of the congruence $x^2 + y^2 - 1 \equiv 0 \pmod{p^r}$, we try to lift $(u \pmod{p^r}, v \pmod{p^r}) \in C(\mathbf{Z}/p^r\mathbf{Z})$ to a solution $\pmod{p^{r+1}}$. In other words, we are looking for $(a, b) \in (\mathbf{Z}/p\mathbf{Z})^2$ such that

$$(u+p^r a)^2+(v+p^r b)^2-1 \stackrel{?}{\equiv} 0 \pmod{p^{r+1}},$$

which is equivalent to

$$(u^2+v^2-1)/p^r + (2ua+2vb) \equiv 0 \pmod{p}. \tag{4.7.1}$$

If $p \neq 2$, then at least one of the coefficients $2u$ or $2v$ is relatively prime to p, which means that (4.7.1) has a solution $(a_0, b_0) \in (\mathbf{Z}/p\mathbf{Z})^2$.

Note that the non-zero linear equation

$$2ux + 2vy \equiv 0 \pmod{p} \tag{4.7.2}$$

is precisely the equation for the $\mathbf{Z}/p\mathbf{Z}$-valued points of the (one-dimensional) tangent space $T_{(\overline{u},\overline{v})}C$ of C at the $\mathbf{Z}/p\mathbf{Z}$-valued point $(\overline{u},\overline{v}) = (u \pmod{p}, v \pmod{p}) \in C(\mathbf{Z}/p\mathbf{Z})$. It follows that the set of all solutions of (4.7.1) is of the form

$$(a_0, b_0) + (T_{(\overline{u},\overline{v})}C)(\mathbf{Z}/p\mathbf{Z}),$$

i.e., it is an affine line over $\mathbf{Z}/p\mathbf{Z}$ with direction $T_{(\overline{u},\overline{v})}C$. In particular, $|C(\mathbf{Z}/p^{r+1}\mathbf{Z})| = p\,|C(\mathbf{Z}/p^r\mathbf{Z})|$.

This argument works for $p \neq 2$ because the tangent space (4.7.2) has dimension $\dim T_{(\overline{u},\overline{v})}C = 1 = \dim(C)$ (geometrically, $(\overline{u},\overline{v})$ is a **smooth point** of C over $\mathbf{Z}/p\mathbf{Z}$), but fails for $p = 2$ (C has no smooth points over fields containing $\mathbf{Z}/2\mathbf{Z}$, since $x^2+y^2-1 = (x+y+1)^2$ over such fields – the circle becomes a doble line).

The same argument proves the following statement (a version of Hensel's Lemma).

(4.8) Proposition. *Let $f \in \mathbf{Z}[x_1, \ldots, x_n]$, let $Z : f(x_1, \ldots, x_n) = 0$. Assume that p is a prime, $r \geq 1$ and that $a = (a_1, \ldots, a_n) \in \mathbf{Z}^n$ is a solution of the congruence $f(x_1, \ldots, x_n) \equiv 0 \pmod{p^r}$ (in other words, $a \pmod{p^r}$ is an element of $Z(\mathbf{Z}/p^r\mathbf{Z})$). If $a \pmod{p} \in Z(\mathbf{Z}/p\mathbf{Z})$ is a smooth point of Z over $\mathbf{Z}/p\mathbf{Z}$ in the sense that $(\partial f/\partial x_j)(a) \not\equiv 0 \pmod{p}$ for some $j = 1, \ldots, n$, then the fibre of the map $Z(\mathbf{Z}/p^{r+1}\mathbf{Z}) \longrightarrow Z(\mathbf{Z}/p^r\mathbf{Z})$ over $a \pmod{p^r}$ is an affine space under the $\mathbf{Z}/p\mathbf{Z}$-valued points of the tangent space*

$$T_{a \pmod{p}}Z : (\partial f/\partial x_1)(a)\, x_1 + \cdots + (\partial f/\partial x_n)(a)\, x_n \equiv 0 \pmod{p}.$$

In particular, if all points in $Z(\mathbf{Z}/p\mathbf{Z})$ are smooth over $\mathbf{Z}/p\mathbf{Z}$, then the map $Z(\mathbf{Z}/p^{r+1}\mathbf{Z}) \longrightarrow Z(\mathbf{Z}/p^r\mathbf{Z})$ is surjective (hence so is the map $Z(\mathbf{Z}_p) \longrightarrow Z(\mathbf{Z}/p\mathbf{Z})$) and its fibres have cardinality p^{n-1}.

(4.9) Exercise. *(i) For each $r \geq 1$ the congruence $x^5 \equiv 10 \pmod{11^r}$ (resp. $x^5 \equiv 10 \pmod{13^r}$) has five solutions in $\mathbf{Z}/11^r\mathbf{Z}$ (resp. one solution in $\mathbf{Z}/13^r\mathbf{Z}$). The equation $x^5 = 10$ has five 11-adic solutions $x \in \mathbf{Z}_{11}$ (resp. one 13-adic solution $x \in \mathbf{Z}_{13}$).*
(ii) What about the equation $x^n = a$ in $\mathbf{Z}_p$, where p is a prime number not dividing integers a and $n \geq 1$?

(4.10) Consider, on the other hand, the equation

$$C' : y^2 - (x^3 + p) = 0,$$

where p is a prime number. The point $(0,0) \in C'(\mathbf{Z}/p\mathbf{Z})$ is not smooth over $\mathbf{Z}/p\mathbf{Z}$, since C' over $\mathbf{Z}/p\mathbf{Z}$ simplifies as $y^2 - x^3 = 0$, which has a singularity at the origin, as remarked already in the Introduction. We see immediately that $(0,0) \in C'(\mathbf{Z}/p\mathbf{Z})$ does not lie in the image of $C'(\mathbf{Z}/p^2\mathbf{Z})$ (in other words, the solution $(0,0)$ of $y^2 - (x^3+p) \equiv 0 \pmod{p}$ does not lift to any solution of $y^2 - (x^3+p) \equiv 0 \pmod{p^2}$).

(4.11) Theorem on implicit functions (example). Consider again the circle $C : f(x,y) = x^2+y^2-1 = 0$, this time over a ring A such that $2 \in A^*$.

The tangent space T_bC to C at the point $b = (0,1)$ is horizontal, since $(\partial f/\partial x)(b) = 0 \neq (\partial f/\partial y)(b)$. In the classical case $A = \mathbf{R}$ the fact that $(\partial f/\partial y)(b) \neq 0$ implies that, in a suitable neighbourhood U of b, the projection of C onto the horizontal line is a diffeomorphism between U and its image.

Let us analyse this situation from a purely algebraic perspective. Write a point "**close to** b" as $(u, 1-v)$, with u, v being "close to 0". The equation of C then becomes $u^2 + v^2 - 2v = 0$, which yields, recursively, a power series expansion for v in terms of u^2:

$$v = \frac{u^2}{2} + \frac{v^2}{2} = \frac{u^2}{2} + \frac{(u^2+v^2)^2}{8} = \frac{u^2}{2} + \frac{u^4}{8} + \frac{u^2(u^2+v^2)^2}{16} + \frac{(u^2+v^2)^4}{2^7} = \frac{u^2}{2} + \frac{u^4}{8} + \frac{u^6}{16} + \cdots \in A[[u^2]].$$

Of course, this is just the standard power series expansion

$$v = 1 - \sqrt{1-u^2} = \sum_{n=1}^{\infty} (-1)^{n-1} \binom{1/2}{n} u^{2n}, \tag{4.11.1}$$

but it is obtained by an iterative procedure of lifting a solution $v \in A[u]/(u^{2r})$ of the congruence

$$v^2 - 2v + u^2 \equiv 0 \pmod{u^{2r}}$$

to a unique solution modulo $u^{2(r+1)}$, starting with the solution $v = u^2/2$ for $r = 2$. In other words, this is a variant of 4.7–4.8, with $\mathbf{Z}_p$ being replaced by $A[[u^2]]$.

(4.12) Exercise (Hensel's Lemma, one variable version). *Let I be an ideal of a ring A, let $f \in A[X]$ and $a \in A$ be such that $f(a) \equiv 0 \pmod{I^n}$ $(n \geq 1)$ and $f'(a) \pmod{I}$ is invertible in A/I.*
(i) There exists $b \in A$ (unique $\pmod{I^{n+1}}$) such that $b \equiv a \pmod{I^n}$ and $f(a) \equiv 0 \pmod{I^{n+1}}$.
(ii) There exists unique $\widehat{a} \in \widehat{A} = \varprojlim_n A/I^n$ such that $f(\widehat{a}) = 0$ and the image of $\widehat{a}$ in A/I^n is equal to a.
(iii) What is the relation to Newton's iterative method $x_{n+1} = x_n - f(x_n)/f'(x_n)$ for finding approximations to the roots of the equation $f(x) = 0$?

(4.13) Exercise. *(i) If $p \neq 2$ is a prime and $u \in \mathbf{Z}_p$ satisfies $u \equiv 0 \pmod{p}$, then the power series (4.11.1) evaluated at u converges to an element $v \in \mathbf{Z}_p$ such that $(u, 1-v) \pmod{p} = (0,1) \in C(\mathbf{Z}/p\mathbf{Z})$. Relate this to the story explained in 4.7–4.8.*
(ii) What happens if $p = 2$?
(iii) Is it obvious that $\binom{1/2}{n} \in \mathbf{Z}[1/2]$ for all $n \geq 0$? Show that $\binom{a/b}{n} \in \mathbf{Z}[1/b]$ for all $a, b \in \mathbf{Z}$, $b \neq 0$ and $n \geq 0$.
[Hint: $\binom{\mathbf{Z}}{n} \subset \mathbf{Z}$ implies that $\forall x \in \mathbf{Z}_p \quad \binom{x}{n} \subset \mathbf{Z}_p$, for all primes p, by continuity and Exercise 3.9(v).]

(4.14) Exercise. *Let $n \geq 1$ be an integer.*
(i) There exists $u_n \in \mathbf{R}[X]$ such that $u_n \equiv X \pmod{(X^2+1)}$ and $u_n^2 + 1 \equiv 0 \pmod{(X^2+1)^n}$.
(ii) The class $c_n = u_n \pmod{(X^2+1)^n} \in \mathbf{R}[X]/(X^2+1)^n$ is unique. Give an explicit formula for c_{n+1} in terms of c_n.
(iii) The formulas

$$\alpha_n : \mathbf{C}[Y] \longrightarrow \mathbf{R}[X]/(X^2+1)^n, \qquad a + bi \mapsto a + bu_n, \qquad Y \mapsto X - c_n$$

define a surjective morphism of $\mathbf{R}$-algebras.
(iv) α_n induces an isomorphism of $\mathbf{R}$-algebras $\overline{\alpha}_n : \mathbf{C}[Y]/(Y^n) \xrightarrow{\sim} \mathbf{R}[X]/(X^2+1)^n$.
(v) For every non-constant polynomial $f \in \mathbf{R}[X]$ there exists an isomorphism of $\mathbf{R}$-algebras

$$\mathbf{R}[X]/(f) \xrightarrow{\sim} \prod_{j=1}^{M} \mathbf{R}[X]/(X^{a_j}) \times \prod_{k=1}^{N} \mathbf{C}[Y]/(Y^{b_k}) \qquad (a_j,\, b_k \geq 1).$$

(vi) What happens if we replace $\mathbf{R}$ in (v) by an arbitrary perfect field (see III.6.2 below)?
(vii)** What happens if we replace $\mathbf{R}$ in (v) by a non-perfect field?

5. Divisibility and unique factorisation

Every non-zero integer n can be written in a unique way as a product $n = \pm p_1^{r_1} \cdots p_k^{r_k}$, where p_i are distinct prime numbers. This unique factorisation property holds in other rings of interest.

(5.1) We are going to study three classes of integral domains:

- **Euclidean rings**, which admit a generalised euclidean algorithm ($\mathbf{Z}$, $\mathbf{Z}[i]$ or $K[X]$ for any field K);
- **Unique factorisation domains (UFD)**, in which factorisation into prime elements exists and is unique;
- **Principal ideal domains (PID)**, in which every ideal is principal,

which have the following properties:

$$A \text{ is euclidean} \Longrightarrow A \text{ is a PID} \Longrightarrow A \text{ is a UFD} \Longrightarrow A[X] \text{ is a UFD.}$$

Note that this implies that $\mathbf{C}[X, Y]$ is a UFD, but not a PID (the ideal (X, Y) is not principal). In fact, the ring $\mathbf{C}[x_1, \ldots, x_n]$ has dimension n, but a PID is at most one-dimensional.

(5.2) Definition. *An integral domain A is a euclidean ring (with respect to a function $\varphi : A \longrightarrow \mathbf{N} = \{0, 1, 2, \ldots\}$) if the following two properties hold: (i) $\varphi(a) = 0 \iff a = 0$.*
(ii) For every $a, b \in A$ with $b \neq 0$ there exist $q, r \in A$ such that $a = qb + r$ and $\varphi(r) < \varphi(b)$.

(5.3) Examples. (i) $A = \mathbf{Z}$, $\varphi(n) = |n|$.
(ii) $A = K[X]$ (where K is a field), $\varphi(f) = \deg(f) + 1$ if $f \neq 0$ (alternatively, $\varphi(f) = 2^{\deg(f)}$ also works).
(iii) $A = \mathbf{Z}[i] = \{x + iy \mid x, y \in \mathbf{Z}\}$, $\varphi(x + iy) = (x + iy)(x - iy) = |x + iy|^2 = x^2 + y^2$. Indeed, if $a, b \in \mathbf{Z}[i]$ and $b \neq 0$, we can write $a/b = u + iv$ and take $q = x + iy \in \mathbf{Z}[i]$, where x (resp. y) is the closest integer to $u \in \mathbf{Q}$ (resp. to $v \in \mathbf{Q}$). In this case $|u - x|, |v - y| \leq 1/2$, which implies that $|a/b - q|^2 = |u - x|^2 + |v - y|^2 \leq (1/2)^2 + (1/2)^2 = 1/2 < 1$, hence $a = bq + r$ with $r \in \mathbf{Z}[i]$ and $\varphi(r) = |r|^2 = |a - qb|^2 < |b|^2 = \varphi(b)$, as required.
(iv) $A = \mathbf{Z}[i\sqrt{2}] = \{x + iy\sqrt{2} \mid x, y \in \mathbf{Z}\}$, $\varphi(x + iy\sqrt{2}) = |x + iy\sqrt{2}|^2 = x^2 + 2y^2$. The argument from (iii) applies, since $(1/2)^2 + 2(1/2)^2 = 3/4 < 1$.
(v) However, this argument breaks down for $A = \mathbf{Z}[i\sqrt{d}]$ and $\varphi(a) = |a|^2$ for integers $d \geq 3$, since $(1/2)^2 + d(1/2)^2 \geq 1$. This is no accident – none of these rings is a UFD.
(vi) Exercise: modify (iii) to show that the rings $\mathbf{Z}[(1 + i\sqrt{d})/2] = \{x + y(1 + i\sqrt{d})/2 \mid x, y \in \mathbf{Z}\}$ with $\varphi(a) = |a|^2$ are euclidean for $d = 3, 7, 11$.

(5.4) Proposition. *A euclidean ring A is a PID.*

Proof. Let $I \subset A$ be an ideal. If $I = 0$, then $I = (0)$. If $I \neq (0)$, then there exists $b \in I \setminus \{0\}$ with minimal value of $\varphi(b) \geq 1$. We have $(b) = bA \subset A$. Conversely, if $a \in I$, then there exist $q, r \in A$ such that $a = qb + r$ (thus $r = a - qb \in I$) and $\varphi(r) < \varphi(b)$. Minimality of $\varphi(b)$ implies that $r \notin I \setminus \{0\}$, hence $r = 0$ and $a = qr \in (b)$. It follows that $I \subset (b)$, and so $I = (b)$.

(5.5) Definition. *Let A be an integral domain.*
(i) For $a, b \in A \setminus \{0\}$ we say that b **divides** *a (notation: $b \mid a$) if there exists $c \in A$ (in fact, $c \neq 0$) such that $a = bc$ (equivalently, $a \in (b) \iff (a) \subset (b)$). [Note that ($b \mid a$ and $a \mid b$) $\iff$ $a = bu$ for some $u \in A^*$.]*
(ii) An element $a \in A$ is **irreducible** *if $a \notin A^* \cup \{0\}$ and whenever $a = bc$ with $b, c \in A$, then $b \in A^*$ or $c \in A^*$ (but not both). Equivalently, $a \notin A^* \cup \{0\}$ is* **not irreducible** *if there exist $b, c \in A \setminus A^*$ with $a = bc$. [In $A = \mathbf{Z}$, irreducible elements are of the form $n = \pm p$, where p is a prime number.]*

(5.6) Definition. *A* **unique factorisation domain (UFD)** *is an integral domain A such that*
(i) Each element $a \in A \setminus \{0\}$ is of the form $a = ux_1 \ldots x_r$, where $u \in A^*$, each x_i is irreducible and $r \geq 0$;
(ii) If $ux_1 \ldots x_r = vy_1 \ldots y_s$, where $u, v \in A^*$, each x_i and y_j is irreducible and $r, s \geq 0$, then $r = s$ and there are invertible elements $u_1, \ldots, u_r \in A^*$ and a permutation $\sigma \in S_r$ such that $x_i = u_i y_{\sigma(i)}$ for all $i = 1, \ldots, r$.

(5.7) Proposition. *An integral domain A is a UFD $\iff$ it satisfies 5.6(i) and the condition*
(ii') If an irreducible element $x \in A$ divides ab ($a, b \in A$), then x divides a or b ("Euclid's lemma" in the case $A = \mathbf{Z}$).

Proof. "$\Longrightarrow$" This implication is immediate: write $ab = xc$, factor each element a, b, c as in 5.6(i) and use 5.6(ii) to conclude that x coincides (up to an invertible element) with an irreducible factor of a or b.
"$\Longleftarrow$" Assume that, in the situation of 5.6(ii), $r \leq s$. If $r = 0$, then $s = 0$. If $r \geq 1$, then $x_1 \mid vy_1 \ldots y_s$, hence x_1 divides some y_j, by (ii'). After renumbering, we can assume that $j = 1$. As y_1 is irreducible, we have $y_1 = u_1 x_1$ with $u_1 \in A^*$. We can divide both sides by x_1 and repeat the same procedure several times, obtaining

(after renumbering the y_j's) equalities $y_j = u_j x_j$ ($j = 1, \ldots, r$, $u_j \in A^*$) and $u = vu_1 \ldots u_r y_{r+1} \ldots y_s$, which implies that $r = s$.

(5.8) Example. Let $A = \mathbf{Z}[i\sqrt{5}] = \{x + iy\sqrt{5} \mid x, y \in \mathbf{Z}\}$. The **norm map** $N : A \longrightarrow \mathbf{N}$ given by $N(x + iy\sqrt{5}) = |x + iy\sqrt{5}|^2 = x^2 + 5y^2$ is multiplicative in the sense that $N(ab) = N(a)N(b)$ and $N(1) = 1$, which implies that $A^* \subset \{a \in A \mid N(a) = 1\}$. However, $x^2 + 5y^2 = 1$ with $x, y \in \mathbf{Z}$ is possible only if $x = \pm 1$ and $y = 0$, hence $A^* \subset \{\pm 1\}$, and so $A^* = \{\pm 1\}$.

There are no elements $a \in A$ with norm $N(a) = 3$ or $N(a) = 7$. It follows that the elements $3, 7, 4 \pm i\sqrt{5} \in A$ are all irreducible. The equality

$$21 = 3 \cdot 7 = (4 + i\sqrt{5})(4 - i\sqrt{5})$$

then shows that A is not a UFD.

(5.9) Theorem. *A PID is a UFD.*

Proof. We must check that the conditions 5.6(i) and 5.7(ii') hold in any PID A.
If $a \in A \setminus \{0\}$ cannot be written as in 5.6(i), then $a \notin A^* \cup \{0\}$ is not irreducible, hence $a = a_1 b_1$ with $a_1, b_1 \notin A^*$ and $(a) \subsetneq (a_1)$. Moreover, at least one of a_1 and b_1 (say, a_1) cannot be written as in 5.6(i), either. We can continue this procedure and obtain factorisations $a_n = a_{n+1} b_{n+1}$ with the same properties for all $n \geq 1$. As a result, we obtain an infinite chain of ideals (corresponding to divisibilities $\cdots \mid a_3 \mid a_2 \mid a_1 \mid a$)

$$(a) \subsetneq (a_1) \subsetneq (a_2) \subsetneq (a_3) \subsetneq \cdots$$

The union of this chain of ideals is again an ideal, necessarily of the form (c) for some $c \in A$. The element c is contained in (a_n) for some $n \geq 1$, which implies that $(a_n) = (a_{n+1})$, which is not true, by construction. This contradiction shows that 5.6(i) holds.
It remains to check 5.7(ii'). Assume that x is an irreducible element of A dividing ab. The ideal (x, b) is of the form (c), where $c \in A \setminus \{0\}$ (morally, c is the greatest common divisor of x and b). By definition, $c \mid x$, hence $x = cd$ for some $d \in A$. Irreducibility of x implies that either $d \in A^*$ ($\Longrightarrow x \mid c \mid b$), or $c \in A^*$ ($\Longrightarrow (x, b) = (1)$, hence $1 = xu + bv$ for some $u, v \in A$, which implies that $a = axu + abv$ is divisible by x).

(5.10) Greatest common divisor. Let A be a UFD. Fix a set of representatives P of prime elements of A modulo A^*. Any $a \in A \setminus \{0\}$ can then be written uniquely as

$$a = u \prod_{x \in P} x^{v_x(a)}, \qquad u \in A^*, \qquad v_x(a) \in \mathbf{N}$$

(where all but finitely many exponents $v_x(a)$ are equal to 0). Uniqueness of this factorisation implies that

$$b \mid a \iff \forall x \in P \quad v_x(b) \leq v_x(a).$$

It follows that, for any $a, b \in A \setminus \{0\}$, the element

$$c = \prod_{x \in P} x^{\min(v_x(a), v_x(b))} \in A \setminus \{0\}$$

divides both a and b (c is a **common divisor of a and b**). In addition, any common divisor of a and b divides c. Note that c depends on the chosen set of representatives P; however, the class of c in $(A \setminus \{0\})/A^*$ does not – we denote this class by $\gcd(a, b)$ (the **greatest common divisor of a and b**).

Alternatively, one can denote by $\gcd(a, b)$ the principal ideal generated by (c) (this is again independent on the choice of P). If A is a PID (but not for a general UFD) the ideal (a, b) coincides with (c).

(5.11) Exercise. *(i) Let A be a UFD. Assume that $a, b, c \in A \setminus \{0\}$ satisfy $\gcd(a, b) = 1$ and $ab = c^n$ ($n \geq 1$). Show that $a = uc_1^n$ and $b = vc_2^n$ for some $u, v \in A^*$ and $c_1, c_2 \in A$.*
(ii) If A is a PID and $a = ux_1^{r_1} \cdots x_k^{r_k}$, where $u \in A^*$ and $x_1, \ldots, x_k$ are distinct irreducible elements of A, then Corollary 3.5 applies, yielding a canonical ring isomorphism

$$A/(a) \xrightarrow{\sim} A/(x_1^{r_1}) \times \cdots \times A/(x_k^{r_k}).$$

(iii) If K is a field and $a_1, \dots, a_n$ are distinct elements of K, then the evaluation maps $\mathrm{ev}_{a_i} : K[X] \longrightarrow K$ $(g \mapsto g(a_i))$ induce a ring isomorphism

$$K[X]/((X-a_1)\cdots(X-a_n)) \xrightarrow{\sim} K \times \cdots \times K, \qquad g \pmod{(X-a_1)\cdots(X-a_n)} \mapsto (g(a_1), \dots, g(a_n)).$$

(iv) Under the isomorphism (iii), the idempotent e_i corresponds to $f(X)/((X-a_i)f'(a_i)) \pmod f$, where $f = (X-a_1)\cdots(X-a_n)$.

(5.12) A diophantine equation. Let us determine all solutions $x, y \in \mathbf{Z}$ of the equation

$$y^2 + 2 = x^3.$$

Firstly, if x is divisible by 2, so is y, but then $2 = x^3 - y^2$ is divisible by 4, which is impossible. Thus $2 \nmid xy$. Secondly, we factorise the original equation

$$(y + i\sqrt{2})(y - i\sqrt{2}) = x^3$$

in the ring $\mathbf{Z}[i\sqrt{2}]$, which is a UFD by Example 5.3(iv) + Proposition 5.4 + Theorem 5.9. We claim that $\gcd(y+i\sqrt{2}, y-i\sqrt{2}) = 1$. If not, then there exists an irreducible element $\pi \in \mathbf{Z}[i\sqrt{2}]$ dividing both $y \pm i\sqrt{2}$, hence also dividing their sum $2y$ and difference $2i\sqrt{2}$. It follows that its norm $N(\pi) = \pi\overline{\pi} \in \mathbf{N}$ divides (in $\mathbf{Z}$) both $(2y)^2$ and $(2i\sqrt{2})(-2i\sqrt{2}) = 8$, hence their gcd $= 4$ (since y is odd). However, $4 = (i\sqrt{2})^4$ and $i\sqrt{2}$ is irreducible in $\mathbf{Z}[i\sqrt{2}]$ (consider its norm); thus $\pi = ui\sqrt{2}$ for some $u \in \mathbf{Z}[i\sqrt{2}]^* = \{\pm 1\}$ (the argument from Example 5.8 applies). However, $i\sqrt{2}$ does not divide $y + i\sqrt{2}$, since y is odd. This contradiction proves that $\gcd(y + i\sqrt{2}, y - i\sqrt{2}) = 1$, as claimed.

Exercise 5.11(i) then shows that $y + i\sqrt{2} = ud^3$ for some $u \in \{\pm 1\}$ (hence $u = u^3$) and $d \in \mathbf{Z}[i\sqrt{2}]$; thus

$$y + i\sqrt{2} = (a + bi\sqrt{2})^3$$

for some $a, b \in \mathbf{Z}$, which is equivalent to

$$y = a(a^2 - 6b^2), \qquad 1 = b(3a^2 - 2b^2).$$

The only possibilities are $b = \pm 1$, hence $3a^2 = 2b^2 \pm 1 = 2 \pm 1$, and so $b = 1$, $a = \pm 1$, $y = \mp 5$, $x = 3$.

Conclusion: the only solutions $x, y \in \mathbf{Z}$ are $x = 3$, $y = \pm 5$.

(5.13) Exercise. *Find all solutions $x, y \in \mathbf{Z}$ of the equation $y^2 + 11 = x^3$ (resp. of $y^2 + 28 = x^3$).*

(5.14) Exercise. *What happens if we try to solve $y^2 + y + 1 = x^3$ $(x, y \in \mathbf{Z})$ by the same method?*

(5.15) Exercise. *Find all solutions $x, y \in \mathbf{C}[T]$ of the equation $y^2 = x^3 + T^2$ (resp. of $y^2 = x^3 + T$).*

6. Prime ideals and maximal ideals

Property 5.7(ii') can be reformulated by saying that, for any irreducible element x of a UFD A, the quotient ring $A/(x)$ is a domain.

(6.1) Definition. *Let I be an ideal of a ring A. We say that I is a* **prime ideal** *(resp. a* **maximal ideal***) if A/I is a domain (resp. a field). The set of all prime ideals (resp. maximal ideals) of A will be denoted by* $\mathrm{Spec}(A)$ *(resp. by* $\mathrm{Max}(A) = \mathrm{Specm}(A) = \mathrm{Specmax}(A)$*). Of course,* $\mathrm{Max}(A) \subset \mathrm{Spec}(A)$.

(6.2) Examples. (i) $(0) \in \mathrm{Spec}(A) \iff$ the ring A is a domain.
(ii) If $A = K$ is a field, then $\mathrm{Spec}(K) = \mathrm{Max}(K) = \{(0)\}$.
(iii) $\mathrm{Max}(\mathbf{Z}) = \{(p) \mid p = \text{prime number}\}$, $\mathrm{Spec}(\mathbf{Z}) = \{(0)\} \cup \mathrm{Max}(\mathbf{Z})$.
(iv) More generally, if A is a PID, then $\mathrm{Max}(A) = \{(x) \mid x = \text{irreducible element}\}$, $\mathrm{Spec}(A) = \{(0)\} \cup \mathrm{Max}(A)$.
(v) In particular, $\mathrm{Max}(\mathbf{C}[X]) = \{(X-a) \mid a \in \mathbf{C}\}$ and $\mathrm{Spec}(\mathbf{C}[X]) = \{(0)\} \cup \mathrm{Max}(\mathbf{C}[X])$. If we interpret $\mathbf{C}[X]$ as the ring of regular functions on a complex line L, then $(X-a)$ (resp. (0)) is the ideal of functions vanishing at the point $a \in L(\mathbf{C})$ (resp. of functions vanishing at all elements of $L(\mathbf{C})$).

(vi) More generally, consider $\mathbf{C}[X,Y]$ as the ring of regular functions on a complex plane. We shall see in IV.6 below that $\mathrm{Max}(\mathbf{C}[X,Y]) = \{(X-a, Y-b) \mid a,b \in \mathbf{C}\}$ and $\mathrm{Spec}(\mathbf{C}[X,Y]) = \{(0)\} \cup \{(f) \mid f \in \mathbf{C}[X,Y] \setminus \mathbf{C},\ f \text{ irreducible}\} \cup \mathrm{Max}(\mathbf{C}[X,Y])$. Geometrically, for each prime ideal $P \in \mathrm{Spec}(\mathbf{C}[X,Y])$, the set of complex points of the plane at which all elements of P vanish is an **irreducible algebraic subvariety of $\mathbf{C}^2$**: the full plane if $P=(0)$, the irreducible curve $\{(x,y) \in \mathbf{C}^2 \mid f(x,y)=0\}$ if $P=(f)$, and the point (a,b) if $P=(X-a,Y-b)$.
(vii) Exercise: describe Spec and Max of $\mathbf{R}[X]$ (and guess the answer for $\mathbf{R}[X,Y]$).

(6.3) Proposition. *An ideal I of A is maximal $\iff$ $I \neq A$ and I is maximal among ideals with this property (i.e., the only ideal J satisfying $I \subset J \neq A$ is $J=I$).*

Proof. Let $\mathrm{pr} : A \longrightarrow A/I$ be the projection. $I \neq A$ has the maximality property as in the proposition iff

$$\forall x \notin I \quad I+(x) = A \iff \forall x \notin I \quad (\mathrm{pr}(x)) = A/I \iff \forall x \notin I \quad \mathrm{pr}(x) \in (A/I)^* \iff (A/I)^* = A/I \setminus \{0\}.$$

(6.4) Functoriality. (i) Any ring homomorphism $f : A \longrightarrow B$ induces a map $f^* : \mathrm{Spec}(B) \longrightarrow \mathrm{Spec}(A)$ given by $f^*(Q) = f^{-1}(Q)$ (note that $A/f^{-1}(Q)$ is a subring of B/Q via $\overline{f}$, hence is also a domain).
(ii) If f is surjective, then f^* injective, $A/f^{-1}(Q) = B/Q$ for any $Q \in \mathrm{Spec}(B)$, $\mathrm{Im}(f^*) = \{P \in \mathrm{Spec}(A) \mid P \supset \mathrm{Ker}(f)\}$ and $f^*(\mathrm{Max}(B)) = \{\mathfrak{m} \in \mathrm{Max}(A) \mid \mathfrak{m} \supset \mathrm{Ker}(f)\}$ (we can assume that $f = \mathrm{pr} : A \longrightarrow A/I$, in which case the last sentence in 2.8 applies).
(iii) In general, $f^*(\mathrm{Max}(B)) \not\subset \mathrm{Max}(A)$ (consider the inclusion $f : \mathbf{Z} \hookrightarrow \mathbf{Q}$: $(0) \in \mathrm{Max}(\mathbf{Q})$, but $f^{-1}((0)) = (0) \notin \mathrm{Max}(\mathbf{Z})$).

(6.5) Reduced and non-reduced rings. Let A be a ring. The set of all nilpotent elements of A

$$\sqrt{(0)} = \{x \in A \mid \exists n \geq 1 \qquad x^n = 0\}$$

is an ideal, called the **nilradical of A**. The ring A is **reduced** if $\sqrt{(0)} = (0)$. In general, $A^{\mathrm{red}} = A/\sqrt{(0)}$ is the biggest reduced quotient ring of A. For any ideal I of A we have $(A/I)^{\mathrm{red}} = A/\sqrt{I}$. For example, in the geometric situation considered in 4.5, $O(Z') = \mathbf{C}[x,y]/(y^2)$ and $O(Z')^{\mathrm{red}} = \mathbf{C}[x,y]/(y) = O(Z)$.

The nilradical $\sqrt{(0)}$ is contained in any prime ideal of A, which implies, by 6.4(ii), that the canonical projection $\mathrm{pr} : A \longrightarrow A^{\mathrm{red}}$ induces bijections

$$\mathrm{pr}^* : \mathrm{Spec}(A^{\mathrm{red}}) \xrightarrow{\sim} \mathrm{Spec}(A), \qquad \mathrm{Max}(A^{\mathrm{red}}) \xrightarrow{\sim} \mathrm{Max}(A).$$

(6.6) Theorem. *If the ring A is non-zero, then* $\mathrm{Max}(A) \neq \emptyset$.

Proof. The set S of all ideals $I \neq A$ is non-empty ($(0) \in S$), partially ordered by inclusion and inductive (every totally ordered subset $\{I_\alpha\} \subset S$ admits an upper bound in S (namely, the union of all I_α)). By Zorn's Lemma – which is equivalent to the axiom of choice, and therefore can be considered mostly harmless – the set S contains a maximal element.

(6.7) Corollary. *Let A be a ring. (i) If $I \neq A$ is an ideal, then there exists a maximal ideal $\mathfrak{m} \supset I$.*
(ii) $A \setminus A^* = \bigcup_{\mathfrak{m} \in \mathrm{Max}(A)} \mathfrak{m}$.

Proof. (i) Take $\mathfrak{m} = \mathrm{pr}^{-1}(\overline{\mathfrak{m}})$, where $\mathrm{pr} : A \longrightarrow A/I \neq 0$ is the projection and $\overline{\mathfrak{m}} \in \mathrm{Max}(A/I)$.
(ii) If $x \in A^*$, then we have, for each $\mathfrak{m} \in \mathrm{Max}(A)$, $x \pmod{\mathfrak{m}} \in (A/\mathfrak{m})^* = (A/\mathfrak{m}) \setminus \{0\}$, hence $x \notin \mathfrak{m}$. Conversely, if $x \notin A^*$, then $(x) \neq A$, hence there exists $\mathfrak{m} \in \mathrm{Max}(A)$ containing (x).

(6.8) Proposition. *Let A be an integral domain, let $x \in A \setminus \{0\}$.*
(i) If $(x) \in \mathrm{Spec}(A)$, then x is an irreducible element of A.
(ii) If A is a UFD and x is an irreducible element of A, then $(x) \in \mathrm{Spec}(A)$.

Proof. (i) Exercise. (ii) This is 5.7(ii').

(6.9) Example. As noted in 5.8, the ring $\mathbf{Z}[i\sqrt{5}]$ is not a UFD, since 21 can be factored

$$3 \cdot 7 = 21 = 4^2 + 5 \cdot 1^2 = (4 + i\sqrt{5})(4 - i\sqrt{5})$$

in two distinct ways into products of irreducible elements. However, these factorisations admit a common "refinement", which was already alluded to in 2.7.

The morphism "evaluation at $i\sqrt{5}$" induces a ring isomorphism

$$\mathbf{Z}[X]/(X^2+5) \xrightarrow{\sim} A = \mathbf{Z}[i\sqrt{5}], \qquad g(X) \mapsto g(i\sqrt{5}),$$

which implies that the ring $A/3A$ is isomorphic to

$$\mathbf{Z}[X]/(3, X^2+5) = \mathbf{F}_3[X]/(X^2+5) = \mathbf{F}_3[X]/(X^2-1) = \mathbf{F}_3[X]/((X-1)(X+1)) \xrightarrow{\sim} \mathbf{F}_3 \times \mathbf{F}_3, \quad (6.9.1)$$

where the last isomorphism is induced by the evaluation maps $X \mapsto 1$ and $X \mapsto -1$, as in 5.11(iii). In particular, (3) is not a prime ideal of A.

Consider the projections pr_1 and pr_2 on the two factors in (6.9.1). The surjective ring homomorphisms

$$\alpha_1 : A \xrightarrow{\sim} \mathbf{Z}[X]/(X^2+5) \longrightarrow \mathbf{Z}[X]/(3, X^2+5) \xrightarrow{\sim} \mathbf{F}_3 \times \mathbf{F}_3 \xrightarrow{\mathrm{pr}_1} \mathbf{F}_3$$
$$\alpha_2 : A \xrightarrow{\sim} \mathbf{Z}[X]/(X^2+5) \longrightarrow \mathbf{Z}[X]/(3, X^2+5) \xrightarrow{\sim} \mathbf{F}_3 \times \mathbf{F}_3 \xrightarrow{\mathrm{pr}_2} \mathbf{F}_3$$

are given by the formulas

$$a + bi\sqrt{5} \mapsto a + bX \mapsto a + b \pmod 3$$
$$a + bi\sqrt{5} \mapsto a + bX \mapsto a - b \pmod 3,$$

which implies that

$$P_1 := \mathrm{Ker}(\alpha_1) = \{a + bi\sqrt{5} \,|\, a, b \in \mathbf{Z},\ a + b \equiv 0 \pmod 3\} = (3,\ i\sqrt{5} - 1)$$
$$P_2 := \mathrm{Ker}(\alpha_2) = \{a + bi\sqrt{5} \,|\, a, b \in \mathbf{Z},\ a - b \equiv 0 \pmod 3\} = (3,\ i\sqrt{5} + 1).$$

The isomorphisms $A/P_i \xrightarrow{\sim} \mathbf{F}_3$ induced by α_1 and α_2 show that $P_1, P_2 \supset (3)$ are maximal ideals of A. They are not principal (since there are no elements of A of norm 3) and their product is equal to

$$P_1P_2 = (3,\ i\sqrt{5} - 1)(3,\ i\sqrt{5} + 1) = (9,\ 3i\sqrt{5} + 3,\ 3i\sqrt{5} - 3,\ -6) = (3,\ 3i\sqrt{5}) = 3A = (3).$$

In other words, 3 is an irreducible element, but the principal ideal (3) admits a non-trivial factorisation.

A similar calculation with 7 replacing 3 yields (exercise!) two non-principal maximal ideals $Q_1, Q_2 \supset (7)$ such that $A/Q_i \xrightarrow{\sim} \mathbf{F}_7$ and

$$(7) = Q_1Q_2, \qquad Q_1 = (7,\ i\sqrt{5} + 3), \qquad Q_2 = (7,\ i\sqrt{5} - 3).$$

Computing the products

$$P_1Q_1 = (3,\ i\sqrt{5} - 1)(7,\ i\sqrt{5} + 3) = (21,\ 3i\sqrt{5} + 9,\ 7i\sqrt{5} - 7,\ 2i\sqrt{5} - 8) = (21,\ 3i\sqrt{5} + 9,\ 2i\sqrt{5} - 8) =$$
$$= (21,\ i\sqrt{5} + 17,\ 2i\sqrt{5} - 8) = (21, i\sqrt{5} - 4) = (i\sqrt{5} - 4)$$
$$P_2Q_2 = (3,\ i\sqrt{5} + 1)(7,\ i\sqrt{5} - 3) = (21,\ 3i\sqrt{5} - 9,\ 7i\sqrt{5} + 7,\ 2i\sqrt{5} + 8) = (21,\ 3i\sqrt{5} - 9,\ 2i\sqrt{5} + 8) =$$
$$= (21,\ i\sqrt{5} - 17,\ 2i\sqrt{5} + 8) = (21, i\sqrt{5} + 4) = (i\sqrt{5} + 4),$$

we obtain a refinement of the factorisation (6.9.1):

$$(3) = P_1P_2, \qquad (7) = Q_1Q_2, \qquad (i\sqrt{5} + 4) = P_2Q_2, \qquad (i\sqrt{5} - 4) = P_1Q_1, \qquad (21) = P_1P_2Q_1Q_2.$$

The above calculations are a special case of the Kummer-Dedekind theorem.

(6.10) One can show that the ring $A = \mathbf{Z}[i\sqrt{5}]$ has the following properties.

Non-zero ideals of A admit **unique** factorisation into products of maximal ideals (A is a **Dedekind ring**) and the square of any ideal of A is principal (exercise: compute P_1^2, Q_1^2). More precisely, if I is a non-principal ideal of A, then IP_1 is principal (the **ideal class group** of A has two elements).

One can deduce from these properties that the equation $y^2 + 5 = x^3$ ($x, y \in \mathbf{Z}$) has no solution (cf. 5.12).

In fact, one can easily determine the factorisation of any prime number $p \neq 2, 5$ in A, as follows. The quadratic reciprocity law implies that for $p \equiv 1, 3, 7, 9 \pmod{20}$ the polynomial $X^2 + 5$ has two distinct roots in $\mathbf{Z}/p\mathbf{Z} = \mathbf{F}_p$, hence $(p) = PP'$ with $A/P \xrightarrow{\sim} A/P' \xrightarrow{\sim} \mathbf{F}_p$, as in the case $p = 3$ or 7. For $p \equiv -1, -3, -7, -9 \pmod{20}$ the polynomial $X^2 + 5$ is irreducible in $\mathbf{F}_p[X]$, hence $(p) \in \mathrm{Max}(A)$ and $A/(p) \xrightarrow{\sim} \mathbf{F}_p[X]/(X^2 + 5) \xrightarrow{\sim} \mathbf{F}_{p^2}$. The ideal classes are distributed as follows: P and P' (resp. PP_1 and $P'P_1$) are principal $\iff$ $p \equiv 1, 9 \pmod{20}$ (resp. if $p \equiv 3, 7 \pmod{20}$), which implies that

$$\exists\, x, y \in \mathbf{Z} \quad p = x^2 + 5y^2 \iff p \equiv 1, 9 \pmod{20},$$
$$\exists\, x, y \in \mathbf{Z} \quad p = 2x^2 + 2xy + 3y^2 \iff p \equiv 3, 7 \pmod{20}.$$

(6.11) Exercise. *Let A be a UFD in which every non-zero prime ideal is maximal ("A has dimension ≤ 1").*
(i) If x is an irreducible element not dividing $a \in A$, then $(a, x^n) = (1)$, for all $n \geq 1$.
(ii) For every $a, b \in A \setminus \{0\}$ the ideal (a, b) is principal, generated by $\gcd(a, b)$.
(iii) Every non-zero ideal I of A contains the greatest common divisor of all elements of $I \setminus \{0\}$.
(iv) A is a PID.

7. Irreducibility of polynomials

In this section we show, following Gauss, that $A[X]$ is a UFD for any UFD A. In particular, $\mathbf{Z}[x_1, \ldots, x_n]$ (and $K[x_1, \ldots, x_n]$, for any field K) is a UFD.

(7.1) Definition. *Let A be a UFD with fraction field $K = \mathrm{Frac}(A)$. The* **content** *of a non-zero polynomial $f \in A[X] \setminus \{0\}$ is the greates common divisor of its coefficients, denoted by $ct(f) \in (A \setminus \{0\})/A^*$. The content of $f \in K[X] \setminus \{0\}$ is defined as $ct(f) = ct(af)/a \in (K \setminus \{0\})/A^*$, for any $a \in A \setminus \{0\}$ such that $af \in A[X]$. Note that $ct(f) \in (A \setminus \{0\})/A^* \iff f \in A[X] \setminus \{0\}$.*

(7.2) Lemma (Gauss). *For any $f, g \in K[X] \setminus \{0\}$ we have $ct(fg) = ct(f)ct(g)$.*

Proof. After replacing f by $f/ct(f)$ and g by $g/ct(g)$ we can assume that $f, g \in A[X] \setminus \{0\}$ and $ct(f) = ct(g) = 1$. We must show that $ct(fg) = 1$; in other words, that for each prime element π of A there exists a coefficient of fg not divisible by π. Write $f = \sum a_i X^i$, $g = \sum b_j X^j$ and $fg = \sum c_k X^k$. If $i \geq 0$ (resp. $j \geq 0$) is the smallest index for which π does not divide a_i (resp. b_j), then $c_{i+j} = a_i b_j + \pi d$ for some $d \in A$. The product $a_i b_j$ is not divisible by π (since neither of b_i and c_j is, and A is a UFD), which implies that π does not divide c_{i+j}, as required.

(7.3) Corollary. *(i) If $f \in A[X] \setminus \{0\}$ is a product $f = gh$ with $g, h \in K[X] \setminus K$, then there exist $g_1, h_1 \in A[X] \setminus A$ such that $f = g_1 h_1$ [This is false, in general, for domains which are not UFD].*
(ii) If $f \in A[X] \setminus \{0\}$ and $ct(f) = 1$, then $fK[X] \cap A[X] = fA[X]$. Equivalently, the canonical ring homomorphism $A[X]/(f) \longrightarrow K[X]/(f)$ is injective.

Proof. (i) We have $f = g_1 h_1$, where $g_1 = g/ct(g)$ and $h_1 = h\, ct(g)$. Both g_1 and h_1 are non-constant and lie in $A[X]$, since $ct(g_1) = 1$ and $ct(h_1) = ct(g)ct(h) = ct(f) \in (A \setminus \{0\})/A^*$.
(ii) If $g \in K[X] \setminus \{0\}$ satisfies $fg \in A[X]$, then $ct(g) = ct(f)ct(g) = ct(fg) \in (A \setminus \{0\})/A^*$, hence $g \in A[X]$.

(7.4) Proposition. *If A is a UFD with fraction field K, then the irreducible elements of $A[X]$ are as follows.*
(i) Irreducible elements of A.
(ii) Non-constant polynomials $f \in A[X] \setminus A$ with $ct(f) = 1$ which are irreducible in $K[X]$.

Proof. Firstly, the fact that A is a domain implies that so is $A[X]$ and that $A[X]^* = A^*$. Assume that f is an irreducible element of $A[X]$.

If $\deg(f) = 0$, then f is necessarily an irreducible element of A. Conversely, for any irreducible $f \in A$ the quotient ring $A[X]/(f) = (A/(f))[X]$ is a domain, since $A/(f)$ is, hence f is irreducible in $A[X]$ (use both parts of Proposition 6.8).

If $\deg(f) > 0$, then $ct(f) = 1$ (otherwise there would be a non-trivial factorisation $f = ct(f)\,(f/ct(f))$) and f is irreducible in $K[X]$, by Corollary 7.3(i). Conversely, if $f \in A[X] \setminus A$ has $ct(f) = 1$ and is irreducible in $K[X]$, then $K[X]/(f)$ is a domain by Proposition 6.8(ii) (since $K[X]$ is a UFD), hence so is $A[X]/(f)$, by Corollary 7.3(ii). It follows that f is irreducible in $A[X]$, by Proposition 6.8(i).

(7.5) Theorem. *If A is a UFD, so is $A[X]$.*

Proof. We must check that $A[X]$ satisfies 5.6(i) and 5.7(ii'). The second condition was verified in the course of the proof of Proposition 7.4, so it is enough to check that any non-zero element f of $A[X]$ is a product of an invertible element and finitely many irreducible elements. If $f \in A$, then 5.6(i) holds in A and we conclude by Proposition 7.4(i). If $f \notin A$, write $f = g_1 \cdots g_r$, where each g_i is an irreducible non-constant element of the UFD $K[X]$ ($K = \mathrm{Frac}(A)$). The rescaled polynomials $h_i = g_i/ct(g_i)$ satisfy $ct(h_i) = 1$, thus $h_i \in A[X] \setminus A$ and h_i is irreducible in $A[X]$, by Proposition 7.4(ii). Writing $f = ct(g_1) \cdots ct(g_r)\, h_1 \cdots h_r = ct(f)\, h_1 \cdots h_r$ and factoring $ct(f) \in A \setminus \{0\}$ into a product of irreducible elements of A we obtain the desired factorisation of f in $A[X]$.

(7.6) Theorem (Eisenstein's irreducibility criterion). *Let A be a domain, let $P \in \mathrm{Spec}(A)$. Any monic polynomial $f = X^n + a_{n-1}X^{n-1} + \cdots + a_0 \in A[X]$ such that $a_0, \ldots, a_{n-1} \in P$ and $a_0 \notin P^2$ is an irreducible element of $A[X]$.*

Proof. If $f = gh$ with $g, h \in A[X] \setminus (\{0\} \cup A^*)$, then $g, h \notin A$ and we can assume that both g and h are monic polynomials. Denote by $\overline{g}$ (resp. $\overline{h}$) the image of g (resp. h) in $(A/P)[X] \subset (\mathrm{Frac}(A/P))[X]$. As $\overline{g}\,\overline{h} = \overline{f} = X^n$, unique factorisation in $(\mathrm{Frac}(A/P))[X]$ implies that $\overline{g} = X^r$ and $\overline{h} = X^{n-r}$ for some $1 \leq r \leq n$. In other words, $g = X^r + b_{r-1}X^{r-1} + \cdots + b_0$ and $h = X^{n-r} + c_{n-r-1}X^{n-r-1} + \cdots + c_0$ with $b_j, c_k \in P$ for all j, k. It follows that $a_0 = b_0 c_0 \in P^2$, which contradicts our assumptions.

(7.7) Corollary. *If, in addition, A is a UFD with fraction field K, then f is irreducible in $K[X]$.*

Proof. Apply Corollary 7.3(i).

(7.8) Exercise. *Let $a(Y) \in \mathbf{C}[Y] \setminus \mathbf{C}$ be a polynomial which has at least one simple root. Show that, for any $n \geq 1$, the polynomial $X^n + a(Y)$ is an irreducible element of $\mathbf{C}[X, Y]$.*

II. Modules

1. Basic concepts

Modules (in particular, ideals) are linear objects. As a result, they are easier to study than rings themselves. On the other hand, modules over rings other than fields tend not to have a basis (i.e., are not free in general).

(1.1) Let R be a ring (commutative and unital, of course). Recal that an **R-module** is an abelian group M equipped with multiplication by elements of R ($R \times M \longrightarrow M$, $(r, m) \mapsto rm$) satisfying the usual properties ($r(m+m') = rm + rm'$, $(r+r')m = rm + r'm$, $r(r'm) = (rr')m$, $1m = m$ for all $r, r' \in R$ and $m, m' \in M$). In particular, if $R = K$ is a field (resp. if $R = \mathbf{Z}$), then M is a K-vector space (resp. an abelian group).

The standard concepts of a submodule, quotient module and the submodule $\langle S \rangle \subset M$ generated by a subset $S \subset M$ work analogously as in the case of vector spaces.

The multiplication in the ring itself makes R into an R-module; its submodules are the ideals of R.

For any ideal I of R the set of elements of M killed by I

$$M[I] := \{m \in M \mid \forall a \in I \quad am = 0\}$$

is a submodule of M. If $I = (x)$ is principal, then $M[(x)] = M[x] := \{m \in M \mid xm = 0\}$. If I' is another ideal, then $M[I] \cap M[I'] = M[I + I']$.

We denote by $IM \subset M$ the submodule generated by the products am, where $a \in I$ and $m \in M$.

(1.2) A **homomorphism** of R-modules is a homomorphism of abelian groups $f : M \longrightarrow N$ satisfying $f(rm) = rf(m)$, for all $r \in R$ and $m \in M$. The set of all such homomorphisms $\mathrm{Hom}_R(M, N)$ is an abelian group with respect to addition $(f + f')(m) = f(m) + f'(m)$, in fact an R-module for the operation $(rf)(m) = f(rm) = rf(m)$ (this is true only for modules over commutative rings). The **kernel** of f

$$\mathrm{Ker}(f) = \{m \in M \mid f(m) = 0\} \subset M$$

is a submodule of M, the **image** of f

$$\mathrm{Im}(f) = \{f(m) \mid m \in M\} \subset N$$

is a submodule of N and the **cokernel** of f, $\mathrm{Coker}(f) = N/\mathrm{Im}(f)$, is a quotient module of N.

(1.3) Change of the ring. If $f : R' \longrightarrow R$ is a ring homomorphism, then any R-module M becomes an R'-module via f, with multiplication given by $r'm := f(r')m$.

In particular, an R/I-module (for an ideal I of R) can be identified with an R-module M such that $M = M[I]$ (via the projection $\mathrm{pr} : R \longrightarrow R/I$). This implies that a $\mathbf{Z}/n\mathbf{Z}$-module is an abelian group M such that $nM = 0$ (for any $n \in \mathbf{Z}$).

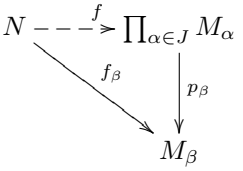
(1.4) Direct sums and products. Let J be a non-empty set. Assume that we are given, for each $\alpha \in J$, an R-module M_α. The **direct product** of these modules is the cartesian product

$$\prod_{\alpha \in J} M_\alpha = \{(m_\alpha)_{\alpha \in J} \mid m_\alpha \in M_\alpha\}$$

with the R-module structure given by

$$(m_\alpha) + (m'_\alpha) = (m_\alpha + m'_\alpha), \qquad r(m_\alpha) = (rm_\alpha).$$

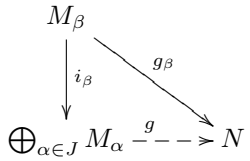
For each $\beta \in J$ the projection map $p_\beta : \prod_{\alpha \in J} M_\alpha \longrightarrow M_\beta$, $p_\beta((m_\alpha)_{\alpha \in J}) = m_\beta$ is a (surjective) module homomorphism. The direct product is the "smallest" module with this property: for each R-module N equipped with module homomorphisms $f_\beta : N \longrightarrow M_\beta$ there exists a unique module homomorphism $f : N \longrightarrow \prod_{\alpha \in J} M_\alpha$ such that $f_\beta = p_\beta \circ f$ for all $\beta \in J$, namely, $f(n) = (f_\alpha(n))_{\alpha \in J}$.



Dually, the **direct sum** of the modules M_α is the submodule

$$\bigoplus_{\alpha\in J} M_\alpha = \{(m_\alpha)_{\alpha\in J} \in \prod_{\alpha\in J} M_\alpha \mid m_\alpha = 0 \text{ for all but finitely many } \alpha \in J\} \subset \prod_{\alpha\in J} M_\alpha$$

(in particular, it coincides with the direct product if the set J is finite). For each $\beta \in J$ the inclusion $i_\beta : M_\beta \hookrightarrow \bigoplus_{\alpha\in J} M_\alpha$, $i_\beta(x) = (m_\alpha)$ with $m_\alpha = x$ for $\alpha = \beta$ and $m_\alpha = 0$ otherwise is an (injective) module homomorphism. Again, the direct sum is the "smallest" module with this property: for each R-module N equipped with module homomorphisms $g_\beta : M_\beta \longrightarrow N$ there exists a unique module homomorphism $g : \bigoplus_{\alpha\in J} M_\alpha \longrightarrow N$ such that $g_\beta = g \circ i_\beta$ for all $\beta \in J$, namely, $g(\sum_{k=1}^s i_{\beta_k}(x_k)) = \sum_{k=1}^s g_{\beta_k}(x_k)$. This uses the fact that each element of the direct sum admits a unique decomposition as a finite sum $i_{\beta_1}(x_1) + \cdots + i_{\beta_s}(x_s)$ for distinct $\beta_1, \ldots, \beta_s \in J$, $x_k \in M_{\beta_k}$ and $s \geq 0$.



The previous discussion can be summed up by saying that the maps $f \mapsto (p_\alpha \circ f)_{\alpha\in J}$ and $g \mapsto (g \circ i_\alpha)_{\alpha\in J}$ induce bijections of sets (in fact, isomorphisms of R-modules)

$$\mathrm{Hom}_R(N, \prod_{\alpha\in J} M_\alpha) \xrightarrow{\sim} \prod_{\alpha\in J} \mathrm{Hom}_R(N, M_\alpha), \qquad \mathrm{Hom}_R(\bigoplus_{\alpha\in J} M_\alpha, N) \xrightarrow{\sim} \prod_{\alpha\in J} \mathrm{Hom}_R(M_\alpha, N). \tag{1.4.1}$$

If J is empty, it is useful to define the corresponding direct sum and product to be the zero module $0 = \{0\}$.

In the special case when $M_\alpha = M$ for all $\alpha \in J$ we are going to use the notation

$$M^{(J)} = \bigoplus_{\alpha\in J} M \subset \prod_{\alpha\in J} M = M^J.$$

The direct sum (= the direct product) of n copies of M will be denoted by M^n. In another special case when each M_α is a submodule of a fixed R-module M there is a canonical morphism

$$\bigoplus_{\alpha\in J} M_\alpha \longrightarrow M, \qquad (m_\alpha) \mapsto \sum_{\alpha\in J} m_\alpha \qquad \text{(a finite sum!)}. \tag{1.4.2}$$

Its image is the submodule $\sum_{\alpha\in J} M_\alpha \subset M$ generated by the M_α. If the morphism (1.4.2) is injective (which is equivalent to $M_\beta \cap \sum_{\alpha\in J\setminus\{\beta\}} M_\alpha = 0$ for all $\beta \in J$), it identifies $\bigoplus_{\alpha\in J} M_\alpha$ with $\sum_{\alpha\in J} M_\alpha \subset M$.

(1.5) Free modules. The **free R-module on a set** J is the R-module $R^{(J)} = \bigoplus_{\alpha\in J} R$. Its elements are finite linear combinations $\sum_{\alpha\in J} r_\alpha e(\alpha)$ (with only finitely many $r_\alpha \in R$ non-zero), where $e(\alpha)$ are the elements of the "canonical basis" of $R^{(J)}$: $e(\alpha)_\beta = 1$ if $\beta = \alpha$ (resp. $e(\alpha)_\beta = 0$ if $\beta \neq \alpha$). There is no R-linear relation between the $e(\alpha)$: for any R-module N and any collection of elements $n(\alpha) \in N$ ($\alpha \in J$) there is a unique homomorphism of R-modules $f : R^{(J)} \longrightarrow N$ such that $f(e(\alpha)) = n(\alpha)$ for all $\alpha \in J$, namely, $f(\sum_{\alpha\in J} r_\alpha e(\alpha)) = \sum_{\alpha\in J} r_\alpha n(\alpha)$ (this is a special case of (1.4.1) for $M_\alpha = R$).

(1.6) Proposition-Definition. *Assume that the ring R is non-zero. An R-module M is* **free** *if there exists an isomorphism of R-modules $f : R^{(J)} \xrightarrow{\sim} M$ for some set J. A* **basis of** *M is the image of the canonical basis of $R^{(J)}$ under any such f. The* **rank** *of the free module M is the cardinality of J; it depends only on M. [In the special case $R = \mathbf{Z}$ we obtain the notion of a free abelian group.]*

Proof. Assume that there exists an isomorphism of R-modules $f : R^{(J)} \xrightarrow{\sim} R^{(J')}$. We must show that J and J' have the same cardinality. According to Theorem I.6.6 there exists a maximal ideal I of R; the quotient ring $R/I = K$ is then a field. The isomorphism f induces an isomorphism $(R/I)^{(J)} = R^{(J)}/IR^{(J)} \xrightarrow{\sim} R^{(J')}/IR^{(J')} = (R/I)^{(J')}$ of R/I-modules, hence an isomorphism $K^{(J)} \xrightarrow{\sim} K^{(J')}$ of K-vector spaces. The statement follows from the fact that two bases of any vector space have the same cardinality.

(1.7) Examples. (i) For a field K, all K-modules (= K-vector spaces) are free.
(ii) Exercise: the additive group of $\mathbf{Q}$ is not contained in any free abelian group.

(1.8) Generators and relations. Let N be an R-module. If we are given elements $n(\alpha) \in N$ ($\alpha \in J$) which generate N (a set of not necessarily distinct **generators** of N), then the module homomorphism $f : R^{(J)} \longrightarrow N$ constructed in 1.5 is surjective. Its kernel consists of finite linear combinations $\sum_{\alpha\in J} r_\alpha e(\alpha)$ for which $\sum_{\alpha\in J} r_\alpha n(\alpha) = 0 \in N$, in other words, of R-linear **relations** between the generators $n(\alpha)$.

(1.9) Definition. *An R-module N is* **finitely generated** *(or is an R-***module of finite type***) if it admits a finite generating set ($\iff$ there exists a surjective homomorphism of R-modules $f : R^a \longrightarrow N$ for some $a \in \mathbf{N}$). It is* **finitely presented** *if there exists f as above for which the module of relations* $\mathrm{Ker}(f)$ *is also finitely generated ($\iff$ N is defined by finitely many generators and finitely many relations). [These two properties are equivalent if the ring R is noetherian; see Proposition 3.6 below.] N is* **cyclic** *if there is a surjective homomorphism of R-modules $f : R \longrightarrow N$ ($\iff$ N is isomorphic to R/I, for some ideal I of R).*

(1.10) Modules over a product. Let $R = R_1 \times \cdots \times R_n$ be a product of rings, let $e_k \in R$ be the corresponding idempotents ($R_k = e_k R$) as in I.3.2. Proposition I.3.3 has an obvious analogue for modules: any R-module M defines R_k-modules $M_k := e_k M \subset M$ with respect to the product $(e_k r)(e_k m) := e_k(rm)$ and the map

$$M \longrightarrow M_1 \oplus \cdots \oplus M_n, \qquad m \mapsto (e_1 m, \ldots, e_n m)$$

is an isomorphism of R-modules (with each M_k considered as an R-module via the projection $p_k : R \longrightarrow R_k$). The inverse map is the sum of the inclusions $M_k \hookrightarrow M$, as in (1.4.2).

(1.11) Torsion modules. Assume that M is a module over a domain R. An element $m \in M$ is **torsion** if there is $a \in R \setminus \{0\}$ such that $am = 0$. The union of all torsion elements is the **torsion submodule** of M:

$$M_{\mathrm{tors}} = \bigcup_{a\neq 0} M[a] = \bigcup_{I\neq 0} M[I] \subset M,$$

where I runs through all non-zero ideals of R. We say that M is a **torsion module** (resp. a **torsion-free module**) if $M = M_{\mathrm{tors}}$ (resp. if $M_{\mathrm{tors}} = 0$). For any M the module M/M_{tors} is torsion-free. Any free module is torsion-free.

The structure theorem for finitely generated abelian groups (= $\mathbf{Z}$-modules) proved in Algebra 1 implies that a finitely generated torsion-free abelian group is free. Example 1.7(ii) shows that this is no longer true for abelian groups which are not finitely generated.

(1.12) Proposition-Definition (Primary decomposition for torsion abelian groups). *Let X be a torsion abelian group. For every prime number p denote by $X(p) := \{x \in X \mid \exists a \geq 1 \;\; p^a x = 0\} = \bigcup_{a\geq 1} X[p^a]$ the p-***primary component** *of X.*
(i) If $X = X[n]$, where $n = p_1^{a_1} \cdots p_r^{a_r}$, $r \geq 0$, $a_i \geq 1$ and p_i are distincts prime numbers, then the inclusions $X[p_i^{a_i}] \hookrightarrow X[n] = X$ give rise – as in (1.4.2) – to an isomorphism of abelian groups

$$X[p_1^{a_1}] \oplus \cdots \oplus X[p_r^{a_r}] = X(p_1) \oplus \cdots \oplus X(p_r) \xrightarrow{\sim} X = X[n] = X[p_1^{a_1} \cdots p_r^{a_r}].$$

In particular, $X(p) = 0$ for every prime number p not dividing n.
(ii) If X is arbitrary, then the inclusions $X(p) \hookrightarrow X$ give rise to an isomorphism

$$\bigoplus_{p \text{ prime}} X(p) \xrightarrow{\sim} X.$$

Proof. (i) This is a consequence of 1.10 for the isomorphism $\mathbf{Z}/n\mathbf{Z} \xrightarrow{\sim} \mathbf{Z}/p_1^{a_1}\mathbf{Z} \times \cdots \times \mathbf{Z}/p_r^{a_r}\mathbf{Z}$. (ii) Every element of the left hand side (resp. right hand side) is contained in some $X[p_1^{a_1}] \oplus \cdots \oplus X[p_r^{a_r}]$ (resp. in some $X[p_1^{a_1} \cdots p_r^{a_r}]$); the statement (i) implies injectivity (resp. surjectivity) of the corresponding map (1.4.2).

(1.13) Exercise. *State and prove a generalisation of 1.12 for torsion modules over a PID.*

(1.14) Exercise. *Let A be a finite abelian group. The set $\{m \in \mathbf{Z} \mid mA = 0\}$ is a non-zero subgroup of $\mathbf{Z}$; its positive generator is called the* **exponent of** *A. Show that:*
(1) The exponent of A divides the order of A.
(2) The exponent of A is equal to the product of the exponents of the p-primary components of A.
(3) There exists an element of A whose order is equal to the exponent of A.

(1.15) Exercise (Duality and the structure theorem for finite abelian groups). *Let A be a finite abelian group. A* **character** *of A is a homomorphism of abelian groups $\chi : A \longrightarrow \mathbf{C}^*$. Show that:*
(1) The characters of A form an abelian group $\widehat{A}$ (the **dual group***) with respect to multiplication : $(\chi\chi')(a) = \chi(a)\chi'(a)$.*
(2) The exponent of $\widehat{A}$ divides the exponent of A.
(3) If A is cyclic of order n, so is $\widehat{A}$.
(4) If $A = B \oplus C$, then $\widehat{A} = \widehat{B} \oplus \widehat{C}$.
[If one admits (9) below, deduce from (3) and (4) that $\widehat{A}$ is isomorphic to A, for every A.]
(5) Every homomorphism of abelian groups $\alpha : B \longrightarrow A$ defines a dual homomorphism $\widehat{\alpha} : \widehat{A} \longrightarrow \widehat{B}$, $\widehat{\alpha}(\chi) = \chi \circ \alpha : B \xrightarrow{\alpha} A \xrightarrow{\chi} \mathbf{C}^*$.
(6) α is surjective $\Longrightarrow$ $\widehat{\alpha}$ is injective.
(7)* α is injective $\Longrightarrow$ $\widehat{\alpha}$ is surjective.
(8) Let $a \in A$ be an element whose order is equal to the exponent of A (see 1.14(3)); denote by B the cyclic subgroup generated by a and by $\alpha : B \longrightarrow A$ the inclusion. Thanks to (3) and (7) there exists $\chi \in \widehat{A}$ for which the order of $\chi(a)$ is equal to the order of a. Show that $A = B \oplus \mathrm{Ker}(\chi)$.
(9) Every finite abelian group is a direct sum of cyclic groups.
(10) The biduality homomorphism $A \longrightarrow \widehat{\widehat{A}}$, $a \mapsto (\chi \mapsto \chi(a))$ is an isomorphism.
(11) Let $B \subset A$ be a subgroup; set $C = A/B$. Show that $\widehat{C}$ is a subgroup of $\widehat{A}$ and $\widehat{B} = \widehat{A}/\widehat{C}$.
(12) The map $B \mapsto \widehat{B}$ defines a bijection between the set of all subgroups of A and the set of all quotient groups of $\widehat{A}$.
(13) Every subgroup of A is isomorphic to a quotient group of A.

(1.16) Exercise. *Let A be a finite abelian group. For every integer $d \geq 1$ denote by $s_d(A)$ (resp. $q_d(A)$) the number of subgroups (resp. of quotient groups) of A of order d. Let $s(A) = \sum_d s_d(A)$ (resp. $q(A) = \sum_d q_d(A)$) be the number of all subgroups (resp. of all quotient groups) of A.*
(1) Determine $s_d(A)$ and $q_d(A)$ if A is a cyclic group of order n.
(2) Show that $s_d(A) = s_d(A[d])$ and $q_d(A) = q_d(A/dA)$.
(3) Show that, if $de = |A|$, then $s_d(A) = q_e(A)$. Deduce from this the equality $s(A) = q(A)$.
(4) If the order of A is relatively prime to the order of B, then $s(A \oplus B) = s(A)s(B)$.
(5) Show that $s(A) = \prod_{p|n} s(A(p))$, where $n = |A|$.
(6) Let p be a prime number. Determine, for every $i \geq 0$,

$$s_{p^i}(\mathbf{Z}/p\mathbf{Z} \oplus \mathbf{Z}/p\mathbf{Z}), \qquad s_{p^i}(\mathbf{Z}/p\mathbf{Z} \oplus \mathbf{Z}/p^2\mathbf{Z}), \qquad s_{p^i}(\mathbf{Z}/p\mathbf{Z} \oplus \mathbf{Z}/p^n\mathbf{Z}) \qquad (n \geq 2).$$

(7) Determine all finite abelian groups A for which $s(A) = 4$ (resp. $s(A) = 5$).
(8) Show that $s_d(A) = q_d(\widehat{A}) = q_d(A)$, where $\widehat{A}$ is the dual group (see 1.15).

2. The language of exact sequences

(2.1) Definition. *A sequence of homomorphisms of R-modules*

$$M_a \longrightarrow \cdots \longrightarrow M_{i-1} \xrightarrow{f_{i-1}} M_i \xrightarrow{f_i} M_{i+1} \longrightarrow \cdots \longrightarrow M_b$$

is **exact at the term** *M_i if $\mathrm{Im}(f_{i-1}) = \mathrm{Ker}(f_i)$ (which implies that $f_i f_{i-1} = 0$ and $\mathrm{Coker}(f_{i-1}) \xrightarrow{\sim} \mathrm{Im}(f_i)$); it is* **exact** *if it is exact at M_i for all $a < i < b$. A* **morphism** *(resp. an* **isomorphism***) between such an exact sequence and another exact sequence*

$$N_a \longrightarrow \cdots \longrightarrow N_{i-1} \xrightarrow{g_{i-1}} N_i \xrightarrow{g_i} N_{i+1} \longrightarrow \cdots \longrightarrow N_b$$

is given by homomorphisms (resp. isomorphisms) of R-modules $u_i : M_i \longrightarrow N_i$ $(a \leq i \leq b)$ *such that* $g_i \circ u_i = u_{i+1} \circ f_i$ *for all* $i = a, \ldots, b-1$.

(2.2) Examples. (1) $0 \longrightarrow M \xrightarrow{f} N$ is exact $\iff \mathrm{Ker}(f) = 0 \iff f$ is injective.
(2) $M \xrightarrow{f} N \longrightarrow 0$ is exact $\iff \mathrm{Im}(f) = N \iff \mathrm{Coker}(f) = 0 \iff f$ is surjective.
(3) $0 \longrightarrow M \xrightarrow{f} N \longrightarrow 0$ is exact $\iff f$ is an isomorphism.
(4) If M is a submodule of N, then $0 \longrightarrow M \xrightarrow{i} N \xrightarrow{p} N/M \longrightarrow 0$ is exact, where i and p denote the inclusion and the canonical projection, respectively.
(5) Any **short exact sequence** $0 \longrightarrow M \xrightarrow{f} N \xrightarrow{g} P \longrightarrow 0$ is naturally isomorphic to a sequence of the form considered in (4), namely, to the sequence $0 \longrightarrow f(M) \longrightarrow N \longrightarrow N/f(M) \longrightarrow 0$.
(6) For any morphism $f : M \longrightarrow N$, the sequence $0 \longrightarrow \mathrm{Ker}(f) \longrightarrow M \xrightarrow{f} N \longrightarrow \mathrm{Coker}(f) \longrightarrow 0$ is exact.
(7) For any short exact sequence $0 \longrightarrow M \xrightarrow{f} N \xrightarrow{g} P \longrightarrow 0$ the following conditions are equivalent (if they are satisfied, we say that the sequence is **split**):
(a) there exists a morphism $s : P \longrightarrow N$ such that $gs = \mathrm{id}_P$;
(b) there exists a morphism $r : N \longrightarrow M$ such that $rf = \mathrm{id}_M$;
(c) the 7-tuple (M, N, P, f, g, r, s) is isomorphic to $(X_1, X_1 \oplus X_2, X_2, i_1, p_2, p_1, i_2)$, where $p_i(x_1, x_2) = x_i$, $i_1(x_1) = (x_1, 0)$ and $i_2(x_2) = (0, x_2)$.
In concrete terms, a splitting s is equivalent to a choice of submodule $M' \subset N$ complementary to $f(M) \subset N$ in the sense that $f(M) \oplus M' = N$ ($M' = \mathrm{Im}(s)$ and s is the inverse of the isomorphism $M' \hookrightarrow N \xrightarrow{g} P$ composed with the inclusion $M' \hookrightarrow N$).
(8) An exact sequence from Definition 2.1 can be cut into short exact sequences

$$0 \longrightarrow P_i \xrightarrow{\alpha_i} M_i \xrightarrow{\beta_i} P_{i+1} \longrightarrow 0 \qquad (a < i < b-1), \qquad P_i = \mathrm{Im}(f_{i-1}) = \mathrm{Ker}(f_i), \qquad \alpha_{i+1}\beta_i = f_i,$$

and vice versa.

(2.3) Remarks. (1) In particular, an R-module M is finitely generated $\iff$ there exists an exact sequence

$$R^n \longrightarrow M \longrightarrow 0.$$

It is finitely presented $\iff$ there exists an exact sequence

$$R^m \longrightarrow R^n \longrightarrow M \longrightarrow 0.$$

It is useful to continue this process and consider exact sequences of the form

$$R^{a_n} \longrightarrow \cdots \longrightarrow R^{a_1} \longrightarrow R^{a_0} \longrightarrow M \longrightarrow 0$$

("relations between relations between relations..."). An important result of Hilbert states that for $R = K[x_1, \ldots, x_n]$ (where K is a field) every finitely generated R-module M sits in an exact sequence

$$0 \longrightarrow R^{a_n} \longrightarrow \cdots \longrightarrow R^{a_1} \longrightarrow R^{a_0} \longrightarrow M \longrightarrow 0$$

(M admits a "free resolution of length n").
(2) Exercise 1.15(6),(7),(11) implies that the "duality functor" which associates to a finite abelian group A its dual group $\widehat{A} = \mathrm{Hom}_{\mathbf{Z}}(A, \mathbf{C}^*)$ is exact: for any exact sequence of finite abelian groups

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

the sequence

$$0 \longrightarrow \widehat{C} \xrightarrow{\widehat{g}} \widehat{B} \xrightarrow{\widehat{f}} \widehat{A} \longrightarrow 0$$

is also exact.
(3) In fact, this property holds for exact sequences of arbitrary abelian groups ($\mathbf{C}^*$ is an "injective abelian group"), but a proof of 1.15(7) in the general situation requires a use of Zorn's Lemma.

(2.4) Snake Lemma. *Let*

$$\begin{array}{ccccccc} (0 \longrightarrow) & A & \xrightarrow{f} & B & \xrightarrow{g} & C & \longrightarrow 0 \\ & \downarrow{\scriptstyle \alpha} & & \downarrow{\scriptstyle \beta} & & \downarrow{\scriptstyle \gamma} & \\ 0 \longrightarrow & A' & \xrightarrow{f'} & B' & \xrightarrow{g'} & C' & (\longrightarrow 0) \end{array}$$

be a commutative diagram with exact rows. Then there is an exact sequence

$$(0 \longrightarrow)\ \operatorname{Ker}(\alpha) \xrightarrow{f} \operatorname{Ker}(\beta) \xrightarrow{g} \operatorname{Ker}(\gamma) \xrightarrow{\Delta} \operatorname{Coker}(\alpha) \xrightarrow{f'} \operatorname{Coker}(\beta) \xrightarrow{g'} \operatorname{Coker}(\gamma)\ (\longrightarrow 0),$$

in which the non-obvious morphism $\Delta : \operatorname{Ker}(\gamma) \longrightarrow \operatorname{Coker}(\alpha)$ *is given by "*$\Delta = f'^{-1}\beta g^{-1}$*".*

Proof. We only give a definition of Δ and leave the verification of exactness as an exercise. Let $c \in \operatorname{Ker}(\gamma)$. As g is surjective, there exists $b \in B$ such that $g(b) = c$. As $g'\beta(b) = \gamma g(b) = \gamma(c) = 0$, there exists $a' \in A'$ (unique, since f' is injective) such that $f'(a') = \beta(b)$. We wish to define $\Delta(c) := a' + \alpha(A) \in \operatorname{Coker}(\alpha)$. In order to check that this definition makes sense we must analyse what happens if we take two different elements $b_1, b_2 \in B$ satisfying $g(b_i) = c$. As $b_1 - b_2 \in \operatorname{Ker}(g) = \operatorname{Im}(f)$, we have $b_1 - b_2 = f(a)$ for some $a \in A$; thus the corresponding elements $a'_i \in A'$ (where $f'(a'_i) = \beta(b_i)$) satisfy $f'(a'_1 - a'_2) = \beta(b_1 - b_2) = \beta f(a) = f'\alpha(a)$, hence $a'_1 - a'_2 = \alpha(a) \in \alpha(A)$ (since f' is injective), which implies that $\Delta(c) := a'_1 + \alpha(A) = a'_2 + \alpha(A) \in \operatorname{Coker}(\alpha)$ is, indeed, independent of the choice of b. As Δ is a composition of possibly multivalued R-linear maps, it is also R-linear.

(2.5) Exercise. *For any homomorphisms of* R*-modules* $X \xrightarrow{f} Y \xrightarrow{g} Z$ *the following sequence is exact.*

$$0 \longrightarrow \operatorname{Ker}(f) \longrightarrow \operatorname{Ker}(g \circ f) \xrightarrow{f} \operatorname{Ker}(g) \longrightarrow \operatorname{Coker}(f) \xrightarrow{g} \operatorname{Coker}(g \circ f) \longrightarrow \operatorname{Coker}(g) \longrightarrow 0.$$

(2.6) Exercise. *If* $0 \longrightarrow M \xrightarrow{f} N \xrightarrow{g} P \longrightarrow 0$ *is an exact sequence, then every module homomorphism* $h : R^{(J)} \longrightarrow P$ *is of the form* $h = g \circ h'$ *for some homomorphism* $h' : R^{(J)} \longrightarrow N$ *("free modules are projective").*

(2.7) Exercise. *(i) Let* $0 \longrightarrow M \longrightarrow N \longrightarrow P \longrightarrow 0$ *be an exact sequence. If* M *and* P *are* R*-modules of finite type, so is* N*.*
(ii) If M *is an* R*-module of finite presentation, then for* **every** *surjective homomorphism* $g : R^b \longrightarrow M$ *(with* $b \in \mathbf{N}$*) the kernel* $\operatorname{Ker}(g)$ *is an* R*-module of finite type.*
(iii) Why did we wait with (ii) until §2?

3. Noetherian (after Emmy Noether) rings and modules

(3.1) Definition. *An* R*-module* M *is* **noetherian** *if it satisfies the following equivalent conditions.*
(i) Every submodule N *of* M *is finitely generated.*
(ii) **Ascending chain condition.** *Every ascending chain of submodules* $M_1 \subset M_2 \subset \cdots \subset M$ *of* M *stabilises: there is an index* j *such that* $M_k = M_j$ *for all* $k \geq j$*.*
(iii) Every non-empty set S *of submodules of* M *contains a maximal element* $P \in S$ *with respect to inclusion (i.e., there is no* $P' \in S$ *such that* $P \subsetneq P'$*).*

Proof (of the fact that the three conditions are equivalent). (i) $\Longrightarrow$ (ii) The union $N = \bigcup_{i=1}^{\infty} M_i$ is a submodule of M, hence generated by a finite set of elements $n_1, \ldots, n_r \in N$, by (i). If $j \geq 1$ is large enough for M_j to contain all $n_1, \ldots, n_r$, then $N = M_j$ and $M_k = M_j$ for all $k \geq j$.
(ii) $\Longrightarrow$ (iii) There exists $M_1 \in S$, since S is non-empty. If S does not have a maximal element, then there exist $M_2 \in S$ such that $M_1 \subsetneq M_2$, $M_3 \in S$ such that $M_2 \subsetneq M_3$ etc., hence a strictly increasing chain of submodules $M_1 \subsetneq M_2 \subsetneq M_3 \cdots \subset M$, in contradiction with (ii).
(iii) $\Longrightarrow$ (i) Let S be the set of finitely generated submodules of a fixed submodule N of M. It is non-empty, since $0 \in S$, hence it contains a maximal element P, by (iii). If $P \neq N$, then there exists $n \in N \setminus P$. The submodule module $P + Rn \subset N$ generated by P and n is again finitely generated, hence $P + Rn \in S$. On the other hand, $P \subsetneq P + Rn$, which contradicts the maximality of P. It follows that $P = N$, hence N is finitely generated.

(3.2) Proposition. *If M is an R-module and N a submodule of M, then M is noetherian $\iff$ both N and M/N are noetherian.*

Proof. Denote by $\mathrm{pr} : M \longrightarrow M/N$ the canonical projection. If M is noetherian, so are N (by 3.1(i)) and M/N (indeed, for any submodule $X \subset M/N$ the submodule $\mathrm{pr}^{-1}(X) \subset M$ is finitely generated, hence so is $X = \mathrm{pr}(\mathrm{pr}^{-1}(X))$). Conversely, if N and M/N are noetherian and $M_1 \subset M_2 \subset \cdots \subset M$ is a chain of submodules, then the chains of submodules

$$M_1 \cap N \subset M_2 \cap N \subset \cdots \subset N, \qquad \mathrm{pr}(M_1 + N) \subset \mathrm{pr}(M_2 + N) \subset \cdots \subset M/N$$

must stabilise: there exists j such that $M_k \cap N = M_j \cap N$ and $\mathrm{pr}(M_k + N) = \mathrm{pr}(M_j + N)$ for all $k \geq j$. Lemma 3.3 below then implies that $M_k = M_j$ for $k \geq j$. Thus M is noetherian, by 3.1(ii).

(3.3) Lemma. *If $X \subset Y \subset M$ are submodules such that $X \cap N = Y \cap N$ and $\mathrm{pr}(X + N) = \mathrm{pr}(Y + N)$, then $X = Y$.*

Proof. Let $y \in Y$; we must show that $y \in X$. The condition $\mathrm{pr}(X+N) = \mathrm{pr}(Y+N)$ implies that there exist $x \in X$ and $n \in N$ such that $y = x+n$. It follows that $n = y-x \in N \cap Y = N \cap X$, hence $y = x+(y-x) \in X$.

(3.4) Definition. *A ring R is* **noetherian** *if R is noetherian as an R module, i.e., if the following equivalent conditions are satisfied.*
(i) Every ideal I of R is finitely generated.
(ii) **Ascending chain condition.** *Every ascending chain of ideals $I_1 \subset I_2 \subset \cdots \subset R$ of R stabilises: there is an index j such that $I_k = I_j$ for all $k \geq j$.*
(iii) Every non-empty set S of ideals of R contains a maximal element $I \in S$ with respect to inclusion (i.e., there is no $I' \in S$ such that $I \subsetneq I'$).

(3.5) Examples. (i) Any PID (in particular, any field) is a noetherian ring.
(ii) The polynomial ring $R = \mathbf{C}[x_1, x_2, \ldots] = \bigcup_{n=1}^{\infty} \mathbf{C}[x_1, \ldots, x_n]$ in an infinite number of variables is not noetherian, since the ideal $I = (x_1, x_2, \ldots)$ is not finitely generated (in other words, $M = R$ is a finitely generated R-module which is not noetherian). Note that R is a UFD.
(iii) According to a theorem of I.S. Cohen ([Ma, Thm. 3.4]), a ring R is noetherian $\iff$ every prime ideal of R is finitely generated.

(3.6) Proposition. *Let R be a noetherian ring. An R-module M is noetherian $\iff$ M is finitely generated. In particular, every finitely generated R-module is finitely presented.*

Proof. The implication "$\Longrightarrow$" is automatic. Conversely, the isomorphisms of R-modules $R^n/R \xrightarrow{\sim} R^{n-1}$ imply, by induction and Proposition 3.2, that each R^n ($n \geq 1$) is a noetherian R-module. A finitely generated R-module is of the form R^n/N for some $n \geq 1$, hence is noetherian, again by Proposition 3.2.

(3.7) Proposition. *Let $f : R \longrightarrow R'$ be a surjective ring homomorphism. If R is noetherian, so is R'.*

Proof. For any chain of ideals $J_1 \subset J_2 \subset \cdots \subset R'$ of R' the chain $f^{-1}(J_1) \subset f^{-1}(J_2) \subset \cdots \subset R$ of ideals of R must stabilise: there exists j such that $f^{-1}(J_k) = f^{-1}(J_j)$ for all $k \geq j$. Surjectivity of f then yields $J_k = f(f^{-1}(J_k)) = f(f^{-1}(J_j)) = J_j$ for all $k \geq j$, which means that R' is noetherian.

(3.8) Theorem ("Hilbert's basis theorem"). *If R is a noetherian ring, so is $R[x_1, \ldots, x_n]$ ($n \geq 1$).*

Proof. By induction, we can assume that $n = 1$. We must show that any ideal $I \subset R[x]$ is finitely generated. For any $i \geq 0$ consider

$$I_i = \{a \in R \mid \exists\, ax^i + a_{i-1}x^{i-1} + \cdots a_0 \in I\} \subset R.$$

This is an ideal of R and these ideals form a chain $I_0 \subset I_1 \subset \cdots \subset R$. The noetherian assumption on R implies that there exists $r \geq 0$ such that $I_k = I_r$ for all $k \geq r$. Moreover, $I_0, \ldots, I_r$ are finitely generated ideals of R. As a results, there exist $f_{ij} \in I$ ($0 \leq i \leq r$, $1 \leq j \leq m$, $\deg(f_{ij}) \leq i$) such that, for each $i = 0, \ldots, r$, the ideal I_i is generated by the coefficients at x^i of the polynomials $f_{i1}, \ldots, f_{im}$. Denote by $J \subset I$ the ideal generated by the (finite set of) polynomials f_{ij} ($0 \leq i \leq r$, $1 \leq j \leq m$). It is enough to

show that any $f \in I \setminus \{0\}$ is contained in J ($\Longrightarrow I = J$ is finitely generated). If $d = \deg(f)$, then there exist $c_1, \ldots, c_m \in R$ such that the polynomial $g_d = (c_1 f_{r1} + \cdots c_m f_{rm}) x^{\max(d-r,0)} \in J$ satisfies $\deg(f - g_d) < d$ (or $f - g_d = 0$). By decreasing induction on d we obtain, after at most $d+1$ steps, $g \in J$ such that $f - g = 0$, hence $f \in J$.

(3.9) Corollary. *If R is a noetherian ring, so is $R[x_1, \ldots, x_n]/I$, for any ideal I of $R[x_1, \ldots, x_n]$.*

Proof. Combine Theorem 3.8 with Proposition 3.7.

(3.10) Exercise. *If R is a noetherian ring, so is the power series ring $R[[X]]$.*

(3.11) Proposition. *In an noetherian ring R, every ideal I contains a suitable product $P_1 \cdots P_r$ ($r \geq 0$) of prime ideals (in particular, the zero ideal (0) is a product of prime ideals).*

Proof. Let S be the set of all ideals of R which do not contain any product of prime ideals. If S is non-empty, then it contains a maximal element I. By definition of S, I is not a prime ideal, hence there exist $x, x' \in R$ such that $x, x' \notin I$ and $xx' \in I$. As $I \subsetneq J = I + (x)$ and $I \subsetneq J' = I + (x')$, there exist prime ideals P_i, P'_j such that $J \supset P_1 \cdots P_r$ and $J' \supset P'_1 \cdots P'_s$, which implies that $I \supset JJ' \supset P_1 \cdots P_r P'_1 \cdots P'_s$. This contradiction shows that S is empty, as claimed.

(3.12) Proposition. *The condition I.5.6(i) is satisfied in any noetherian integral domain A.*

Proof. As in the proof of Theorem I.5.9, if I.5.6(i) does not hold, then there exists an infinite chain of principal ideals $(a_1) \subsetneq (a_2) \subsetneq (a_3) \subsetneq \cdots \subseteq A$ contradicting the noetherian assumption.

(3.13) Corollary. *A noetherian integral domain A is a UFD $\iff$ the principal ideal generated by any irreducible element of A is a prime ideal.*

Proof. Combine Proposition I.5.7 with Proposition 3.12.

4. Finitely generated modules over principal ideal domains

(4.1) Matrices. Let R be a ring. If we write elements of R^n as column vectors with entries in R, then a module homomorphism $f : R^m \longrightarrow R^n$ can be identified with a matrix $A \in M_{n \times m}(R)$: $f(x) = Ax$ for any $x \in R^m$.

The morphism f is invertible $\iff$ $m = n$ (by Proposition 1.6) and the matrix A is invertible, i.e., there exists a matrix $B \in M_n(R)$ such that $AB = BA = I_n$. The latter condition implies that $\det(A) \in R^*$. Conversely, the identity

$$A \cdot \mathrm{adj}(A) = \mathrm{adj}(A) \cdot A = \det(A) I_n$$

("Cramer's rule for solving a system of linear equations") satisfied by the adjoint matrix $\mathrm{adj}(A) \in M_n(R)$ (where $(-1)^{i+j}\mathrm{adj}(A)_{ij}$ is the determinant of the matrix obtained by removing from A the i-th column and the j-th row) shows that A is invertible if $\det(A) \in R^*$, with inverse $B = \det(A)^{-1}\mathrm{adj}(A)$. To sum up, isomorphisms $R^n \xrightarrow{\sim} R^n$ correspond to matrices in $GL_n(R) = \{A \in M_n(R) \mid \det(A) \in R^*\}$.

(4.2) Submodules of R^n. Let R be a noetherian ring, let X be a free R-module of finite rank n. Any submodule $Y \subset X$ is finitely generated, by Proposition 3.6. Choose an isomorphism $\alpha : X \xrightarrow{\sim} R^n$ (i.e., a basis of X) and a surjective homomorphism $R^m \longrightarrow Y$ (i.e., a system of m generators of Y). The composite homomorphism (where i is the inclusion of Y into X)

$$A = \alpha \circ i \circ \beta : R^m \longrightarrow Y \hookrightarrow X \xrightarrow{\sim} R^n, \qquad A \in M_{n \times m}(R)$$

is identified with a matrix A, as in 4.1. In concrete terms, the columns of A are the coordinates of the fixed set of generators of Y in the fixed basis of X. A choice of another basis of X is equivalent to replacing α by $\alpha' = P \circ \alpha$ for $P \in GL_n(R)$. Similarly, if we replace β by $\beta' = \beta \circ Q$ with $Q \in GL_m(R)$, then we obtain another system of m generators of Y. The matrix A is then replaced by

$$A' = PAQ, \qquad P \in GL_n(R), \qquad Q \in GL_m(R). \tag{4.2.1}$$

(4.3) If $R = K$ is a field, the Gauss elimination method yields, for any $A \in M_{n\times m}(K)$, matrices P and Q (obtained as a composition of elementary row and column operations, respectively) such that

$$A' = PAQ = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}, \qquad r = \mathrm{rk}(A).$$

It was shown in the course Algebra 1 that a variant of this method works for $R = \mathbf{Z}$ (in fact, for any euclidean ring) and yields, for any $A \in M_{n\times m}(\mathbf{Z})$, matrices $P \in GL_n(\mathbf{Z})$ and $Q \in GL_m(\mathbf{Z})$ such that

$$A' = PAQ = \begin{pmatrix} d_1 & 0 & \cdots & 0 & \cdots\cdots & 0 \\ 0 & d_2 & \cdots & 0 & \cdots\cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \cdots\cdots & \vdots \\ 0 & 0 & \cdots & d_r & \cdots\cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots\cdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots\cdots & 0 \end{pmatrix}, \tag{4.3.1}$$

where $d_1 \mid \cdots \mid d_r$ are positive integers depending only on A (and $0 \le r \le \min(m,n)$). The structure theorems for subgroups of $\mathbf{Z}^n$ and for finitely generated abelian groups are immediate consequences of this statement.

We are going to show that an analogue of (4.3.1) holds if R is an arbitrary PID (but the matrices P and Q cannot be expressed as products of elementray operations, in general).

(4.4) Theorem. *Let $A \in M_{n\times m}(R)$, where R is a PID.*
(i) There exist matrices $P \in GL_n(R)$ and $Q \in GL_m(R)$ such that $A' = PAQ$ is of the form (4.3.1), where $d_1, \ldots, d_r$ are non-zero elements of R such that $d_1 \mid \cdots \mid d_r$ (and $0 \le r \le \min(m,n)$).
(ii) The integer r and the ideals $(d_1), \ldots, (d_r)$ depend only on A: r is the rank of A (considered as a matrix with entries in the fraction field of R) and $d_1 \cdots d_k$ is the greatest common divisor of all $k \times k$ minors of A.

Proof. (i) We can assume that $A \neq 0$. By induction, it is enough to transform A into

$$\begin{pmatrix} d_1 & 0 \\ 0 & d_1\, B \end{pmatrix}, \qquad B \in M_{n-1,m-1}(R) \tag{$\star$}$$

by applying row operations $A \mapsto gA$, $g \in GL_n(R)$ (resp. column operations $A \mapsto Ah$, $h \in GL_m(R)$). In particular, we can permute the rows (resp. the columns).

We use the following observation: if a column C (resp. a row L) of A contains $a, b \in R \setminus \{0\}$, then there exists a row (resp. a column) operation which replaces the couple a, b by $d, 0$ ($d = \gcd(a,b)$), but which does not change the remaining elements of C (resp. of L).

Indeed, there exist $u, v \in R$ such that $au + bv = d$ ($\Longrightarrow \gcd(u,v) = 1 \Longrightarrow$ there exist $u', v' \in R$ such that $\begin{vmatrix} u & v \\ u' & v' \end{vmatrix} = 1$), which yields the following row operations

$$\begin{pmatrix} u & v \\ u' & v' \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} d \\ d' \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 \\ -d'/d & 1 \end{pmatrix} \begin{pmatrix} d \\ d' \end{pmatrix} = \begin{pmatrix} d \\ 0 \end{pmatrix}$$

($d' = au' + bv'$ is divisible by d).

This observation allows us (after permuting the columns) to replace the first column by

$$C_1 = \begin{pmatrix} d \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad d \neq 0.$$

If d divides all elements of A, it is easy to obtain $(\star)$ by subtracting from each column a multiple of C_1. If d does not divide all elements of A, we must distinguish two cases:
(a) there exists an element of the first row L_1 which is not divisible by d;
(b) d divides all elements of L_1, but does not divide an element of the i-th row L_i.
After replacing L_1 by $L_1 + L_i$ we can assume (a). After applying the above observation to L_1 and then to the first column, we replace C_1 by

$$C_1' = \begin{pmatrix} d' \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

where $d' \mid d$ is a **proper divisor** of d ($d/d' \notin R^*$). If d' does not divide all elements of A, we repeat the same procedure. Since d has only finitely many divisors (up to elements of R^*), we obtain $(\star)$ after a finite number of steps.
(ii) We have $r = \mathrm{rk}(A') = \mathrm{rk}(A)$. The set of all $k \times k$ minors of A (up to elements of R^*) does not change if we replace A by $A' = PAQ$. For A' as in (4.3.1) with $d_1 \mid \cdots \mid d_r$ the gcd of the $k \times k$ minors is equal to $d_1 \cdots d_k$, for all $k \leq r$.

(4.5) Theorem on elementary divisors. *Let R be a PID, let X be a free module of rank n over R and $Y \subset X$ a submodule.*
(i) $Y \subset X$ is free of rank $r \leq n$.
(ii) There exist non-zero elements $d_1, \ldots, d_r$ of R such that $d_1 \mid \cdots \mid d_r$ and a basis $e_1, \ldots, e_n$ of X such that $d_1 e_1, \ldots, d_r e_r$ is a basis of Y.
(iii) The quotient module X/Y is isomorphic to $R^{n-r} \oplus R/(d_1) \oplus \cdots \oplus R/(d_r)$. [Of course, if $d_1, \ldots, d_i \in R^*$, then $R/(d_1) \oplus \cdots \oplus R/(d_i) = 0$ and we can remove these terms.]
(iv) The ideals $(d_1), \ldots, (d_r)$ depend only on the pair $Y \subset X$.

Proof. A choice of a basis of X and of a (finite) system of generators of Y yields a matrix $A \in M_{n \times m}(R)$, as in 4.2. Theorem 4.4(i) implies that there is another basis $e_1, \ldots, e_n$ of X and non-zero elements $d_1, \ldots, d_r$ of R such that $d_1 \mid \cdots \mid d_r$ (and $r \leq n$) for which Y is generated by $d_1 e_1, \ldots, d_r e_r$. There is no R-linear relation between these r elements of X, hence they form a basis of Y (which is then free of rank r). This proves (i) and (ii). The statement (iii) is an immediate consequence of (ii). The remaining statement (iv) follows from Theorem 4.6 below (the reader may check that our reasoning is not circular).

(4.6) Theorem. *Let R be a PID, let M be a finitely generated module over R. There is an isomorphism $M \xrightarrow{\sim} R^a \oplus R/(d_1) \oplus \cdots \oplus R/(d_r)$, where $d_1, \ldots, d_r$ ($r \geq 0$) are non-zero non-invertible elements of R such that $d_1 \mid \cdots \mid d_r$. The integer $a \in \mathbf{N}$ and the ideals $(d_1), \ldots, (d_r)$ depend only on the isomorphism class of M. [One can use 5.11(i) to further decompose each term $R/(d_i)$.]*

Proof. By assumption, there exists a surjective homomorphism of R-modules $f : X \longrightarrow M$, where X is a free R-module of finite rank. Thus $M \xrightarrow{\sim} X/Y$, where $Y = \mathrm{Ker}(f) \subset X$. Applying Theorem 4.5(iii) we obtain the desired isomorphism.

We have $M_{\mathrm{tors}} \xrightarrow{\sim} R/(d_1) \oplus \cdots \oplus R/(d_r)$ and $M/M_{\mathrm{tors}} \xrightarrow{\sim} R^a$, which implies that a depends only on the isomorphism class of the R-module M (by Proposition 1.6). In order to prove the uniqueness of the ideals $(d_1), \ldots, (d_r)$ we must analyse, for each irreducible element x of R, the sequence of exponents

$$0 \leq v_x(d_1) \leq \cdots \leq v_x(d_r) \tag{4.6.1}$$

with which x occurs in $d_1, \ldots, d_r$. The goal is to show that they are determined by M_{tors}. For $k \geq 0$ set $r_k(x) = |\{i \mid v_x(d_i) \geq k\}|$ (in particular, $r_0(x) = r$). In other words, the sequence (4.6.1) contains each integer $k \geq 0$ with multiplicity $r_k(x) - r_{k+1}(x)$. The sequence of submodules

$$M[x] \supset xM[x^2] \supset x^2 M[x^3] \supset \cdots \tag{4.6.2}$$

is isomorphic to

$$(R/(x))^{r_1(x)} \supset (R/(x))^{r_2(x)} \supset (R/(x))^{r_3(x)} \supset \cdots, \tag{4.6.3}$$

which implies that the non-zero exponents in (4.6.1) depend only on the isomorphism class of M_{tors}.

As d_1 is not invertible, it is divisible by some irreducible element x_0, which means that $r = r_1(x_0) = \max(r_1(x))$. This determines the value of r, hence also the number of exponents equal to zero in (4.6.1), namely $r - r_1(x)$. Theorem is proved.

(4.7) In the special case $R = \mathbf{Z}$, Theorems 4.4–4.6 were proved in the course Algebra 1.

We are now going to discuss another important special case $R = K[T]$, where K is a field. An R-module M is the same thing as a K-vector space M equipped with a K-linear endomorphism $f : M \longrightarrow M$ (= the action of T). A general polynomial $P(T) \in K[T]$ acts on M as $P(f)$.

If $M \neq 0$ is finite-dimensional over K, then it is finitely generated as a $K[T]$-module. According to Theorem 4.6 it is isomorphic to a direct sum

$$M \xrightarrow{\sim} K[T]/(P_1) \oplus \cdots K[T]/(P_r), \tag{4.7.1}$$

where $r \geq 0$ and $P_1 \mid \cdots \mid P_r$ are non-constant monic polynomials (the integer a in Theorem 4.6 is equal to zero, since $R = K[T]$ has infinite dimension over K). We have

$$\sum_{i=1}^{r} \deg(P_i) = \sum_{i=1}^{r} \dim_K(K[T]/(P_i)) = \dim_K(M) = m.$$

We can choose a K-basis of M and consider $f \in M_m(K)$ as a matrix. The uniqueness statement in Theorem 4.6 then says that two matrices $f, f' \in M_m(K)$ give rise to the same sequence of polynomials $P_1 \mid \cdots \mid P_r$ $\iff$ $f' = gfg^{-1}$ for some $g \in GL_m(K)$.

The formula (4.7.3) below implies that the characteristic polynomial of f is equal to $\det(X \cdot I - f) = P_1 \cdots P_r$. On the other hand, the minimal polynomial of f (the monic polynomial $P \in K[T]$ of smallest degree such that $P(T)M = 0$) is equal to P_r. In particular,

$$M \xrightarrow{\sim} K[T]/(P_1) \iff \text{the minimal and the characteristic polynomials of } f \text{ coincide.} \tag{4.7.2}$$

Let us consider a special case when $M \neq 0$ is a cyclic R-module (which we assume until the end of 4.7): $M \xrightarrow{\sim} K[T]/(Q)$ for some non-constant monic polynomial $Q = T^d + a_{d-1}T^{d-1} + \cdots + a_0$ ($d \geq 1$) and $1, T, \ldots, T^{d-1} \pmod{Q}$ is a basis of M over K. The matrix of $f : M \longrightarrow M$ (= multiplication by T) in this basis is equal to

$$C(Q) = \begin{pmatrix} 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & \ldots & 0 & -a_1 \\ 0 & 1 & \ldots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -a_{d-1} \end{pmatrix}$$

and the characteristic polynomial of f is equal to

$$\det(X \cdot I - C(Q)) = Q. \tag{4.7.3}$$

Write $Q = Q_1^{n_1} \cdots Q_k^{n_k}$, where $n_i \geq 1$ and Q_i are distinct irreducible non-constant monic polynomials; then

$$M \xrightarrow{\sim} K[T]/(Q_1^{n_1}) \oplus \cdots \oplus K[T]/(Q_k^{n_k})$$

and the matrix of f in the union of the bases $1, T, \ldots, T^{n_i \deg(Q_i)-1} \pmod{Q_i^{n_i}}$ is given by a block matrix

$$\begin{pmatrix} C(Q_1^{n_1}) & 0 & \cdots & 0 \\ 0 & C(Q_2^{n_2}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & C(Q_k^{n_k}) \end{pmatrix}$$

If the field K is algebraically closed, then $Q_i = T - \lambda_i$ for some $\lambda_i \in K$. In the basis $1, (T-\lambda_i), \ldots, (T-\lambda_i)^{n_i-1} \pmod{(T-\lambda_i)^{n_i}}$ multiplication by T has matrix

$$\begin{pmatrix} \lambda_i & 0 & 0 & \cdots & 0 & 0 \\ 1 & \lambda_i & 0 & \cdots & 0 & 0 \\ 0 & 1 & \lambda_i & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i & 0 \\ 0 & 0 & 0 & \cdots & 1 & \lambda_i \end{pmatrix}$$

and we recover from Theorem 4.6 the existence and uniqueness of the Jordan normal form of a matrix over an algebraically closed field.

(4.8) Exercise. *For any square matrix $A \in M_n(\mathbf{Z})$, the index of the subgroup $A\mathbf{Z}^n \subset \mathbf{Z}^n$ (= the subgroup generated by the columns of A) is equal to*

$$(\mathbf{Z}^n : A\mathbf{Z}^n) = \begin{cases} \infty, & \det(A) = 0 \\ |\det(A)|, & \det(A) \neq 0. \end{cases}$$

(4.9) Exercise (Subgroups of finite index of $\mathbf{Z}^2$). *For an abelian group A and an integer $n \geq 1$ denote by $S(A,n)$ the set of all subgroups $X \subset A$ of index $(A:X) = n$.*
(i) If $(A:X) = n$, then $nA \subset X$.
(ii) There is a natural bijection between $S(A,n)$ and $S(A/nA, n)$.
(iii) If $m \geq 1$ and $\gcd(m,n) = 1$, then there is a natural bijection between $S((\mathbf{Z}/mn\mathbf{Z})^N, mn)$ and $S((\mathbf{Z}/m\mathbf{Z})^N, m) \times S((\mathbf{Z}/n\mathbf{Z})^N, n)$, for all $N \geq 1$.
(iv) $S(\mathbf{Z}^2, 2)$ has three elements, namely

$$\mathbf{Z}\begin{pmatrix}1\\0\end{pmatrix} \oplus \mathbf{Z}\begin{pmatrix}0\\2\end{pmatrix}, \qquad \mathbf{Z}\begin{pmatrix}2\\0\end{pmatrix} \oplus \mathbf{Z}\begin{pmatrix}0\\1\end{pmatrix}, \qquad \mathbf{Z}\begin{pmatrix}2\\0\end{pmatrix} \oplus \mathbf{Z}\begin{pmatrix}1\\1\end{pmatrix}.$$

(v) For each positive divisor $a \mid n$ there exist exactly a elements $X \in S(\mathbf{Z}^2, n)$ such that $X \cap (\mathbf{Z} \oplus 0) = a\mathbf{Z} \oplus 0$; describe them explicitly. Deduce that

$$|S(\mathbf{Z}^2, n)| = \sum_{a|n} a$$

(vi) Give an explicit formula for the generating series

$$\sum_{r=0}^{\infty} |S(\mathbf{Z}^2, p^r)|\, T^r, \qquad \sum_{n=1}^{\infty} |S(\mathbf{Z}^2, n)|\, n^{-s},$$

where p is a prime number.
(vii) What happens if $\mathbf{Z}^2$ is replaced by $\mathbf{Z}^3$ (or by $\mathbf{Z}^m$)?

(4.10) Exercise (Hecke operators and the Bruhat-Tits tree). *Let $X = \mathbf{Z}^2$, let p be a prime number.*
(1) There exist precisely $p+1$ subgroups $Y \subset X$ of index $(X:Y) = p$ ($\iff X/Y \xrightarrow{\sim} \mathbf{Z}/p\mathbf{Z}$). [Hint:

consider $Y/pX \subset X/pX$.]
(2) If $Y \subset X$ is a subgroup of index $(X : Y) = p^n$ for some $n \geq 0$, then there exists a unique integer $a \geq 0$ such that $Y = p^a Y'$ and X/Y' is a cyclic group (in fact, isomorphic to $\mathbf{Z}/p^{n-2a}\mathbf{Z}$).
(3) Consider the graph with vertices $\{Y$ subgroup of $X \mid X/Y \xrightarrow{\sim} \mathbf{Z}/p^m\mathbf{Z}$ for some $m \geq 0\}$, with vertices Y, Y' joined by an edge $\iff Y/Y' \xrightarrow{\sim} \mathbf{Z}/p\mathbf{Z}$ or $Y'/Y \xrightarrow{\sim} \mathbf{Z}/p\mathbf{Z}$. Show that this graph is an infinite tree, in which each vertex has degree $p + 1$.
(4) For each $n \geq 1$ determine the number of subgroups $Y \subset X$ such that $X/Y \xrightarrow{\sim} \mathbf{Z}/p^n\mathbf{Z}$ (resp. such that $(X : Y) = p^n$).
(5) Let A be the free abelian group on symbols $[Y]$, where $Y \subset X$ runs through all subgroups of X of index p^m (for all possible values of $m \in \mathbf{N}$). For each $n \geq 0$ consider the homomorphisms of abelian groups $T(p^n), S(p^n) : A \longrightarrow A$ defined on the basis elements as follows.

$$T(p^n) : [Y] \mapsto \sum_{(Y:Y')=p^n} [Y'], \qquad S(p^n) = S(p)^n : [Y] \mapsto [p^n Y].$$

Show that

$$\forall n \geq 1 \qquad T(p)T(p^n) = T(p^{n+1}) + pS(p)T(p^{n-1}).$$

(6) Deduce that all homomorphisms $T(p^n)$ $(n \geq 0)$ commute with each other (and with $S(p)$) and that there is an equality of formal generating series

$$\sum_{n=0}^{\infty} T(p^n)u^n = (1 - T(p)u + pS(p)u^2)^{-1},$$

where u is a formal variable.
(7)** What happens for subgroups of $\mathbf{Z}^3$?

III. Field extensions and Galois theory

1. Solving equations of degree 2, 3 and 4

(1.1) Consider a polynomial equation

$$x^n + a_1 x^{n-1} + \cdots + a_n = 0 \tag{1.1.1}$$

of degree $n \geq 1$with complex coefficients $a_1, \ldots, a_n \in \mathbf{C}$. According to a fundamental result of Gauss, the equation (1.1.1) has n complex roots $x_1, \ldots, x_n \in \mathbf{C}$ (not necessarily distinct) for which the polynomial (1.1.1) splits as

$$x^n + a_1 x^{n-1} + \cdots + a_n = (x - x_1) \cdots (x - x_n) \tag{1.1.2}$$

in $\mathbf{C}[x]$. After comparing the coefficients of both sides of (1.1.2) we obtain

$$a_1 = -\sigma_1, \qquad a_2 = \sigma_2 \qquad \ldots \qquad a_n = (-1)^n \sigma_n, \tag{1.1.3}$$

where

$$\begin{aligned} \sigma_1 &= x_1 + \cdots + x_n = \sum_i x_i \\ \sigma_2 &= x_1 x_2 + x_1 x_3 + \cdots + x_{n-1} x_n = \sum_{i<j} x_i x_j \\ &\cdots \\ \sigma_k &= \sum_{1 \leq i_1 < \cdots < i_k \leq n} x_{i_1} \cdots x_{i_k} \\ &\cdots \\ \sigma_n &= x_1 \cdots x_n \end{aligned} \tag{1.1.4}$$

are the **elementary symmetric functions** of the roots $x_1, \ldots, x_n$. In other words, solving the equation (1.1.1) amounts to solving the system of equations (1.1.4) for $\sigma_k = (-1)^k a_k$.

The expressions (1.1.4) are symmetric functions of the roots $x_1, \ldots, x_n$ in the sense that they do not change if we permute the roots. A fundamental idea of Lagrange was to try to solve (1.1.4) by breaking this symmetry.

In this section we recall the classical approach to solving equations of degree $n \leq 4$. After that we reformulate it using Lagrange's idea of resolvents and show that this method does not permit to solve equations of degree $n \geq 5$. The history of the subject can be found in [Ti 2].

(1.2) Quadratic equations. A quadratic equation

$$x^2 + px + q = 0, \tag{1.2.1}$$

is solved by completing the square:

$$0 = \left(x + \frac{p}{2}\right)^2 + q - \left(\frac{p}{2}\right)^2.$$

Equivalently, one can write $x = u + v$, which yields

$$0 = (u+v)^2 + p(u+v) + q = u^2 + u(2v+p) + v^2 + pv + q. \tag{1.2.2}$$

For v such that $2v + p = 0$ the equation (1.2.2) simplifies as

$$0 = u^2 + \left(-\frac{p}{2}\right)^2 + p\left(-\frac{p}{2}\right) + q = u^2 + q - \left(\frac{p}{2}\right)^2. \tag{1.2.3}$$

The solutions x_1, x_2 of (1.2.1) can be written in terms of the solutions $u_1, u_2 = -u_1$ of (1.2.3) as

$$x_1 = -\frac{p}{2} + u_1, \qquad x_2 = -\frac{p}{2} + u_2 = -\frac{p}{2} - u_1. \tag{1.2.4}$$

Conversely,

$$2u_1 = x_1 - x_2, \qquad 2u_2 = x_2 - x_1. \tag{1.2.5}$$

(1.3) Cubic equations. A general cubic equation

$$x^3 + ax^2 + bx + c = 0$$

can be transformed to a simpler equation

$$x^3 + px + q = 0 \tag{1.3.1}$$

by completing the cube, i.e., by replacing $x + a/3$ by x. In order to solve (1.3.1) one writes $x = u + v$ (as in 1.2), which gives

$$0 = (u+v)^3 + p(u+v) + q = u^3 + v^3 + (3uv + p)(u+v) + q. \tag{1.3.2}$$

If $3uv + p = 0$, then the equation (1.3.2) simplifies as

$$u^3 + v^3 + q = 0,$$

hence

$$u^6 + qu^3 - \left(\frac{p}{3}\right)^3 = 0, \qquad v^6 + qv^3 - \left(\frac{p}{3}\right)^3 = 0.$$

In other words, the two roots T_1, T_2 of an auxiliary quadratic equation

$$T^2 + qT - \left(\frac{p}{3}\right)^3 = 0 \tag{1.3.3}$$

are equal to u^3 et v^3 (of course, $T_1 T_2 = (-p/3)^3 = u^3 v^3$).

If $p = 0$, then the roots of 1.3.1 are the cubic roots of $-q$. If $p \neq 0$, then each cubic root u_1, $u_2 = \rho u_1$, $u_3 = \rho^2 u_1$ of $T_1 \neq 0$ (where $\rho = e^{2\pi i/3}$) determines a unique cubic root $v_j = -3p/u_j$ of T_2 ($v_2 = \rho^2 v_1$, $v_3 = \rho v_1$) for which $x_j = u_j + v_j$ is a root of (1.3.1):

$$x_1 = u_1 + v_1, \qquad x_2 = \rho u_1 + \rho^2 v_1, \qquad x_3 = \rho^2 u_1 + \rho v_1, \tag{1.3.4}$$

hence

$$3u_1 = x_1 + \rho^2 x_2 + \rho x_3, \qquad 3v_1 = x_1 + \rho x_2 + \rho^2 x_3. \tag{1.3.5}$$

(1.4) Quartic equations. Let us consider a general quartic equation

$$x^4 + ax^3 + bx^2 + cx + d = 0.$$

As above, it can be transformed to a simpler equation

$$x^4 + px^2 + qx + r = 0 \tag{1.4.1}$$

by replacing $x + a/4$ by x. The trick used in 1.2 and 1.3 ($x = u + v$) does not lead to a significant simplification of (1.4.1). Instead, one can try to write the quartic polynomial in (1.4.1) as a product of two quadratic polynomials:

$$\begin{aligned} x^4+px^2+qx+r &= (x^2+ax+b)(x^2+cx+d) \\ &= (x^2+ax+b)(x^2-ax+\frac{r}{b}) \\ &= x^4+\big(b+\frac{r}{b}-a^2\big)x^2+a\big(\frac{r}{b}-b\big)x+r, \end{aligned} \tag{1.4.2}$$

which is equivalent (under the assumption that $ab\neq 0$; the case $ab=0$ is left to the reader) to

$$\frac{r}{b}+b=p+a^2, \qquad \frac{r}{b}-b=\frac{q}{a} \quad\Longleftrightarrow\quad 2b=p+a^2-\frac{q}{a}, \qquad \frac{2r}{b}=p+a^2+\frac{q}{a},$$

where

$$4r=\big(p+a^2-\frac{q}{a}\big)\big(p+a^2+\frac{q}{a}\big)=(p+a^2)^2-\frac{q^2}{a^2}. \tag{1.4.3}$$

After multiplying (1.4.3) by a^2 we obtain an auxiliary cubic equation for a^2:

$$(a^2)^3+2p(a^2)^2+(p^2-4r)a^2-q^2=0. \tag{1.4.4}$$

Conversely, each root a of (1.4.4) gives rise to a factorisation as in (1.4.2), with $b=\frac{1}{2}(p+a^2-\frac{q}{a})$.

Let us investigate the relationship between the roots $\pm a_1,\pm a_2,\pm a_3$ of (1.4.4) and the roots $x_1,\dots,x_4$ of the original equation (1.4.1). The two quadratic factors

$$\begin{aligned} x^2+ax+b &= (x-x_i)(x-x_j)=x^2-(x_i+x_j)x+x_ix_j \\ x^2+cx+d &= (x-x_k)(x-x_l)=x^2-(x_k+x_l)x+x_kx_l \end{aligned}$$

correspond to a choice of indices such that $\{1,2,3,4\}=\{i,j\}\cup\{k,l\}$ (of course, $x_1+x_2+x_3+x_4=0$, hence $x_k+x_l=-(x_i+x_j)$).

In particular, the three roots a_1^2,a_2^2,a_3^2 of the cubic equation

$$T^3+2pT^2+(p^2-4r)T-q^2=0 \tag{1.4.5}$$

are equal to

$$\begin{aligned} a_1^2 &= (x_1+x_2)^2=-(x_1+x_2)(x_3+x_4)=-y_2-y_3=y_1-p \\ a_2^2 &= (x_1+x_3)^2=-(x_1+x_3)(x_2+x_4)=-y_1-y_3=y_2-p \\ a_3^2 &= (x_1+x_4)^2=-(x_1+x_4)(x_2+x_3)=-y_1-y_2=y_3-p, \end{aligned}$$

where

$$y_1=x_1x_2+x_3x_4, \qquad y_2=x_1x_3+x_2x_4, \qquad y_3=x_1x_4+x_2x_3. \tag{1.4.6}$$

Note that y_1,y_2,y_3 are the roots of the cubic equation

$$0=(y-p)^3+2p(y-p)^2+(p^2-4r)(y-p)-q^2=y^3-py^2-4ry+(4pr-q^2). \tag{1.4.7}$$

In order to determine the roots x_1,x_2,x_3,x_4 one must first solve (1.4.4) and then use the following relations:

$$\begin{gathered} x_1+x_2+x_3+x_4=0 \\ \{\pm a_1,\pm a_2,\pm a_3\}=\{x_1+x_2,x_1+x_3,x_1+x_4,x_2+x_3,x_2+x_4,x_3+x_4\} \end{gathered}$$

The formulas

$$\begin{aligned} (x_1+x_2)+(x_1+x_3)+(x_1+x_4) &= 2x_1 \\ (x_1+x_2)+(x_1+x_3)+(x_2+x_3) &= -2x_4 \end{aligned} \tag{1.4.8}$$

show that the sum $\pm a_1 \pm a_2 \pm a_3$ is equal, for any choice of the signs, to $\pm 2x_i$ $(i = 1, \dots, 4)$.

We have $(a_1a_2a_3)^2 = q^2$ and

$$\begin{aligned} -q &= x_1x_2x_3 + (x_1x_2 + x_1x_3 + x_2x_3)x_4 = x_1x_2x_3 - (x_1x_2 + x_1x_3 + x_2x_3)(x_1 + x_2 + x_3) \\ &= -2x_1x_2x_3 - (x_1^2x_2 + \cdots + x_2x_3^2) = (x_1 + x_2)(x_1 + x_3)(x_1 + x_4) = -(x_1 + x_2)(x_1 + x_3)(x_2 + x_3), \end{aligned}$$

which allows us to distinguish between the two cases in (1.4.8): for an arbitrary choice of square roots a_1, a_2, a_3 of the roots $T_j = a_j^2$ of the cubic equation (1.4.5) there exists a root x_i of (1.4.1) such that

$$a_1 + a_2 + a_3 = \begin{cases} 2x_i, & \text{if} \quad a_1a_2a_3 = -q \\ -2x_i, & \text{if} \quad a_1a_2a_3 = q. \end{cases} \tag{1.4.9}$$

(1.5) Let us now reconsider equations of degree $n \leq 4$ from a more scientific point of view. All coefficients $a_k = (-1)^k\sigma_k$ of the equation

$$x^n + a_1x^{n-1} + \cdots + a_n = 0$$

are symmetric functions of the roots $x_1, \dots, x_n$. The idea of Lagrange was to break this symmetry step by step, by considering expressions which are slightly less symmetric.

(1.6) Quadratic equations. In order to solve the system

$$x_1 + x_2 = -a_1, \qquad x_1x_2 = a_2 \tag{1.6.1}$$

we consider the function

$$y = x_1 - x_2, \tag{1.6.2}$$

which is not symmetric in x_1 in x_2, but its square is:

$$y^2 = (x_1 - x_2)^2 = x_1^2 + x_2^2 - 2x_1x_2 = (x_1 + x_2)^2 - 4x_1x_2 = \sigma_1^2 - 4\sigma_2 = a_1^2 - 4a_2, \tag{1.6.3}$$

which gives the standard formulas

$$y = \pm\sqrt{a_1^2 - 4a_2}, \qquad x_1, x_2 = \frac{1}{2}((x_1 + x_2) \pm y) = \frac{1}{2}(-a_1 \pm \sqrt{a_1^2 - 4a_2}). \tag{1.6.4}$$

(1.7) Cubic equations. A general cubic equation

$$x^3 + a_1x^2 + a_2x + a_3 = 0 \tag{1.7.1}$$

is equivalent to

$$\sigma_1 = x_1 + x_2 + x_3 = -a_1, \qquad \sigma_2 = x_1x_2 + x_1x_3 + x_2x_3 = a_2, \qquad \sigma_3 = x_1x_2x_3 = -a_3. \tag{1.7.2}$$

It is natural to try to generalise (1.6.2) by considering "Lagrange's resolvents"

$$\begin{aligned} y_1 &= x_1 + \rho x_2 + \rho^2 x_3 \\ y_2 &= x_1 + \rho^2 x_2 + \rho x_3, \end{aligned} \tag{1.7.3}$$

where

$$\rho = e^{2\pi i/3} = \tfrac{-1+i\sqrt{3}}{2}, \qquad \rho^2 = \rho^{-1} = e^{-2\pi i/3} = \tfrac{-1-i\sqrt{3}}{2} = -1 - \rho$$

are as in 1.3. What are the symmetries of the functions y_1, y_2? If we permute x_1 and x_2 (resp. x_2 and x_3), the transformation rules are

$$y_1 \mapsto x_2 + \rho x_1 + \rho^2 x_3 = \rho y_2, \qquad y_2 \mapsto x_2 + \rho^2 x_1 + \rho x_3 = \rho^2 y_1$$

resp.

$$y_1 \mapsto x_1 + \rho x_3 + \rho^2 x_2 = y_2, \qquad y_2 \mapsto x_1 + \rho^2 x_3 + \rho x_2 = y_1.$$

It follows that the functions $y_1 y_2$ and $y_1^3 + y_2^3$ are symmetric en x_1, x_2, x_3:

$$\begin{aligned} y_1 y_2 &= x_1^2 + x_2^2 + x_3^2 - x_1 x_2 - x_1 x_3 - x_2 x_3 = s_2 - \sigma_2 \\ y_1^3 + y_2^3 &= 2(x_1^3 + x_2^3 + x_3^3) + 12 x_1 x_2 x_3 - 3(x_1^2 x_2 + x_1 x_2^2 + x_1^2 x_3 + x_1 x_3^2 + x_2^2 x_3 + x_2 x_3^2) \\ &= 2s_3 + 12\sigma_3 - 3s_{2,1} \end{aligned} \tag{1.7.4}$$

We wish to express them (equivalently, the functions $s_{2,1}$ and $s_k = x_1^k + x_2^k + x_3^k$ for $k = 2$ and 3) in terms of σ_1, σ_2 and σ_3. This can be done as follows:

$$\begin{aligned} \sigma_1^2 &= (x_1 + x_2 + x_3)^2 = x_1^2 + x_2^2 + x_3^2 + 2(x_1 x_2 + x_1 x_3 + x_2 x_3) = s_2 + 2\sigma_2 \\ \sigma_1 \sigma_2 &= (x_1 + x_2 + x_3)(x_1 x_2 + x_1 x_3 + x_2 x_3) = s_{2,1} + 3 x_1 x_2 x_3 = s_{2,1} + 3\sigma_3 \\ \sigma_1 s_2 &= (x_1 + x_2 + x_3)(x_1^2 + x_2^2 + x_3^2) = (x_1^3 + x_2^3 + x_3^3) + s_{2,1} = s_3 + s_{2,1}, \end{aligned} \tag{1.7.5}$$

which implies that

$$s_2 = \sigma_1^2 - 2\sigma_2, \qquad s_{2,1} = \sigma_1 \sigma_2 - 3\sigma_3, \qquad s_3 = \sigma_1 s_2 - s_{2,1} = \sigma_1^3 - 3\sigma_1 \sigma_2 + 3\sigma_3 \tag{1.7.6}$$

and

$$y_1 y_2 = \sigma_1^2 - 3\sigma_2, \qquad y_1^3 + y_2^3 = 2\sigma_1^3 - 9\sigma_1 \sigma_2 + 27\sigma_3 \tag{1.7.7}$$

(we shall see in Theorem 2.7 below that any symmetric polynomial $F(x_1, \ldots, x_n)$ can be written in terms of $\sigma_1, \ldots, \sigma_n$).

To sum up, the cubes y_1^3, y_2^3 are the roots of an auxiliary quadratic equation

$$(t - y_1^3)(t - y_2^3) = t^2 - (2\sigma_1^3 - 9\sigma_1 \sigma_2 + 27\sigma_3)t + (\sigma_1^2 - 3\sigma_2)^3 = 0 \tag{1.7.8}$$

and

$$y_1 y_2 = \sigma_1^2 - 3\sigma_2. \tag{1.7.9}$$

The roots of the original equation (1.7.1) are given by

$$3x_1 = y_1 + y_2 + \sigma_1, \qquad 3x_2 = \rho^2 y_1 + \rho y_2 + \sigma_1, \qquad 3x_3 = \rho y_1 + \rho^2 y_2 + \sigma_1. \tag{1.7.10}$$

In the special case of the equation (1.3.1) we recover the formulas (1.3.4) and (1.3.5).

(1.8) Quartic equations. Consider the simplified quartic equation (1.4.1)

$$x^4 + px^2 + qx + r = 0, \tag{1.8.1}$$

for which

$$\sigma_1 = 0, \qquad \sigma_2 = p, \qquad \sigma_3 = -q, \qquad \sigma_4 = r. \tag{1.8.2}$$

Let us try to generalise (1.7.3). A natural guess would be to consider the following linear expressions

$$x_1 + ix_2 - x_3 - ix_4, \qquad x_1 - x_2 + x_3 - x_4, \qquad x_1 - ix_2 - x_3 + ix_4. \tag{1.8.3}$$

We must study their behaviour under all permutations of the roots $x_1, \ldots, x_4$.

The first and the third polynomial in (1.8.3) do not behave well, even if we raise them to the fourth power: they give rise to six different expressions. On the other hand, the second polynomial in (1.8.3) works: the set of the following three polynomials

$$u_1 = x_1 + x_2 - x_3 - x_4, \qquad u_2 = x_1 + x_3 - x_2 - x_4, \qquad u_3 = x_1 + x_4 - x_2 - x_3$$

is preserved – up to a sign – by arbitrary permutations of the roots. As a result, the coefficients of an auxiliary cubic polynomial

$$(u - u_1^2)(u - u_2^2)(u - u_3^2) \tag{1.8.4}$$

are symmetric in x_1, x_2, x_3, x_4 and one can write them down explicitly by a calculation similar to that in (1.7.5). The resulting auxiliary cubic equation is given, up to rescaling the variable, by the formula (1.4.4), since $u_j^2 = 4a_j^2$ for all $j = 1, 2, 3$.

Alternatively, one can consider the cubic polynomial whose roots are given by the expressions (1.4.6)

$$y_1 = x_1x_2 + x_3x_4 = p + (u_1/2)^2, \qquad y_2 = x_1x_3 + x_2x_4 = p + (u_2/2)^2, \qquad y_3 = x_1x_4 + x_2x_3 = p + (u_3/2)^2.$$

These three expressions are permuted under arbitrary permutations of the roots x_1, x_2, x_3, x_4 and we can again compute the coefficients of

$$(y - y_1)(y - y_2)(y - y_3)$$

in terms of p, q and r as in (1.7.5), arriving at the formula (1.4.7).

(1.9) The general mechanism should now be clear. Given an equation (1.1.1), the goal is to choose in an intelligent way an auxiliary polynomial $g(x_1, \dots, x_n) = g_1$ in the roots $x_1, \dots, x_n$ (a "resolvent") which will be a root of a new polynomial equation (the "resolvent equation")

$$(y - g_1) \cdots (y - g_d) = 0, \tag{1.9.1}$$

hopefully simpler than the original equation. The distinct roots of (1.9.1) are obtained from g by applying to it all possible permutations of the roots $x_1, \dots, x_n$ (but keeping each expression obtained in this way only once, disregarding multiplicities with which they appear). The coefficients of (1.9.1) are symmetric under all permutations of $x_1, \dots, x_n$, and therefore expressible in terms of the coefficients of the original equation (1.1.1).

For $n = 2, 3, 4$ we can take, respectively, $g_1 = (x_1 - x_2)^2, (x_1 + \rho x_2 + \rho x_3)^3$ and $x_1x_2 + x_3x_4$ (or $(x_1 + x_2 - x_3 - x_4)^2$), obtaining resolvent equations for g_1 of respective degrees $d = 1, 2$ and 3.

(1.10) Question. What happens for $n \geq 5$?

2. Symmetric functions and resolvents

In this section we first show that any polynomial in $x_1, \dots, x_n$ which is symmetric (in the sense of 2.3 below) can be expressed as a polynomial in $\sigma_1, \dots, \sigma_n$. After that we briefly discuss Lagrange's theory of resolvents and answer Question 1.10.

(2.1) The symmetric group. Recall that a **permutation** of a set X is a bijection $\sigma : X \longrightarrow X$. The permutations of X form a group S_X with respect to composition $\sigma\tau = \sigma \circ \tau$, $(\sigma\tau)(x) = \sigma(\tau(x))$. For an integer $n \geq 1$ the **symmetric group** S_n is defined to be $S_n = S_X$ for $X = \{1, \dots, n\}$. The **sign** of a permutation $\sigma \in S_n$ ($n \geq 2$) is defined as

$$\operatorname{sgn}(\sigma) = (-1)^{|\{(i,j)\,|\,1 \leq i < j \leq n,\, \sigma(i) > \sigma(j)\}|}.$$

The map

$$\operatorname{sgn} : S_n \longrightarrow \{\pm 1\}$$

is a group homomorphism. Its kernel is the **alternating group** $A_n = \mathrm{Ker}(\mathrm{sgn}) \subset S_n$. We have $|S_n| = n!$, $|A_n| = n!/2$.

There are two types of notation for elements of S_n. One can write $\sigma \in S_n$ either as

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix},$$

or as a product of disjoint cycles (= orbits under the action of σ on $\{1, \ldots, n\}$). For example, the element

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 5 & 3 & 6 & 2 & 1 \end{pmatrix} \in S_6$$

is a product of the following disjoint cycles:

$$1 \mapsto 4 \mapsto 6 \mapsto 1, \qquad 2 \mapsto 5 \mapsto 2, \qquad 3 \mapsto 3,$$

hence

$$\sigma = (146)(25)(3).$$

(2.2) Definition (Action of S_n on polynomials). *Let R be a ring (commutative and unital, as usual). The polynomial ring $R[x_1, \ldots, x_n]$ has a natural left action of S_n, given by*

$$(\sigma \cdot f)(x_1, \ldots, x_n) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) \qquad (\sigma \in S_n,\ f \in R[x_1, \ldots, x_n]).$$

If $R = K$ is a field, the same formula defines an action of S_n on the field of rational functions $K(x_1, \ldots, x_n) = \mathrm{Frac}(K[x_1, \ldots, x_n])$.

(2.3) Definition. *A polynomial $f \in R[x_1, \ldots, x_n]$ (resp. a rational function $f \in K(x_1, \ldots, x_n)$, if $R = K$ is a field) is* **symmetric** *if $\forall \sigma \in S_n \quad \sigma \cdot f = f$. They form a subring $R[x_1, \ldots, x_n]^{S_n}$ (resp. a subfield $K(x_1, \ldots, x_n)^{S_n}$) of $R[x_1, \ldots, x_n]$ (resp. of $K(x_1, \ldots, x_n)$).*

(2.4) Example. The polynomials $x_1x_2x_3$ and $x_1^7 + x_2^7 + x_3^7 \in R[x_1, x_2, x_3]$ are symmetric, but $x_1^2x_2 + x_2^2x_3 + x_3^2x_1$ is not.

(2.5) Symmetrised monomials. For any set $I = (i_1, \ldots, i_n)$ of integers $i_1 \geq i_2 \geq \cdots \geq i_n \geq 0$ we define

$$s_I = s_{i_1, \ldots, i_n} = \sum_{f \in A_I} f \in R[x_1, \ldots, x_n]^{S_n}, \qquad \text{where} \qquad A_I = \{\sigma \cdot (x_1^{i_1} \cdots x_n^{i_n}) \,|\, \sigma \in S_n\}.$$

We often omit the values $i_k = 0$. For example,

$$s_1 = \sigma_1, \qquad s_{1,1} = \sigma_2, \qquad s_{1,1,1} = \sigma_3, \qquad \ldots$$

are the elementary symmetric functions from (1.1.4) and

$$s_k = x_1^k + \cdots + x_n^k, \qquad s_{2,1} = x_1^2x_2 + x_1x_2^2 + x_1^2x_3 + x_1x_3^2 + \cdots + x_{n-1}^2x_n + x_{n-1}x_n^2.$$

For $I = (i_1, \ldots, i_n)$ and $J = (j_1, \ldots, j_n)$ (where $i_1 \geq i_2 \geq \cdots \geq i_n \geq 0$ and $j_1 \geq j_2 \geq \cdots \geq j_n \geq 0$) we define

$$I + J = (i_1 + j_1, \ldots, i_n + j_n).$$

If $I \neq J$, we say that $I < J$ (resp. $I > J$) if $i_1 = j_1, \ldots, i_k = j_k$ and $i_{k+1} < j_{k+1}$ (resp. and $i_{k+1} > j_{k+1}$), $0 \leq k < n$.

(2.6) Proposition. *(i) Each symmetric polynomial $f \in R[x_1, \dots, x_n]^{S_n}$ can be written as a finite sum $f = \sum c_I s_I$, $c_I \in R$.*
(ii) For every I, J as in (2.5) we have

$$s_I s_J = s_{I+J} + \sum_{K<I+J} c_K s_K \qquad (c_K \in R).$$

Proof. (i) If f contains a monomial $cx_1^{i_1} \cdots x_n^{i_n}$ (with exponents not necessarily ordered), it contains all monomials $cx_{\sigma(1)}^{i_1} \cdots x_{\sigma(n)}^{i_n}$.
(ii) This is a general version of (1.7.5).

(2.7) Theorem on symmetric functions. *We have $R[\sigma_1, \dots, \sigma_n] = R[x_1, \dots, x_n]^{S_n}$ and there is no polynomial relation between the elements $\sigma_1, \dots, \sigma_n$. In other words, every symmetric polynomial over R can be written in a unique way as a polynomial over R in $\sigma_1, \dots, \sigma_n$.*

Proof. Existence: one can generalise the calculations in (1.7.5), as follows. Thanks to Proposition 2.6(i) it is enough to show that $(\forall I)\ s_I \in R[\sigma_1, \dots, \sigma_n]$. As $s_{0,\dots,0} = 1$, we can assume, by induction, that $I = (i_1 = \cdots i_k > i_{k+1} \geq \cdots \geq i_n \geq 0\ (1 \leq k \leq n)$ and that we already know that $s_K \in R[\sigma_1, \dots, \sigma_n]$ for all $K < I$. We can write $I = I' + J$, where $I' = (1, \dots, 1, 0, \dots, 0)$ (and 1 appears k times). It follows from Proposition 2.6(ii) that

$$s_I = \sigma_k s_J + \sum_{K<I} c_K s_K \qquad (c_K \in R),$$

which lies in $R[\sigma_1, \dots, \sigma_n]$, by the induction hypothesis.
Uniqueness: for every set of exponents $A = (a_1, \dots, a_n)$, $a_1, \dots, a_n \geq 0$, we have

$$\sigma^A := \sigma_1^{a_1} \cdots \sigma_n^{a_n} = s_I + \sum_{J<I} c_J s_J, \qquad I = I(A) = (a_1 + \cdots + a_n, a_2 + \cdots + a_n, \dots, a_n), \qquad (c_J \in R). \quad (2.7.1)$$

Let

$$g(y_1, \dots, y_n) = \sum g_{a_1,\dots,a_n} y_1^{a_1} \cdots y_n^{a_n} = \sum_A g_A\, y^A$$

be a non-zero polynomial. The set $\{I(A) \,|\, g_A \neq 0\}$ contains the biggest element (necessarily unique!) with respect to the order "$<$" $I = I(A)$ (for a unique valued of A, $g_A \neq 0$). It follows from de (2.7.1) that $g(\sigma_1, \dots, \sigma_n)$ contains the monomial $g_A s_I$, hence $g(\sigma_1, \dots, \sigma_n) \neq 0$.

(2.8) Example. Let us write $s_{2,2} = x_1^2 x_2^2 + x_1^2 x_3^2 + \cdots + x_{n-1}^2 x_n^2$ in terms of $\sigma_1, \dots, \sigma_n$. As

$$\begin{gathered}\sigma_2^2 = s_{1,1}^2 = (x_1x_2 + x_1x_3 + \cdots + x_{n-1}x_n)^2 = (x_1^2x_2^2 + \cdots) + 2(x_1^2x_2x_3 + \cdots) + 6(x_1x_2x_3x_4 + \cdots) = \\ = s_{2,2} + 2s_{2,1,1} + 6\sigma_4, \\ \sigma_1\sigma_3 = s_1 s_{1,1,1} = (x_1 + \cdots + x_n)(x_1x_2x_3 + \cdots) = (x_1^2x_2x_3 + \cdots) + 4(x_1x_2x_3x_4 + \cdots) = \\ = s_{2,1,1} + 4\sigma_4,\end{gathered}$$

we have

$$s_{2,1,1} = \sigma_1\sigma_3 - 4\sigma_4, \qquad s_{2,2} = \sigma_2^2 - 2\sigma_1\sigma_3 + 2\sigma_4.$$

(2.9) Corollary. *If K is a field, then $K(\sigma_1, \dots, \sigma_n) = K(x_1, \dots, x_n)^{S_n}$.*

Proof. If $f, g \in K[x_1, \dots, x_n]$ $(g \neq 0)$ and $f/g \in K(x_1, \dots, x_n)^{S_n}$, then

$$h_1 := \prod_{\sigma\in S_n} (\sigma \cdot g) \in K[x_1,\dots,x_n]^{S_n} \setminus \{0\} = K[\sigma_1,\dots,\sigma_n] \setminus \{0\}$$

and

$$h_2 := h_1 f/g = f \prod_{\sigma\in S_n\setminus\{1\}} (\sigma \cdot g) \in K[x_1,\dots,x_n]^{S_n} = K[\sigma_1,\dots,\sigma_n],$$

hence $f/g = h_2/h_1 \in K(\sigma_1,\dots,\sigma_n)$.

(2.10) Exercise (Newton's formulas). *The polynomials $s_k = x_1^k + \cdots + x_n^k$ satisfy recursive relations*

$$s_k - \sigma_1 s_{k-1} + \cdots + (-1)^{k-1}\sigma_{k-1}s_1 + (-1)^k k\sigma_k = 0 \qquad (k \geq 1)$$

(of course, $\sigma_k = 0$ for $k > n$).

(2.11) Discriminant. Let $n \geq 2$. The polynomial

$$\Delta := \prod_{i<j}(x_i - x_j) \in \mathbf{Z}[x_1,\dots,x_n]$$

is not symmetric, since

$$\forall \tau \in S_n \quad \tau \cdot \Delta = \operatorname{sgn}(\tau)\Delta,$$

but its square

$$\Delta^2 = \prod_{i<j}(x_i - x_j)^2 \in \mathbf{Z}[x_1,\dots,x_n]^{S_n} = \mathbf{Z}[\sigma_1,\dots,\sigma_n]$$

is. Writing Δ^2 in terms of the coefficients $a_k = (-1)^k\sigma_k$ of the polynomial

$$f = x^n + a_1x^{n-1} + \cdots + a_n = (x - x_1)\cdots(x - x_n) \in \mathbf{Z}[a_1,\dots,a_n][t],$$

we obtain the **discriminant** $\operatorname{disc}(f) \in \mathbf{Z}[a_1,\dots,a_n]$ of f.

(2.12) Exercise. *(i) Compute the discriminants* $\operatorname{disc}(x^n + ax + b)$ *for $n = 2, 3, 4$. [Hint: relate the discriminant of a cubic (resp. quartic) polynomial to the discriminant of its quadratic (resp. cubic) resolvent (1.7.8) (resp. (1.4.7)).]*
(ii) What happens for general $n \geq 2$?

(2.13) Exercise. *Let K be a field such that $2 \in K^*$. Show that, for any $n \geq 2$, the field of rational functions invariant under A_n is equal to*

$$K(x_1,\dots,x_n)^{A_n} = \{f + g\Delta \mid f, g \in K(\sigma_1,\dots,\sigma_n)\}.$$

(2.14) Resolvents revisited. *Fix a base field K and consider (1.1.1) as a "generic equation", i.e., let $x_1,\dots,x_n$ be variables. Fix a polynomial $u = u(x_1,\dots,x_n) \in K[x_1,\dots,x_n]$ (a "resolvent"). Its orbit $O(u) = \{\tau \cdot u \mid \tau \in S_n\}$ under the action of the symmetric group is in a canonical bijection with the coset space S_n/H, where $H = \{\tau \in S_n \mid \tau \cdot u = u\}$ is the stabiliser of u ($\tau H \in S_n/H$ corresponds to $\tau \cdot u$).*

Let y be a new variable. The polynomial

$$U(y) = \prod_{v\in O(g)} (y - v) = \prod_{v\in O(u)} (y - v(x_1,\dots,x_n)) = \prod_{\tau H\in S_n/H} (y - u(x_{\tau(1)},\dots,x_{\tau(n)}))$$

lies in

$$K[x_1,\dots,x_n]^{S_n}[y] = K[\sigma_1,\dots,\sigma_n][y].$$

In other words, $u(x_1, \dots, x_n)$ is a root of the polynomial $U(y) = U(y; \sigma_1, \dots, \sigma_n)$, whose coefficients can be written in terms of the coefficients of the original polynomial $f(x) = (x - x_1) \cdots (x - x_n)$. The degree of $U(y)$ is equal to

$$\deg_y(U) = |O(u)| = (S_n : H).$$

The resolvents studied in 1.6–1.8 correspond to the following polynomials.

(2.14.1) $n = 2$, $K \supset \mathbf{Q}$, $u = (x_1 - x_2)^2$, $H = S_2$.
(2.14.2) $n = 3$, $K \supset \mathbf{Q}(\rho)$, $u = (x_1 + \rho x_2 + \rho^2 x_3)^3$, $H = A_3 \subset S_3$.
(2.14.3) $n = 4$, $K \supset \mathbf{Q}$, $u = x_1x_2 + x_3x_4$ (or $(x_1 - x_2 + x_3 - x_4)^2$), $H = D_8 \subset S_4$.

More generally, **Lagrange's resolvents** are given by $u = (\sum_{j=1}^n \zeta^j x_j)^m$, where $m \mid n$ and $\zeta^m = 1$. If $m = n$ and ζ is a primitive m-th root of unity (i.e., if $\zeta^a \neq 1$ for $0 < a < m$), then $H = C_n$ is the cyclic group generated by $(12 \cdots n)$. If $m \mid n$ is arbitrary and ζ is a primitive m-th root of unity, then H is a semi-direct product $H = (S_{n/m})^m \rtimes C_m$. The case $n = m = 3$ (resp. $n = 4$, $m = 2$) corresponds to (2.14.2) (resp. to the second resolvent in (2.14.3)).

If we take $u = \Delta = \prod_{1 \leq i < j \leq n}(x_i - x_j)$ (assuming that $2 \in K^*$ and $n \geq 2$), then $H = A_n$ and $O(\Delta) = \{\Delta, -\Delta\}$, which implies that

$$U(y) = (y - \Delta)(y + \Delta) = y^2 - \Delta^2 = y^2 - \mathrm{disc}(f).$$

Conversely, Exercise 2.13 implies that Δ is, essentially, the only resolvent for which $H = A_n$.

(2.15) Question. If $n \geq 5$, is there a resolvent u (apart from $u = \Delta$) for which $1 \neq \deg_y(U) < n$?

(2.16) Answer (Lagrange and his followers): no. This follows from the discussion in 2.14 and Proposition 2.17 below.

(2.17) Proposition. *Let $H \subset S_n$ be a subgroup.*
(i) If $(S_n : H) = 2$, then $H = A_n$.
(ii) If $n \geq 5$ and $H \lhd S_n$ is a normal subgroup of S_n, then $H = \{e\}, A_n$ or S_n.
(iii) If $n \geq 5$ and $H \neq A_n, S_n$, then $(S_n : H) \geq n$.
(iv) If $n \geq 5$ and $H \subsetneq A_n$, then $(A_n : H) \geq n$.

Proof. (i) Exercise. (ii) We use the fact that A_n is a simple group for $n \geq 5$ ([De 1], Thm. I.5.1), which implies that $H \cap A_n$ is equal to A_n ($\Longrightarrow H = A_n, S_n$) or to $\{e\}$ ($\Longrightarrow H \hookrightarrow S_n \longrightarrow S_n/A_n \xrightarrow{\sim} \{\pm 1\}$ is injective $\Longrightarrow |H| \leq 2$; a subgroup of order two is never normal in S_n, hence $H = \{e\}$).
(iii) The action of S_n on $X = S_n/H$ gives rise to a group homomorphism $\alpha : S_n \longrightarrow S_X$. We claim that this action is faithful (i.e., that α is injective). Indeed,

$$\mathrm{Ker}(\alpha) = \bigcap_{\tau \in S_n} \tau H \tau^{-1} \lhd S_n, \qquad \mathrm{Ker}(\alpha) \subset H \neq A_n, S_n,$$

which implies, by (ii), that $\mathrm{Ker}(\alpha) = \{e\}$. It follows that $n \leq |X| = (S_n : H)$.
(iv) The same argument as in (iii) shows that the action of A_n on $Y = A_n/H$ is faithful, hence $|A_n| = n!/2 \leq |S_Y| = (A_n : H)!$, which implies that $(A_n : H) \geq n$.

3. Field extensions (basic properties)

We saw in §1 classical formulas for solving cubic and quartic equations discovered in the 16th century and in §2 Lagrange's reformulation (and generalisation) in terms of symmetric functions. This theory covers general equations (1.1.1), in which the coefficients a_i appear as independent variables. Galois realised that it was possible – and extremely useful – to study symmetries of the equation (1.1.1) even in the case when the coefficients have numerical values. In somewhat vague terms, the Galois group of the equation (1.1.1) is the subgroup $\mathrm{Gal}(f) \subset S_n$ which preserves all polynomial relations between the roots of f. For example, for the equation $x^n - 1 = 0$ ("division of the circle in n parts") there are many relations between the roots, which

impose rather severe restrictions on $\mathrm{Gal}(f)$. This group was implicitly computed by Gauss (in the case when $n = p$ is a prime number) before Galois was even born. The corresponding problem for the division of the lemniscate was also considered by Gauss (unpublished) and Abel.

The "modern" formulation of what we call Galois theory is due to E. Artin. One considers the equation (1.1.1) for a polynomial f with coefficients in an arbitrary field K. The fundamental object is not the set of roots $\alpha_1, \ldots, \alpha_n$ of (1.1.1) (whose existence somewhere in the mathematical universe needs to be proved first – see Theorem 3.26 below), but the set of all expressions such as $\alpha_1/(\alpha_1^3 - 5\alpha_2\alpha_3)$, in other words, the field $L = K(\alpha_1, \ldots, \alpha_n)$ generated over K by the roots of f. The Galois group in this context depends only on the extension of fields $K \hookrightarrow L$ (strictly speaking, this works in full generality only for fields containing $\mathbf{Q}$; for fields containing $\mathbf{F}_p$ one needs to be more careful; see §6 below). The fundamental objects of study will be, therefore, field extensions $K \hookrightarrow L$ of finite degree.

(3.1) Fields. Let K be a field. The kernel of the canonical ring morphism $i : \mathbf{Z} \longrightarrow K$ (see I.4.3) is a principal ideal $\mathrm{Ker}(i) \subset \mathbf{Z}$.

If $\mathrm{Ker}(i) = (0)$, then K contains $i(\mathbf{Z}) = \mathbf{Z}$, hence also $\mathrm{Frac}(\mathbf{Z}) = \mathbf{Q}$. We say that K is a **field of characteristic zero** (notation: $char(K) = 0$).

If $\mathrm{Ker}(i) = (n)$ for some $n \geq 1$, then $n = p$ is a prime number, since $\mathrm{Im}(i) \xrightarrow{\sim} \mathbf{Z}/n\mathbf{Z} \subset K$ is a domain. We say that K is a **field of characteristic** p (notation: $char(K) = p$). The field K then contains the finite field $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$ with p elements.

(3.2) Homomorphisms (= embeddings) of fields. Let K and L be fields. A ring homomorphism $f : K \longrightarrow L$ is called a homomorphism of fields. Its kernel $\mathrm{Ker}(f) \neq (1)$ is an ideal of K, hence equal to (0), which means that f is automatically injective. We consider, therefore, f as an embedding of fields $f : K \hookrightarrow L$ (equivalently, we say that $f : K \hookrightarrow L$ is a **field extension**). Note that, for fixed K and L, there can be many different field embeddings $K \hookrightarrow L$ (example: the identity map and the complex conjugation $\mathbf{C} \hookrightarrow \mathbf{C}$). In general, the subfield $f(K) \subset L$ depends on f, not just on K and L (see Example 3.12(iv) below). We say that the field extension $f : K \hookrightarrow L$ (sometimes denoted simply as L/K) is of **finite type** if there exist finitely many elements $\alpha_1, \ldots, \alpha_n \in L$ such that $L = K(\alpha_1, \ldots, \alpha_n)$. A **simple extension** is a field extension generated by one element: $L/K = K(\alpha)/K$.

(3.3) Proposition. *Let K be a field. Any finite subgroup $A \subset K^*$ of the multiplicative group of K is cyclic.*

Proof. We can assume that $|A| = p_1^{n_1} \cdots p_r^{n_r} > 1$. Fix a prime number $p = p_i$ dividing $|A|$. The p-primary part of A is isomorphic to

$$A(p) \xrightarrow{\sim} \mathbf{Z}/p^{a_1}\mathbf{Z} \times \cdots \times \mathbf{Z}/p^{a_k}\mathbf{Z} \qquad (a_i \geq 1),$$

which implies that

$$p^k = |\{x \in A(p) \,|\, x^p = 1\}| \leq |\{x \in A \,|\, x^p = 1\}| \leq |\{x \in K \,|\, x^p - 1 = 0\}| \leq \deg(X^p - 1) = p,$$

hence $k = 1$, which means that the group $A(p) = A(p_i) \xrightarrow{\sim} \mathbf{Z}/p_i^{n_i}\mathbf{Z}$ is cyclic. As a result, the group $A \xrightarrow{\sim} A(p_1) \oplus \cdots \oplus A(p_r) \xrightarrow{\sim} \mathbf{Z}/p_1^{n_1}\mathbf{Z} \oplus \cdots \oplus \mathbf{Z}/p_r^{n_r}\mathbf{Z} \xrightarrow{\sim} \mathbf{Z}/|A|\mathbf{Z}$ is cyclic, too.

(3.4) Definition. *Let $K \hookrightarrow L$ be a field extension. A* **basis** *(resp. the* **degree***) of this extension is any basis of L (resp. the dimension $[L : K] := \dim_K(L) \in \mathbf{N} \cup \{\infty\}$ of L) considered as a K-vector space. If $[L : K] < \infty$, we say that $K \hookrightarrow L$ (or L/K) is a* **finite** *extension.*

(3.5) Examples. $[\mathbf{C} : \mathbf{R}] = 2$ and $\{1, i\}$ (or $\{2 - 3i, i + 17\}$, for example) is a basis of $\mathbf{C}/\mathbf{R}$. On the other hand, $[\mathbf{C} : \mathbf{Q}] = \infty$ and $[K(X) : X] = \infty$, for any field K.

(3.6) Proposition (multiplicativity of the degree). *Let $K \hookrightarrow L \hookrightarrow M$ be fields, let $\{\ell_i\}_{i \in I}$ (resp. $\{m_j\}_{j \in J}$) be a basis of L/K (resp. of M/L). Then $S = \{\ell_i m_j\}_{(i,j) \in I \times J}$ is a basis of M/K. In particular,*

$$[M : K] = |I \times J| = |I| \cdot |J| = [L : K][M : L].$$

Proof. The set S generates M as a K-vector space, since each $m \in M$ can be written as

$$m = \sum_{j \in J} y_j m_j \qquad (y_j \in L)$$

and each y_j as

$$y_j = \sum_{i \in I} x_{ij} \ell_i \qquad (x_{ij} \in K),$$

which yields

$$m = \sum_{i \in I} \sum_{j \in J} x_{ij} \ell_i m_j \qquad (x_{ij} \in K).$$

Conversely, the set S is linearly independent over K, since any relation

$$\sum_{i \in I} \sum_{j \in J} x_{ij} \ell_i m_j = \sum_{j \in J} \left(\sum_{i \in I} x_{ij} \ell_i \right) m_j = 0 \qquad (x_{ij} \in K),$$

implies

$$(\forall j \in J) \qquad \sum_{i \in I} x_{ij} \ell_i = 0$$

(since the m_j are linearly independent over L), hence $x_{ij} = 0$ for all i, j (since the ℓ_i are linearly independent over K).

(3.7) Corollary. *The extension M/K is finite $\iff$ both extensions L/K and M/L are finite.*

(3.8) Definition. *Let $K \hookrightarrow L$ be a field extension. An element $\alpha \in L$ is* **algebraic over** *K (resp.* **transcendental over** *K) if there exists a non-zero polynomial $f \in K[X]$ such that $f(\alpha) = 0$ (resp. if $f(\alpha) \neq 0$ for all non-zero $f \in K[X]$). The* **extension** *L/K* **is algebraic** *if all elements of L are algebraic over K.*

(3.9) This definition has a useful reformulation in terms of the map "evaluation at α":

$$\mathrm{ev}_\alpha : K[X] \longrightarrow L, \qquad g(X) \mapsto g(\alpha). \tag{3.9.1}$$

This is a homomorphism of K-algebras, whose image $\mathrm{Im}(\mathrm{ev}_\alpha) \xrightarrow{\sim} K[X]/\operatorname{Ker}(\mathrm{ev}_\alpha)$ is equal to $K[\alpha] \subset K(\alpha) \subset L$ (hence is a domain).

If α is transcendental over K, then $\operatorname{Ker}(\mathrm{ev}_\alpha) = (0)$ and the map (3.9.1) induces isomorphisms of K-algebras $K[X] \xrightarrow{\sim} K[\alpha]$ and $K(X) \xrightarrow{\sim} K(\alpha)$. In particular, $[K(\alpha) : K] = \infty$ in this case.

(3.10) Proposition. *If $\alpha \in L$ is algebraic over a subfield $K \hookrightarrow L$, then:*
(1) $\operatorname{Ker}(\mathrm{ev}_\alpha) = (f)$*, for a unique non-constant monic polynomial $f \in K[X]$ of minimal degree satisfying $f(\alpha) = 0$. We say that f is the* **minimal polynomial of α over K** *and its degree $n = \deg(f) \geq 1$ is the* **degree of α over K.**
(2) The map (3.9.1) induces an isomorphism of K-algebras $\overline{\mathrm{ev}}_\alpha : K[X]/(f) \xrightarrow{\sim} K[\alpha]$.
(3) The K-algebra $K[\alpha]$ is a domain of dimension $\dim_K K[\alpha] = n$. The elements $1, \alpha, \ldots, \alpha^{n-1}$ form a basis of $K[\alpha]$ as a K-vector space.
(4) The polynomial f is irreducible in $K[X]$ (conversely, if $g(\alpha) = 0$ for an irreducible monic polynomial $g \in K[X]$, then $g = f$).
(5) The K-algebra $K[\alpha]$ is a field; thus $K[\alpha] = K(\alpha)$ and $[K(\alpha) : K] = n$.
(6) Conversely, if $f \in K[X]$ is an irreducible monic polynomial of degree $n \geq 1$, then the K-algebra $L = K[X]/(f)$ is a field extension of K, the element $\alpha = X \pmod f \in L$ is algebraic over K, its minimal polynomial over K is equal to f and $L = K[\alpha] = K(\alpha)$ (in particular, $[L : K] = n$).

Proof. The statement (1) (resp. (2)) follows from the fact that $K[X]$ is a PID and $(1) \neq \operatorname{Ker}(\mathrm{ev}_\alpha) \neq (0)$ (resp. is automatic). The statement (3) is a consequence of (2) and the fact that the elements $1, \overline{X}, \ldots, \overline{X}^{n-1}$

(where $\overline{X} = X \pmod f \in K[X]/(f)$) form a basis of $K[X]/(f)$ as a K-vector space, since $\overline{\mathrm{ev}}_\alpha(\overline{X}) = \alpha$ and $K[X]/(f) = K[\overline{X}]$. The statement (5) (resp. the first part of (4)) is a special case of Lemma 3.11 below (resp. of Proposition I.6.8(i)). The second part of (4) follows from the fact that $f \mid g$, hence g is a constant multiple of f (since both are irreducible). In (6), $K[X]/(f)$ is a field, by Lemma 3.11 below, of degree $\deg(f)$ over K (as in (3)). Finally, the element α satisfies $f(\alpha) = f(\overline{X}) = \overline{f(X)} = 0$, so we conclude by the second part of (4).

(3.11) Lemma. *If R is a PID and $f \in R$ is an irreducible element, then $R/(f)$ is a field. [This was used implicitly in I.6.2(iv).]*

Proof. A non-zero element of the domain $R/(f)$ can be written as $\overline{g} = g \pmod f$, where $g \in R$, $g \notin (f)$. We must show that $\overline{g}$ is invertible. The ideal $(f, g) = (h)$ is principal, with $h \mid g$ and $h \mid f$. Irreducibility of f implies that $h = uf$ or $h = u$, for some $u \in R^*$. If $h = uf$, then $f \mid h \mid g$, which contradicts our assumptions; thus $h \in R^*$ and $(f, g) = (1)$. In particular, there exist $a, b \in R$ such that $af + bg = 1$, which implies that $\overline{b} = b \pmod f$ is a multiplicative inverse of $\overline{g}$ in $R/(f)$.

(3.12) Examples. (i) For each $n \geq 1$, the polynomial $X^n - 2$ is irreducible in $\mathbf{Q}[X]$, thanks to Eisenstein's criterion (Corollary I.7.7 for $A = \mathbf{Z}$ and $P = (2)$). Let $\sqrt[n]{2}$ be its unique positive real root. Then

$$\mathbf{Q}(\sqrt[n]{2}) = \mathbf{Q}[\sqrt[n]{2}] = \{a_0 + a_1 \sqrt[n]{2} + \cdots + a_{n-1}(\sqrt[n]{2})^{n-1} \mid a_j \in \mathbf{Q}\}$$

is a subfield of $\mathbf{R}$ of degree $[\mathbf{Q}(\sqrt[n]{2}) : \mathbf{Q}] = n$ over $\mathbf{Q}$.
(ii) The quadratic polynomial $f = X^2 + X + 1 \in \mathbf{F}_2[X]$ has no root in $\mathbf{F}_2$ (since $f(0) = f(1) = 1 \in \mathbf{F}_2$), hence $\mathbf{F}_2[X]/(X^2 + X + 1)$ is a field with 4 elements (namely, $0, 1, \alpha, \alpha + 1$, where $\alpha = X \pmod f$).
(iii) According to Proposition 3.15(4) below, the set of algebraic numbers

$$\overline{\mathbf{Q}} = \{\alpha \in \mathbf{C} \mid \alpha \text{ is algebraic over } \mathbf{Q}\}$$

is a subfield of $\mathbf{C}$. The example (i) shows that $[\overline{\mathbf{Q}} : \mathbf{Q}] = \infty$. On the other hand, the set of polynomials $\mathbf{Q}[X]$ is countable, which implies that $\overline{\mathbf{Q}}$ is countable, too. In other words, "most" complex numbers are transcendental (over $\mathbf{Q}$).
(iv) The polynomial $f = X^3 - 2 \in \mathbf{Q}[X]$ has three distinct complex roots, namely, $\alpha = \alpha_1 = \sqrt[3]{2} \in \mathbf{R}$, $\alpha_2 = \rho\alpha_1$ and $\alpha_3 = \rho^2\alpha_1$, where $\rho = e^{2\pi i/3}$.

The three subfields $\mathbf{Q}(\alpha_j) \subset \mathbf{C}$ are physically distinct, but they are all isomorphic to the abstract field $\mathbf{Q}[X]/(X^3 - 2)$, via the evaluation map $\overline{\mathrm{ev}}_{\alpha_j}$. In particular, $[\mathbf{Q}(\alpha_j) : \mathbf{Q}] = 3$ for each $j = 1, 2, 3$. The field $L = \mathbf{Q}(\alpha_1, \alpha_2, \alpha_3) \subset \mathbf{C}$ is equal to $\mathbf{Q}(\alpha_1, \rho)$, hence $[L : \mathbf{Q}] = [L : \mathbf{Q}(\alpha_1)][\mathbf{Q}(\alpha_1) : \mathbf{Q}] = 2 \cdot 3 = 6$ (the degree of the first extension is equal to 2, since $\rho \in L$ satisfies a quadratic equation $\rho^2 + \rho + 1 = 0$ over $\mathbf{Q}(\alpha_1)$, but $\rho \notin \mathbf{Q}(\alpha_1) \subset \mathbf{R}$).

(3.13) Exercise (Liouville's Theorem). *Let $\alpha \in \overline{\mathbf{Q}} \subset \mathbf{C}$ be an algebraic number of degree $n \geq 2$ (over $\mathbf{Q}$). Show that there exists a constant $c > 0$ such that*

$$\forall p/q \in \mathbf{Q} \qquad \left|\alpha - \frac{p}{q}\right| \geq \frac{c}{|q|^n} \qquad\qquad (p, q \in \mathbf{Z}).$$

Deduce that the number $\sum_{k=0}^{\infty} 10^{-k!} \in \mathbf{C}$ is transcendental (over $\mathbf{Q}$). [Hint: consider $f(p/q) = f(p/q) - f(\alpha) = (p/q - \alpha)g(p/q)$, where f is the minimal polynomial of α over $\mathbf{Q}$ and $g(X) = f(X)/(X - \alpha)$.]

(3.14) For $n \geq 3$ the exponent n in Liouville's theorem was first improved by Thue to $n/2 + 1 + \varepsilon$ (for any $\varepsilon > 0$ and $c = c(\varepsilon) > 0$ depending on ε). Thue's result implies, among other things, that equations such as $x^n - 2y^n = c$ have only finitely many solutions $x, y \in \mathbf{Z}$, for fixed $n \geq 3$ and $c \in \mathbf{Z} \setminus \{0\}$ (exercise: explain the relevance of Thue's theorem to I.5.14). The exponent $n/2 + 1 + \varepsilon$ was subsequently improved by Siegel and others; the optimal exponent $2 + \varepsilon$ was obtained by K. Roth (Fields medal).

(3.15) Proposition. *Let $K \hookrightarrow L$ be a field extension.*
(1) L/K is a finite extension $\Longrightarrow$ L/K is algebraic and of finite type.
(2) If $L = K(\alpha_1, \ldots, \alpha_n)$, where each α_i is algebraic over K, then the extension L/K is finite, hence algebraic.

(3) The implication "$\Longrightarrow$" in (1) is an equivalence.
(4) $K' = \{\alpha \in L \mid \alpha$ is algebraic over $K\}$ is a subfield of L.

Proof. (1) If $\alpha_1, \ldots, \alpha_d$ is a basis of L/K ($d = [L : K] < \infty$), then $L = K(\alpha_1, \ldots, \alpha_d)$; thus L/K is a field extension of finite type. If $\beta \in L$, then the $d+1$ elements $1, \beta, \ldots, \beta^d$ are linearly dependent over K, which implies that there exist $a_0, \ldots, a_d \in K$ (not all zero) such that $a_0 + a_1\beta + \cdots + a_d\beta^d = 0$.
(2) Consider the tower of simple extensions

$$K = K_0 \subset K_1 = K(\alpha_1) \subset \cdots \subset K_{i-1} \subset K_i = K_{i-1}(\alpha_i) = K(\alpha_1, \ldots, \alpha_i) \subset \cdots \subset K_n = K(\alpha_1, \ldots, \alpha_n).$$

Let $f_i \in K[X]$ (resp. $g_i \in K_{i-1}[X]$) be the minimal polynomial of α over K (resp. over K_{i-1}), for each $i = 1, \ldots, n$. The divisibility $g_i \mid f_i$ in $K_{i-1}[X]$ implies that

$$[K_i : K_{i-1}] = [K_{i-1}(\alpha_i) : K_{i-1}] = \deg(g_i) \leq \deg(f_i) = [K(\alpha_i) : K],$$

hence

$$[L : K] = \prod_{i=1}^{n} [K_i : K_{i-1}] \leq \prod_{i=1}^{n} [K(\alpha_i) : K] < \infty.$$

(3) The implication "$\Longleftarrow$" was proved in (2).
(4) For any $\alpha, \beta \in K'$ the subfield $K(\alpha, \beta) \subset L$ is a finite extension of K, by (2), hence is algebraic over K, by (1). In particular, the elements $\alpha \pm \beta$, $\alpha\beta$ and $\alpha\beta^{-1}$ (if $\beta \neq 0$) also lie in K'.

(3.16) Corollary. *Let $K \hookrightarrow L \hookrightarrow M$ be field extensions. The extension M/K is algebraic $\Longleftrightarrow$ both extensions L/K and M/L are algebraic.*

Proof. The implication "$\Longrightarrow$" is automatic. Conversely, if both L/K and M/L are algebraic, then each β is a root of a suitable polynomial $f = X^n + a_1X^{n-1} + \cdots + a_n \in L[X]$ ($n \geq 1$). The subfield $L' = K(a_1, \ldots, a_n) \subset L$ is a finite extension of K, by (2). Moreover, $[L'(\beta) : L'] \leq \deg(f) < \infty$, since $f \in L'[X]$. It follows that

$$[K(\beta) : K] \leq [L'(\beta) : K] = [L'(\beta) : L'][L' : K] < \infty,$$

hence β is algebraic over K.

(3.17) Computing the inverse. The ring $K[X]$ is euclidean for any field K, which means that the proof of Proposition 3.10(5) can be made completely algorithmic. As an example (for $K = \mathbf{Q}$), let us compute the inverse of

$$\beta = 3 - 2\sqrt[3]{2} + \sqrt[3]{4} = g(\sqrt[3]{2}), \qquad g(X) = X^2 - 2X + 3.$$

An application of Euclid's algorithm to the polynomials $f(X) = X^3 - 2$ and $g(X)$ yields

$$\begin{aligned} X^3 - 2 &= (X^2 - 2X + 3)(X + 2) + (X - 8) \\ X^2 - 2X + 3 &= (X - 8)(X + 6) + 51, \end{aligned}$$

hence

$$51 = (X^2 + 8X + 13)(X^2 - 2X + 3) - (X + 6)(X^3 - 2) = h(X)g(X) - (X + 6)f(X).$$

For $X = \sqrt[3]{2}$, we obtain

$$\beta\, h(\sqrt[3]{2}) = g(\sqrt[3]{2})h(\sqrt[3]{2}) = 51 \implies \beta^{-1} = \frac{h(\sqrt[3]{2})}{51} = \frac{13 + 8\sqrt[3]{2} + \sqrt[3]{4}}{51}.$$

(3.18) Explicit equations and characteristic polynomials. The proof of Proposition 3.15(1) does not give an effective method for finding a polynomial equation for $\beta \in L$. Instead, one can proceed as follows.

If L/K is a finite extension and $\alpha_1, \ldots, \alpha_n$ is a set of generators of L as a K-vector space, then we can write, for each $j = 1, \ldots, n$, $\beta\alpha_j = \sum_{i=1}^n M_{ij}\alpha_i$, where $M_{ij} \in K$. This system of equations can be written in matrix terms as

$$\beta \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = M \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \qquad M = (M_{ij})_{1 \le i,j \le n} \in M_n(K),$$

which means that β is an eigenvalue of the matrix M. In particular, it is a root of the characteristic polynomial $\det(X \cdot I - M) \in K[X]$ of M.

(3.19) In the special case when $\alpha_1, \ldots, \alpha_n$ is a **basis** of L/K (hence $n = [L : K]$), the matrix $M = M(\beta)$ is the matrix of the K-linear map

$$m(\beta) : L \longrightarrow L, \qquad x \mapsto \beta x$$

given by multiplication by β in the basis $\alpha_1, \ldots, \alpha_n$. The **characteristic polynomial**

$$P_{L/K,\beta}(X) = \det(X \cdot I - M(\beta)) \in K[X]$$

does not depend on the choice of a basis. As a result, we obtain a canonical polynomial equation of degree $n = [L : K]$ with coefficients in K (for any $\beta \in L$):

$$P_{L/K,\beta}(\beta) = 0. \tag{3.19.1}$$

The matrices $M(\beta) \in M_n(K)$ ($\beta \in L$) have the following properties:

$$M(\alpha)+M(\beta) = M(\alpha+\beta), \qquad M(\alpha)M(\beta) = M(\alpha\beta), \qquad \forall \alpha \in K \quad M(\alpha) = \alpha \cdot I, \qquad \beta \neq 0 \Longrightarrow M(\beta) \neq 0 \tag{3.19.2}$$

(since $(\alpha + \beta)x = \alpha x + \beta x$ and $\alpha(\beta x) = (\alpha\beta)x$ for all $x \in L$). In other words, we have constructed an injective homomorphism of K-algebras ("**the regular representation of** L **over** K")

$$L \hookrightarrow M_n(K), \qquad \beta \mapsto M(\beta) \tag{3.19.3}$$

(the matrix ring $M_n(K)$ is the only non-commutative ring appearing in this course; the structure map $K \longrightarrow M_n(K)$ identifies K with the set of scalar matrices $\{\lambda I \mid \lambda \in K\}$).

(3.20) Definition. *Let $K \hookrightarrow L$ be a field extension of degree $n = [L : K] < \infty$. The* **norm** *(resp. the* **trace***) of an element $\beta \in L$ is defined as*

$$N_{L/K}(\beta) = \det(M(\beta)) \in K, \qquad \mathrm{Tr}_{L/K}(\beta) = \mathrm{Tr}(M(\beta)) \in K.$$

(3.21) Example. *If $L = K(\sqrt{d})$, where $d \in K \setminus K^2$, then $[L : K] = 2$ and $1, \sqrt{d}$ is a basis of L/K. Multiplication by $\beta = a + b\sqrt{d}$ $(a, b \in K)$ in this basis is given by*

$$m_\beta : \begin{array}{rcl} 1 \mapsto \beta & = & a \cdot 1 + b \cdot \sqrt{d} \\ \sqrt{d} \mapsto \beta\sqrt{d} & = & db \cdot 1 + a \cdot \sqrt{d}, \end{array}$$

hence

$$M(a + b\sqrt{d}) = \begin{pmatrix} a & db \\ b & a \end{pmatrix}, \qquad P_{L/K,\beta}(X) = X^2 - 2aX + (a^2 - db^2) = (X - \beta)(X - \beta'),$$

where $\beta = a + b\sqrt{d}$ and $\beta' = a - b\sqrt{d}$. In particular,

$$N_{L/K}(a+b\sqrt{d}) = a^2 - db^2 = \beta\beta', \qquad \mathrm{Tr}_{L/K}(a+b\sqrt{d}) = 2a = \beta + \beta'.$$

(3.22) Theorem (minimal polynomial and the characteristic polynomial). *Let $K \hookrightarrow L$ be a finite extension, let $\beta \in L$. Denote by $f \in K[X]$ the minimal polynomial of β over K; then $P_{L/K,\beta} = f^{[L:K(\beta)]}$.*

Proof. Let $\alpha_1, \ldots, \alpha_d$ (resp. $\omega_1, \ldots, \omega_m$) be a basis of $K(\beta)/K$ (resp. of $L/K(\beta)$). The matrix $M(\beta)$ in the basis $\alpha_i\omega_j$ of L/K is a block matrix

$$M(\beta) = \begin{pmatrix} A & 0 & \ldots & 0 \\ 0 & A & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & A \end{pmatrix},$$

where A is the matrix of multiplication by β with respect to the basis $\alpha_1, \ldots, \alpha_d$ of $K(\beta)/K$. In particular,

$$P_{L/K,\beta}(X) = P_{K(\beta)/K,\beta}(X)^m.$$

Both polynomials $P_{K(\beta)/K,\beta}(X), f(X) \in K[X]$ are monic, have the same degree $[K(\beta) : K]$ and satisfy $P_{K(\beta)/K,\beta}(\beta) = f(\beta) = 0$. It follows that $P_{K(\beta)/K,\beta}(X) = f(X)$, by uniqueness of the minimal polynomial.

One can also compute explicitly the matrix A for $\alpha_1, \ldots, \alpha_d = 1, \beta, \ldots, \beta^{d-1}$; in the notation of II.4.7 we have $A = C(f)$ and a short calculation shows that $\det(X \cdot I - C(f)) = f(X)$.

(3.23) Exercise. *Under the assumptions of 3.20,*
(i) $\forall \alpha, \beta \in L \quad \mathrm{Tr}_{L/K}(\alpha+\beta) = \mathrm{Tr}_{L/K}(\alpha) + \mathrm{Tr}_{L/K}(\beta), \quad N_{L/K}(\alpha\beta) = N_{L/K}(\alpha)N_{L/K}(\beta)$. *In other words, the trace (resp. the norm) is a group homomorphism $\mathrm{Tr}_{L/K} : L \longrightarrow K$ (resp. $N_{L/K} : L^* \longrightarrow K^*$).*
(ii) $\forall \alpha \in K, \forall \beta \in L \quad \mathrm{Tr}_{L/K}(\alpha\beta) = \alpha \mathrm{Tr}_{L/K}(\beta), \quad N_{L/K}(\alpha\beta) = \alpha^{[L:K]} N_{L/K}(\beta)$.
(iii) If M/L is a finite extension, then $N_{L/K} \circ N_{M/L} = N_{M/K}, \quad \mathrm{Tr}_{L/K} \circ \mathrm{Tr}_{M/L} = \mathrm{Tr}_{M/K}$.

(3.24) We are now ready to make sense of the statement "let L be the field generated over K by the roots of a given polynomial $f \in K[X]$".

(3.25) Definition. *Let K be a field, let $f \in K[X]$ be a polynomial of degree $n \geq 1$. A* **splitting field of f over** *K is a field extension $K \hookrightarrow L$ such that f splits in $L[X]$ as $f(X) = c(X-\alpha_1)\cdots(X-\alpha_n)$ (where $c \in K^*$ and $\alpha_1, \ldots, \alpha_n \in L$) and $L = K(\alpha_1, \ldots, \alpha_n)$ (according to Proposition 3.15(2), L/K is a finite extension).*

(3.26) Theorem. *A splitting field of any non-constant polynomial $f \in K[X]$ exists. Two splitting fields of f are isomorphic as K-algebras. [As a result, we can use the terminology "the splitting field of f over K".]*

Proof. Existence: fix an irreducible factor $g \mid f$, $g \in K[X]$. The quotient ring $K_1 = K[X]/(g)$ is a field containing K and $K_1 = K(\alpha_1)$, where $\alpha_1 = X \pmod g \in K_1$ satisfies $g(\alpha_1) = 0$, hence $f(\alpha_1) = 0$. This means that $f(X) = (X-\alpha_1)f_1(X)$, $f_1 \in K_1[X]$. After replacing (K, f) by (K_1, f_1) and repeating this procedure we obtain a tower of simple extensions

$$K = K_0 \subset K_1 = K(\alpha_1) \subset \cdots \subset K_{i-1} \subset K_i = K_{i-1}(\alpha_i) = K(\alpha_1, \ldots, \alpha_i) \subset \cdots \subset K_n = K(\alpha_1, \ldots, \alpha_n)$$

such that $f(X) = c(X-\alpha_1)\cdots(X-\alpha_n) \in K_n[X]$. The field K_n is then a splitting field of f over K.

Uniqueness: let $K \hookrightarrow L$ and $K \hookrightarrow L'$ be two splitting fields of f over K. Fix an irreducible factor $g \mid f$, $g \in K[X]$. By definition of a splitting field, the polynomial g splits both in L and L', which means that there exist $\alpha \in L$ and $\alpha' \in L'$ such that $g(\alpha) = 0$ and $g(\alpha') = 0$. The evaluation morphisms ev_α and $\mathrm{ev}_{\alpha'}$ give rise to isomorphisms of K-algebras

$$\overline{\mathrm{ev}}_\alpha : K[X]/(g) \xrightarrow{\sim} K(\alpha) = K_1 \subset L, \qquad \overline{\mathrm{ev}}_{\alpha'} : K[X]/(g) \xrightarrow{\sim} K(\alpha') = K_1' \subset L'$$

and $\sigma = \overline{\mathrm{ev}}_{\alpha'} \circ (\overline{\mathrm{ev}}_{\alpha})^{-1} : K_1 \xrightarrow{\sim} K_1'$. Consider K_1 as a subfield of both L (via the inclusion) and L' (via σ and the inclusion $K_1' \subset L'$). As above, $f = (X-\alpha)f_1$, where $f_1 \in K_1[X]$ and L, L' are splitting fields of f_1 over K_1. After replacing (K, f) by (K_1, f_1) and repeating this procedure we obtain two towers of simple extensions

$$K = K_0 \subset K_1 = K(\alpha_1) \subset \cdots \subset K_{i-1} \subset K_i = K_{i-1}(\alpha_i) = K(\alpha_1, \ldots, \alpha_i) \subset \cdots \subset K_n = K(\alpha_1, \ldots, \alpha_n) \subset L$$
$$K = K_0 \subset K_1' = K(\alpha_1') \subset \cdots \subset K_{i-1}' \subset K_i' = K_{i-1}'(\alpha_i') = K(\alpha_1', \ldots, \alpha_i') \subset \cdots \subset K_n' = K(\alpha_1', \ldots, \alpha_n') \subset L'$$

such that $f = c(X-\alpha_1)\cdots(X-\alpha_n) \in K_n[X]$ and $f = c(X-\alpha_1')\cdots(X-\alpha_n') \in K_n'[X]$ (which implies that $L = K_n$ and $L' = K_n'$), and field isomorphisms $\sigma_i : K_i \xrightarrow{\sim} K_i'$ satisfying $\sigma_i|_{K_{i-1}} = \sigma_{i-1}$. In particular, $\sigma_n : L \xrightarrow{\sim} L'$ is an isomorphism of K-algebras.

(3.27) Proposition-Definition. *Let $K \hookrightarrow L$ be a field extension, let $\alpha \in L$. The* **derivative** *of a polynomial $f = \sum_{i=0}^n a_i X^i \in K[X]$ is defined to be $f' = \sum_{i=1}^n i a_i X^{i-1} \in K[X]$ (where $ia_i = (i \cdot 1_K)a_i$).*
(1) The polynomial $f(X) - f(\alpha) - (X-\alpha)f'(\alpha) \in L[X]$ is divisible in $L[X]$ by $(X-\alpha)^2$.
(2) α is a simple root of f (i.e., $(X-\alpha) \mid f$ and $(X-\alpha)^2 \nmid f$ in $L[X]$) $\iff$ $f(\alpha) = 0 \neq f'(\alpha)$.
(3) All roots of a non-zero polynomial f (in its splitting field) are simple $\iff$ $(f, f') = (1)$ in $K[X]$.

Proof. (1) By linearity, it is sufficient to consider the case $K = L$ and $f = (X-\alpha)^n$ ($n \in \mathbf{N}$), when everything is explicit.
(2) This is an immediate consequence of (1).
(3) The ideal $(f, f') = (g) \subset K[X]$ is principal. If $\deg(g) \geq 1$, then g has a root α in some extension $L \supset K$; thus $f(\alpha) = f'(\alpha) = 0$. If $\deg(g) = 0$, then $(f, f') = (1)$ and there exist $a, b \in K[X]$ such that $a(X)f(X) + b(X)f'(X) = 1$. It follows that $b(\alpha)f'(\alpha) = 1$ ($\Longrightarrow f'(\alpha) \neq 0$) for each root α of f.

(3.28) Exercise. *Let K be a field of characteristic $char(K) \neq 2$.*
(1) If $[L : K] = 2$, show that $L = K(\sqrt{a})$ for some $a \in K \setminus K^2$.
(2) If $a \in K \setminus K^2$, then $K(\sqrt{a})^{*2} \cap K^* = K^{*2} \cup aK^{*2}$.
(3) If $a, b \in K \setminus K^2$, then $K(\sqrt{a}) = K(\sqrt{b}) \iff a/b \in K^2$.
(4) If $a, b \in K$, then $[K(\sqrt{a}, \sqrt{b}) : K] = 4 \iff a, b, ab \notin K^2$. If this is the case, then $1, \sqrt{a}, \sqrt{b}, \sqrt{ab}$ is a basis of $K(\sqrt{a}, \sqrt{b})/K$ and the only intermediate fields $K \subsetneq L \subsetneq K(\sqrt{a}, \sqrt{b})$ are $L = K(\sqrt{a}), K(\sqrt{b}), K(\sqrt{ab})$.

(3.29) Exercise. *Let K be a field of characteristic $char(K) \neq 2$, let $a, b, c \in K^*$, $c \notin K^{*2}$. Consider the fields $K_1 = K(\sqrt{c})$ and $L = K(\alpha)$ $(= K(\sqrt{a + b\sqrt{c}}))$, where $\alpha^2 = a + b\sqrt{c}$.*
(1) Show that $L = K_1 \iff$ there exists $d \in K^*$ such that $a^2 - b^2 c = d^2$ and $2(a+d) \in K^{*2}$.
(2) Show that there exists $\beta \in L$ such that $\beta^2 = a - b\sqrt{c} \iff a^2 - b^2 c \in K^{*2} \cup cK^{*2}$. [Hint: consider $K_1^* \cap L^{*2}$.]
(3) Determine $K^* \cap L^{*2}$.
(4) Show that there is $c' \in K^*$ such that $L = K(\sqrt{c}, \sqrt{c'}) \iff a^2 - b^2 c \in K^{*2}$. Explicitly, if $a^2 - b^2 c = d^2$ with $d \in K^*$, then $(\alpha \pm \beta)^2 = 2(a \pm d)$, hence $L = K(\sqrt{c}, \sqrt{2(a+d)}) = K(\sqrt{c}, \sqrt{2(a-d)})$.
[Example: $\sqrt{2 + \sqrt{3}} = (1 + \sqrt{3})/\sqrt{2}$.]

(3.30) Exercise. *Let $K \hookrightarrow L_1 \hookrightarrow M$, $K \hookrightarrow L_2 \hookrightarrow M$ be fields such that the only subfield of M containing both L_1 and L_2 is M itself. If $d_i = [L_i : K] < \infty$ for $i = 1, 2$ and if $\gcd(d_1, d_2) = 1$, then $[M : K] = d_1 d_2$.*

4. Finite felds

(4.1) Proposition. *(1) A field K has finitely many elements $\iff$ $char(K) = p > 0$ and $n = [K : \mathbf{F}_p] < \infty$. If this is the case, then $|K| = p^n = q$ and each element of K^* (resp. of K) is a root of the polynomial $X^{q-1} - 1 \in \mathbf{F}_p[X]$ (resp. of $X^q - X \in \mathbf{F}_p[X]$).*
(2) For each $n \geq 1$ the polynomial $X^{p^n} - X \in \mathbf{F}_p[X]$ has simple roots.
(3) If K is a field with $|K| = p^n = q$ elements, then K is a splitting field of the polynomial $X^{p^n} - X$ over $\mathbf{F}_p$. [In particular, it is unique up to isomorphism, by Theorem 3.26.]

(4) If $f \in \mathbf{F}_p[X]$ is an irreducible polynomial of degree $n \geq 1$, then the ring $K = \mathbf{F}_p[X]/(f)$ is a field with $|K| = p^n$ elements and the polynomial f divides $X^{p^n} - X$ in $\mathbf{F}_p[X]$.
(5) Every finite field is isomorphic to a field obtained by the construction in (4).

Proof. (1) If $char(K) = p$ and $n = [K : \mathbf{F}_p] < \infty$, then K is isomorphic to $\mathbf{F}_p^n$ as an $\mathbf{F}_p$-vector space, hence $|K| = |\mathbf{F}_p|^n = p^n$. If $char(F) = p$ and $[K : \mathbf{F}_p] = \infty$ (resp. if $char(K) = 0$), then K contains $\mathbf{F}_p^n$ for all $n \geq 1$ (resp. K contains $\mathbf{Q}$), hence $|K| = \infty$.

If $|K| = q < \infty$, then K^* is a finite group of order $q - 1$, hence $\forall a \in K^*\ a^{q-1} = 1$ ($\Longrightarrow a^q = a$). If $a \in K \setminus K^*$, then $a = 0 \Longrightarrow a^q = a$.
(2) The polynomial $g(X) = X^{p^n} - X \in \mathbf{F}_p[X]$ has derivative $g'(X) = p^n X^{p^n-1} - 1 = -1 \in \mathbf{F}_p[X]$; thus $(g, g') = (1)$. The statement follows from Proposition 3.27(3).
(3) It follows from (1) and (2) that the elements of K are precisely the p^n distinct roots of the polynomial $g(X) = X^{p^n} - X \in \mathbf{F}_p[X]$ in its splitting field L over $\mathbf{F}_p$. As L is generated over $\mathbf{F}_p$ by these roots, it coincides with K.
(4) The ring $K = \mathbf{F}_p[X]/(f)$ is a field and $[K : \mathbf{F}_p] = n$, by Proposition 3.10(6). If we denote by $\overline{X}$ the image of X in K, then (1) implies that $\overline{X}^{p^n} - \overline{X} = 0$, which is equivalent to the divisibility $f|(X^{p^n} - X)$.
(5) According to Proposition 3.3, the multiplicative group K^* is cyclic of order $q - 1$, where $q = |K|$. If α is any generator of K^*, then $K = \mathbf{F}_p(\alpha) \xrightarrow{\sim} \mathbf{F}_p[X]/(f)$, where $f \in \mathbf{F}_p[X]$ is the minimal polynomial of α over $\mathbf{F}_p$, by Proposition 3.10(2),(5).

(4.2) Definition (Frobenius morphism). *Let p be a prime number, let R be any $\mathbf{F}_p$-algebra ($\Longleftrightarrow p \cdot 1 = 0$ in R). The formula*

$$\varphi : R \longrightarrow R, \qquad \varphi(x) = x^p$$

defines a homomorphism of $\mathbf{F}_p$-algebras ($\varphi(1) = 1$, $\varphi(xy) = \varphi(x)\varphi(y)$, $\varphi(x \pm y) = \varphi(x) \pm \varphi(y)$ for all $x, y \in R$, $\varphi(a) = a$ for each $a \in \mathbf{F}_p$), called **the Frobenius morphism**. *For $n \geq 1$ set $q = p^n$ and $\varphi_q = \varphi^n = \underbrace{\varphi \circ \cdots \circ \varphi}_{n-\text{times}} : R \longrightarrow R$; then $\varphi_q(x) = x^q$.*

(4.3) Proposition. *Let p be a prime number, let $E \supset \mathbf{F}_p$ be any field containing a splitting field of the polynomial $g = X^q - X \in \mathbf{F}_p[X]$ over $\mathbf{F}_p$.*
(1) If K is a field with $|K| = p^n = q$ elements, then $K \xrightarrow{\sim} E^{\varphi_q = 1} = \{a \in E \mid (\varphi_q - 1)(a) = 0\}$ (above, "1" denotes the identity map $E \longrightarrow E$, $x \mapsto x$).
(2) Conversely, $E^{\varphi_q=1} = \{a \in E \mid (\varphi_q - 1)(a) = 0\}$ is a field with q elements.

Proof. (1) This was already observed in the course of the proof of Proposition 4.1(3).
(2) If $x, y \in E$ satisfy $\varphi_q(x) = x$ and $\varphi_q(y) = y$, then $\varphi_q(x \pm y) = x \pm y$, $\varphi_q(xy) = xy$ and $\varphi_q(x/y) = x/y$ (if $y \neq 0$), since $\varphi_q : E \longrightarrow E$ is a homomorphism of fields. As a result, $E^{\varphi_q=1}$ is a subfield of E. By definition, its elements are precisely the roots of the polynomial g contained in E; there are $q = \deg(g)$ of them, since E contains a splitting field of g and g has no multiple roots, by Proposition 4.1(2).

(4.4) Theorem. *Let p be a prime number. (1) For each $n \geq 1$ there exists a field with p^n elements; it is unique up to isomorphism. It is usually denoted by $\mathbf{F}_{p^n}$ (or $GF(p^n)$).*
(2) Let $m, n, r \geq 1$ be integers, let $q = p^r$. There exists a field homomorphism $\sigma : \mathbf{F}_{q^m} \hookrightarrow \mathbf{F}_{q^n} \iff m|n$. If this is the case, then $\sigma(\mathbf{F}_{q^m}) = \{x \in \mathbf{F}_{q^n} \mid x^{q^m} = x\}$.

Proof. (1) The uniqueness (resp. the existence) of $\mathbf{F}_{p^n}$ was proved in Proposition 4.1(3) (resp. in Proposition 4.3(2)).
(2) If $m \mid n$, then $(q^m - 1) \mid (q^n - 1)$, hence the polynomial $X^{q^m} - X$ divides $X^{q^n} - X$. In particular, the splitting field of $X^{q^n} - X$ contains that of $X^{q^m} - X$. Conversely, if there exists a field embedding $\sigma : \mathbf{F}_{q^m} \hookrightarrow \mathbf{F}_{q^n}$, then $K = \sigma(\mathbf{F}_{q^m})$ is a subfield of $\mathbf{F}_{q^n}$ and $q^n = |\mathbf{F}_{q^n}| = |K|^d = q^{md}$, where $d = [\mathbf{F}_{q^n} : K]$; thus $n = md$. The equality $\sigma(\mathbf{F}_{q^m}) = \{x \in \mathbf{F}_{q^n} \mid x^{q^m} = x\}$ follows from Proposition 4.3(1) applied to $K = \sigma(\mathbf{F}_{q^m})$ and $E = \mathbf{F}_{q^n}$.

(4.5) Corollary. *(1) If $f \in \mathbf{F}_q[X]$ is an irreducible polynomial of degree $m \geq 1$, then f divides $X^{q^n} - X$ in $\mathbf{F}_q[X]$ if and only if $m \mid n$.*

(2) For $m \geq 1$, denote by A_m the set of monic irreducible polynomials $f \in \mathbf{F}_q[X]$ of degree $\deg(f) = m$. The set A_m is non-empty and

$$\forall n \geq 1 \quad X^{q^n} - X = \prod_{m|n} \prod_{f \in A_m} f, \qquad q^n = \sum_{m|n} m|A_m|.$$

Proof. (1) Let $f \in A_m$ ($m \geq 1$). If $m \mid n$, then $f \mid (X^{q^m} - X) \mid (X^{q^n} - X)$, by Proposition 4.1(4). Conversely, if $f \mid (X^{q^n} - X)$, then $K = \mathbf{F}_q[X]/(f) = \mathbf{F}_{q^m}$ is contained in the splitting field of the polynomial $X^{q^n} - X$ over $\mathbf{F}_q$ (i.e., in $\mathbf{F}_{q^n}$), hence $m \mid n$, thanks to Theorem 4.4(2).
(2) The factorisation of $X^{q^n} - X$ follows from (1) and the fact that the polynomial $X^{q^n} - X$ has no multiple roots; one concludes by counting the degrees.

(4.6) Example. If $p = q = n = 2$, then $X^4 - X = X(X-1)(X^2+X+1) \in \mathbf{F}_2[X]$, $\mathbf{F}_2[X]/(X^2+X+1) = \mathbf{F}_4$.

(4.7) Factorisation in $\mathbf{F}_q[X]$. A combination of the Frobenius morphism with the Chinese remainder theorem leads to an efficient factorisation algorithm for polynomials with coefficients in finite fields. The main idea is very simple:

(4.8) Proposition. *Let $f \in \mathbf{F}_q[X]$ (where $q = p^n$) be a non-constant monic polynomial* **without multiple roots**, *$f = f_1 \cdots f_k$, where $f_i \in \mathbf{F}_q[X]$ are distinct irreducible non-constant monic polynomials. The map*

$$\varphi_q - 1 : \mathbf{F}_q[X]/(f) \longrightarrow \mathbf{F}_q[X]/(f), \qquad h \pmod f \mapsto h^q - h \pmod f$$

is $\mathbf{F}_q$-linear and

$$\mathbf{F}_q \subseteq \operatorname{Ker}(\varphi_q - 1) = (\mathbf{F}_q[X]/(f))^{\varphi_q = 1}, \qquad \dim_{\mathbf{F}_q} \operatorname{Ker}(\varphi_q - 1) = k,$$
$$f \text{ is irreducible} \iff \operatorname{Ker}(\varphi_q - 1) = \mathbf{F}_q.$$

Proof. For each $a \in \mathbf{F}_q$ we have $\varphi_q(a) = a^q = a$, which implies that the map $\varphi_q - 1 : \mathbf{F}_q[X]/(f) \longrightarrow \mathbf{F}_q[X]/(f)$ satisfies $(\varphi_q - 1)(ah) = (ah)^q - ah = a(h^q - h) = a(\varphi_q - 1)(h)$, for each $a \in \mathbf{F}_q$ and $h \in \mathbf{F}_q[X]/(f)$. The Chinese remainder theorem I.5.11(ii) then yields

$$\mathbf{F}_q[X]/(f) \xrightarrow{\sim} \mathbf{F}_q[X]/(f_1) \times \cdots \times \mathbf{F}_q[X]/(f_k) = \mathbf{F}_{q^{d_1}} \times \cdots \times \mathbf{F}_{q^{d_k}} \qquad (d_i = \deg(f_i)),$$
$$(\mathbf{F}_q[X]/(f))^{\varphi_q = 1} \xrightarrow{\sim} (\mathbf{F}_{q^{d_1}})^{\varphi_q = 1} \times \cdots \times (\mathbf{F}_{q^{d_k}})^{\varphi_q = 1} = \prod_{i=1}^{k} \mathbf{F}_q,$$

hence $\dim_{\mathbf{F}_q} \operatorname{Ker}(\varphi_q - \mathrm{id}) = k$.

(4.9) Corollary (Berlekamp's algorithm). *Let $f \in \mathbf{F}_q[X]$ be a monic polynomial* **without multiple roots** *of degree $n > 1$.*
(1) Compute the matrix $A \in M_n(\mathbf{F}_q)$ of the $\mathbf{F}_q$-linear map $\varphi_q - 1 : \mathbf{F}_q[X]/(f) \longrightarrow \mathbf{F}_q[X]/(f)$ in the basis $1, \overline{X}, \ldots, \overline{X}^{n-1}$, where $\overline{X} = X \pmod f$.
(2) Solve the system of linear equations $Au = 0$ ($u \in \mathbf{F}_q^n$).
(3) If the space of solutions is equal to $\mathbf{F}_q \cdot \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, then f is irreducible in $\mathbf{F}_q[X]$.
(4) If, on the other hand, there exists another solution $u \in \mathbf{F}_q^n$, it corresponds to a non-constant polynomial $h \in \mathbf{F}_q[X]$ of degree $\deg(h) < n$ such that $h^q - h \pmod f = 0$. In other words, $f \mid h^q - h = \prod_{a \in \mathbf{F}_q}(h - a)$.
(5) The polynomials $h - a$ ($a \in \mathbf{F}_q$) in the above product are pairwise relatively prime and satisfy $\deg(h-a) < \deg(f)$, which implies that the factorisation

$$f = \prod_{a \in \mathbf{F}_q} h_a, \qquad h_a = \operatorname{pgcd}(f, h - a), \qquad \deg(h_a) < \deg(f)$$

is non-trivial.

(4.10) Example. Let us factor the polynomial $f = X^5 + X + 1 \in \mathbf{F}_2[X]$ by working in the ring

$$\mathbf{F}_2[X]/(f) = \mathbf{F}_2 \cdot 1 + \mathbf{F}_2 \cdot \overline{X} + \mathbf{F}_2 \cdot \overline{X}^2 + \mathbf{F}_2 \cdot \overline{X}^3 + \mathbf{F}_2 \cdot \overline{X}^4 \qquad (\overline{X} = X \pmod{(X^5 + X + 1)}).$$

The formulas

$$\overline{X}^5 = \overline{X} + 1, \qquad \overline{X}^6 = \overline{X}^2 + \overline{X}, \qquad \overline{X}^7 = \overline{X}^3 + \overline{X}^2, \qquad \overline{X}^8 = \overline{X}^4 + \overline{X}^3$$

imply that

$$\varphi_2 - 1 : 1 \mapsto 1^2 - 1 = 0, \qquad \overline{X} \mapsto \overline{X}^2 - \overline{X}, \qquad \overline{X}^2 \mapsto \overline{X}^4 - \overline{X}^2,$$
$$\overline{X}^3 \mapsto \overline{X}^6 - \overline{X}^3 = \overline{X}^3 + \overline{X}^2 + \overline{X}, \qquad \overline{X}^4 \mapsto \overline{X}^8 - \overline{X}^4 = \overline{X}^3.$$

We find a non-constant solution

$$(\varphi_2 - 1)(\overline{X}^4 + \overline{X}^3 + \overline{X}) = 0 \implies f \mid h^2 - h = h(h-1), \qquad h = X^4 + X^3 + X \in \mathbf{F}_2[X].$$

Euclid's algorithm in $\mathbf{F}_2[X]$ yields

$$h_0 = \operatorname{pgcd}(f, h) = X^3 + X^2 + 1, \qquad h_1 = \operatorname{pgcd}(f, h-1) = X^2 + X + 1,$$
$$X^5 + X + 1 = (X^3 + X^2 + 1)(X^2 + X + 1) \in \mathbf{F}_2[X].$$

(4.11) Exercise. *What if the polynomial $f \in \mathbf{F}_q[T]$ does have a multiple root?*

(4.12) Exercise. *Let $p \neq 2, 3$ be a prime number, let $L \supset \mathbf{F}_p$ be a field.*
(1) The polynomial $\Phi_3(X) = X^2 + X + 1 \in \mathbf{F}_p[X]$ has a root $\alpha \in \mathbf{F}_p \iff p \equiv 1 \pmod 3$.
(2) $\alpha \in L$ satisfies $\Phi_3(\alpha) = 0 \iff \beta = 2\alpha + 1 \in L$ satisfies $\beta^2 = -3 \in L$.
(3) There exists $x \in \mathbf{Z}$ such that $x^2 \equiv -3 \pmod p \iff p \equiv 1 \pmod 3$.

(4.13) Exercise. *Let $p \neq 5$ be a prime number. Denote by L a splitting field of the polynomial $\Phi_5(X) = X^4 + X^3 + X^2 + X + 1 \in \mathbf{F}_p[X]$ over $\mathbf{F}_p$.*
(1) $\Phi_5(X)$ has no multiple root in L.
(2) $\alpha \in L$ is a root of $\Phi_5(X) \iff \alpha^5 = 1$ and $\alpha \neq 1$.
(3) The degree of $L/\mathbf{F}_p$ is equal to

$$[L : \mathbf{F}_p] = \begin{cases} 1, & p \equiv +1 \pmod 5 \\ 2, & p \equiv -1 \pmod 5 \\ 4, & p \equiv \pm 2 \pmod 5 \end{cases}$$

(4) Fix a root $\zeta \in L$ of $\Phi_5(X)$. Write down a polynomial $g \in \mathbf{F}_p[X]$ of degree $\deg(g) = 2$ such that $g(\zeta + \zeta^{-1}) = 0$. Deduce that $(2\beta + 1)^2 = 5 \in L$.
(5) Show that $\beta = \zeta + \zeta^{-1} \in \mathbf{F}_p \iff [L : \mathbf{F}_p] \leq 2 \iff p \equiv \pm 1 \pmod 5$.
(6) If $p \neq 2$, then there exists $x \in \mathbf{Z}$ such that $x^2 \equiv 5 \pmod p \iff p \equiv \pm 1 \pmod 5$.

(4.14) Exercise. *Let F be a field of characteristic $p > 0$, let $k \geq 1$ be an integer.*
(1) If $\alpha \in F$ satisfies $(\varphi^k - 1)(\alpha) = \alpha^{p^k} - \alpha \in \mathbf{F}_p$, then $\alpha \in \mathbf{F}_{p^{kp}}$. [Hint: $\alpha \in \mathbf{F}_{p^m} \iff (\varphi^m - 1)(\alpha) = 0$.]
(2) Let $c \in \mathbf{F}_p$. If the polynomial $X^{p^k} - X - c$ is irreducible in $\mathbf{F}_p[X]$, then $p^k \mid kp$.

(3) Determine all values of p, k and $c \in \mathbf{F}_p$ for which the polynomial $X^{p^k} - X - c$ is irreducible in $\mathbf{F}_p[X]$.
(4) If $k > 1$ and $a \in \mathbf{F}_{p^k}$, show that the polynomial $X^{p^k} - X - a$ is reducible in $\mathbf{F}_{p^k}[X]$.

5. Algebraic closure

It is often useful to collect "all" finite extensions of a given field K in a common field extension of K, called an algebraic closure of K. Strictly speaking, the whole theory of field extensions can be developed without introducing this concept, replacing everywhere "an algebraic closure of K" by "a sufficiently large algebraic extension of K".

(5.1) Proposition. *Let $K(\alpha)/K$ be a simple algebraic extension, let $f \in K[X]$ be the minimal polynomial of α over K. For every field extension $K \hookrightarrow M$ the map*

$$\mathrm{Hom}_{K-\mathrm{Alg}}(K(\alpha), M) \longrightarrow \{\beta \in M \mid f(\beta) = 0\}, \qquad \lambda \mapsto \lambda(\alpha)$$

is a bijection; its inverse is given by

$$\beta \mapsto (\lambda : g(\alpha) \mapsto g(\beta)), \qquad g \in K[X].$$

Proof. This is a combination of the isomorphism of K-algebras $\overline{\mathrm{ev}}_\alpha : K[X]/(f) \xrightarrow{\sim} K(\alpha)$ with Proposition I.4.4(iii). In concrete terms, the map is well-defined, since $f(\lambda(\alpha)) = \lambda(f(\alpha)) = 0$. It is injective, since $K(\alpha) = K[\alpha]$ and $\lambda(g(\alpha)) = g(\lambda(\alpha))$ for all $g \in K[X]$. Conversely, if $\beta \in M$ satisfies $f(\beta) = 0$, then the composite homomorphism of K-algebras $\mathrm{ev}_\beta : K[X] \longrightarrow M$ factors through the canonical projection $\mathrm{pr} : K[X] \longrightarrow K[X]/(f)$ as $K[X] \longrightarrow K[X]/(f) \xrightarrow{\sim} K(\alpha) \longrightarrow M$, which proves the surjectivity.

(5.2) Definition. *A field L is* **algebraically closed** *if each non-constant polynomial $f \in L[X]$ has a root in L ($\iff$ f splits in $L[X]$ as $f = c(X - \alpha_1) \cdots (X - \alpha_n)$, where $c \in L^*$ and $\alpha_1, \ldots, \alpha_n \in L$). In other words, the only algebraic extension of L is L itself. For example, the field $\mathbf{C}$ is algebraically closed.*

(5.3) Definition. *A field extension $K \hookrightarrow L$ is an* **algebraic closure of** *K if it is an algebraic extension and L is algebraically closed.*

(5.4) Proposition. *If $K \hookrightarrow L$ is an algebraic extension and if every non-constant polynomial $f \in K[X]$ splits in $L[X]$ as $f = c(X - \alpha_1) \cdots (X - \alpha_n)$, where $c \in K^*$ and $\alpha_1, \ldots, \alpha_n \in L$, then L is an algebraic closure of K.*

Proof. Assume that α is alebraic over L; let $f \in K[X]$ (resp. $g \in L[X]$) be its minimal polynomial over K (resp. over L). The polynomial g divides f in $L[X]$, which means that it splits in $L[X]$ (since f does), and so its root α must lie in L.

(5.5) Corollary. *If $K \hookrightarrow L$ is a field extension and L is algebraically closed, then the subfield $\{\alpha \in L \mid \alpha$ is algebraic over $K\}$ is an algebraic closure of K. In particular, $\overline{\mathbf{Q}} \subset \mathbf{C}$ is an algebraic closure of $\mathbf{Q}$.*

(5.6) Theorem. *Let K be a field. (1) An algebraic closure of K exists.*
(2) If Ω is an algebraically closed extension field of K, then for every algebraic extension $K \hookrightarrow L$ there exists a morphism of K-algebras $L \hookrightarrow \Omega$.
*(3) Two algebraic closures of K are isomorphic as K-algebras (***notation:** *a fixed algebraic closure of K is usually denoted by $\overline{K}$).*

Proof. (1) If the field K is (at most) countable, so is the set P of irreducible non-constant monic polynomials in $K[X]$. We define inductively $K_1 = K$ and $K_{n+1} \supset K_n$ to be a splitting field of the n-th element of P over K_n. The union $L = \bigcup_{n \geq 1} K_n$ is an algebraic extension of K satisfying the assumptions of Proposition 5.4. Therefore, L is an algebraic closure of K.

For a general field K, we begin by showing that there exists an algebraic field extension $K \hookrightarrow E(K)$ with the property that each $f \in P$ has a root in $E(K)$. This field is defined as a quotient of two monstrously big objects; the key point is to show that the quotient is non-zero.

Let $A = K[x_f]_{f \in P}$ be a polynomial ring in variables x_f, one variable for each element $f \in P$. Let $I \subset A$ be the ideal generated by $f(x_f)$, for all $f \in P$.

(5.7) Lemma. *The ideal $I \subset A$ is not equal to A.*

Proof of Lemma. If $I = A$, then there exists a relation

$$\sum_{i=1}^{n} a_i\, f_i(x_{f_i}) = 1 \qquad\qquad (a_i \in A), \tag{5.7.1}$$

where $f_1, \ldots, f_n \in P$ are distinct elements of P. Let L be a splitting field of $f_1 \cdots f_n$ over K; for each $i = 1, \ldots, n$ fix a root $\beta_i \in L$ of f_i. Define a morphism of K-algebras $\lambda : A \longrightarrow L$ sending each x_{f_i} $(i = 1, \ldots, n)$ to β_i and each x_f $(f \neq f_1, \ldots, f_n)$ to 0. Applying λ to (5.7.1) we obtain $0 = 1$. This contradiction proves the lemma.

We can now continue the proof of Theorem 5.6(1). There exists a maximal ideal $\mathfrak{m} \in \mathrm{Max}(A)$ containing I, by Corollary I.6.7(ii). The quotient ring $A/\mathfrak{m} = E(K)$ is a field extension of K, via the composite homomorphism

$$K \hookrightarrow A \longrightarrow A/I \longrightarrow A/\mathfrak{m} = E(K).$$

For each $f \in P$, the image $\alpha_f = x_f \pmod{\mathfrak{m}} \in E(K)$ of x_f satisfies $f(\alpha_f) = 0$. It follows that $E(K) = K[\alpha_f]_{f \in P}$ is an algebraic extension of K and that each non-constant polynomial in $K[X]$ has a root in $E(K)$. Iterating this construction, we obtain a sequence of algebraic field extensions

$$K \hookrightarrow E(K) \hookrightarrow E^2(K) = E(E(K)) \hookrightarrow \cdots \hookrightarrow E^\infty(K) = \bigcup_{n \geq 1} E^n(E) = \Omega.$$

Any non-constant polynomial $g \in \Omega[X]$ lies in some $E^n(K)[X]$; therefore it has a root in $E^{n+1}(K) \subset \Omega$. It follows that Ω is algebraically closed.
(2) (a) If $L = K(\alpha)$ is a simple algebraic extension of K, then the minimal polynomial $f \in K[X]$ of α over K has a root in Ω, which implies that $\mathrm{Hom}_{K-\mathrm{Alg}}(L, \Omega) \neq \emptyset$, by Proposition 5.1.
(b) If $L \supset K$ is an arbitrary algebraic extension of K, then the set of pairs (M, σ), where M is an intermediate field $K \subset M \subset L$ and $\sigma \in \mathrm{Hom}_{K-\mathrm{Alg}}(M, \Omega)$, has a natural partial order ($(M, \sigma) \leq (M', \sigma')$ if $M \subset M'$ and $\sigma'|_M = \sigma$). This partially ordered set is non-empty (it contains $M = K$ and the inclusion $K \hookrightarrow \Omega$) and inductive (given a totally ordered subset, its union gives an upper bound). Zorn's Lemma implies that it contains a maximal element (M, σ). If $M \neq L$, then σ can be extended to a bigger subfield $M(\alpha)$, for any $\alpha \in L \setminus M$, as in (a), which contradicts maximality. It follows that $M = L$, as required.
(3) If $K \hookrightarrow \Omega$ and $K \hookrightarrow \Omega'$ are algebraic closures of K, the statement (2) implies that there exists a morphism of K-algebras $\sigma : \Omega' \hookrightarrow \Omega$. The field Ω is an algebraic extension of K, hence also of $\sigma(\Omega')$. However, the field $\sigma(\Omega')$ is isomorphic to Ω', which means that it is algebraically closed; thus $\sigma(\Omega') = \Omega$ and σ is an isomorphism of K-algebras.

6. Separable extensions

(6.1) Question: when is a finite extension L/K simple ($L = K(\alpha)$)? Here are a few examples.
(i) For $K = \mathbf{Q}$ and $L = \mathbf{Q}(\sqrt{2}, \sqrt{3})$ the element $\alpha = \sqrt{2} + \sqrt{3} \in L$ satisfies $\alpha^{-1} = \sqrt{3} - \sqrt{2}$, $\sqrt{2} = (\alpha - \alpha^{-1})/2$ and $\sqrt{3} = (\alpha + \alpha^{-1})/2$, which implies that $L = K(\alpha)$.
(ii) If $K = \mathbf{F}_q$ and $L = \mathbf{F}_{q^n}$ are finite fields, then $L = K(\alpha)$ for any generator α of the finite cyclic group L^*.
(iii) If $L = \mathbf{F}_p(a, b)$ and $K = \mathbf{F}_p(a^p, b^p)$ (where a, b are independent variables), then the elements $a^i b^j$ $(0 \leq i, j < p)$ form a basis of L/K, $[L : K] = p^2$ and $\alpha^p \in K$ for each $\alpha = \sum c_{ij} a^i b^p \in L$ (since $\alpha^p = \sum c_{ij}^p a^{pi} b^{pj}$). In particular, $[K(\alpha) : K] \leq p$, hence $K(\alpha) \neq L$.

What is the difference between (i) and (ii) on one hand, and (iii) on the other hand? The key point in (iii) is that the minimal polynomial of $a \in L$ over K is equal to $X^p - a^p = (X - a)^p \in L[X]$; in particular, it has a multiple root (and similarly for the minimal polynomial of $b \in L$ over K)! This can never happen in characteristic zero, by Proposition 6.4(2) below. Let us investigate this phenomenon in more detail.

(6.2) Definition. *Let K be a field. A non-constant polynomial $f \in K[X]$ is* **separable** *if it has no multiple root (in its splitting field). According to Proposition 3.27(3), this condition is equivalent to $(f, f') = (1)$. If $K \hookrightarrow L$ is a field extension and $\alpha \in L$ is algebraic over K, we say that α* **is separable over** *K if its minimal polynomial over K is separable. An algebraic extension $K \hookrightarrow L$ is* **separable** *if each element $\alpha \in L$ is separable over K. The field K is* **perfect** *if each algebraic extension of K is separable over K ($\iff$ every irreducible polynomial $f \in K[X]$ is separable).*

(6.3) Proposition. *If K is a field of characteristic $p > 0$ and if $a \in K \setminus K^p$ is not a p-th power, then the polynomial $f = X^{p^n} - a \in K[X]$ $(n \geq 1)$ is irreducible. If α is a root of f, then $f(X) = (X-\alpha)^{p^n} \in K(\alpha)[X]$. The extension $K(\alpha)/K$ is not separable and $[K(\alpha) : K] = p^n$.*

Proof. Consider first the case $n = 1$. If $g \in K[X]$ is a non-constant polynomial dividing f, then $g = (X - \alpha)^m = X^m - m\alpha X^{m-1} + \cdots$ for some $1 \leq m \leq p$. In particular, $m\alpha \in K$. If $m < p$, then $m \in K^*$, hence $\alpha \in K$ and $a = \alpha^p \in K^p$, which is false. Therefore $m = p$ and f is irreducible.

For general $n \geq 1$ define $a_i = \alpha^{p^{n-i}}$ $(0 \leq i \leq n)$ and $K_i = K(a_i)$; then $K = K_0$ and $K_i = K_{i-1}(a_i)$ with $a_i^p = a_{i-1} \in K_{i-1}$, hence $K_i^p \subset K_{i-1}$ (as in 6.1(iii)). By assumption, $a = a_0 \notin K^p = K_0^p$. If $a_i \in K_i^p$, then $a_{i-1} = a_i^p \in (K_i^p)^p \subset K_{i-1}^p$; thus $a_i \notin K_i^p$ for all $i = 0, \ldots, n$, by induction. In particular, $[K_i : K_{i-1}] = p$ by the case $n = 1$ proved earlier, hence $[K(\alpha) : K] = [K_n : K_0] = p^n = \deg(f)$, which shows that f is irreducible.

(6.4) Proposition. *Let K be a field.*
(1) A non-constant irreducible polynomial $f \in K[X]$ is separable $\iff$ $f' \neq 0$. This is always true if $char(K) = 0$. If $char(K) = p > 0$, then f is separable $\iff$ $f(X) \neq g(X^p)$ for all $g \in K[X]$.
(2) If $char(K) = 0$, then K is perfect.
(3) If $char(K) = p$, then the field K is perfect $\iff$ $K = K^p$.
(4) Every finite field is perfect.

Proof. (1) The ideal $(f, f') \subset K[X]$ is principal, equal to (h) for a monic polynomial $h \in K[X]$ dividing both f and f'. If $f' = 0$, then $h = f$ and f is not separable, by Proposition 3.27(3). If $f' \neq 0$, then $\deg(h) \leq \deg(f') < \deg(f)$, which implies that $h = 1$, by irreducibility of f; thus f is separable, again by Proposition 3.27(3). If $char(K) = 0$, then $f' \neq 0$ for every non-constant polynomial f. If $char(K) = p > 0$, then $f' = 0 \iff f(X) = g(X^p)$ for some $g \in K[X]$.
(2) This is an immediate consequence of (1).
(3) If $K \neq K^p$, then K admits non-separable extensions constructed in Proposition 6.3. Conversely, if $K = K^p$ and if $g(X) = \sum_{i=0}^n a_i X^i \in K[X]$, then there exist $b_i \in K$ such that $a_i = b_i^p$ for all $i = 0, \ldots, n$, hence $g(X^p) = \sum_{i=0}^n b_i^p X^{pi} = h(X)^p$ for $h(X) = \sum_{i=0}^n b_i X^n \in K[X]$. In particular, an irreducible polynomial $f \in K[X]$ cannot be of the form $g(X^p)$, hence f is separable, by (1), and K is perfect.
(4) If $K = \mathbf{F}_{p^n} = \mathbf{F}_q$, then each element $a \in K$ satisfies $a = a^q$, hence is a p-th power: $a = b^p$ for $b = a^{p^{n-1}} \in K$.

(6.5) Theorem on the primitive element. *Let K be a field, let $L = K(\alpha, \beta_1, \ldots, \beta_n)$ be a finite extension of K. If all elements $\beta_1, \ldots, \beta_n$ are separable over K, then there exists $\gamma \in L$ (a "primitive element") such that $L = K(\gamma)$.*

Proof. If $|K| < \infty$, then $L = K(\gamma)$ for any generator γ of the finite cyclic group L^*. We can assume, therefore, that $|K| = \infty$. By induction on n, it is enough to treat the case $L = K(\alpha, \beta)$, where β is separable over K. Denote by $f \in K[X]$ (resp. $g \in K[X]$) the minimal polynomial of α (resp. of β) over K. We have

$$f(X) = \prod_{i=1}^{m}(X - \alpha_i), \qquad g(X) = \prod_{j=1}^{n}(X - \beta_j),$$

with α_i and β_j contained in a fixed splitting field of fg. Moreover, $\alpha = \alpha_1$, $\beta = \beta_1$ and the roots β_j of $g(X)$ are distinct.

We want to choose $t \in K$ for which $\gamma = \alpha + t\beta \in L$ will satisfy $L = K(\gamma)$. The idea is to try to express β in terms of γ using the two relations $f(\gamma - t\beta) = 0 = g(\beta)$. The polynomial $h(X) = f(\gamma - tX) \in K(\gamma)[X]$ has roots $(\gamma - \alpha_i)/t = \beta + (\alpha - \alpha_i)/t$. If $t \in K$ is chosen in such a way that $t(\beta - \beta_j) \neq \alpha_i - \alpha$ for all

$i = 1, \ldots, m$ and all $j = 2, \ldots, n$ (which is possible, since the field K is infinite and $\beta - \beta_j \neq 0$ for $j > 1$), then the only common root of g and h will be β, with multiplicity one. Both polynomials g and h have coefficients in $K(\gamma)$, which implies that their $\gcd(g, h) = X - \beta$ also has coefficients in $K(\gamma)$; thus $\beta \in K(\gamma)$. It follows that $\alpha = \gamma - t\beta \in K(\gamma)$, too, hence $L = K(\alpha, \beta) = K(\gamma)$.

(6.6) Theorem. *Let $K \subset L$ be a finite extension. There exists $\gamma \in L$ such that $L = K(\gamma) \iff$ there are only finitely many intermediate fields $K \subset M \subset L$.*

Proof. The statement is clear if $|K| < \infty$, by Theorem 4.4(2) and Proposition 4.1(5). Assume that $|K| = \infty$.

Proof of "$\Longrightarrow$": let $f \in K[X]$ be the minimal polynomial of γ over K. For a field M such that $K \subset M \subset L$ let $f_M \in M[X]$ be the minimal polynomial of γ over M and let $M' \subset M$ be the subfield generated over K by the coefficients of f_M. The polynomial f_M is irreducible in $M[X]$, hence also in $M'[X]$, which implies that $[L : M] = [M(\gamma) : M] = \deg(f_M) = [M'(\gamma) : M'] = [L : M']$; thus $M = M'$. There are only finitely many possibilities for f_M (they all divide f in $\overline{K}[X]$) and each f_M determines M, by the previous discussion.

Proof of "$\Longleftarrow$": the extension L/K is of finite type, $L = K(\alpha_1, \ldots, \alpha_n)$. By induction, it is enough to show that for any $\alpha, \beta \in L$ there exists $\gamma \in L$ such that $K(\alpha, \beta) = K(\gamma)$. The field K is infinite, but there are only finitely many possibilities for the fields $K \subset K(\alpha + t\beta) \subset L$ ($t \in K$). Therefore there exist $t \neq t' \in K$ such that $K(\alpha + t\beta) = K(\alpha + t'\beta) = M$. In particular, $\alpha + t\beta, \alpha + t'\beta \in M \Longrightarrow (t - t')\beta \in M \Longrightarrow \beta \in M \Longrightarrow \alpha = (\alpha + t\beta) - t\beta \in M$, hence $K(\alpha, \beta) = M = K(\alpha + t\beta)$.

(6.7) We would like to formulate a numerical criterion of separability. The idea is to use Proposition 5.1, which implies thst, for every field extension $K \hookrightarrow K'$ and every element α algebraic over K (with minimal polynomial $f \in K[X]$ over K), we have

$$|\mathrm{Hom}_{K-\mathrm{Alg}}(K(\alpha), K')| = |\{\beta \in K' \mid f(\beta) = 0\}| \leq \deg(f) = [K(\alpha) : K]. \tag{6.7.1}$$

Moreover, we have an equality for a suitable (finite) extension K' of $K \iff$ all roots of f are distinct $\iff$ α is separable over K.

We need to investigate analogues of (6.7.1) for more general finite extensions L/K, which are not necessarily of the form $K(\alpha)/K$.

(6.8) Proposition-Definition. *Let $K \hookrightarrow L$ be a finite extension, let $\overline{K}$ be a fixed algebraic closure of K. The* **separable degree** *of the extension L/K is defined to be*

$$[L : K]_s := \max_{[K':K] < \infty} |\mathrm{Hom}_{K-\mathrm{Alg}}(L, K')| = \left|\mathrm{Hom}_{K-\mathrm{Alg}}(L, \overline{K})\right|.$$

The discussion in 6.7 can be reformulated as $[K(\alpha) : K]_s \leq [K(\alpha) : K]$, with equality $\iff$ α is separable over K.

Proof. Every finite extension K' of K admits an embedding $K' \hookrightarrow \overline{K}$ (a homomorphism of K-algebras), by Theorem 5.6(2); this proves the inequality "$\leq$". Conversely, if $L = K(\alpha_1, \ldots, \alpha_n)$ and if $K' \subset \overline{K}$ is the splitting field of the product of the minimal polynomials of $\alpha_1, \ldots, \alpha_n$ over K, then the image of every homomorphism of K-algebras $\sigma : L \longrightarrow \overline{K}$ is contained in K', which proves the opposite inequality "$\geq$".

(6.9) Proposition. *If $K \hookrightarrow L \hookrightarrow M$ are finite extensions, then $[M : K]_s = [M : L]_s [L : K]_s$ and $[L : K]_s \leq [L : K]$.*

Proof. The restriction map

$$\mathrm{res} : \mathrm{Hom}_{K-\mathrm{Alg}}(M, \overline{K}) \longrightarrow \mathrm{Hom}_{K-\mathrm{Alg}}(L, \overline{K}), \qquad \sigma \mapsto \sigma|_L$$

is surjective, by Theorem 5.6(2). For fixed $\tau \in \mathrm{Hom}_{K-\mathrm{Alg}}(L, \overline{K})$, the fibre of res above τ is equal to $\mathrm{res}^{-1}(\tau) = \mathrm{Hom}_{L-\mathrm{Alg}}(M, \overline{L})$, where we consider $\overline{L} = \overline{K}$ as an L-algebra via τ. In particular,

$$\left|\mathrm{Hom}_{K-\mathrm{Alg}}(M, \overline{K})\right| = \sum_{\tau} \left|\mathrm{res}^{-1}(\tau)\right| = \left|\mathrm{Hom}_{L-\mathrm{Alg}}(M, \overline{L})\right| \left|\mathrm{Hom}_{K-\mathrm{Alg}}(L, \overline{K})\right|.$$

Wrriting $K \hookrightarrow L$ as a tower of simple extensions

$$K = K_0 \subset K_1 = K(\alpha_1) \subset \cdots \subset K_{i-1} \subset K_i = K_{i-1}(\alpha_i) = K(\alpha_1, \ldots, \alpha_i) \subset \cdots \subset K_n = K(\alpha_1, \ldots, \alpha_n) = L, \tag{6.9.1}$$

we obtain

$$[L : K]_s = \prod_{i=1}^{n} [K_i : K_{i-1}]_s = \prod_{i=1}^{n} [K_{i-1}(\alpha_i) : K_{i-1}]_s \leq \prod_{i=1}^{n} [K_{i-1}(\alpha_i) : K_{i-1}] = \prod_{i=1}^{n} [K_i : K_{i-1}] = [L : K]. \tag{6.9.2}$$

(6.10) Proposition. *Let $K \hookrightarrow L$ be a finite extension. The following properties are equivalent:*
(i) $L = K(\alpha_1, \ldots, \alpha_n)$, where each α_i is separable over K.
(ii) $[L : K]_s = [L : K]$.
(iii) The extension L/K is separable.
(iv) $L = K(\alpha)$, where α is separable over K.

Proof. The implication (iii) $\Longrightarrow$ (i) is automatic.
(i) $\Longrightarrow$ (ii): consider the tower (6.9.1). Each α_i is separable over K, hence over K_{i-1}; it follows that $[K_i : K_{i-1}]_s = [K_i : K_{i-1}]$ for all $i = 1, \ldots, n$, hence $[L : K]_s = [L : K]$ (cf. (6.9.2)).
(ii) $\Longrightarrow$ (iii): if there is $\alpha \in L$ which is not separable over K, then the tower of extensions $K \subset K(\alpha) \subset L$ satisfies $[K(\alpha) : K]_s < [K(\alpha) : K]$ and $[L : K(\alpha)]_s \leq [L : K(\alpha)]$, hence $[L : K]_s = [L : K(\alpha)]_s [K(\alpha) : K]_s < [L : K(\alpha)][K(\alpha) : K] = [L : K]$.
The implication (iv) $\Longrightarrow$ (i) is automatic.
(i) $\Longrightarrow$ (iv): Theorem 6.5 implies that $L = K(\alpha)$ for some $\alpha \in L$; α is separable over K by (iii).

(6.11) Corollary. *Let $K \hookrightarrow L \hookrightarrow M$ be algebraic extensions.*
(1) If $[M : K] < \infty$, then the extension M/K is separable $\iff$ M/L and L/K are separable.
(2) If L/K is a separable extension and α is separable over L, then α is separable over K.
(3) The extension M/K is separable $\iff$ M/L and L/K are separable.

Proof. (1) M/K is separable $\iff$ $[M : K]_s = [M : L]_s [L : K]_s = [M : K] = [M : L][L : K]$ $\iff$ $[M : L]_s = [M : L]$ and $[L : K]_s = [L : K]$ (since $[M : L]_s \leq [M : L]$ and $[L : K]_s \leq [L : K]$).
(2) Let $L' = K(a_1, \ldots, a_d) \subset L$ be the subfield generated over K by the coefficients $a_1, \ldots, a_d \in L$ of the minimal polynomial of α over L. By assumption, L'/K and $L'(\alpha)/L'$ are finite separable extensions, hence $L'(\alpha)/K$ is also separable, by (1).
(3) The implication "$\Longrightarrow$" is automatic. The converse implication follows from (2).

(6.12) Proposition. *If $K \hookrightarrow L$ is a field extension, then $\{\alpha \in L \mid \alpha$ is algebraic and separable over $K\}$ is a subfield of L. In particular, $K^{\mathrm{sep}} = \{\alpha \in \overline{K} \mid \alpha$ is algebraic and separable over $K\}$ is a subfield of $\overline{K}$, called the* **separable closure of** *K. Of course, $K^{\mathrm{sep}} = \overline{K}$ if K is perfect (in particular, if $char(K) = 0$).*

Proof. If $\alpha, \beta \in L$ are algebraic and separable over K, then the extension $K(\alpha, \beta)/K$ is separable by Proposition 6.10, hence its elements $\alpha \pm \beta, \alpha\beta, \alpha/\beta$ (if $\beta \neq 0$) are separable over K.

(6.13) Proposition. *If L/K is a finite extension, then the field $L_s = \{\alpha \in L \mid \alpha$ is separable over $K\}$ $(K \hookrightarrow L_s \hookrightarrow L)$ has the following properties.*
(1) The extension L_s/K is separable.
(2) If $\beta \in L$ is separable over L_s, then $\beta \in L_s$.
(3) Assume that $char(K) = p > 0$. For each $\alpha \in L$ there exists $n \geq 0$ such that $\alpha^{p^n} \in L_s$ (the extension L/L_s is **purely inseparable***).*
(4) $[L : L_s]_s = 1$ and $[L_s : K] = [L : K]_s$.

Proof. The statements (1) and (2) follow from Proposition 6.10 and Corollary 6.11(2), respectively. In (3), let $f \in K[X]$ be the minimal polynomial of α over K. There exists $n \geq 0$ such that $f(X) = g(X^{p^n})$, where $g \in K[X]$ and $g(X) \neq h(X^p)$ for any $h \in K[X]$. The polynomial g is irreducible in $K[X]$ (since f is), hence

separable, by Proposition 6.4(1). As a result, $\beta = \alpha^{p^n} \in L$ lies in L_s, since $g(\beta) = f(\alpha) = 0$. In (4) there is nothing to prove if $char(K) = 0$ (when $L_s = L$). If $char(K) = p > 0$, fix an embedding $L_s \hookrightarrow \overline{L}_s = \overline{K}$. Let $\sigma \in \mathrm{Hom}_{L_s-\mathrm{Alg}}(L, \overline{L}_s)$. For any $\alpha \in L$ and $\beta = \alpha^{p^n} \in L_s$ as in (3) we have $\sigma(\alpha)^{p^n} = \beta$, which uniquely determines $\sigma(\beta) \in \overline{L}_s$ (cf. Proposition 6.3). It follows that σ is unique, hence $[L : L_s]_s = 1$. Finally, $[L : K]_s = [L : L_s]_s[L_s : K]_s = 1 \cdot [L_s : K] = [L_s : K]$, since L_s/K is separable.

(6.14) Exercise. *Consider the tower of extensions $K = \mathbf{F}_2(a, b) \subset K(\alpha) \subset K(\beta) = L$, where a, b are variables, $\alpha^2 + a\alpha + b = 0$ and $\beta^2 = \alpha$.*
(1) $[L : K] = 4$, $[L : K]_s = 2$ and $L_s = K(\alpha)$.
(2) An element $\gamma \in L$ satisfies $\gamma^2 \in K \iff \gamma \in K$.
(3) There is no intermediate field $K \subsetneq F \subset L$ purely inseparable over K.

(6.15) Exercise. *Let $K \hookrightarrow L$ be a finite extension of fields of characteristic $p > 0$.*
(1) $\alpha \in L$ is separable over $K \iff K(\alpha^p) = K(\alpha)$.
(2) $L_s = \bigcap_{n \geq 0} \varphi^n(L)$.

7. Separability, norm, trace and discriminants

In this section we investigate in more detail characteristic polynomials for separable extensions.

(7.1) Example 3.21, in which we discussed the regular representation of $L = K(\sqrt{d})$ over K in the basis $\{1, \sqrt{d}\}$, can be reformulated as follows. For each $\beta = a + b\sqrt{d} \in L$ $(a, b \in K)$ the matrix $M(\beta) = \begin{pmatrix} a & db \\ b & a \end{pmatrix}$ has eigenvalues $\beta = a + b\sqrt{d}$ and $\beta' = a - b\sqrt{d}$.
(i) If $char(K) = 2$, then the extension L/K is not separable and $\beta = \beta'$. In particular, for each $\beta \in L \setminus K$ the matrix $M(\beta)$ is not diagonalisable.
(ii) If $char(K) \neq 2$, then the extension L/K is separable and for every field extension $L \hookrightarrow K'$ there exist two homomorphisms of K-algebras $\sigma, \sigma' : L \longrightarrow K'$; namely, $\sigma(\beta) = \beta$ and $\sigma'(\beta) = \beta'$.

The eigenvectors $\begin{pmatrix} \sqrt{d} \\ 1 \end{pmatrix}, \begin{pmatrix} -\sqrt{d} \\ 1 \end{pmatrix}$ of $M(\beta)$ are the same for all $\beta \in L$. As a result, we obtain a simultaneous diagonalisation of all $M(\beta)$ over any field $K' \supset L$:

$$\forall \beta \in L \quad S^{-1}M(\beta)S = \begin{pmatrix} \beta & 0 \\ 0 & \beta' \end{pmatrix}, \qquad S = \begin{pmatrix} \sqrt{d} & -\sqrt{d} \\ 1 & 1 \end{pmatrix} \in GL_2(K'). \tag{7.1.1}$$

(7.2) Proposition. *Let $K \hookrightarrow L$ be a finite separable extension, let $K \hookrightarrow K'$ be a field extension such that $|\mathrm{Hom}_{K-\mathrm{Alg}}(L, K')| = [L : K] = n$ (for example, $K' = K^{\mathrm{sep}}$). Then we have, for each $\beta \in L$,*

$$P_{L/K,\beta}(X) = \prod_{i=1}^{n}(X - \sigma_i(\beta)), \qquad \mathrm{Tr}_{L/K}(\beta) = \sum_{i=1}^{n} \sigma_i(\beta), \qquad N_{L/K}(\beta) = \prod_{i=1}^{n} \sigma_i(\beta),$$

where $\{\sigma_1, \ldots, \sigma_n\} = \mathrm{Hom}_{K-\mathrm{Alg}}(L, K')$.

Proof. Let $f \in K[X]$ be the minimal polynomial of β over K. According to Theorem 3.22,

$$P_{L/K,\beta}(X) = P_{K(\beta)/K,\beta}(X)^e = f(X)^e,$$

where $e = [L : K(\beta)] = n/d$, $d = \deg(f) = [K(\beta) : K]$. The polynomial f has d distinct roots $\beta_1, \ldots, \beta_d \in K'$:

$$f(X) = (X - \beta_1) \cdots (X - \beta_d),$$

where $\beta_j = \tau_j(\beta)$, $\mathrm{Hom}_{K-\mathrm{Alg}}(K(\beta), K') = \{\tau_1, \ldots, \tau_d\}$. As in the proof of Proposition 6.9, the restriction map $\mathrm{res} : \mathrm{Hom}_{K-\mathrm{Alg}}(L, K') \longrightarrow \mathrm{Hom}_{K-\mathrm{Alg}}(K(\beta), K')$ is surjective and each fibre $\mathrm{res}^{-1}(\tau_j)$ has cardinality e. It follows that each root β_j $(1 \leq j \leq d)$ appears with multiplicity e among the $n = de$ values $\sigma_1(\beta), \ldots, \sigma_n(\beta)$, hence

$$P_{L/K,\beta}(X) = (X-\beta_1)^e \cdots (X-\beta_d)^e = \prod_{i=1}^{n}(X-\sigma_i(\beta)) =$$
$$= X^n - \mathrm{Tr}_{L/K}(\beta)X^{n-1} + \cdots + (-1)^n N_{L/K}(\beta).$$

(7.3) The field embeddings $\sigma_i : L \hookrightarrow K'$ in Proposition 7.2 can be described explicitly as follows. According to Proposition 6.10 we have $L = K(\alpha)$ for some $\alpha \in L$, separable over K. Let $f \in K[X]$ ($\deg(f) = n$) be the minimal polynomial of α over K. If K' is any extension of K containing a splitting field of f over K, then

$$\mathrm{Hom}_{K-\mathrm{Alg}}(L, K') = \{\sigma_1, \ldots, \sigma_n\}, \qquad \forall g(X) \in K[X] \quad \sigma_i : g(\alpha) \mapsto g(\alpha_i),$$

where $f(X) = \prod_{i=1}^{n}(X-\alpha_i)$, $\alpha_i \in K'$. In particular,

$$\forall g(X) \in K[X] \qquad \mathrm{Tr}_{L/K}(g(\alpha)) = \sum_{i=1}^{n} g(\alpha_i), \qquad N_{L/K}(g(\alpha)) = \prod_{i=1}^{n} g(\alpha_i). \tag{7.3.1}$$

(7.4) Proposition. *Let K be a field, let $f \in K[X]$ be an irreducible separable monic polynomial of degree $\deg(f) = n \geq 1$, let α be a root of f. Then*

$$N_{K(\alpha)/K}(f'(\alpha)) = (-1)^{n(n-1)/2}\,\mathrm{disc}(f).$$

Proof. Let K' and $\mathrm{Hom}_{K-\mathrm{Alg}}(L, K') = \{\sigma_1, \ldots, \sigma_n\}$ be as in 7.3. The formula (7.3.1) for $g = f'$ yields

$$N_{K(\alpha)/K}\,(f'(\alpha)) = \prod_{i=1}^{n} f'(\alpha_i) = \prod_{i=1}^{n}\prod_{j\neq i}(\alpha_i - \alpha_j) = (-1)^{n(n-1)/2}\prod_{i<j}(\alpha_i-\alpha_j)^2 = (-1)^{n(n-1)/2}\,\mathrm{disc}(f).$$

(7.5) Example. Let a, b be variables, $K = \mathbf{Q}(a,b)$, $f(X) = X^3 + aX + b \in K[X]$, $K(\alpha) = K[X]/(f)$, $\alpha = \overline{X}$ (= the image of X). In the basis $1, \alpha, \alpha^2$ of $K(\alpha)/K$, the matrix $M(f'(\alpha)) = M(3\alpha^2 + a)$ is equal to

$$\begin{pmatrix} a & -3b & 0 \\ 0 & -2a & -3b \\ 3 & 0 & -2a \end{pmatrix},$$

since

$$(3\alpha^2+a)\cdot 1 = a\cdot 1 + 0\cdot\alpha + 3\cdot\alpha^2$$
$$(3\alpha^2+a)\cdot\alpha = 3\alpha^3 + a\alpha = 3(-a\alpha - b) + a\alpha = -3b\cdot 1 - 2a\cdot\alpha + 0\cdot\alpha^2$$
$$(3\alpha^2+a)\cdot\alpha^2 = 0\cdot 1 - 3b\cdot\alpha - 2a\cdot\alpha^2;$$

thus

$$\mathrm{disc}(X^3+aX+b) = (-1)^{3\cdot 2/2}\begin{vmatrix} a & -3b & 0 \\ 0 & -2a & -3b \\ 3 & 0 & -2a \end{vmatrix} = -4a^3 - 27b^2.$$

(7.6) The trace form. Let $K \hookrightarrow L$ be a finite extension. The trace form attached to L/K is the symmetric K-bilinear form

$$L \times L \longrightarrow K, \qquad x, y \mapsto \mathrm{Tr}_{L/K}(xy). \tag{7.6.1}$$

It is non-degenerate $\iff$ $\mathrm{Tr}_{L/K} : L \longrightarrow K$ is a non-zero map (if $\mathrm{Tr}_{L/K}(b) \neq 0$, then $\mathrm{Tr}_{L/K}(x \cdot bx^{-1}) \neq 0$, for any $x \in L^*$). In a fixed basis $\omega_1, \ldots, \omega_n$ of L/K, the matrix of (7.6.1) is equal to

$$M = (M_{ij}) \in M_n(K), \qquad M_{ij} = \mathrm{Tr}_{L/K}(\omega_i \omega_j). \tag{7.6.2}$$

The **discriminant of** $\omega_1, \ldots, \omega_n$ is defined as the determinant

$$D(\omega_1, \ldots, \omega_n) := \det(M) \in K.$$

The following result is very important (see Theorem IV.4.2 below).

(7.7) Theorem. *Let $K \hookrightarrow L$ be an extension of degree $[L : K] = n < \infty$.*
(1) If L/K is not separable, then $\mathrm{Tr}_{L/K}$ is the zero map and the trace form (7.6.1) is identically zero.
(2) If L/K is separable, fix $\alpha \in L$ such that $L = K(\alpha)$ and let $f \in K[X]$ be the minimal polynomial of α over K. Then

$$D(1, \alpha, \ldots, \alpha^{n-1}) = \mathrm{disc}(f) \neq 0.$$

In particular, the trace form (7.6.1) is non-degenerate.

Proof. (1) Fix $\alpha \in L$ which is not separable over K. It follows from Proposition 6.13 that $[K(\alpha) : K(\alpha^p)] = p$ and that $1, \alpha, \ldots, \alpha^{p-1}$ is a basis of $K(\alpha)/K(\alpha^p)$. The transitivity rule

$$\mathrm{Tr}_{L/K} = \mathrm{Tr}_{K(\alpha^p)/K} \circ \mathrm{Tr}_{K(\alpha)/K(\alpha^p)} \circ \mathrm{Tr}_{L/K(\alpha)}$$

(cf. Exercise 3.23(iii)) implies that it is enough to show that $\mathrm{Tr}_{K(\alpha)/K(\alpha^p)} = 0$. Let $\beta = a_0 + a_1\alpha + \cdots + a_{p-1}\alpha^{p-1} \in K(\alpha)$ ($a_i \in K(\alpha^p)$). The diagonal elements of the matrix $M(\beta)$ (in the basis $1, \alpha, \ldots, \alpha^{p-1}$) are all equal to a_0, hence $\mathrm{Tr}_{K(\alpha)/K(\alpha^p)}(\beta) = pa_0 = 0$.
(2) Let K' be as in 7.3; then $f(X) = \prod_{k=1}^{n}(X - \alpha_k) \in K'[X]$. The entries of the matrix M from (7.6.2) in the basis $1, \alpha, \ldots, \alpha^{n-1}$ are equal to $\mathrm{Tr}_{L/K}(\alpha^i \cdot \alpha^j) = \sum_{k=1}^{n} \alpha_k^i \cdot \alpha_k^j$, by (7.3.1). As a result, $M = A \cdot {}^tA$, where $A \in M_n(K')$ is the matrix with entries $A_{ik} = \alpha_k^{i-1}$ ($1 \leq i, k \leq n$). The Vandermonde formula states that

$$\det(A) = \pm \prod_{1 \leq k < l \leq n} (\alpha_k - \alpha_l),$$

hence

$$\det(M) = \det(A)^2 = \mathrm{disc}(f).$$

(7.8) Tensor products of K-algebras. Let K be a field. Recall from the course Algebra 1 the notion of a tensor product $A \otimes_K B$ of two K-vector spaces A and B. A general element of $A \otimes_K B$ is a finite sum of decomposable elements $a \otimes b$, where

$$A \times B \longrightarrow A \otimes_K B, \qquad (a, b) \mapsto a \otimes b$$

is a universal K-bilinear map. In particular, $(\lambda a) \otimes b = a \otimes (\lambda b) = \lambda(a \otimes b)$ for all $\lambda \in K$, $a \in A$ and $b \in B$.

If we write $A = \bigcup A_\alpha$ as a union of its finite-dimensional K-subspaces, then $A \otimes_K B = \bigcup (A_\alpha \otimes_K B)$.

If A and B are (commutative) K-algebras, so is $A \otimes_K B$, with product given by $(a \otimes b)(a' \otimes b') = aa' \otimes bb'$. Note that $A \otimes_K B$ is simultaneously an A-algebra via $A \longrightarrow A \otimes_K B$ ($a \mapsto a \otimes 1$) and a B-algebra via $B \longrightarrow A \otimes_K B$ ($b \mapsto 1 \otimes b$). The two maps coincide on $K = K \cdot 1 \otimes 1$, since $\lambda \otimes 1 = 1 \otimes \lambda$ for all $\lambda \in K$.

Our aim is to generalise the following isomorphism.

(7.9) Exercise. *The map $\mathbf{C} \otimes_{\mathbf{R}} \mathbf{C} \longrightarrow \mathbf{C} \times \mathbf{C}$, $a \otimes b \mapsto (ab, \overline{a}b)$ defines an isomorphism of $\mathbf{R}$-algebras.*

(7.10) Tensor products of fields (the separable case). Let $K \hookrightarrow L$ be a finite **separable** extension. Fix $\alpha \in L$ such that $L = K(\alpha)$ and let $f \in K[X]$ be the minimal polynomial of α over K (of course, f is separable). We identify $L = K(\alpha)$ with $K[X]/(f)$ via the evaluation isomorphism of K-algebras $\overline{\mathrm{ev}}_\alpha : K[X]/(f) \xrightarrow{\sim} K(\alpha)$.

For any field extension $K \hookrightarrow K'$ we obtain an isomorphism of K'-algebras

$$L \otimes_K K' = (K[X] \otimes_K K')/(f \otimes 1) = K'[X]/(f) \xrightarrow{\sim} \prod_i K'[X]/(f_i), \tag{7.10.1}$$

where we have decomposed $f = \prod_i f_i \in K'[X]$ into a product of at most $[L : K] = \deg(f) = n$ **distinct** monic irreducible polynomials $f_i \in K'[X]$. Each quotient $K'[X]/(f_i)$ is a field, finite over K'. In particular, the ring $L \otimes_K K'$ is reduced.

If K' contains a splitting field over K of f, then $f_i = X - \alpha_i$, where $\alpha_1, \ldots, \alpha_n \in K'$ are the (distinct) roots of f. Composing (7.10.1) with the evaluation isomorphisms $\overline{\mathrm{ev}}_{\alpha_i} : K'[X]/(X - \alpha_i) \xrightarrow{\sim} K'$ we obtain an isomorphism of K'-algebras

$$L \otimes_K K' \xrightarrow{\sim} \prod_{i=1}^{n} K', \qquad \forall g(X) \in K[X]\, \forall b \in K' \quad g(\alpha) \otimes b \mapsto (g(\alpha_i) \otimes b)_{1 \le i \le n}. \tag{7.10.2}$$

The latter isomorphism can be reformulated in terms of the embeddings $\{\sigma_1, \ldots, \sigma_n\} = \mathrm{Hom}_{K-\mathrm{Alg}}(L, K')$ from 7.3 ($\sigma_i(\alpha) = \alpha_i$) as

$$L \otimes_K K' \xrightarrow{\sim} (K')^{\mathrm{Hom}_{K-\mathrm{Alg}}(L,K')}, \qquad a \otimes b \mapsto (\sigma \mapsto \sigma(a)b) \tag{7.10.3}$$

(in the special case $K = \mathbf{R}$ and $L = K' = \mathbf{C}$ we obtain the map from Exercise 7.9). The isomorphism (7.10.3) gives a simultaneous diagonalisation of all matrices $M(\beta) \in M_n(K)$ ($\beta \in L$) with respect to a fixed basis $\omega_1, \ldots, \omega_n$ of L/K. Indeed, if $S \in GL_n(K')$ is the matrix which interchanges the K'-basis $\omega_1 \otimes 1, \ldots, \omega_n \otimes 1$ of $L \otimes_K K'$ with the basis consisting of the idempotents $e_{\sigma_1}, \ldots, e_{\sigma_n}$ corresponding to the decomposition of the R.H.S. of (7.10.3), then

$$\forall \beta \in L \quad S^{-1} M(\beta) S = \begin{pmatrix} \sigma_1(\beta) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_n(\beta) \end{pmatrix}.$$

(7.11) Tensor products of fields (the non-separable case). Let $K \hookrightarrow L$ be a non-separable finite extension, $p = char(K)$. As in the proof of Theorem 7.7(1), there exists $\alpha \in L$ such that $K \subset K_1 = K(\alpha^p) \subsetneq K(\alpha) \subset L$. In concrete terms, $\alpha^p = a \in K_1 \setminus K_1^p$ and we identify $K(\alpha)$ with $K_1[X]/(X^p - a)$, $\alpha = \overline{X} = X \pmod{(X^p - a)}$.

The tensor product $K(\alpha) \otimes_K K(\alpha)$ is a subring of $L \otimes_K L$ and the canonical ring homomorphism $K(\alpha) \otimes_K K(\alpha) \longrightarrow K(\alpha) \otimes_{K_1} K(\alpha)$ ($a \otimes b \mapsto a \otimes b$) is surjective. We have

$$K(\alpha) \otimes_{K_1} K(\alpha) = K_1[X, Y]/(X^p - a, Y^p - a) = \bigoplus_{i,j=0}^{p-1} K_1 \cdot \overline{X}^i \overline{Y}^j,$$

where $\overline{X} = \alpha \otimes 1$ (resp. $\overline{Y} = 1 \otimes \alpha$) is the image of X (resp. Y). In particular, $\overline{X} - \overline{Y} = \alpha \otimes 1 - 1 \otimes \alpha$ is a non-zero element of $K(\alpha) \otimes_{K_1} K(\alpha)$ satisfying

$$(\overline{X} - \overline{Y})^p = (\alpha \otimes 1 - 1 \otimes \alpha)^p = \alpha^p \otimes 1 - 1 \otimes \alpha^p = a \otimes 1 - 1 \otimes a = 0;$$

thus $\overline{X} - \overline{Y}$ is a nilpotent element and the ring $K(\alpha) \otimes_{K_1} K(\alpha)$ is not reduced (which implies that $L \otimes_K L$ is not reduced, either).

(7.12) Proposition. *Let $K \hookrightarrow L$ be an algebraic extension.*
(1) If L/K is separable, then the ring $L \otimes_K K'$ is reduced for any field K' containing K; it is a finite product of fields of finite degree over K' if $[L : K] < \infty$.
(2) If L/K is not separable, then the ring $L \otimes_K K'$ is not reduced for any field K' containing L (in particular, for $K' = \overline{K}$).

Proof. If $[L : K] < \infty$, the statement (1) (resp. (2)) follows from 7.10 (resp. from 7.11). In general write $L = \bigcup L_\alpha$ as a union of its subfields which have finite degree over K and use the fact that $L \otimes_K K' = \bigcup (L_\alpha \otimes_K K')$.

8. Normal extensions

(8.1) We need an abstract characterisation of field extensions of the form $K \hookrightarrow K(\alpha_1, \ldots, \alpha_n) = L$, where $f(X) = (X - \alpha_1) \cdots (X - \alpha_n) \in K[X] \setminus K$. The example we should keep in mind is that of 3.12(iv), when $K = \mathbf{Q}$, $f = X^3 - 2$, $\alpha_1 = \sqrt[3]{2} \in \mathbf{R}$, $\alpha_2 = \rho\alpha_1$, $\alpha_3 = \rho^2\alpha_1$, $\rho = e^{2\pi i/3}$ and $L = \mathbf{Q}(\alpha_1, \rho)$. The three subfields $\mathbf{Q}(\alpha_j) \subset L$ are distinct, but isomorphic to each other.

(8.2) Definition. *Let $K \hookrightarrow L$ be a field extension, let $\alpha \in L$ be algebraic over K. The* **conjugates of α over** *K are the roots (in some splitting field of f) of the minimal polynomial $f \in K[X]$ of α over K.*

(8.3) Examples. (i) If $K = \mathbf{Q}$, $L = \mathbf{C}$ and $\alpha = \sqrt[4]{2}$, then $f = X^4 - 2$ and the conjugates of $\sqrt[4]{2}$ over $\mathbf{Q}$ are $\pm\sqrt[4]{2}, \pm i\sqrt[4]{2}$.
(ii) If $K = \mathbf{Q}(\sqrt{2})$, $L = \mathbf{C}$ and $\alpha = \sqrt[4]{2}$, then $f = X^2 - \sqrt{2}$ and the conjugates of $\sqrt[4]{2}$ over $\mathbf{Q}(\sqrt{2})$ are $\pm\sqrt[4]{2}$.

(8.4) Proposition-Definition. *A finite extension $K \hookrightarrow L$ is* **normal** *if the following equivalent conditions hold.*
(i) Any irreducible monic polynomial $f \in K[X]$ with a root $\alpha \in L$ splits in $L[X]$: $f = (X - \alpha_1) \cdots (X - \alpha_n)$, $\alpha_i \in L$, $\alpha = \alpha_1$ ["L contains with each element all of its conjugates over K."]
(ii) L is a splitting field over K of some polynomial $g \in K[X] \setminus K$.
(iii) For any field extension $K \hookrightarrow K'$ and any pair of homomorphisms $\sigma, \tau \in \mathrm{Hom}_{K-\mathrm{Alg}}(L, K')$ we have $\sigma(L) = \tau(L)$.
(iv) The same property for $K' = \overline{K}$ (a fixed algebraic closure of K).

Proof. (i) $\Longrightarrow$ (ii): let $\alpha_1, \ldots, \alpha_d$ be a basis of L/K, let $f_i \in K[X]$ be the minimal polynomial of α_i over K, let $g = f_1 \cdots f_d$. The condition (i) implies that all roots of g are contained in $L = K(\alpha_1, \ldots, \alpha_d)$, which means that L is a splitting field of g over K.
(ii) $\Longrightarrow$ (iii): Assume that $L = K(\alpha_1, \ldots, \alpha_n)$, where $(X - \alpha_1) \cdots (X - \alpha_n) = g(X) \in K[X]$ and $n \geq 1$ (note that the roots α_i are not necessarily distinct). If we apply σ and τ to g, we obtain

$$\prod_{j=1}^{n} (X - \sigma(\alpha_j)) = {}^{\sigma}g(X) = g(X) = {}^{\tau}g(X) = \prod_{j=1}^{n} (X - \tau(\alpha_j)) \in K'[X],$$

which means that the sets of roots (possibly with multiplicities) $\{\sigma(\alpha_j)\} \subset K'$ and $\{\tau(\alpha_j)\} \subset K'$ coincide. As a result, $\sigma(L) = K(\sigma(\alpha_1), \ldots, \sigma(\alpha_n)) = K(\tau(\alpha_1), \ldots, \tau(\alpha_n)) = \tau(L)$.
The implication (iii) $\Longrightarrow$ (iv) is automatic, so it remains to prove that (iv) $\Longrightarrow$ (i). Let $\alpha \in L$; its minimal polynomial $f \in K[X]$ over K splits in $\overline{K}[X]$ as $f = \prod_{j=1}^{n}(X - \sigma(\alpha_i))$. For each of its roots $\alpha_i \in \overline{K}$ there exists a (unique) morphism of K-algebras $\sigma_i : K(\alpha) \longrightarrow \overline{K}$ such that $\sigma_i(\alpha) = \alpha_i$, by Proposition 5.1. Furthermore, there exists a morphism of K-algebras $\tau_i : L \longrightarrow \overline{K}$ such that $\tau_i|_{K(\alpha)} = \sigma_i$, by Theorem 5.6(2). The condition (iv) implies that $\tau_i(L) = \tau_1(L)$, hence $\alpha_i = \tau_i(\alpha) \in \tau_1(L)$ for all i.

(8.5) Proposition. *If L is a splitting field over K of a polynomial $f \in K[X]$ of degree $n \geq 1$, then the degree $[L : K]$ divides $n!$.*

Proof. We argue by induction on n, the case $n = 1$ being trivial. Assume that $n > 1$. If f is irreducible in $K[X]$, let $\alpha \in L$ be a root of f. We have $[K(\alpha) : K] = n$ and L is a splitting field of the polynomial $f(X)/(X - \alpha)$ over $K(\alpha)$. By induction, $[L : K(\alpha)]$ divides $(n-1)!$, hence $[L : K]$ divides $n \cdot (n-1)! = n!$.
If $f = gh$ is reducible in $K[X]$ ($g, h \in K[X]$, $1 \leq \deg(g) = d < n$), let $F \subset L$ be a splitting field of g over K; then L is a splitting field of h over F. By induction, $[F : K]$ divides $d!$ and $[L : F]$ divides $(n-d)!$, hence $[L : K]$ divides $d!(n-d)!$, which in turn divides $n!$.

(8.6) Proposition. *Let $K \hookrightarrow L \hookrightarrow \Omega$ be field extensions, with L/K finite and Ω algebraically closed. There exists a finite extension of L contained in Ω which is normal over K. The intersection of all such extensions is the smallest finite extension of L in Ω which is normal over K; it is called* **the normal closure of L/K in Ω.**

Proof. If $\alpha_1, \ldots, \alpha_d$ is a basis of L/K and $f_i \in K[X]$ is the minimal polynomial of α_i over K, then the subfield of Ω generated over K by the roots of $g = f_1 \cdots f_d$ in Ω is a finite extension of L which is normal over K. By construction, it is contained in any other finite extension of L inside Ω which is normal over K.

9. Galois extensions

(9.1) Proposition-Definition. *Let $K \hookrightarrow L$ be a field extension.*
(1) The set of **field automorphisms of L over K**, $\mathrm{Aut}(L/K) = \{\sigma \in \mathrm{Hom}_{K-\mathrm{Alg}}(L,L) \mid \sigma(L) = L\}$, *is a group with respect to composition $\sigma\tau = \sigma \circ \tau$.*
(2) For any subgroup $G \subset \mathrm{Aut}(L/K)$ the set of fixed points $L^G = \{\alpha \in L \mid \forall g \in G \quad g(\alpha) = \alpha\}$ is a subfield of L containing K; we say that L^G is the **fixed field of** *G.*
(3) If $K \hookrightarrow L$ is an algebraic extension, then $\mathrm{Aut}(L/K) = \mathrm{Hom}_{K-\mathrm{Alg}}(L,L)$.
(4) If $K \hookrightarrow L$ is a finite extension, then $|\mathrm{Aut}(L/K)| \leq [L:K]$.

Proof. (2) By definition, $K \subset L^G$. If $\alpha, \beta \in L^G$ and $g \in G$, then $g(\alpha \pm \beta) = g(\alpha) \pm g(\beta) = \alpha \pm \beta$, $g(\alpha\beta) = g(\alpha)g(\beta) = \alpha\beta$ and $g(\alpha/\beta) = g(\alpha)/g(\beta)$ (if $\beta \neq 0$), hence $\alpha \pm \beta$, $\alpha\beta$ (and α/β if $\beta \neq 0$) all belong to L^G.
(3) We must show that any $\sigma \in \mathrm{Hom}_{K-\mathrm{Alg}}(L,L)$ (which is always injective) is surjective. Let $\alpha \in L$, let $f \in K[X]$ be the minimal polynomial of α over K. The set of roots $S = \{\beta \in L \mid f(\beta) = 0\} \subset L$ is finite and $\sigma(S) \subset S$. Injectivity of σ implies that $\sigma(S) = S$. Therefore $\alpha \in \sigma(L)$.
(4) Fix an embedding of K-algebras $L \hookrightarrow \overline{K}$; then

$$|\mathrm{Aut}(L/K)| \leq |\mathrm{Hom}_{K-\mathrm{Alg}}(L, \overline{K})| = [L:K]_s \leq [L:K].$$

(9.2) Exercise. *Let K be a field.*
(1) The map

$$\mathrm{Hom}_{K-\mathrm{Alg}}(K(X), K(X)) \longrightarrow K(X) \setminus K, \qquad \lambda \mapsto \lambda(X)$$

is bijective.
(2) For any $\lambda \in \mathrm{Hom}_{K-\mathrm{Alg}}(K(X), K(X))$ we have

$$[K(X) : \lambda(K(X))] = [K(X) : K(\lambda(X))] = \max(\deg(f), \deg(g)),$$

where $\lambda(X) = f/g$, $f, g \in K[X]$, $g \neq 0$, $\gcd(f,g) = (1)$.
(3) $\lambda \in \mathrm{Aut}(K(X)/K) \iff \lambda(X) = (aX+b)/(cX+d)$, $a,b,c,d \in K$, $ad - bc \neq 0$.
(4) $\mathrm{Aut}(K(X)/K) \xrightarrow{\sim} PGL_2(K)$ and $K(X)^{\mathrm{Aut}(K(X)/K)} = K$.

(9.3) Proposition-Definition. *A finite extension $K \hookrightarrow L$ is a* **Galois extension** *(with* **Galois group** $\mathrm{Gal}(L/K) := \mathrm{Aut}(L/K)$*) if the following equivalent conditions hold.*
(i) $|\mathrm{Aut}(L/K)| = [L:K]$.
(ii) The extension L/K is normal and separable.
(iii) L is a splitting field of a separable polynomial $f \in K[X]$.
(iv) There is an isomorphism of K-algebras $L \otimes_K L \xrightarrow{\sim} L^{[L:K]}$.
(v) $L^{\mathrm{Aut}(L/K)} = K$.

Proof. (i) $\iff$ (ii): as in the proof of Proposition 9.1(4), fix an embedding of K-algebras $L \hookrightarrow \overline{K}$. We have an equality $\mathrm{Aut}(L/K) = \mathrm{Hom}_{K-\mathrm{Alg}}(L,L)$ (by Proposition 9.1(3)) and two inequalities

$$|\mathrm{Hom}_{K-\mathrm{Alg}}(L,L)| \leq |\mathrm{Hom}_{K-\mathrm{Alg}}(L, \overline{K})| = [L:K]_s \tag{a}$$

$$[L:K]_s \leq [L:K], \tag{b}$$

with (a) (resp. (b)) being an equality $\iff$ the extension L/K is normal (resp. separable), by Proposition 8.4 (resp. by Proposition 6.10).
(ii) $\iff$ (iii): combine Proposition 8.4 with Proposition 6.10.
(ii) $\iff$ (iv): the calculations in 7.10–7.11 show that the ring $L \otimes_K L$ is reduced $\iff$ L/K is separable. If this is the case, then $L \xrightarrow{\sim} K[X]/(f)$ for an irreducible separable polynomial $f \in K[X]$ and we have $L \otimes_K L \xrightarrow{\sim} \prod_i L[X]/(f_i)$, where $f_i \in L[X]$ are distinct monic irreducible polynomials in $L[X]$ dividing f. In particular, each $L[X]/(f_i)$ is isomorphic to L $\iff$ all roots of f lie in L $\iff$ L/K is a normal extension.

(i) $\Longrightarrow$ (v): the field $K' = L^{\mathrm{Aut}(L/K)}$ contains K and satisfies $\mathrm{Aut}(L/K) = \mathrm{Aut}(L/K')$. We deduce from

$$[L : K'] \geq |\,\mathrm{Aut}(L/K')| = |\,\mathrm{Aut}(L/K)| \overset{(i)}{=} [L : K] \geq [L : K']$$

that $[L : K] = [L : K']$, hence $K = K'$.
(v) $\Longrightarrow$ (i): Artin's theorem 10.1 below (which does not use the implication (v) $\Longrightarrow$ (i)) states that $L/L^{\mathrm{Aut}(L/K)}$ is a Galois extension.

(9.4) Galois groups as permutation groups. If we write a (finite) Galois extension as a splitting field of a separable polynomial f, then the corresponding Galois group will have an old-fashioned description as a group of permutations of the set of roots of f. However, such a description is not canonical (see 9.6 below). The formal definition is as follows.

(9.5) Proposition-Definition. *Let K be a field, let $f \in K[X]$ be a* **separable** *polynomial of degree $\deg(f) = n \geq 1$. Fix a splitting field $L = K(A)$ of f over K, where A denotes the set of roots of f in L: $f = \prod_{\alpha \in A}(X - \alpha) \in L[X]$.*
(1) L/K is a Galois extension; denote by $G = \mathrm{Gal}(L/K)$ its Galois group.
(2) Each element $g \in G$ maps A to itself. The assignement

$$G \longrightarrow S_A, \qquad g \mapsto g|_A$$

given by the restriction to A defines an injective group homomorphism; denote its image by $\mathrm{Gal}(f) \subset S_A$.
(3) A choice of a numbering of the set of roots of f (i.e., a choice of a bijection $A \xrightarrow{\sim} \{1, 2, \ldots, n\}$) identifies S_A with S_n and G with a subgroup $\mathrm{Gal}(f) \subset S_n$. If we choose another numbering of the set of roots, $\mathrm{Gal}(f) \subset S_n$ will be replaced by a conjugate subgroup $\sigma\,\mathrm{Gal}(f)\,\sigma^{-1} \subset S_n$, where $\sigma \in S_n$ interchanges the two numberings.
(4) The group $\mathrm{Gal}(f) \subset S_A$ acts transitively on $A \iff$ the polynomial f is irreducible in $K[X]$.

Proof. (1) See Proposition 9.3. (2) If $\alpha \in A$ and $g \in G$, then $0 = g(f(\alpha)) = {}^g f(g(\alpha)) = f(g(\alpha))$, hence $g(\alpha) \in A$. The map $g : L \longrightarrow L$ is injective, and so is its restriction $g|_A : A \longrightarrow A$; thus $g|_A$ is bijective (since $|A| < \infty$). Finally, $(g \circ h)|_A = g|_A \circ h|_A$, which means that the restriction map is a group homomorphism.
(3) Any bijection $\tau : A \xrightarrow{\sim} \{1, 2, \ldots, n\}$ induces a group isomorphism $S_A \xrightarrow{\sim} S_n$, $g \mapsto \tau \circ g \circ \tau^{-1}$. If we choose another bijection $\tau' : A \xrightarrow{\sim} \{1, 2, \ldots, n\}$, then $\tau' \circ g \circ \tau'^{-1} = \sigma \circ \tau \circ g \circ \tau^{-1} \circ \sigma^{-1}$, where $\sigma = \tau' \circ \tau^{-1} \in S_n$.
(4) If $f = gh$, where $g, h \in K[X] \setminus K$, then $A = B \cup C$ and $B \cap C = \emptyset$, where B (resp. C) is the (non-empty) set of roots of g (resp. of h) in L. According to (2) the action of each $g \in G$ satisfies $g(B) = B$ and $g(C) = C$; thus $\mathrm{Gal}(f) \subset S_B \times S_C \subset S_A$ does not act transitively on A.

If f is irreducible in $K[X]$, we must show that for each $\alpha, \beta \in A$ there exists $g \in G$ such that $g(\alpha) = \beta$. The evaluation isomorphisms $\overline{\mathrm{ev}}_\alpha : K[X]/(f) \xrightarrow{\sim} K(\alpha) \subset L$ and $\overline{\mathrm{ev}}_\beta : K[X]/(f) \xrightarrow{\sim} K(\beta) \subset L$ define an isomorphism of K-algebras $\sigma = \overline{\mathrm{ev}}_\beta \circ (\overline{\mathrm{ev}}_\alpha)^{-1} : K(\alpha) \xrightarrow{\sim} K(\beta)$. Fix an embedding $L \hookrightarrow \overline{K}$ to an algebraic closure of K. According to Theorem 5.6(2), the embedding $K(\alpha) \xrightarrow{\sigma} K(\beta) \subset L \hookrightarrow \overline{K}$ can be extended to an embedding $\tau : L \longrightarrow \overline{K}$. The extension L/K is normal, which means that $\tau(L) = L \subset \overline{K}$. As a result, $\tau \in \mathrm{Hom}_{K-\mathrm{Alg}}(L, L) = G$. By construction, $\tau|_{K(\alpha)} = \sigma$, which implies that $\tau(\alpha) = \sigma(\alpha) = \beta$.

(9.6) Examples. (i) $K = \mathbf{Q}$, $f = X^3 - 2$, $n = 3$. Fix a complex root $\alpha_1 = \sqrt[3]{2} \in \mathbf{C}$ of f and set $\alpha_2 = \rho\alpha_1 = \rho\sqrt[3]{2}$, $\alpha_3 = \rho^2\alpha_1 = \rho^2\sqrt[3]{2}$ ($\rho = e^{2\pi i/3}$). The splitting field of f over $\mathbf{Q}$ inside $\mathbf{C}$ is equal to $L = \mathbf{Q}(\alpha_1, \alpha_2, \alpha_3) = \mathbf{Q}(\alpha_1, \rho) = \mathbf{Q}(\sqrt[3]{2}, \rho) \subset \mathbf{C}$.

The Galois group $\mathrm{Gal}(L/\mathbf{Q}) \xrightarrow{\sim} \mathrm{Gal}(f) \subset S_3$ has order dividing $|S_3| = 3! = 6$. On the other hand, $[L : \mathbf{Q}] = 6$, by Example 3.12(iv). The equality $|\mathrm{Gal}(L/\mathbf{Q})| = [L : \mathbf{Q}]$ then implies that $\mathrm{Gal}(f) = S_3$.

Denote by

$$\sigma_1 = (1)(23), \qquad \sigma_2 = (2)(13), \qquad \sigma_3 = (3)(12)$$

the three elements of S_3 of order 2 and by $H_j = \{1, \sigma_j\} \subset S_3$ the corresponding cyclic subgroups of S_3 of order two. For each $j = 1, 2, 3$ we have $H_j \subset \mathrm{Aut}(L/\mathbf{Q}(\alpha_j))$ and $2 = |H_j| \leq |\,\mathrm{Aut}(L/\mathbf{Q}(\alpha_j))| \leq [L : \mathbf{Q}(\alpha_j)] = 2$; thus $L/\mathbf{Q}(\alpha_j)$ is a Galois extension with Galois group $\mathrm{Gal}(L/\mathbf{Q}(\alpha_j)) = H_j$. On the other hand, $\mathbf{Q}(\alpha_j)/\mathbf{Q}$ is not a Galois extension, since $\alpha_k \notin \mathbf{Q}(\alpha_j)/\mathbf{Q}$ for $k \neq j$. The formulas

$$\sigma_1(\rho) = \sigma_1(\alpha_2)/\sigma_1(\alpha_1) = \alpha_3/\alpha_1 = \rho^2, \qquad \sigma_2(\rho) = \sigma_2(\alpha_2)/\sigma_2(\alpha_1) = \alpha_2/\alpha_3 = \rho^2 \implies \sigma_1\sigma_2(\rho) = \rho$$

imply that the subgroup $H = \{1, \tau, \tau^2\} = A_3 \subset S_3$, where

$$\tau = \sigma_1\sigma_2 = (123), \qquad \tau^2 = \sigma_2\sigma_1 = (132),$$

is contained in $\mathrm{Aut}(L/\mathbf{Q}(\rho))$. As $3 = |H| \leq |\mathrm{Aut}(L/\mathbf{Q}(\rho))| \leq [L : \mathbf{Q}(\rho)] = 3$, the extension $L/\mathbf{Q}(\rho)$ is also a Galois extension, with Galois group equal to $H = \mathrm{Gal}(L/\mathbf{Q}(\rho))$. Moreover, $\mathbf{Q}(\rho)/\mathbf{Q} = \mathbf{Q}(\sqrt{-3})/\mathbf{Q}$ is also a Galois extension, being the splitting field of the polynomial $X^2 + X + 1$ over $\mathbf{Q}$.
(ii) Under the assumptions of (i), let $\beta_1 = \sqrt[3]{2} + \rho$. The images

$$\{g(\beta_1) \in L \,|\, g \in \mathrm{Gal}(L/\mathbf{Q}) = S_3\} = \{\sqrt[3]{2}+\rho, \sqrt[3]{2}+\rho^2, \rho\sqrt[3]{2}+\rho, \rho\sqrt[3]{2}+\rho^2, \rho^2\sqrt[3]{2}+\rho, \rho^2\sqrt[3]{2}+\rho^2\} = \{\beta_1, \ldots, \beta_6\}$$

are distinct, which means that β_1 has at least $6 = [L : \mathbf{Q}]$ conjugates, hence $[\mathbf{Q}(\beta_1) : \mathbf{Q}] \geq 6$. It follows that $L = \mathbf{Q}(\beta_1)$, that $h(X) = (X - \beta_1) \cdots (X - \beta_6)$ is the minimal polynomial of β_j $(j = 1, \ldots, 6)$ over $\mathbf{Q}$ and that L is its splitting field over $\mathbf{Q}$. The action of $G = \mathrm{Gal}(L/\mathbf{Q})$ on the roots of $h(X)$ gives a realisation of $G = \mathrm{Gal}(L/\mathbf{Q})$ as a subgroup $\mathrm{Gal}(h) \subset S_6$; of course, $\mathrm{Gal}(h)$ is isomorphic to S_3, by (i).
(iii) $K = \mathbf{Q}$, $f = (X^2 - 2)(X^2 - 3)$, $n = 4$. The complex roots of f are $\alpha_1 = \sqrt{2}, \alpha_2 = -\sqrt{2}, \alpha_3 = \sqrt{3}, \alpha_4 = -\sqrt{3}$. The splitting field of f over $\mathbf{Q}$ is equal to $L = \mathbf{Q}(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = \mathbf{Q}(\sqrt{2}, \sqrt{3})$. Exercise 3.28(2) implies that $[L : \mathbf{Q}] = 4$, hence $|\mathrm{Gal}(L/\mathbf{Q})| = 4$. The action of any element $g \in \mathrm{Gal}(L/\mathbf{Q})$ is determined by the values $g(\sqrt{2}) = \pm\sqrt{2}$ and $g(\sqrt{3}) = \pm\sqrt{3}$; thus

$$\mathrm{Gal}(L/\mathbf{Q}) = \{g_{ab}, \,|\, a, b \in \mathbf{Z}/2\mathbf{Z}\}, \qquad g_{ab}(\sqrt{2}) = (-1)^a\sqrt{2}, \qquad g_{ab}(\sqrt{3}) = (-1)^b\sqrt{3}.$$

The formulas $g_{ab}^2 = 1$ and $g_{ab}g_{cd} = g_{a+c,b+d}$ imply that the map

$$\mathrm{Gal}(L/\mathbf{Q}) \xrightarrow{\sim} \mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}, \qquad g_{ab} \mapsto (a, b)$$

is a group isomorphism. Using the above numbering of roots we have $\mathrm{Gal}(L/\mathbf{Q}) \xrightarrow{\sim} S_{\{1,2\}} \times S_{\{3,4\}} \xrightarrow{\sim} S_2 \times S_2 \subset S_4$ and

$$g_{00} = (1)(2)(3)(4), \qquad g_{10} = (12)(3)(4), \qquad g_{01} = (1)(2)(34), \qquad g_{11} = (12)(34).$$

(iv) We saw in 6.1(i) that the same field $L = \mathbf{Q}(\sqrt{2}, \sqrt{3})$ can be written as $L = \mathbf{Q}(\sqrt{2} + \sqrt{3})$. In other words, L is the splitting field (over $\mathbf{Q}$) of the minimal polynomial $h(X) = X^4 - 10X^2 + 1$ (over $\mathbf{Q}$) of $\beta_1 = \sqrt{2} + \sqrt{3} \in L$; we have $h(X) = (X - \beta_1) \cdots (X - \beta_4)$, where $\beta_2 = -\sqrt{2} + \sqrt{3} = 1/\beta_1$, $\beta_3 = -\sqrt{2} - \sqrt{3} = -\beta_1$, $\beta_4 = \sqrt{2} - \sqrt{3} = -\beta_2 = -1/\beta_1$. The formulas

$$\begin{array}{lll} g_{10} : \beta_1 \longleftrightarrow \beta_2, & g_{10} : \beta_3 \longleftrightarrow \beta_4, & g_{10} = (12)(34) \\ g_{01} : \beta_1 \longleftrightarrow \beta_4, & g_{01} : \beta_2 \longleftrightarrow \beta_3, & g_{01} = (14)(23) \\ g_{11} : \beta_1 \longleftrightarrow \beta_3, & g_{11} : \beta_2 \longleftrightarrow \beta_4, & g_{11} = (13)(24) \end{array}$$

give another realisation of the Galois group $\mathrm{Gal}(L/\mathbf{Q})$ as a subgroup $\mathrm{Gal}(h) \xrightarrow{\sim} S_4$. The two subgroups $\mathrm{Gal}(f), \mathrm{Gal}(h) \subset S_4$ are both isomorphic to $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$, but they are not conjugate inside S_4, since the subgroup $\mathrm{Gal}(h) \subset S_4$ acts transitively on $\{1, 2, 3, 4\}$, but the subgroup $\mathrm{Gal}(f) \subset S_4$ from (iii) does not, as predicted by Proposition 9.5(4).

(9.7) Proposition. *Let $F \hookrightarrow K$ be an algebraic extension, let $\overline{F}$ be an algebraic closure of F. For any $\sigma, \tau \in \mathrm{Hom}_{F-\mathrm{Alg}}(K, \overline{F})$ there exists $g \in \mathrm{Aut}(\overline{F}/F)$ such that $\tau = g \circ \sigma$. In other words, $\mathrm{Aut}(\overline{F}/F)$ acts transitively on $\mathrm{Hom}_{F-\mathrm{Alg}}(K, \overline{F})$.*

Proof. Theorem 5.6(2) for $\sigma : K \hookrightarrow L = \overline{F}$ and $\tau : K \hookrightarrow \Omega = \overline{F}$ tells us that there exists a field homomorphism $g : \overline{F} \longrightarrow \overline{F}$ such that $g \circ \sigma = \tau$ (which implies that $g \in \mathrm{Hom}_{F-\mathrm{Alg}}(\overline{F}, \overline{F}) = \mathrm{Aut}(\overline{F}/F)$, using Proposition 9.1(3)).

10. The Galois correspondence

(10.1) Theorem (E. Artin). *Let $K \hookrightarrow L$ be a field extension, let $G \subset \mathrm{Aut}(L/K)$ be a finite group of field automorphisms of L over K. Then $[L : L^G] = |G|$, the extension L/L^G is a Galois extension and $G = \mathrm{Gal}(L/L^G)$.*

Proof. For any $\alpha \in L$ consider its orbit $O(\alpha) = \{g(\alpha) \mid g \in G\}$. The polynomial $f_\alpha(X) = \prod_{\beta \in O(\alpha)}(X - \beta) \in L[X]$ is separable and lies in $L^G[X]$, since the action of any $g \in G$ permutes the elements of $O(\alpha)$. In addition, $f_\alpha(\alpha) = 0$ and $\deg(f_\alpha) = |O(\alpha)| \leq |G|$. In particular, L/L^G is a separable algebraic extension and $[L^G(\alpha) : L^G] \leq |G|$ for every $\alpha \in L$. Lemma 10.2 below implies that L/L^G is a finite extension, of degree $[L : L^G] \leq |G|$. On the other hand, $K \subset L^G$, hence $G \subset \mathrm{Aut}(L/L^G) \subset \mathrm{Aut}(L/K)$ and $[L : L^G] \leq |G| \leq |\mathrm{Aut}(L/L^G)|$. Proposition 9.1(4) yields $|\mathrm{Aut}(L/L^G)| \leq [L : L^G]$, which means that we have equalities $|G| = |\mathrm{Aut}(L/L^G)| = [L : L^G]$ and $G = \mathrm{Aut}(L/L^G)$.

(10.2) Lemma. *Let L/F be a separable algebraic extension. If $n := \max\{[F(\alpha) : F] \mid \alpha \in L\} \in \mathbf{N} \cup \{\infty\}$ is finite, then $[L : F] = n$.*

Proof. Fix $\alpha \in L$ such that $[F(\alpha) : F] = n$. For each $\beta \in L$ there exists $\gamma \in F(\alpha, \beta)$ such that $F(\alpha, \beta) = F(\gamma)$ (thanks to Theorem 6.5). It follows that $[F(\alpha, \beta) : F] = [F(\gamma) : F] \leq n = [F(\alpha) : F]$, hence $\beta \in F(\alpha)$. Therefore $L = F(\alpha)$ and $[L : F] = n$.

(10.3) Example (The general polynomial equation of degree n). Let K be a field, $L = K(x_1, \ldots, x_n)$ the field of rational functions in n variables over K and $G = S_n \subset \mathrm{Aut}(L/K)$ acting on L as in 2.2. According to Corollary 2.9 we have $L^G = K(\sigma_1, \ldots, \sigma_n)$. Theorem 10.1 implies that $K(x_1, \ldots, x_n)/K(\sigma_1, \ldots, \sigma_n)$ is a Galois extension of degree $|S_n| = n!$ and its Galois group is equal to S_n.

(10.4) Proposition (Every finite group is a Galois group). *Let K be a field. For every finite group G there exists a Galois extension $F \hookrightarrow L$ with $F \supset K$ and $\mathrm{Gal}(L/F) \xrightarrow{\sim} G$.*

Proof. As observed by Cayley, the action of G on itself by left multiplication defines an injective group homomorphism $G \hookrightarrow S_G \xrightarrow{\sim} S_n$ (where $|G| = n$). As in 10.3, we let $L = K(x_1, \ldots, x_n)$ and $F = L^G$.

(10.5) Main Theorem of Galois theory. *Let L/K be a (finite) Galois extension, with Galois group $G = \mathrm{Gal}(L/K)$.*
(1) The formulas $F \mapsto H = \mathrm{Gal}(L/F)$, $H \mapsto F = L^H$ define mutually inverse bijections

$$\{F \text{ field} \mid K \subset F \subset L\} \longleftrightarrow \{\text{subgroups } H \subset G\}$$

(the "Galois correspondence"). In particular, L/F is a Galois extension.
(2) If F corresponds to H, then $[L : F] = |H|$, $[F : K] = |G|/|H| = (G : H)$.
(3) If F_1 (resp. F_2) corresponds to H_1 (resp. to H_2), then

$$F_1 \subset F_2 \iff H_1 \supset H_2.$$

(4) If F corresponds to H, then $\forall g \in G$ $g(F)$ corresponds to gHg^{-1}.
(5) If F corresponds to H, then

$$F/K \text{ is a Galois extension} \iff H \text{ is a normal subgroup of } G.$$

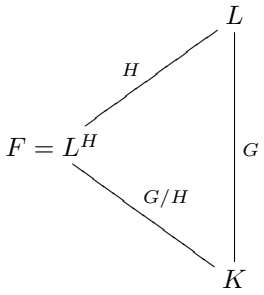
If this is the case, the map "restriction to F" defines a surjective group homomorphism

$$G = \mathrm{Gal}(L/K) \longrightarrow \mathrm{Gal}(F/K), \qquad g \mapsto g_{|F}$$

with kernel $H = \mathrm{Gal}(L/F)$, hence a group isomorphism

$$G/H = \mathrm{Gal}(L/K)/\mathrm{Gal}(L/F) \xrightarrow{\sim} \mathrm{Gal}(F/K).$$

This can be summed up by the following diagramme:



Proof. (1),(2) If $H \subset G$ is a subgroup and $F = L^H$, then $K = L^G \subset L^H = F \subset L$. According to Theorem 10.1, L/F is a Galois extension, $H = \mathrm{Gal}(L/F)$ and $[L : F] = |H|$. Conversely, if F a field such that $K \subset F \subset L$, then the extension L/F is separable and normal (since L/K is), so it is a Galois extension. Its Galois group $H = \mathrm{Gal}(L/F) = \mathrm{Aut}(L/F) \subset \mathrm{Aut}(L/K) = G$ satisfies $L^H = F$ and $[L : F] = |H|$. The degree $[F : K]$ is equal to $[L : K]/[L : F] = |G|/|H| = (G : H)$.

(3) This is clear.

(4) Let $g, h \in G$. An element $x \in L$ is fixed by $h \iff g(x)$ is fixed by ghg^{-1}, since $ghg^{-1}g(x) = gh(x)$; it follows that

$$L^{gHg^{-1}} = g(L^H).$$

(5) If $F/K = L^H/K$ is a Galois extension, then it is normal, which implies that $\forall g \in G\ g(F) = \mathrm{id}(F) = F$; it follows from (4) that $\forall g \in G\ \ gHg^{-1} = H$. Conversely, if H is a normal subgroup of G and $F = L^H$, then for each $g \in G$ the restriction $g_{|F} : F \longrightarrow g(F) \overset{(4)}{=} F$ is an element of $\mathrm{Aut}(F/K)$. The map $g \mapsto g|_F$ is a group homomorphism $r : G \longrightarrow \mathrm{Aut}(F/K)$ with kernel $\mathrm{Ker}(r) = \mathrm{Aut}(L/F) = H$. It follows that

$$[F : K] \geq |\,\mathrm{Aut}(F/K)| \geq |\mathrm{Im}(r)| = |G|/|\,\mathrm{Ker}(r)| = |G|/|H| = [F : K],$$

hence $[F : K] = |\,\mathrm{Aut}(F/K)| = |\mathrm{Im}(r)|$, which proves (5).

(10.6) Corollary. *If $F = L^H$, fix an embedding $L \hookrightarrow \overline{L} = \overline{K}$ and $g_1, \ldots, g_m \in G$ ($m = (G : H) = |G|/|H| = [F : K]$) such that $G = g_1H \cup \cdots \cup g_mH$; then*

$$\mathrm{Hom}_{F-\mathrm{Alg}}(L, \overline{L}) = \mathrm{Hom}_{F-\mathrm{Alg}}(L, L) = H, \qquad \mathrm{Hom}_{K-\mathrm{Alg}}(F, \overline{L}) = \mathrm{Hom}_{K-\mathrm{Alg}}(F, L) = \{g_1|_F, \ldots, g_m|_F\}$$

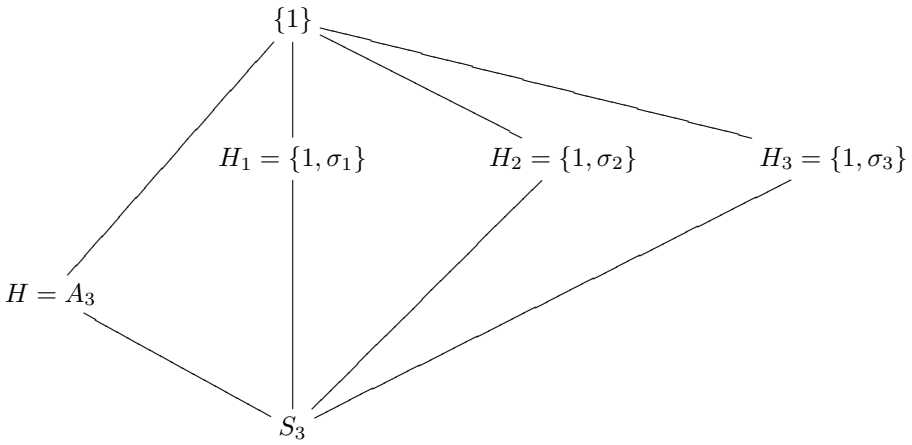
and

$$\forall \beta \in L \qquad \mathrm{Tr}_{L/F}(\beta) = \sum_{h \in H} h(\beta), \qquad N_{L/F}(\beta) = \prod_{h \in H} h(\beta)$$

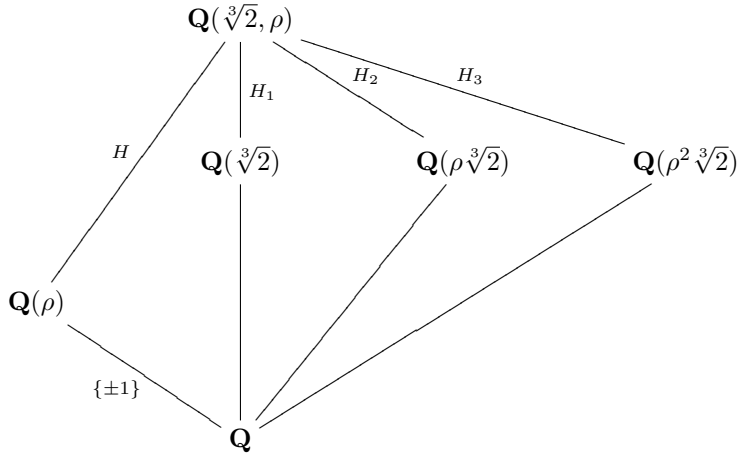
$$\forall \alpha \in F \qquad \mathrm{Tr}_{F/K}(\alpha) = \sum_{i=1}^{m} g_i(\alpha), \qquad N_{F/K}(\alpha) = \prod_{i=1}^{m} g_i(\alpha).$$

Proof. Combine Theorem 10.5 with Proposition 7.2.

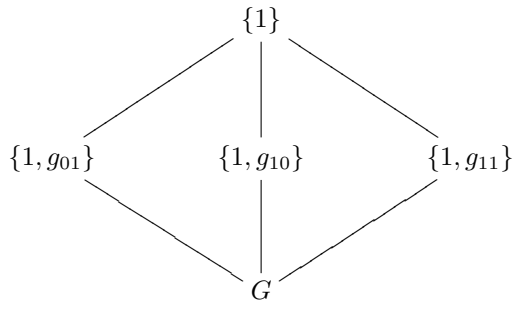
(10.7) Examples. (i) According to Example 9.6(i), the subgroups of S_3

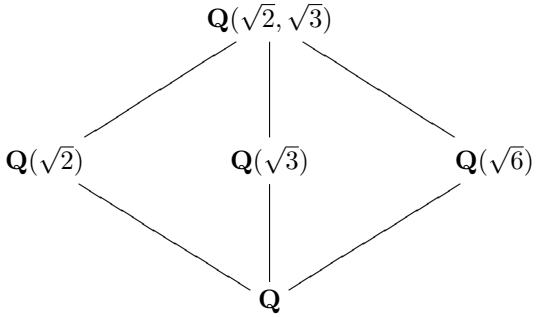


correspond to subfields $\mathbf{Q} \subset F \subset \mathbf{Q}(\sqrt[3]{2}, \rho)$:



We have $H = A_3 \lhd S_3$, but $H_j \not\lhd S_3$ $(j = 1, 2, 3)$; the map sgn : $S_3 \longrightarrow \{\pm 1\}$ induces a group isomorphism $\mathrm{Gal}(\mathbf{Q}(\rho)/\mathbf{Q}) = S_3/H = S_3/A_3 \xrightarrow{\sim} \{\pm 1\}$.
(ii) Let $L/K = \mathbf{Q}(\sqrt{2}, \sqrt{3})/\mathbf{Q}$. The subgroups of $G = \mathrm{Gal}(\mathbf{Q}(\sqrt{2}, \sqrt{3})/\mathbf{Q}) = \{g_{00}, g_{01}, g_{10}, g_{11}\} \xrightarrow{\sim} \mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$ (cf. Example 9.6(iii)) correspond to subfields $\mathbf{Q} \subset F \subset \mathbf{Q}(\sqrt{2}, \sqrt{3})$:





(10.8) Galois groups and fundamental groups. There is a close analogy between Galois theory and theory of coverings in topology:

Algebra	Topology
field K	"nice" connected topological space X
separable field extension $K \hookrightarrow L$	covering $\pi : Y \longrightarrow X$
$G = \mathrm{Aut}(L/K)$	$G = \mathrm{Aut}(Y/X)$
Galois extension: $K = L^G$	Galois covering: $G\backslash Y = X$
Galois correspondence: $K \hookrightarrow F = L^H \hookrightarrow L$	$Y \longrightarrow H\backslash Y \longrightarrow X$, $H \subset G$ subgroup
separable closure $K \hookrightarrow K^{\mathrm{sep}}$	universal covering $\widetilde{X} \longrightarrow X$
absolute Galois group $G_K = \mathrm{Aut}(K^{\mathrm{sep}}/K)$	fundamental group $\pi_1(X) \xrightarrow{\sim} \mathrm{Aut}(\widetilde{X}/X)$

Recall that a **covering** is a continuous map between topological spaces $\pi : Y \longrightarrow X$ such that Y is locally a product of the base with a discrete topological space: there exist open subsets $U_\alpha \subset X$ such that $X = \bigcup U_\alpha$ and $(\pi^{-1}(U_\alpha) \longrightarrow U_\alpha) \xrightarrow{\sim} (\mathrm{pr} : U_\alpha \times (\text{discrete space}) \longrightarrow U_\alpha)$. The **automorphism group** of a covering $\pi : Y \longrightarrow X$ is the group $\mathrm{Aut}(Y/X)$ of all continuous maps $f : Y \longrightarrow Y$ satisfying $\pi \circ f = \pi$.

A typical example of a covering is the "infinite staircase" $\exp : \mathbf{C} \longrightarrow \mathbf{C}^*$, which is a universal covering of $\mathbf{C}^*$. Its automorphism group is equal to $\{z \mapsto z + a \mid a \in 2\pi i\mathbf{Z}\} \xrightarrow{\sim} 2\pi i\mathbf{Z}$.

(10.9) Exercise. *Let $n \geq 2$ be an integer.*
(i) There are no intermediate subgroups $S_n \supsetneq H \supsetneq S_{n-1} = \{\sigma \in S_n \mid \sigma(n) = n\}$.
(ii) If K is a field and $f \in K[X]$ is a separable polynomial of degree $\deg(f) = n$ such that $\mathrm{Gal}(f) = S_n$, then there are no intermediate subfields $K \subsetneq F \subsetneq K(\alpha)$, for any root α of f.
(iii) Does the same result hold if $\mathrm{Gal}(f) = A_n$?

(10.10) Proposition ($\overline{\mathbf{R}} = \mathbf{R}(i) = \mathbf{C}$). *(1) Any polynomial $f \in \mathbf{R}[X]$ with $2 \nmid \deg(f)$ has a root $\alpha \in \mathbf{R}$.*
(2) If $L/\mathbf{R}$ is a finite extension with $2 \nmid [L : \mathbf{R}]$, then $L = \mathbf{R}$.
(3) Any polynomial $g \in \mathbf{C}[X]$ with $\deg(g) = 2$ has a root $\beta \in \mathbf{C}$.
(4) If $L/\mathbf{C}$ is a finite extension, then $[L : \mathbf{C}] \neq 2$.
(5) The field $\mathbf{C} = \mathbf{R}(i)$ is an algebraic closure of $\mathbf{R}$.

Proof. (1) We can assume that f is monic. In that case $f(t) < 0$ (resp. $f(t) > 0$) if $t << 0$ (resp. $t >> 0$); the existence of α follows from the fact that $f : \mathbf{R} \longrightarrow \mathbf{R}$ is a continuous function.
(2) There exists $\gamma \in L$ such that $L = \mathbf{R}(\gamma)$, by Theorem 6.5. The minimal polynomial $f \in \mathbf{R}[X]$ of γ over $\mathbf{R}$ has odd degree $\deg(f) = [L : \mathbf{R}]$, hence has a root in $\mathbf{R}$, by (1). Irreducibility of f in $\mathbf{R}[X]$ implies that $\deg(f) = 1$; thus $L = \mathbf{R}$.
(3) ($\iff$ (4)) This follows from the fact that any $z = a + bi \in \mathbf{C}$ has a square root in $\mathbf{C}$, namely

$$\sqrt{(c+a)/2} + i\,\mathrm{sgn}(b)\sqrt{(c-a)/2}, \qquad c = \sqrt{a^2 + b^2}.$$

(5) The following argument is due to E. Artin. Let α (lying in some field containing $\mathbf{R}$) be algebraic over $\mathbf{R}$, let f be its minimal polynomial over $\mathbf{R}$, let $L \supset \mathbf{C} = \mathbf{R}(i)$ be a splitting field of $(X^2 + 1)f(X)$ over $\mathbf{R}$. The extension $L/\mathbf{R}$ is a Galois extension; let $G = \mathrm{Gal}(L/\mathbf{R})$ be its Galois group. Let $H \subset G$ be a 2-Sylow subgroup of G. The fixed field $L^H \subset L$ satisfies $2 \nmid [L^H : \mathbf{R}] = (G : H)$, hence $L^H = \mathbf{R} = L^G$, by (2). Therefore $G = H$ is a 2-group, and so is $G_1 = \mathrm{Gal}(L/\mathbf{C}) \subset G$. If $|G_1| \neq 1$, then there exists a subgroup $G_2 \subset G_1$ such that $(G_1 : G_2) = 2$ (see Corollary 14.4 below), which gives an extension $\mathbf{C} = L^{G_1} \subset L^{G_2}$ of degree $[L^{G_2} : \mathbf{C}] = (G_1 : G_2) = 2$. This contradiction with (4) implies that $|G_1| = 1$ and $\alpha \in L = \mathbf{C}$.

(10.11) Proposition. *Let $K \hookrightarrow L$ be finite extension.*
(1) The restriction map $\mathrm{res} : \mathrm{Aut}(L/K) \longrightarrow \mathrm{Aut}(L_s/K)$ is injective.
(2) $|\mathrm{Aut}(L/K)|$ divides $[L : K]_s$.

Proof. (1) If $g \in \mathrm{Ker}(\mathrm{res})$ and $\alpha \in L$, then $\alpha^{p^r} \in L_s$ for some $r \geq 0$, hence $(g(\alpha) - \alpha)^{p^r} = g(\alpha)^{p^r} - \alpha^{p^r} = g(\alpha^{p^r}) - \alpha^{p^r} = 0$; thus g acts trivially on L.
(2) The fixed field $K \subset F = L_s^G \subset L_s$ of $G = \mathrm{Aut}(L_s/K)$ satisfies $[L_s : F] = |G|$, by Theorem 10.1. It follows that $[L : K]_s = [L_s : K]$ is divisible by $|G|$, hence by $|\mathrm{Aut}(L/K)|$.

(10.12) Proposition. *Let $K \hookrightarrow L$ be a finite* **normal** *extension.*
(1) For every field $K \hookrightarrow F \hookrightarrow L$ and every $\sigma \in \mathrm{Aut}(F/K)$ there exists $\tau \in \mathrm{Aut}(L/K)$ such that $\tau(F) = F$ and $\tau|_F = \sigma$. In particular, if F/K is normal, then the restriction map $\mathrm{res} : \mathrm{Aut}(L/K) \longrightarrow \mathrm{Aut}(F/K)$ is surjective.
(2) L_s is a Galois extension of K.
(3) The restriction map $\mathrm{Aut}(L/K) \longrightarrow \mathrm{Gal}(L_s/K)$ is an isomorphism.
(4) $|\mathrm{Aut}(L/K)| = [L : K]_s$.
(5) $|\mathrm{Aut}(L/K)| = 1 \iff$ the extension $K \hookrightarrow L$ is purely inseparable.
(6) The extension $K \hookrightarrow F = L^{\mathrm{Aut}(L/K)}$ is normal and purely inseparable, while $F \hookrightarrow L$ is a Galois extension. [Cf. Exercise 6.14.]

Proof. Fix an algebraic closure $\overline{K}$ and a homomorphism of K-algebras $L \hookrightarrow \overline{K}$.
(1) For each $\sigma \in \mathrm{Aut}(F/K)$ the composite map $F \xrightarrow{\sigma} F \hookrightarrow L \hookrightarrow \overline{K}$ extends to a homomorphism of K-algebras $\tau : L \hookrightarrow \overline{K}$, by Theorem 5.6(2). As L/K is normal, $\tau(L) = L$, hence $\tau \in \mathrm{Hom}_{K-\mathrm{Alg}}(L, \overline{K}) = \mathrm{Hom}_{K-\mathrm{Alg}}(L, L) = \mathrm{Aut}(L/K)$. By definition, $\tau|_F = \sigma$. Conversely, if F/K is a normal extension, then $\tau(F) = F$ for each $\tau \in \mathrm{Aut}(L/K) = \mathrm{Hom}_{K-\mathrm{Alg}}(L, \overline{K})$; thus res is well-defined (and surjective, by the previous argument).
(2) It is enough to show that L_s/K is a normal extension. As in (1), each $\sigma \in \mathrm{Hom}_{K-\mathrm{Alg}}(L_s, \overline{K})$ extends to $\tau \in \mathrm{Hom}_{K-\mathrm{Alg}}(L, \overline{K})$ satisfying $\tau(L) = L$. The separable extension $K \hookrightarrow \sigma(L_s) = \tau(L_s)$ is a subextension of $K \hookrightarrow \tau(L) = L$, hence $\sigma(L_s) \subset L_s$ and $\sigma(L_s) = L_s$ (by comparing the degrees).
(3) Injectivity (resp. surjectivity) was proved in Proposition 10.11(1) (resp. in (1)).
(4) (resp. (5)) is an immediate consequence of (3) (resp. of (4)).
(6) As above, each $\sigma \in \mathrm{Hom}_{K-\mathrm{Alg}}(F, \overline{K})$ extends to $\tau \in \mathrm{Hom}_{K-\mathrm{Alg}}(L, \overline{K})$ satisfying $\tau(L) = L$. As $\tau \in \mathrm{Aut}(L/K)$, it acts trivially on F. In particular, $\sigma = \mathrm{id}$ and $\sigma(F) = F$, implying that $K \hookrightarrow F$ is a normal extension. By definition, the image of the restriction map $\mathrm{res} : \mathrm{Aut}(L/K) \longrightarrow \mathrm{Aut}(F/K)$ is equal to $\{\mathrm{id}\}$. The statement (1) (resp. (5) applied to F/K) shows that $\mathrm{Aut}(F/K) = \{\mathrm{id}\}$ (resp. that $K \hookrightarrow F$ is purely inseparable). The remaining statement follows from Theorem 10.1.

(10.13) Exercise. *Let $K \hookrightarrow L_1 \hookrightarrow M$, $K \hookrightarrow L_2 \hookrightarrow M$ be finite field extensions. Denote by L_1L_2 the intersection of all subfields of M containing both L_1 and L_2; it is again a field.*
(1) If L_1/K is a Galois extension, so is L_1L_2/L_2 and the restriction map $g \mapsto g|_{L_1}$ defines an injective group homomorphism $\mathrm{Gal}(L_1L_2/L_2) \hookrightarrow \mathrm{Gal}(L_1/K)$ with image equal to $\mathrm{Gal}(L_1/L_1 \cap L_2)$.
(2) If both L_1/K and L_2/K are Galois extensions, so is L_1L_2/K and the restriction maps $g \mapsto g|_{L_i}$ define an injective group homomorphism $\mathrm{Gal}(L_1L_2/K) \hookrightarrow \mathrm{Gal}(L_1/K) \times \mathrm{Gal}(L_2/K)$.

11. Examples of Galois groups

(11.1) Theorem (Extensions of finite fields). *Let $q = p^r$, where p is a prime number and $r \geq 1$.*
(1) For every integer $n \geq 1$ the extension $\mathbf{F}_{q^n}/\mathbf{F}_q$ is a Galois extension. Its Galois group is cyclic of order n, generated by the Frobenius morphism $\varphi_q(x) = x^q$: $\mathrm{Gal}(\mathbf{F}_{q^n}/\mathbf{F}_q) = \{\varphi_q, \varphi_q^2, \ldots, \varphi_q^n = 1\}$.
(2) Every subgroup of $\mathrm{Gal}(\mathbf{F}_{q^n}/\mathbf{F}_q)$ is cyclic, of order n/m (where m is a divisor of n), generated by $\varphi_q^m = \varphi_{q^m} : x \mapsto x^{q^m}$. Its fixed field is equal to $\mathbf{F}_{q^m}$ and $\mathrm{Gal}(\mathbf{F}_{q^n}/\mathbf{F}_{q^m}) = \{\varphi_q^m, \varphi_q^{2m}, \ldots, \varphi_q^{\frac{n}{m} \cdot m} = 1\}$.

Proof. (1) For each $m = 1, \ldots, n-1$ we have

$$|\{x \in \mathbf{F}_{q^n} \mid \varphi_q^m(x) = x\}| = |\{\text{roots of } X^{q^m} - X \text{ in } \mathbf{F}_{q^n}\}| \leq \deg(X^{q^m} - X) = q^m < q^n,$$

which implies that $\varphi_q, \varphi_q^2, \ldots, \varphi_q^{n-1} \neq 1 \in \mathrm{Aut}(\mathbf{F}_{q^n}/\mathbf{F}_q)$. It follows that

$$n = |\{\varphi_q, \varphi_q^2, \ldots, \varphi_q^n = 1\}| \leq |\mathrm{Aut}(\mathbf{F}_{q^n}/\mathbf{F}_q)| \leq [\mathbf{F}_{q^n} : \mathbf{F}_q)] = n,$$

which yields equalities $\{\varphi_q, \varphi_q^2, \ldots, \varphi_q^n = 1\} = \mathrm{Aut}(\mathbf{F}_{q^n}/\mathbf{F}_q) = \mathrm{Gal}(\mathbf{F}_{q^n}/\mathbf{F}_q)$.
(2) It is well-known that any subgroup H of the cyclic group of order n generated by φ_q is also cyclic, generated by φ_q^m (for some $m \mid n$). According to Theorem 4.4(2), the fixed field of H is equal to

$$\mathbf{F}_{q^n}^H = \{x \in \mathbf{F}_{q^n} \mid x^{q^m} = x\} = \mathbf{F}_{q^m},$$

hence $\mathrm{Gal}(\mathbf{F}_{q^n}/\mathbf{F}_{q^m}) = H$, thanks to Theorem 10.1.

(11.2) Corollary. *Let $q = p^r$ be as in Theorem 11.1. For each $n \geq 1$, the norm map $N_{\mathbf{F}_{q^n}/\mathbf{F}_q} : \mathbf{F}_{q^n}^* \longrightarrow \mathbf{F}_q^*$ is given by the formula $N_{\mathbf{F}_{q^n}/\mathbf{F}_q}(x) = x^{1+q+\cdots+q^{n-1}} = x^{(q^n-1)/(q-1)}$. If x is a generator of the cyclic group $\mathbf{F}_{q^n}^*$, then its norm $N_{\mathbf{F}_{q^n}/\mathbf{F}_q}(x)$ is a generator of $\mathbf{F}_q^*$. In particular, the norm map $N_{\mathbf{F}_{q^n}/\mathbf{F}_q} : \mathbf{F}_{q^n}^* \longrightarrow \mathbf{F}_q^*$ is surjective.*

Proof. The formula for the norm follows from a combination of Corollary 10.6 with Theorem 11.1. If the order of $x \in \mathbf{F}_{q^n}^*$ is equal to $q^n - 1$, then the order of $x^{(q^n-1)/(q-1)}$ is equal to $q - 1$.

(11.3) Question: when is $\mathrm{Gal}(f) \subset H$? Given a subgroup $H \subset S_n$, it is natural to ask under what conditions the Galois group of a separable polynomial $f \in K[X]$ of degree $\deg(f) = n$ is contained in a conjugate subgroup $\sigma H \sigma^{-1}$, for some $\sigma \in S_n$ (the conjugation is necessary, since $\mathrm{Gal}(f) \subset S_n$ itself is defined only up to conjugation). We first give the answer for $H = A_n$ and then apply the machinery of resolvents (see 2.14) to treat the general case.

(11.4) Proposition. *Let K be a field, let $f \in K[X]$ be a separable monic polynomial of degree $n \geq 2$.*
(1) If $\mathrm{Gal}(f) \subset A_n$, then $\mathrm{disc}(f) \in K^{*2}$.
(2) If $char(K) \neq 2$ and $\mathrm{disc}(f) \in K^{*2}$, then $\mathrm{Gal}(f) \subset A_n$.
(3) If $char(K) = 2$, then $\mathrm{disc}(f) \in K^{*2}$.

Proof. Let $L = K(\alpha_1, \ldots, \alpha_n)$ be a splitting field of f over K, where $\alpha_1, \ldots, \alpha_n$ are the (distinct) roots of f in L. We have

$$\mathrm{disc}(f) = d^2, \qquad d = \prod_{i<j}(\alpha_i - \alpha_j) \in L^*, \qquad \forall g \in \mathrm{Gal}(f) \quad g(d) = \mathrm{sgn}(g)d. \tag{11.4.1}$$

(1) If $\mathrm{Gal}(f) \subset A_n$, then (11.4.1) implies that $d \in (L^*)^{\mathrm{Gal}(f)} = K^*$, hence $\mathrm{disc}(f) = d^2 \in K^{*2}$.
(2) Conversely, if $\mathrm{disc}(f) \in K^{*2}$, then $d \in K^*$. However, if $char(K) \neq 2$ and $g \in \mathrm{Gal}(f) \setminus A_n$, then $g(d) = -d \neq d$, which contradicts the fact that $d \in K^*$. Therefore $\mathrm{Gal}(f) \subset A_n$ in this case.
(3) If $char(K) = 2$, then $d = \prod_{i<j}(\alpha_i + \alpha_j) \in (L^*)^{\mathrm{Gal}(f)} = K^*$, hence $\mathrm{disc}(f) = d^2 \in K^{*2}$.

(11.5) Let us consider the general case. Let K be a field, let

$$f(X) = X^n + a_1 X^{n-1} + \cdots + a_n \in K[X]$$

be a monic separable polynomial of degree $n \geq 1$. Denote by $L = K(\alpha_1, \ldots, \alpha_n)$ a splitting field of f over K, where $\alpha_1, \ldots, \alpha_n$ are the (distinct) roots of f in L.

Fix an auxiliary polynomial $u = u(x_1, \ldots, x_n) \in K[x_1, \ldots, x_n]$. Its stabiliser

$$H = \{\tau \in S_n \mid \tau * u = u\} \subset S_n \tag{11.5.1}$$

is a subgroup of S_n. Conversely, a polynomial u satisfying (11.5.1) exists for any given subgroup $H \subset S_n$, thanks to Theorem 6.5 applied to the extension $K(x_1, \ldots, x_n)^H / K(x_1, \ldots, x_n)^{S_n}$.

As in 2.14 we have a polynomial $U(y; \sigma_1, \ldots, \sigma_n) \in K[\sigma_1, \ldots, \sigma_n][y]$ which splits in $K[x_1, \ldots, x_n][y]$ as follows:

$$U(y; \sigma_1, \ldots, \sigma_n) = \prod_{\tau H \in S_n/H} \left(y - u(x_{\tau(1)}, \ldots, x_{\tau(n)})\right). \tag{11.5.2}$$

After applying to (11.5.2) the morphism of K-algebras

$$\lambda : K[x_1, \ldots, x_n] \longrightarrow L, \qquad \lambda(P(x_1, \ldots, x_n)) = P(\alpha_1, \ldots, \alpha_n)$$

which sends each x_i to α_i (hence each σ_i to $\lambda(\sigma_i) = (-1)^i a_i \in K$, $i = 1, \ldots, n$), we obtain the "resolvent polynomial of f"

$$R_f(y) = U(y; -a_1, a_2, \ldots, (-1)^n a_n) = \prod_{\tau H \in S_n/H} \big(y - \rho(u(x_{\tau(1)}, \ldots, x_{\tau(n)}))\big)$$
$$= \prod_{\tau H \in S_n/H} \big(y - u(\alpha_{\tau(1)}, \ldots, \alpha_{\tau(n)})\big) \in K[y]. \tag{11.5.3}$$

This polynomial does not depend on the numbering of the set of roots of f. Its coefficients lie in K, its degree is equal to $\deg(R_f) = |O(u)| = (S_n : H)$ and its roots are equal to $(\tau * u)(\alpha_1, \ldots, \alpha_n) = u(\alpha_{\tau(1)}, \ldots, \alpha_{\tau(n)})$ $(\tau H \in S_n/H)$. Note that the roots of R_f are not necessarily distinct.

(11.6) Proposition. *Let K be a field, let $u \in K[x_1, \ldots, x_n]$ $(n \geq 1)$ be a fixed auxiliary polynomial as in 11.5; let $H \subset S_n$ be its stabiliser as in (11.5.1). Let $f \in K[X]$ be a monic separable polynomial of degree n.*
(1) If $\mathrm{Gal}(f) \subset \sigma H \sigma^{-1}$ for some $\sigma \in S_n$, then the resolvent polynomial $R_f(y)$ has a root $\beta \in K$.
(2) If $R_f(y)$ has a **simple** *root $\beta \in K$, then $\mathrm{Gal}(f) \subset \sigma H \sigma^{-1}$ for some $\sigma \in S_n$ (which is equivalent to the existence of a numbering of the set of roots of f for which $\mathrm{Gal}(f) \subset H$).*

Proof. (cf. the proof of Proposition 11.4). (1) After renumbering the roots of f we can assume that $\mathrm{Gal}(f) \subset H$. In this case $\beta = u(\alpha_1, \ldots, \alpha_n)$ is a root of $R_f(y)$ and

$$\forall g \in \mathrm{Gal}(f) \subset H \subset S_n \qquad g(\beta) = u(\alpha_{g(1)}, \ldots, \alpha_{g(n)}) = (g * u)(\alpha_1, \ldots, \alpha_n) = u(\alpha_1, \ldots, \alpha_n) = \beta,$$

hence $\beta \in L^{\mathrm{Gal}(f)} = K$.
(2) Again, after renumbering we can assume that the simple root of R_f in question is $\beta = u(\alpha_1, \ldots, \alpha_n) \in K$. As before, we have

$$\forall g \in \mathrm{Gal}(f) \subset S_n \quad \beta = g(\beta) = u(\alpha_{g(1)}, \ldots, \alpha_{g(n)}) = (g * u)(\alpha_1, \ldots, \alpha_n).$$

On the other hand, if $g \in S_n$ but $g \notin H$, then $u(\alpha_{g(1)}, \ldots, \alpha_{g(n)}) = (g * u)(\alpha_1, \ldots, \alpha_n)$ is a root of the polynomial $R_f(y)/(y - \beta)$, hence is different from β, since β is a simple root of R_f, by assumption. As a result,

$$\forall g \in S_n \setminus H \qquad u(\alpha_{g(1)}, \ldots, \alpha_{g(n)}) \neq \beta,$$

which means that $S_n \setminus H \subset S_n \setminus \mathrm{Gal}(f)$, hence $\mathrm{Gal}(f) \subset H$.

(11.7) Examples. (i) If $n \geq 2$, $char(K) \neq 2$ and $u = \Delta = \prod_{i<j}(x_i - x_j)$, then $H = A_n$, $R_f(y) = y^2 - \mathrm{disc}(f)$ and we recover Proposition 11.4.

(ii) If $n = 4$ and $u = x_1x_2 + x_3x_4$, then $H = D_8 \subset S_4$ and $\mathrm{disc}(R_f) = \mathrm{disc}(f)$. Proposition 11.6 then states that, for any separable polynomial $f(X) = (X - \alpha_1) \cdots (X - \alpha_4) \in K[X]$, the Galois group $\mathrm{Gal}(f) \subset S_4$ is conjugate to a subgroup of $D_8 \subset S_4 \iff$ the cubic resolvent

$$(y - (\alpha_1\alpha_2 + \alpha_3\alpha_4))(y - (\alpha_1\alpha_3 + \alpha_2\alpha_4))(y - (\alpha_1\alpha_4 + \alpha_2\alpha_3)) \in K[y]$$

has a root $\beta \in K$.

(11.8) Exercise (Lagrange). *If $u_1, u_2 \in K[x_1, \ldots, x_n]$ are two auxiliary polynomials with respective stabilisers $H_i = \{\tau \in S_n \mid \tau * u_i = u_i\}$, then we have*

$$H_1 \supset H_2 \iff u_1 \in K(\sigma_1, \ldots, \sigma_n)(u_2).$$

(11.9) In practical calculations of Galois groups one applies Proposition 11.6 to irreducible polynomials, in which case $\mathrm{Gal}(f)$ acts transitively on $\{1, \ldots, n\}$, by Proposition 9.5(4) (we say that $\mathrm{Gal}(f)$ is a **transitive subgroup of** S_n). What is required is a list of transitive subgroups $H \subset S_n$ (up to conjugation), for each H an auxilary polynomial u with stabiliser H and an explicit formula for the corresponding resolvent R_f in terms of f. See [Co, ch. 13] for examples.

(11.10) Exercise. *(1) Every transitive subgroup of S_4 is conjugate to S_4, A_4, D_8, C_4 or to $C_2 \times C_2$ from Example 9.6(iv).*
(2) A subgroup $G \subset S_p$, where p is a prime number, is transitive $\iff$ the order of G is divisible by p $\iff$ G contains a p-cycle.
(3) If p is a prime number and $G \subset S_p$ is a subgroup containing a p-cycle and a transposition ($= 2$-cycle), then $G = S_p$. (cf. [De 1], Ex. I.2.7).

(11.11) Proposition. *Let p be a prime number, K a subfield of $\mathbf{R}$ and $f \in K[X]$ an irreducible polynomial of degree p. Let $L = K(\alpha_1, \dots, \alpha_p) \subset \mathbf{C}$, where $\alpha_1, \dots, \alpha_p \in \mathbf{C}$ are the complex roots of f. If $\alpha_1, \alpha_2 \notin \mathbf{R}$ and $\alpha_3, \dots, \alpha_p \in \mathbf{R}$, then $\mathrm{Gal}(f) = \mathrm{Gal}(L/K) = S_p$. [This applies, for example, to $K = \mathbf{Q}$ and $f = X^5 - 6X + 2$.]*

Proof. The action of $\mathrm{Gal}(f)$ on $\{1, \dots, p\}$ is transitive, since f is irreducible; the group $\mathrm{Gal}(f) \subset S_p$ contains a p-cycle, by 11.10(2). As $K \subset \mathbf{R}$, the complex conjugation $c(a + bi) = a - bi$ $(a, b \in \mathbf{R})$ is an element of $c \in \mathrm{Hom}_{K-\mathrm{Alg}}(L, c(L)) = \mathrm{Hom}_{K-\mathrm{Alg}}(L, L) = \mathrm{Gal}(L/K) = \mathrm{Gal}(f)$ ($c(L) = L$, since L/K is a normal extension). The assumptions on the roots of f imply that $c = (12) \subset S_p$ acts on $\{1, \dots, p\}$ as a 2-cycle; we conclude by 11.10(3).

(11.12) Theorem (Dedekind). *Let $f \in \mathbf{Z}[X] \subset \mathbf{Q}[X]$ be a monic polynomial of degree $n \geq 1$. Let p be a prime number; denote by $\overline{f} = f \pmod p \in \mathbf{F}_p[X]$ the reduction of f modulo p. Assume that $\overline{f}$ is* **separable** *($\iff$ $p \nmid \mathrm{disc}(f)$) and write $\overline{f} = \overline{f}_1 \cdots \overline{f}_r$, where $\overline{f}_1, \dots, \overline{f}_r \in \mathbf{F}_p[X]$ are distinct monic irreducible polynomials of respective degrees n_i ($n_1 + \cdots + n_r = n$). The polynomial f is then separable and $\mathrm{Gal}(f) \subset S_n$ contains an element $c_1 \cdots c_r$, where $c_1, \dots, c_r$ are disjoint cycles of lengths $n_1, \dots, n_r$.*

Proof. Later.

(11.13) Example. Let $K = \mathbf{Q}$, $f(X) = X^5 - X + 1 \in \mathbf{Q}[X]$. The factorisation established in 4.10 $f \pmod 2 = (X^2 + X + 1)(X^3 + X^2 + 1) \in \mathbf{F}_2[X]$ shows that $G = \mathrm{Gal}(f) \subset S_5$ contains an element of the form $g = (ab)(cde)$, hence a 2-cycle $g^3 = (ab)$. One can check that the polynomial $f \pmod 3 \in \mathbf{F}_3[X]$ is irreducible (for example, by computing $\gcd(f \pmod 3, X^9 - X) = 1$, which implies that f has no irreducible factor of degree 1 or 2, by Corollary 4.5). It follows from 11.12 that G contains a 5-cycle, hence $G = S_5$, by 11.10(3). One can also check that $f \pmod 5 \in \mathbf{F}_5[X]$ is irreducible:

(11.14) Exercise. *If p is a prime number and $a \in \mathbf{F}_p^*$, then the polynomial $f(X) = X^p - X - a$ is irreducible in $\mathbf{F}_p[X]$. [Hint: if $f(\alpha) = (\varphi - 1)(\alpha) - a = 0$, then $(\varphi^m - 1)(\alpha) \neq 0$, for all $0 < m < p$.]*

(11.15) Exercise. *Interpret Exercise 3.29 in terms of Galois theory: determine the normal closure M over K of the field $L = K(\sqrt{a + b\sqrt{c}})$, the Galois group $G = \mathrm{Gal}(M/K)$ and all intermediate fields $K \subsetneq F \subsetneq M$.*

(11.16) Exercise. *Let K be a field of characteristic $\mathrm{char}(K) \neq 2$, let L/K be an extension of degree $[L : K] = 2$. Show that: there exists a Galois extension M/K such that $M \supset L$ and $\mathrm{Gal}(M/K) \xrightarrow{\sim} \mathbf{Z}/4\mathbf{Z}$ $\iff$ $-1 \in N_{L/K}(L^*)$.*

(11.17) Exercise. *Let p be a prime number.*
(1) Let $G \subset S_p$ be a subgroup; for each $i = 1, \dots, p$ denote by $G_i = \{g \in G \mid g(i) = i\} \subset G$ its stabiliser. If G is a transitive subgroup of S_p for which $G_1 \subset G_2$, then $G = C_p$ is a cyclic group generated by a p-cycle and $|G_i| = 1$ for each i.
(2) Let K be a field, let $f \in K[X]$ be an irreducible separable polynomial of degree $\deg(f) = p$, let $L = K(\alpha_1, \dots, \alpha_p)$ be its splitting field, where $\alpha_1, \dots, \alpha_p$ are the roots of f in L. Show that: $\alpha_2 \in K(\alpha_1)$ $\iff$ $L = K(\alpha_1)$ $\iff$ $\mathrm{Gal}(L/K)$ is cyclic of order p.

12. Roots of unity

Historically, the equation $X^n - 1 = 0$ ("division of the circle in n parts of equal length") played an extremely important rôle in the development of both Galois theory and arithmetic. The first systematic treatment of this equation (for $n = p$ a prime number) was given by Gauss.

(12.1) For each $n \geq 1$, the set of complex n-th roots of unity

$$\mu_n(\mathbf{C}) = \{\zeta \in \mathbf{C} \mid \zeta^n = 1\}$$

is a cyclic subgroup of $\mathbf{C}^*$, generated by $\zeta_n = e^{2\pi i/n}$. The polynomials

$$X^n - 1 = \prod_{\zeta \in \mu_n} (X - \zeta)$$

are highly reducible in $\mathbf{Q}[X]$. For example,

$$X^{12} - 1 = (X-1)(X+1)(X^2+X+1)(X^2-X+1)(X^4-X^2+1) = \Phi_1\Phi_2\Phi_3\Phi_6\Phi_{12}.$$

The general pattern is as follows. The set of generators of $\mu_n(\mathbf{C})$ (= the set of primitive complex n-th roots of unity) is equal to

$$\mu_n^0(\mathbf{C}) = \{\zeta_n^a \mid a \in (\mathbf{Z}/n\mathbf{Z})^*\} \tag{12.1.1}$$

and we have

$$\mu_n(\mathbf{C}) = \bigcup_{d|n} \mu_d^0(\mathbf{C}). \tag{12.1.2}$$

(12.2) Proposition-Definition. *For $n \geq 1$ define the n-th* **cyclotomic polynomial** *to be*

$$\Phi_n(X) = \prod_{\zeta \in \mu_n^0} (X - \zeta) \in \mathbf{C}[X].$$

These monic polynomials have the following properties.
(1) $\deg(\Phi_n) = \varphi(n)$.
(2) $\prod_{d|n} \Phi_d(X) = X^n - 1$.
(3) $\Phi_n(X) \in \mathbf{Z}[X]$.
(4) If p is a prime number and $k \geq 1$, then

$$\Phi_p(X) = (X^p - 1)/(X - 1) = X^{p-1} + X^{p-2} + \cdots + 1, \qquad \Phi_{p^k}(X) = \Phi_p(X^{p^{k-1}}) = (X^{p^k} - 1)/(X^{p^{k-1}} - 1).$$

Proof. (1) and (2) follow from (12.1.1-2). The Möbius inversion formula 12.3(3) applied to (2) shows that the complex polynomial

$$\Phi_n(X) = \prod_{d|n} (X^{n/d} - 1)^{\mu(d)}$$

is a quotient of two monic polynomials with coefficients in $\mathbf{Z}$, proving (3) and (4).

(12.3) Exercise (The Möbius inversion formula). *Let $f, g : \mathbf{N} \setminus \{0\} \longrightarrow \mathbf{C}$ be functions related by*

$$\forall n \geq 1 \quad g(n) = \sum_{d|n} f(d).$$

(1) The generating functions $Z_f(s) = \sum_{n=1}^{\infty} f(n)/n^s$ and $Z_g(s) = \sum_{n=1}^{\infty} g(n)/n^s$ are related by $Z_g(s) = Z_f(s)\zeta(s)$, where $\zeta(s) = \sum_{n=1}^{\infty} 1/n^s$.
(2) $1/\zeta(s) = \prod_p (1 - 1/p^s) = \sum_{n=1}^{\infty} \mu(n)/n^s$, where μ is the Möbius function: $\mu(p_1 \cdots p_r) = (-1)^r$, if $r \geq 0$ and p_i are distinct primes; $\mu(n) = 0$ if n is divisible by the square of a prime.
(3) The function f is given by

$$\forall n \geq 1 \quad f(n) = \sum_{d|n} \mu(d) f(n/d).$$

(12.4) *Let K be an arbitrary field. For every integer $n \geq 1$ denote by*

$$\mu_n(K) = \{a \in K \mid a^n = 1\}$$

the set of all n-th roots of unity lying in K. If $char(K) = p > 0$ and $n = p^k \cdot m$, $p \nmid m$, then the formula $X^n - 1 = (X^m - 1)^{p^k} \in K[X]$ shows that $\mu_n(K) = \mu_m(K)$.

We are going to consider the cyclotomic polynomials $\Phi_n(X) \in \mathbf{Z}[X]$ as elements of $K[X]$; the formula 12.2(2) still holds in $K[X]$:

$$\forall n \geq 1 \quad \prod_{d|n} \Phi_d(X) = X^n - 1 \in K[X]. \tag{12.4.1}$$

Fix a splitting field $K(\mu_n)$ of the polynomial $X^n - 1$ over K and set $\mu_n = \mu_n(K(\mu_n))$.

(12.5) Proposition. *Let K be a field and $n \geq 1$ an integer such that $char(K) \nmid n$.*
(1) The polynomial $f(X) = X^n - 1 \in K[X]$ is separable.
(2) For every field $F \supset K(\mu_n)$ the group $\mu_n(F)$ is cyclic of order n; denote by $\mu_n^0(F)$ the set of its generators. For any $\zeta_n \in \mu_n^0(F)$ we have

$$\mu_n^0(F) = \{\zeta_n^a \mid a \in (\mathbf{Z}/n\mathbf{Z})^*\} = \{\text{roots of } \Phi_n(X) \text{ in } F\}.$$

(3) The field $K(\mu_n)$ is equal to $K(\zeta_n)$, for any generator ζ_n of μ_n. Moreover, it is a splitting field of the polynomial $\Phi_n(X)$ over K.
(4) $K(\mu_n)/K$ is a Galois extension. There is a canonical injective group homomorphism

$$\chi = \chi_{n,K} : \mathrm{Gal}(K(\mu_n)/K) \longrightarrow (\mathbf{Z}/n\mathbf{Z})^* = GL_1(\mathbf{Z}/n\mathbf{Z})$$

(the **cyclotomic character***) such that*

$$\forall \zeta \in \mu_n \ \forall g \in \mathrm{Gal}(K(\mu_n)/K) \quad g(\zeta) = \zeta^{\chi(g)}.$$

(5) The Galois group $\mathrm{Gal}(K(\mu_n)/K)$ is abelian.

Proof. (1) The assumption $char(K) \nmid n$ implies that $n \in K^*$, hence $1 = X^n - (X^n - 1) = n^{-1}Xf'(X) - f(X) \in (f, f')$.
(2) The group $\mu_n(F)$ is cyclic, by Proposition 3.3. It follows from (1) that the roots of f are distinct, hence $|\mu_n(F)| = \deg(f) = n$. If $\zeta_n \in \mu_n^0(F)$, then we deduce from (12.4.1) that

$$\forall m = 1, \ldots, n-1 \quad \zeta_n^m - 1 \neq 0 \implies \forall m = 1, \ldots, n-1 \quad \Phi_m(\zeta_n) \neq 0;$$

thus $\Phi_n(\zeta_n) = 0$. We have just shown that $\mu_n^0(F) = \{\zeta_n^a \mid a \in (\mathbf{Z}/n\mathbf{Z})^*\}$ is contained in the set of all roots of $\Phi_n(X)$ in F; these two sets have the same cardinality $|\mu_n^0(F)| = \varphi(n) = \deg(\Phi_n)$, hence they are equal.
(3) The equality $\mu_n = \{\zeta_n, \zeta_n^2, \ldots, \zeta_n^n = 1\}$ implies that $K(\mu_n) = K(\zeta_n) = K(\mu_n^0)$; the latter field is a splitting field of $\Phi_n(X)$ by K, by (2).
(4) The field $K(\mu_n)$ is a splitting field over K of a separable polynomial $X^n - 1$, which means that $K(\mu_n)/K$ is a Galois extension. Fix $\zeta_n \in \mu_n^0$ (= a root of $\Phi_n(X)$ in $K(\mu_n)$). For each $g \in \mathrm{Gal}(K(\mu_n)/K)$ we have

$$\Phi_n(g(\zeta_n)) = g(\Phi_n(\zeta_n)) = g(0) = 0,$$

hence there is a unique $a \in (\mathbf{Z}/n\mathbf{Z})^*$ such that $g(\zeta_n) = \zeta_n^a$. Each element $\zeta \in \mu_n$ is of the form $\zeta = \zeta_n^b$ $(1 \leq b \leq n)$; it follows that

$$g(\zeta) = g(\zeta_n^b) = g(\zeta_n)^b = \zeta_n^{ab} = \zeta^a.$$

In particulier, the exponent

$$\chi(g) := a \in (\mathbf{Z}/n\mathbf{Z})^*$$

does not depend on the choice of $\zeta_n \in \mu_n^0$. The map

$$\chi : \mathrm{Gal}(K(\mu_n)/K) \longrightarrow (\mathbf{Z}/n\mathbf{Z})^*$$

is a group homomorphism, since

$$(gg')(\zeta) = g(g'(\zeta)) = g(\zeta^{\chi(g')}) = g(\zeta)^{\chi(g')} = (\zeta^{\chi(g)})^{\chi(g')} = \zeta^{\chi(g)\chi(g')} \implies \chi(gg') = \chi(g)\chi(g')$$

holds for all $g, g' \in \mathrm{Gal}(K(\mu_n)/K)$ and $\zeta \in \mu_n$.

If $g \in \mathrm{Ker}(\chi)$, then $g(\zeta_n) = \zeta_n$, which implies that $g(\alpha) = \alpha$ for all $\alpha \in K(\zeta_n) = K(\mu_n)$. As a result, χ is injective.

One can reformulate the above argument in more scientific terms by saying that the restriction map

$$\mathrm{Aut}(K(\mu_n)/K) \longrightarrow \mathrm{Hom}_{\mathbf{Z}}(K^*, K^*) \longrightarrow \mathrm{Hom}_{\mathbf{Z}}(\mu_n, \mu_n) = \mathrm{Hom}_{\mathbf{Z}/n\mathbf{Z}}(\mu_n, \mu_n) = (\mathbf{Z}/n\mathbf{Z})^*$$

is a group homomorphism for trivial reasons. It is injective, since μ_n generates $K(\mu_n)$ over K.
(5) $\mathrm{Gal}(K(\mu_n)/K)$ is isomorphic to a subgroup $\chi(\mathrm{Gal}(K(\mu_n)/K))$ of an abelian group $(\mathbf{Z}/n\mathbf{Z})^*$.

(12.6) The key point in Proposition 12.5 is the group homomorphism $\chi : \mathrm{Gal}(K(\mu_n)/K) \longrightarrow GL_1(\mathbf{Z}/n\mathbf{Z})$. Historically, this was the first example of a Galois representation.

(12.7) Proposition. *If $K = \mathbf{Q}$ and $n \geq 1$, then:*
(1) The polynomial $\Phi_n(X)$ is irreducible in $\mathbf{Q}[X]$.
(2) The homomorphism $\chi_{n,\mathbf{Q}} : \mathrm{Gal}(\mathbf{Q}(\mu_n)/\mathbf{Q}) \longrightarrow (\mathbf{Z}/n\mathbf{Z})^*$ is an isomorphism.
(3) For any $m|n$ the map $\chi_{n,\mathbf{Q}}$ induces a group isomorphism

$$\mathrm{Gal}(\mathbf{Q}(\mu_n)/\mathbf{Q}(\mu_m)) \xrightarrow{\sim} \{a \in (\mathbf{Z}/n\mathbf{Z})^* \mid a \equiv 1 \pmod m\}.$$

Proof. (1) If $n = p^k$ (p prime), then $\Phi_{p^k}(1+X) = ((1+X)^{p^k} - 1)/((1+X)^{p^{k-1}} - 1)$ is an Eisenstein polynomial with respect to p. In general, assume that $f(X) \in \mathbf{Z}[X]$ is an irreducible monic factor of $\Phi_n(X) = f(X)g(X)$ ($\deg(f) \geq 1$); fix a root $\zeta \in \mathbf{Q}(\mu_n)$ of $f(X)$. Let $p \nmid n$ be a prime number; we are going to show that $f(\zeta^p) = 0$. Indeed, if $f(\zeta^p) \neq 0$, then $g(\zeta^p) = 0$ (since $\Phi_n(\zeta^p) = 0$). Consider $h(X) = g(X^p) \in \mathbf{Z}[X]$; then f divides h in $\mathbf{Q}(X)$, since $h(\zeta) = 0$ and f is the minimal polynomial of ζ over $\mathbf{Q}$. The polynomial $f \in \mathbf{Z}[X]$ is monic, which implies that f divides h in $\mathbf{Z}[X]$. As a result, the reduction modulo p $\overline{f} = f \pmod p \in \mathbf{F}_p[X]$ divides $\overline{h}(X) = \overline{g}(X^p) = \overline{g}(X)^p$ in $\mathbf{F}_p[X]$; in particular, if $r \in \mathbf{F}_p[X]$ is an irreducible non-constant factor of de $\overline{f}$, then $r^2 \mid \overline{f}\overline{g} = \overline{\Phi}_n \in \mathbf{F}_p[X]$, which contradicts the separability of $\overline{\Phi}_n \in \mathbf{F}_p[X]$.

If $a \geq 1$ is an integer prime to n, then $a = p_1 \cdots p_k$ is a product of prime numbers $p_j \nmid n$, hence $f(\zeta^a) = 0$, by induction on k. This shows that each element of μ_n^0 is a root of f, hence $f = \Phi_n$.
(2) According to (1), $\Phi_n(X)$ is the minimal polynomial of ζ_n over $\mathbf{Q}$. The homomorphism $\chi_{n,\mathbf{Q}}$ is an injective map between two sets of the same cardinality

$$|(\mathbf{Z}/n\mathbf{Z})^*| = \varphi(n) = \deg(\Phi_n) = [\mathbf{Q}(\zeta_n) : \mathbf{Q}] = [\mathbf{Q}(\mu_n) : \mathbf{Q}] = |\mathrm{Gal}(\mathbf{Q}(\mu_n)/\mathbf{Q})|,$$

which means that it is bijective. The statement (3) follows from (2).

(12.8) Example (Gauss sums). For any prime number $p > 2$ the Galois group $G = \mathrm{Gal}(\mathbf{Q}(\mu_p)/\mathbf{Q})$ is cyclic of order $p - 1$, which implies that there is a unique subgroup $H \subset G$ of index $(G : H) = 2$, hence a unique field $\mathbf{Q} \subset F \subset \mathbf{Q}(\mu_p)$ such that $[F : \mathbf{Q}] = 2$, namely $F = \mathbf{Q}(\mu_p)^H$. Fix a generator $g \in G$; then H is generated by g^2. Write $\zeta_p = e^{2\pi i/p}$ and consider the following **quadratic Gauss sum**

$$\tau_p = \sum_{j=1}^{p-1} (-1)^j \, g^j(\zeta_p) = \sum_{j=1}^{p-1} (-1)^j \, \zeta_p^{(a^j)} \in \mathbf{Q}(\mu_p) \qquad\qquad (a = \chi_p(g) \in (\mathbf{Z}/p\mathbf{Z})^*)$$

(it is easy to see that τ_p does not depend on the choice of g). The formula

$$g(\tau_p) = \sum_{j=1}^{p-1} (-1)^j \, g^{j+1}(\zeta_p) = -\sum_{k=2}^{p} (-1)^k \, g^k(\zeta_p) = -\tau_p$$

implies that

$$g^2(\tau_p) = \tau_p \implies \tau_p \in \mathbf{Q}(\mu_p)^H = F$$
$$g(\tau_p^2) = \tau_p^2 \implies \tau_p^2 \in \mathbf{Q}(\mu_p)^G = \mathbf{Q}.$$

For example, one can compute by brute force the values

$$\tau_3 = \zeta_3 - \zeta_3^2 = i\sqrt{3}, \qquad \tau_5 = \zeta_5 - \zeta_5^2 - \zeta_5^3 + \zeta_5^4 = \sqrt{5}, \qquad \tau_7 = \zeta_7 + \zeta_7^2 - \zeta_7^3 + \zeta_7^4 - \zeta_7^5 - \zeta_7^6 = i\sqrt{7}.$$

In fact, it is relatively easy to show that $\tau_p^2 = (-1)^{(p-1)/2} p$, but much more difficult to determine the exact square root. This was done by Gauss, who proved that

$$\tau_p = \begin{cases} +\sqrt{p} & p \equiv 1 \pmod 4 \\ +i\sqrt{p} & p \equiv 3 \pmod 4). \end{cases}$$

Quadratic Gauss sums can be used to give an elegant proof of the quadratic reciprocity law.

More generally, if $r \mid (p-1)$ and $\alpha \in \mu_r^0$, one can consider the sum

$$\tau_p(\alpha, g) = \sum_{j=1}^{p-1} \alpha^j \, g^j(\zeta_p) \in \mathbf{Q}(\mu_p, \mu_r) = \mathbf{Q}(\mu_{pr}).$$

The same computation as before shows that

$$\tau_p(\alpha, g)^r \in \mathbf{Q}(\mu_r).$$

(12.9) Exercise (quadratic subfields of cyclotomic fields). *(1) Let $p > 2$ be a prime number. Compute the discriminant $\operatorname{disc}(\Phi_p)$ and use the fact that $\mathbf{Q}(\sqrt{\operatorname{disc}(\Phi_p)}) \subset \mathbf{Q}(\mu_p)$ to show that the unique subfield $F \subset \mathbf{Q}(\mu_p)$ with $[F : \mathbf{Q}] = 2$ is equal to $F = \mathbf{Q}(\sqrt{p^*})$. where $p^* = (-1)^{(p-1)/2} p$.*
(2) Show that $\mathbf{Q}(\mu_8) = \mathbf{Q}(i, \sqrt{2})$.
(3) If $[K : \mathbf{Q}] = 2$, then there is a unique square-free integer $d \in \mathbf{Z}$ ($d \neq 0, 1$) such that $K = \mathbf{Q}(\sqrt{d})$. Write $|d|$ as a product of distinct primes and use (1)-(2) to show that $K \subset \mathbf{Q}(\mu_{|D|})$, where $D = d$ if $d \equiv 1 \pmod 4$ (resp. $D = 4d$ if $d \not\equiv 1 \pmod 4$).
(4) Show that $K \not\subset \mathbf{Q}(\mu_n)$ if $n < |D|$.

13. Euclidean constructibility

We identify the Euclidean plane $\mathbf{R}^2$ with $\mathbf{C}$ in the usual way (the point $\begin{pmatrix} x \\ y \end{pmatrix}$ corresponds to $x + iy$).

(13.1) Definition. *A complex number $z \in \mathbf{C}$ (= the corresponding point) is* **constructible** *(by ruler and compass) if it can be obtained from the points 0 by 1 by applying successively the following constructions: drawing a line through two already constructed points; drawing a circle whose centre and one point have already been constructed; intersecting two lines (resp. a line and a circle, resp. two circles) which have already been constructed.*

(13.2) Proposition. *(1) The set $E \subset \mathbf{C}$ of constructible complex numbers is a field.*
(2) $z \in E \iff \operatorname{Re}(z), \operatorname{Im}(z) \in E$.
(3) $z \in E \implies \pm\sqrt{z} \in E$.

Proof. Exercise in elementary geometry.

(13.3) Theorem. *Let $\alpha \in \mathbf{C}$. The following properties of α are equivalent.*
(1) $\alpha \in E$.
(2) There exists a tower of field extensions $\mathbf{Q} = K_0 \subset K_1 \subset \cdots \subset K_n$ such that $\forall j \ \ [K_{j+1} : K_j] = 2$ and $\alpha \in K_n$.
(3) There exists a tower of field extensions $\mathbf{Q} = K_0 \subset K_1 \subset \cdots \subset K_n = \mathbf{Q}(\alpha)$ such that $\forall j \ \ [K_{j+1} : K_j] = 2$.
(4) α is an algebraic number and $[\mathbf{Q}(\alpha_1, \ldots, \alpha_k) : \mathbf{Q}] = 2^m$ for some $m \geq 0$, where $\alpha_1, \ldots, \alpha_k$ are the conjugates of $\alpha = \alpha_1$ over $\mathbf{Q}$.

Proof. (1) $\iff$ (2): Exercise in elementary geometry.
(2) $\implies$ (4): For each α_j there exists a field embedding $\tau_j : \mathbf{Q}(\alpha) \hookrightarrow \mathbf{C}$ such that $\tau_j(\alpha) = \alpha_j$. If we put the towers $\mathbf{Q} = \tau_j(K_0) \subset \tau_j(K_1) \subset \cdots \subset \tau_j(K_n) \ni \alpha_j$ (for $j = 1, \ldots, k$) on top of each other, we obtain a bigger tower $\mathbf{Q} = L_0 \subset L_1 \subset \cdots \subset L_b$ in which $[L_{a+1} : L_a] \leq 2$ for each a and $\alpha_1, \ldots, \alpha_k \in L_b$. As a result, $[\mathbf{Q}(\alpha_1, \ldots, \alpha_k) : \mathbf{Q}]$ divides $[L_b : \mathbf{Q}] = 2^r$.
(4) $\implies$ (3): $L/\mathbf{Q} = \mathbf{Q}(\alpha_1, \ldots, \alpha_k)/\mathbf{Q}$ is a Galois extension and the order of $G = \mathrm{Gal}(L/\mathbf{Q})$ is a power of 2. According to Proposition 14.7 below, for each subfield $\mathbf{Q} \neq K \subset L$ there exists a subfield $K' \subset K$ such that $[K : K'] = 2$ (if $K = L^H$, then $K' = L^{H'}$). We take $K = \mathbf{Q}(\alpha) = K_n$ (where $[K : \mathbf{Q}] = 2^n$) and obtain, by induction, a chain of subfields $K_n \supset K_{n-1} \supset \cdots \supset K_0 = \mathbf{Q}$ such that $[K_{j+1} : K_j] = 2$ for all j.
(3) $\implies$ (2): This is automatic.

(13.4) Corollary (Gauss). *All vertices of a regular polygon with $n \geq 3$ sides inscribed to the unit circle are constructible $\iff$ $\zeta_n = e^{2\pi i/n}$ is constructible $\iff$ there exists $m \geq 0$ such that $[\mathbf{Q}(\zeta_n) : \mathbf{Q}] = \varphi(n) = 2^m$ $\iff$ $n = 2^a p_1 \cdots p_k$, where $p_1, \ldots, p_k$ are distinct prime numbers such that $p_j = 2^{m_j} + 1$ ($m_j \geq 1$).*

(13.5) Exercise. *If $p = 2^m + 1$ is a prime number, then $m = 2^b$, hence $p = 2^{2^b} + 1 = F_b$ ("Fermat's prime number"). At present (April 2014) the only known Fermat's prime numbers are $F_0 = 3$, $F_1 = 5$, $F_2 = 17$, $F_3 = 257$ and $F_4 = 65537$ (for example, F_5 is divisible by 641, as discovered by Euler).*

(13.6) Example: constructibility of a regular 17-gon (Gauss).
The Galois group $G = \mathrm{Gal}(\mathbf{Q}(\mu_{17})/\mathbf{Q}) \xrightarrow{\sim} (\mathbf{Z}/17\mathbf{Z})^*$ is cyclic of order $16 = 2^4$, generated by $g : \zeta \mapsto \zeta^3$ ($\zeta \in \mu_{17}$). The lattice of all subgroups of G

$$G = H_0 \supset H_1 \supset H_2 \supset H_3 \supset H_4 = \{1\},$$

where $H_j = \{g_j, g_j^2, \ldots, g_j^{2^{4-j}} = 1\}$ is generated by $g_j = g^{2^j}$, corresponds to a tower of fields as in Theorem 13.3:

$$\mathbf{Q} = K_0 \subset K_1 \subset K_2 \subset K_3 \subset K_4 = \mathbf{Q}(\mu_{17}) = \mathbf{Q}(\zeta),$$

where $K_j = \mathbf{Q}(\mu_{17})^{H_j}$ and $\zeta = \zeta_{17} = e^{2\pi i/17}$. If we write

$$a_j = \mathrm{Tr}_{K_4/K_j}(\zeta) = \sum_{\sigma \in H_j} \sigma(\zeta) = \sum_{k=1}^{2^{4-j}} \zeta^{\left(3^{(k \cdot 2^j)}\right)},$$

then $K_j = \mathbf{Q}(a_j)$ and $[K_j : K_{j-1}] = 2$. The conjugates of a_j over K_{j-1} are a_j and $a_j' = g_{j-1}(a_j)$, since $H_{j-1} = H_j \cup g_{j-1} H_j$. Explicitly,

$$\begin{aligned}
a_0 &= -1 & a_0' &= -1 \\
a_1 &= \zeta + \zeta^9 + \zeta^{-4} + \zeta^{-2} + \zeta^{-1} + \zeta^{-9} + \zeta^4 + \zeta^2 & a_1' &= \zeta^3 + \zeta^{-7} + \zeta^5 + \zeta^{-6} + \zeta^{-3} + \zeta^7 + \zeta^{-5} + \zeta^6 \\
a_2 &= \zeta + \zeta^4 + \zeta^{-4} + \zeta^{-1} & a_2' &= \zeta^9 + \zeta^2 + \zeta^{-2} + \zeta^{-9} \\
a_3 &= \zeta + \zeta^{-1} = 2\cos\tfrac{2\pi}{17} & a_3' &= \zeta^{-4} + \zeta^4 \\
a_4 &= \zeta & a_4' &= \zeta^{-1}
\end{aligned}$$

In particular, $a_j + a_j' = a_{j-1}$; one can compute explicitly the products $b_{j-1} = a_j a_j' \in K_{j-1}$, hence obtain the minimal polynomial $(X - a_j)(X - a_j') = X^2 - a_{j-1}X + b_{j-1}$ of a_j over K_{j-1} (see [Es], p. 149–150). The final result states that

$$16\cos\tfrac{2\pi}{17} = -1 + \sqrt{17} + \sqrt{34 - 2\sqrt{17}} + \sqrt{68 + 12\sqrt{17} - 4\sqrt{34 - 2\sqrt{17}} - 8\sqrt{34 + 2\sqrt{17}}}.$$

14. Basic properties of p-groups

Let p be a prime number. In this section we recall basic properties of p-groups (i.e., of finite groups of p-power order).

(14.1) Proposition. *If a finite group G of order $|G| = p^r$ $(r \geq 0)$ acts on a finite set X, then*

$$|X| \equiv |X^G| \pmod{p}.$$

Proof. The stabiliser of any $x \in X \setminus X^G$ is a proper subgroup $G_x \subsetneq G$, which implies that the cardinality of the orbit of x is a non-trivial power of p: $|O(x)| = (G : G_x) = p^a$ $(a \geq 1)$. The set $X \setminus X^G$ is a disjoint union of such orbits; as a result, its cardinality is divisible by p.

(14.2) Corollary (Cauchy). *If the order of a finite group H is divisible by p, then H contains an element of order p.*

Proof. Consider the action of $G = \mathbf{Z}/p\mathbf{Z}$ by cyclic permutations on $X = \{(h_1, \ldots, h_p) \in H^p \mid h_1 \cdots h_p = e\}$. The fixed point set $X^G = \{(h, \ldots, h) \mid h \in H,\ h^p = e\}$ contains $(e, \ldots, e)$ and

$$|\{h \in G \mid h^p = e \neq h\}| = |X^G| - 1 \equiv |X| - 1 = |H|^{p-1} - 1 \equiv -1 \not\equiv 0 \pmod{p}.$$

(14.3) Corollary. *The centre $Z(G) = \{h \in G \mid \forall g \in H \quad gh = hg\}$ of a finite group G of order $|G| = p^r$ $(r \geq 1)$ is non-trivial.*

Proof. The centre $Z(G) = X^G$ is the fixed point set of the action of G on $X = G$ by conjugation: $g * h = ghg^{-1}$. Therefore $|Z(G)| \equiv |G| \equiv 0 \pmod{p}$.

(14.4) Corollary. *For every finite group G of order $|G| = p^n$ $(n \geq 1)$ there exists a chain of subgroups $G = G_0 \supsetneq G_1 \supsetneq \cdots \supsetneq G_n = \{e\}$ such that each $G_i \lhd G$ is a normal subgroup of G and each quotient group G_i/G_{i+1} is cyclic of order p. [In particular, G is a nilpotent group.]*

Proof. We define, inductively, G_{n-1} to be any subgroup of order p of $Z(G)$, then G_{n-2} to be the inverse image in G of any subgroup of order p of $Z(G/G_{n-1}) \subset G/G_{n-1}$, etc.

(14.5) Definition. *Let H be a subgroup of a group G. The* **normaliser** *of H in G*

$$N_G(H) = \{g \in G \mid gHg^{-1} = H\}$$

is the biggest subgroup of G in which H is a normal subgroup.

(14.6) Proposition. *Let H, H' be subgroups of a group G. Consider the usual action of G on $X = G/H$.*
(1) The fixed point set of H' is equal to

$$(G/H)^{H'} = \{gH \mid g \in G,\ H' \subset gHg^{-1}\}.$$

(2) If $|H| = |H'| < \infty$, then

$$(G/H)^{H'} = \{gH \mid g \in G,\ H' = gHg^{-1}\}.$$

(3) If $H = H'$ and $|H| < \infty$, then

$$(G/H)^{H} = N/H, \qquad N = N_G(H).$$

Proof. (1) The group H' fixes $gH \in X \iff H'gH = gH \iff g^{-1}H'gH = H \iff g^{-1}H'g \subset H$.
(2) The equality $|H'| = |H| = |gHg^{-1}|$ implies that the inclusion $H' \subset gHg^{-1}$ is equivalent to $H' = gHg^{-1}$.
(3) This is a special case $H' = H$ of (2).

(14.7) Proposition. *If G is finite group and $H \subset G$ is a subgroup such that $|H| = p^r$ $(r \geq 0)$ and $p \mid (G : H)$, then there exists a subgroup $H \lhd H' \subset G$ such that $(H' : H) = p$ ($\iff$ $H'/H \xrightarrow{\sim} \mathbf{Z}/p\mathbf{Z}$).*

Proof. Let $N = N_G(H)$. According to Proposition 14.6(3) and 14.1 we have

$$|N/H| = |(G/H)^H| \equiv |(G/H)| = (G : H) \equiv 0 \pmod{p}.$$

Furthermore, Corollary 14.2 implies that there exists $n \in N$ whose image $nH \in N/H$ in N/H has order p. The subgroup H' of N generated by H and n has the required property.

(14.8) Definition. *Let G be a finite group whose order is divisible by p. A subgroup $H \subset G$ is a* **p-Sylow subgroup of** *G if $|H| = p^r$ $(r \geq 1)$ and $p \nmid (G : H)$ (in other words, if $|G| = p^r m$, $p \nmid m$ and $|H| = p^r$). [Example : D_8 is a 2-Sylow subgroup of S_4.]*

(14.9) Theorem (Sylow). *Let G be a finite group whose order is divisible by p.*
(1) A p-Sylow subgroup H of G exists.
(2) Two p-Sylow subgroups H, H' of G are conjugate: there exists $g \in G$ such that $H' = gHg^{-1}$.
(3) The number d of p-Sylow subgroups of G divides $|G|$ and satisfies $d \equiv 1 \pmod{p}$.

Proof. (1) According to Corollary 14.2 there exists a subgroup $H_1 \subset G$ of order p. Proposition 14.7 implies, inductively, that there exist subgroups $H_1 \subset H_2 \subset \cdots \subset H_r \subset G$ such that $|H_i| = p^i$ and $p \nmid (G : H_r)$. The last group H_r is a p-Sylow subgroup of G.
(2) According to Proposition 14.6(3) and 14.1, the set $\{gH \mid g \in G,\ H' = gHg^{-1}\} = (G/H)^{H'}$ is non-empty, since

$$|(G/H)^{H'}| \equiv |(G/H)| \not\equiv 0 \pmod{p}.$$

(3) If we let G act by conjugation on the set Y of all p-Sylow subgroups of G, we deduce from (2) that Y is in bijection with G/N, where $N = N_G(H)$. In particular, $d = |Y| = (G : N)$ divides $|G|$. The congruence used in the proof of (2) states (for $H' = H$) that

$$(N : H) = |(G/H)^H| \equiv |(G/H)| = d(N : H) \pmod{p}.$$

The index $(N : H) \mid (G : H)$ is prime to p, which implies that $1 \equiv d \pmod{p}$.

15. Kummer theory

Kummer theory describes explicitly Galois groups of all extensions $K(\sqrt[n]{a_1}, \ldots, \sqrt[n]{a_r})/K$, where $a_i \in K^*$, $\mu_n \subset K$ and $char(K) \nmid n$.

(15.1) Example. The isomorphism

$$G = \mathrm{Gal}(\mathbf{Q}(\sqrt{2}, \sqrt{3})/\mathbf{Q}) \xrightarrow{\sim} \{\pm 1\} \times \{\pm 1\}, \qquad \sigma \mapsto \left(\frac{\sigma(\sqrt{2})}{\sqrt{2}}, \frac{\sigma(\sqrt{3})}{\sqrt{3}}\right) \tag{15.1.1}$$

from Example 9.6(iii) can be succinctly reformulated by saying that G is dual to the group $\Delta \subset \mathbf{Q}^*/\mathbf{Q}^{*2}$ generated by the images of 2 and 3. Main theorem of Kummer theory (Theorem 15.7 below) is a generalisation of this fact.

(15.2) The general setup. Fix an integer $n \geq 2$ and a field K such that $|\mu_n(K)| = n$ (this condition implies that $char(K) \nmid n$). We abbreviate $\mu_n = \mu_n(K)$.

Given $a_1, \ldots, a_r \in K^*$, let L be a splitting field over K of the polynomial

$$(T^n - a_1) \cdots (T^n - a_r) \in K[T].$$

For any fixed root $\alpha_i = \sqrt[n]{a_i} \in L$ of the polynomial $T^n - a_i$ we have

$$T^n - a_i = \prod_{\zeta \in \mu_n} (T - \zeta\alpha_i).$$

In particular, this polynomial is separable, which implies that L/K is a Galois extension. Denote by $G = \mathrm{Gal}(L/K)$ its Galois group. Note that

$$L = K(\alpha_1, \dots, \alpha_n) = K(\sqrt[n]{a_1}, \dots, \sqrt[n]{a_r}) \quad \text{for any fixed root } \alpha_i = \sqrt[n]{a_i} \in L.$$

(15.3) Two finite abelian groups. For $a \in K^*$ denote by $\overline{a}$ its image in K^*/K^{*n}.
Let $\Delta \subset K^*/K^{*n}$ be the subgroup generated by $\overline{a}_1, \dots \overline{a}_r$.
In Example 15.1 we have $n = 2$, $K = \mathbf{Q}$, $a_1 = 2$ and $a_2 = 3$; therefore $\Delta = \{\overline{1}, \overline{2}, \overline{3}, \overline{6}\} \subset \mathbf{Q}^*/\mathbf{Q}^{*2}$.
In general, Δ is a subgroup of

$$\Delta' = \mathrm{Ker}\,(K^*/K^{*n} \longrightarrow L^*/L^{*n}) = (K^* \cap L^{*n})/K^{*n}.$$

The key point of the whole theory is the following generalisation of the map (15.1.1).

(15.4) Proposition-Definition (Construction of a pairing). *(1) The formula*

$$\sigma \times \overline{a} \mapsto (\sigma, \overline{a}) = \frac{\sigma(\alpha)}{\alpha} \ \left(= \frac{\sigma(\sqrt[n]{a})}{\sqrt[n]{a}} \right) \in \mu_n,$$

where $\sigma \in G$, $a \in K^*$, $\overline{a} \in \Delta'$, $\alpha \in L^*$, $\alpha^n = a$ (α is any n-th root $\sqrt[n]{a} \in L^*$) defines a map

$$(\ ,\ ) : G \times \Delta' \longrightarrow \mu_n.$$

(2) Linearity in the second argument: $(\sigma, \overline{a}\overline{b}) = (\sigma, \overline{a})(\sigma, \overline{b})$.
(3) Linearity in the first argument: $(\sigma\tau, \overline{a}) = (\sigma, \overline{a})(\tau, \overline{a})$.
(4) Non-degeneracy on the left: let $\sigma \in G$. If $(\sigma, \overline{a}) = 1$ for all $\overline{a} \in \Delta$, then $\sigma = \mathrm{id}$.
(5) Non-degeneracy on the right: let $\overline{a} \in \Delta'$. If $(\sigma, \overline{a}) = 1$ for all $\sigma \in G$, then $\overline{a} = 1$.

Proof. (1) The value of $(\sigma, \overline{a})$ lies in μ_n, since $(\sigma(\alpha)/\alpha)^n = \sigma(\alpha^n)/\alpha^n = \sigma(a)/a = 1$. We must show that $\sigma(\alpha)/\alpha$ depends only on σ and $\overline{a}$. If $a, b \in K^*$ and $\overline{a} = \overline{b} \in \Delta'$, then there are $\alpha, \beta \in L^*$ and $c \in K^*$ such that $a = \alpha^n$, $b = \beta^n$ and $a = bc^n$. It follows that $\alpha/\beta c \in \mu_n$ and $\alpha/\beta \in K$. In particular,

$$\frac{\sigma(\alpha)}{\alpha} = \frac{\sigma(\beta)}{\beta} \frac{\sigma(\alpha/\beta)}{\alpha/\beta} = \frac{\sigma(\beta)}{\beta}.$$

(2) If $a = \alpha^n$ and $b = \beta^n$, then

$$\frac{\sigma(\alpha\beta)}{\alpha\beta} = \frac{\sigma(\alpha)}{\alpha} \frac{\sigma(\beta)}{\beta}.$$

(3) We have $\sigma(\zeta) = \zeta$ for all $\zeta \in \mu_n$, hence

$$\frac{\sigma(\tau(\alpha))}{\alpha} = \frac{\sigma(\alpha)}{\alpha} \sigma\left(\frac{\tau(\alpha)}{\alpha}\right) = \frac{\sigma(\alpha)}{\alpha} \frac{\tau(\alpha)}{\alpha}.$$

(4) If $\sigma(\alpha_i)/\alpha_i = 1$ for all $i = 1, \dots, r$, then $\sigma = \mathrm{id}$ on $K(\alpha_1, \dots, \alpha_r) = L$.
(5) If $a = \alpha^n$ and $\sigma(\alpha)/\alpha = 1$ for all $\sigma \in G$, then $\alpha \in (L^*)^G = K^*$, hence $a \in K^{*n}$ and $\overline{a} = 1$.

(15.5) Corollary. *(1) For fixed $\sigma \in G$ the map $\overline{a} \mapsto (\sigma, \overline{a})$ is a homomorphism of abelian groups $\Delta' \longrightarrow \mu_n$.*
(2) The map $\sigma \mapsto (\overline{a} \mapsto (\sigma, \overline{a}))$ is group homomorphism $G \longrightarrow \mathrm{Hom}_{\mathbf{Z}}(\Delta', \mu_n)$.
(3) The group homomorphism $G \xrightarrow{(2)} \mathrm{Hom}_{\mathbf{Z}}(\Delta', \mu_n) \xrightarrow{\mathrm{res}} \mathrm{Hom}_{\mathbf{Z}}(\Delta, \mu_n)$ is injective.
(4) G is a finite abelian group satisfying $G = G[n]$.
(5) For fixed $\overline{a} \in \Delta'$ the map $\sigma \mapsto (\sigma, \overline{a})$ is a homomorphism of abelian groups $G \longrightarrow \mu_n$.
(6) The map $\overline{a} \mapsto (\sigma \mapsto (\sigma, \overline{a}))$ is a homomorphism of abelian groups $\Delta' \longrightarrow \mathrm{Hom}_{\mathbf{Z}}(G, \mu_n)$.
(7) The homomorphism (6) is injective.

Proof. (1) and (6) (resp. (5) and (2)) are consequences of Proposition 15.4(2) (resp. of Proposition 15.4(3)). (3) (resp. (7)) is a consequence of Proposiiton 15.4(4) (resp. of Proposition 15.4(5)). Finally, (4) follows from (3).

15.6. Duality (see Exercise II.1.15). Let A be a finite abelian group satisfying $A = A[n]$. Its dual group $D(A) = \mathrm{Hom}_{\mathbf{Z}}(A, \mu_n)$ is a finite abelian group having the same property. If A is cyclic of order $d \mid n$, so is $D(A)$. Combined with the additivity property $D(A \oplus B) = D(A) \oplus D(B)$, this implies that $D(A)$ is isomorphic to A, for any A (since A is isomorphic to a direct sum of cyclic groups). The evaluation map $D(A) \times A \longrightarrow \mu_n$ is non-degenerate on both sides (since the statement holds for A cyclic).

(15.7) Main Theorem of Kummer theory. *We have $\Delta' = \Delta$ and the pairing defined in Proposition 15.4 gives rise to isomorphisms of finite abelian groups*

$$G \xrightarrow{\sim} \mathrm{Hom}_{\mathbf{Z}}(\Delta, \mu_n), \qquad \Delta \xrightarrow{\sim} \mathrm{Hom}_{\mathbf{Z}}(G, \mu_n).$$

In particular, $G = \mathrm{Gal}(L/K)$ is (non-canonically) isomorphic to Δ and $|\Delta| = |G| = [L : K]$.

Proof. Putting together Proposition 15.5(3) and 15.5(7), we obtain

$$|G| \leq |D(\Delta)| = |\Delta|, \qquad |\Delta'| \leq |D(G)| = |G|.$$

However, $\Delta \subset \Delta'$, which implies that there are equalities everywhere: $\Delta = \Delta'$ and $|G| = |\Delta|$. As a result, the injective homomorphisms in Proposition 15.5(3) and 15.5(6) are both isomorphisms.

(15.8) Exercise. *Show that any (finite) linear relation*

$$\sum_{0<a\in\mathbf{Q}} u(a)\,\sqrt{a} = 0 \qquad\qquad (u(a) \in \mathbf{Q})$$

is a sum of tautological relations

$$u\sqrt{b^2c} - ub\sqrt{c} = 0 \qquad\qquad (u, b, c \in \mathbf{Q};\ b, c > 0).$$

(15.9) Theorem. *Assume that K is a field such that $|\mu_n(K)| = n$. Let L'/K be a Galois extension with Galois group $G' = \mathrm{Gal}(L'/K)$.*
(1) For each group homomorphism $\chi : G' \longrightarrow \mu_n = \mu_n(K)$ there exists $\alpha \in L'^*$ such that $\alpha^n = a \in K^*$ and $\chi(\sigma) = \sigma(\alpha)/\sigma$ for all $\sigma \in G'$.
(2) If G' is abelian and satisfies $G' = G'[n]$, then there exist $a_1, \ldots, a_r \in K^*$ (where r is the minimal number of generators of G') such that $L' = K(\sqrt[n]{a_1}, \ldots, \sqrt[n]{a_r})$.

Proof. (1) For each $\beta \in L'$ the "Lagrange resolvent"

$$\alpha = \sum_{\tau\in G'} \chi(\tau)^{-1}\,\tau(\beta) \in L'$$

satisfies

$$\forall \sigma \in G' \quad \sigma(\alpha) = \sum_{\tau\in G'} \chi(\tau)^{-1}\,(\sigma\tau)(\beta) = \sum_{\tau'\in G'} \chi(\sigma^{-1}\tau')^{-1}\,\tau'(\beta) = \chi(\sigma) \sum_{\tau'\in G'} \chi(\tau')^{-1}\,\tau'(\beta) = \chi(\sigma)\alpha,$$

hence $a = \alpha^n \in (L')^{G'} = K$. It follows from Corollary 15.11 below that there exists $\beta \in L'$ for which $\alpha \neq 0$, which concludes the proof.

Note that the homomorphism χ factors through $G'/\mathrm{Ker}(\chi) \hookrightarrow \mu_n$, which means that we could have replaced right at the beginning of the argument L' by $(L')^{\mathrm{Ker}(\chi)}$ and $\mathrm{Gal}(L'/K) = G'$ by $\mathrm{Gal}((L')^{\mathrm{Ker}(\chi)}/K) = G'/\mathrm{Ker}(\chi)$, hence assume that G' is abelian and satisfies $G' = G'[n]$. However, this was not necessary.
(2) Fix an isomorphism $G' \xrightarrow{\sim} G_1 \oplus \cdots \oplus G_r$, where each G_i is a cyclic group, of order $d_i \mid n$. For $i = 1, \ldots, r$ denote by $\mathrm{pr}_i : G' \longrightarrow G_i$ the projection and fix an injective group homomorphism $\chi_i : G_i \hookrightarrow \mu_n$. According to (1) there exists $\alpha_i \in L'^*$ such that $\alpha_i^n = a_i \in K^*$ and $\chi_i(\mathrm{pr}_i(\sigma)) = \sigma(\alpha_i)/\alpha_i$, for each $\sigma \in G'$.

We claim that the field $L = K(\alpha_1, \ldots, \alpha_r) \subset L'$ coincides with L'. Indeed, if $\sigma \in \mathrm{Gal}(L'/L) \subset G'$, then $\chi_i(\mathrm{pr}_i(\sigma)) = 1$, which means that $pr_i(\sigma) = 0$ for each i; thus $\sigma = \mathrm{id}$, $\mathrm{Gal}(L'/L) = \{\mathrm{id}\}$ and $L' = L$.

(15.10) Proposition (Linear independence of characters). *For any abelian group A and any field M, any finite set of (distinct) homomorphisms of abelian groups $\chi_1, \dots, \chi_n : A \longrightarrow M^*$ is linearly independent over M. In other words, if $\lambda_1, \dots, \lambda_n \in M$ and*

$$\forall a \in A \quad \sum_{i=1}^{n} \lambda_i \, \chi_i(a) = 0,$$

then $\lambda_1 = \cdots = \lambda_n = 0$.

Proof. Induction on n. The case $n = 1$ is trivial. If $n \geq 2$, then we have, for all $a, b \in A$,

$$\sum_{i=1}^{n} \lambda_i \, \chi_i(ab) = 0, \qquad \chi_n(a) \sum_{i=1}^{n} \lambda_i \, \chi_i(b) = 0,$$

hence

$$\sum_{i=1}^{n-1} \lambda_i (\chi_i(a) - \chi_n(a)) \, \chi_i(b) = 0.$$

The induction hypothesis implies that

$$\forall a \in A \; \forall i = 1, \dots, n-1 \quad \lambda_i(\chi_i(a) - \chi_n(a)) = 0.$$

The homomorphisms χ_j are distinct, which means that for each $i < n$ there exists $a \in A$ such that $\chi_i(a) - \chi_n(a) \neq 0$, hence $\lambda_i = 0$.

(15.11) Corollary (Linear independence of field embeddings). *Let $L \supset K \subset M$ be fields. Any finite set of (distinct) homomorphisms of K-algebras $\sigma_1, \dots, \sigma_n \in \mathrm{Hom}_{K-\mathrm{Alg}}(L, M)$ is linearly independent over M. In other words, if $\lambda_1, \dots, \lambda_n \in M$ and*

$$\forall y \in L \quad \sum_{i=1}^{n} \lambda_i \, \sigma_i(y) = 0,$$

then $\lambda_1 = \cdots = \lambda_n = 0$.

Proof. Apply Proposition 15.10 to $A = L^*$ and $\chi_i = \sigma_i|_{L^*}$.

(15.12) Representations of finite groups. Let K be a field, let G be a finite group. Recall that a **representation of G over K** is a group homomorphism $\rho : G \longrightarrow \mathrm{Aut}_K(V)$, where V is a finite-dimensional K-vector space. After choosing a basis of V, this becomes a group homomorphism $\rho : G \longrightarrow GL_m(K)$, where $m = \dim_K(V)$.

The **group ring of G over K** is the unital K-algebra $K[G] = \{\sum_{g \in G} a_g \, g \mid a_g \in K\}$ with operations

$$\sum_{g \in G} a_g \, g + \sum_{g \in G} b_g \, g = \sum_{g \in G} (a_g + b_g) \, g, \qquad \big(\sum_{g \in G} a_g \, g\big)\big(\sum_{h \in G} b_h \, h\big) = \sum_{g,h \in G} (a_g b_h) \, gh.$$

It is commutative $\iff$ G is an abelian group.

A representation $\rho : G \longrightarrow \mathrm{Aut}_K(V)$ defines on V a structure of a (left) $K[G]$-module

$$\big(\sum_{g \in G} a_g \, g\big) v = \sum_{g \in G} a_g \, \rho(g) v$$

and vice versa: $\rho(g) \in \mathrm{Aut}_K(V) \subset \mathrm{End}_K(V)$ is given by the action of the invertible element $1 \cdot g \in K[G]$.

In particular, $V = K[G]$ with action given by left multiplication is a representation of G over K, called the **regular representation of G**.

If $char(K) \nmid |G|$, then everything works very much as in the classical case $K = \mathbf{C}$:

(15.12.1) each representation of G over K is isomorphic to a direct sum of irreducible representations ("complete irreducibility");

(15.12.2) each irreducible representation of G over K is isomorphic to a subrepresentation of the regular representation $K[G]$;

(15.12.3) if $K = \overline{K}$ is algebraically closed, then each irreducible representation of G over K occurs in $K[G]$ with multiplicity equal to its dimension $\dim_K(\rho)$.

(15.13) Kummer theory and Galois module structure. Let K, $L = L' = K(\sqrt[n]{a_1}, \ldots, \sqrt[n]{a_r})$ and $G = G' = \mathrm{Gal}(L/K) = G[n]$ be as in Theorem 15.9(2). In this case $char(K) \nmid |G|$, since $char(K) \nmid n$.

The group G is abelian, which means that its irreducible representations over $\overline{K}$ are one-dimensional, given by group homomorphisms $\chi : G \longrightarrow \overline{K}^*$. As $G = G[n]$, the image of χ is contained in $\mu_n(\overline{K}) = \mu_n \subset K^*$; in particular, χ is defined over K. As a result, all representations of G over $\overline{K}$ are defined over K and irreducible representations are precisely the elements of $D(G) = \mathrm{Hom}_{\mathbf{Z}}(G, \mu_n)$. The statement (15.12.3) implies that there is an isomorphism of $K[G]$-modules

$$K[G] \xrightarrow{\sim} \bigoplus_{\chi \in D(G)} \chi.$$

On the other hand, the $K[G]$-module L is isomorphic to

$$L \xrightarrow{\sim} \bigoplus_{\chi \in D(G)} \chi^{\oplus m(\chi)},$$

for suitable multiplicities $m(\chi) \geq 0$, by (15.12.1). Theorem 15.9(1) states that $m(\chi) \geq 1$ for all $\chi \in D(G)$. As $\dim_K(L) = |G| = |D(G)| = \dim_K(K[G])$, it follows that $m(\chi) = 1$ for each χ. Therefore we obtain an isomorphism of $K[G]$-modules

$$K[G] \xrightarrow{\sim} L \tag{15.13.1}$$

(which is, essentially, equivalent to Theorem 15.9(1), by the previous discussion). As we shall see in Theorem 18.4 below, the isomorphism (15.13.1) holds for arbitrary Galois extensions, even if $char(K) \mid |G|$.

(15.14) Exercise. *Let $n \geq 2$, let K be a field such that $|\mu_n(K)| = n$, let $a \in K^*$ be an element for which the field $L = K(\sqrt[n]{a})$ satisfies $[L : K] = n$. Show that: there exists a Galois extension M/K such that $M \supset L$ and $\mathrm{Gal}(M/K) \xrightarrow{\sim} \mathbf{Z}/n^2\mathbf{Z} \iff \mu_n \subset N_{L/K}(L^*)$.*

16. Generalised Kummer theory

In this section we investigate splitting fields of polynomials $X^n - a$ over fields of characteristic not dividing n, without assuming that the base field contains all n-th roots of unity.

(16.1) Fix an integer $n \geq 2$. Let K be a field such that $char(K) \nmid n$. Fix $a \in K^*$ and let L be a splitting field over K of the polynomial $f = X^n - a$. The assumption $n \in K^*$ implies that f is separable, since $(f, f') = (X^n - a, nX^{n-1}) = (X^n - a, X^{n-1}) = (a) = (1)$. For any root $\alpha \in L$ of f, the polynomial f splits in $L[X]$ as

$$f(X) = X^n - a = \prod_{\zeta \in \mu_n} (X - \zeta\alpha),$$

where $\mu_n = \mu_n(L)$ (with $|\mu_n| = n$, by Proposition 12.5(2)). It follows that there is a tower of extensions

$$K \hookrightarrow K(\mu_n) = K(\zeta_n) \hookrightarrow L = K(\mu_n, \alpha) = K(\zeta_n, \alpha),$$

for any generator ζ_n of the cyclic group μ_n.

(16.2) Denote by $G = \mathrm{Gal}(L/K) = \mathrm{Gal}(f)$ the Galois group of f. The intermediate Galois groups

$$H = \mathrm{Gal}(L/K(\mu_n)), \qquad G/H = \mathrm{Gal}(K(\mu_n)/K)$$

were studied, respectively, in §15 and §12. The cyclotomic character $\chi = \chi_{n,K}$ from Proposition 12.5(4) defines a canonical injective group homomorphism

$$\chi : G/H \hookrightarrow (\mathbf{Z}/n\mathbf{Z})^*, \qquad \forall \zeta \in \mu_n \quad g(\zeta) = \zeta^{\chi(g)}$$

into the **multiplicative group** of the ring $\mathbf{Z}/n\mathbf{Z}$, while the pairing 15.4 gives rise (after fixing a group isomorphism $\mu_n \xrightarrow{\sim} \mathbf{Z}/n\mathbf{Z}$, i.e., after fixing a generator $\zeta_n \in \mu_n$ corresponding to $1 \in \mathbf{Z}/n\mathbf{Z}$) to an injective group homomorphism

$$H \hookrightarrow \mu_n \ (\xrightarrow{\sim} \mathbf{Z}/n\mathbf{Z}), \qquad h \mapsto h(\alpha)/\alpha$$

into the **additive group** of the ring $\mathbf{Z}/n\mathbf{Z}$.

How do these two homomorphisms fit together? The action of any $g \in G$ on L is determined by its action on the generating elements ζ_n and α:

$$\forall g \in G \qquad \begin{cases} g : \zeta_n \mapsto \zeta_n^{\chi(g)}, & \chi(g) \in (\mathbf{Z}/n\mathbf{Z})^* \\ g : \alpha \mapsto \alpha\, \zeta_n^{c(g)}, & c(g) \in \mathbf{Z}/n\mathbf{Z}. \end{cases}$$

The transitivity rule $g(g'(x)) = (gg')(x)$ amounts to checking, for all $g, g' \in G$, that the action of $g \circ g'$

$$\begin{cases} g \circ g' : \zeta_n \mapsto \zeta_n^{\chi(g')} \mapsto (\zeta_n^{\chi(g)})^{\chi(g')} = \zeta_n^{\chi(g)\chi(g')} \\ g \circ g' : \alpha \mapsto \alpha\, \zeta_n^{c(g')} \mapsto (\alpha\, \zeta_n^{c(g)})(\zeta_n^{\chi(g)})^{c(g')} = \alpha\, \zeta_n^{\chi(g)c(g')+c(g)} \end{cases}$$

coincides with the action of gg'

$$\begin{cases} gg' : \zeta_n \mapsto \zeta_n^{\chi(gg')} \\ gg' : \alpha \mapsto \alpha\, \zeta_n^{c(gg')}, \end{cases}$$

which is equivalent to

$$\forall g, g' \in G \qquad \chi(gg') = \chi(g)\chi(g'), \qquad c(gg') = \chi(g)c(g') + c(g). \tag{16.2.1}$$

These formulas can be written in a matrix form in terms of the following injective map

$$\rho : G \hookrightarrow GL_2(\mathbf{Z}/n\mathbf{Z}), \qquad g \mapsto \begin{pmatrix} \chi(g) & c(g) \\ 0 & 1 \end{pmatrix}$$

as

$$\rho(gg') = \begin{pmatrix} \chi(gg') & c(gg') \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \chi(g) & c(g) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \chi(g') & c(g') \\ 0 & 1 \end{pmatrix} = \rho(g)\rho(g').$$

In other words, ρ is an injective group homomorphism! As in 12.5, the Galois group G admits a natural Galois representation (depending on our choice of ζ_n), this time into the $\mathbf{Z}/n\mathbf{Z}$-valued points of the subgroup

$$GA_1 = \begin{pmatrix} * & * \\ 0 & 1 \end{pmatrix} \subset GL_2,$$

$$\rho : G = \mathrm{Gal}(K(\mu_n, \sqrt[n]{a})/K) \hookrightarrow GA_1(\mathbf{Z}/n\mathbf{Z}), \qquad g \mapsto \begin{pmatrix} \chi(g) & c(g) \\ 0 & 1 \end{pmatrix}. \tag{16.2.2}$$

What kind of a subgroup is GA_1? It is the affine group in dimension one.

(16.3) Definition. *Let R be a (commutative) ring, let $m \geq 1$ be an integer. The* **affine group of** R^m *is the group*

$$GA_m(R) = \left\{ \begin{pmatrix} A & a \\ 0 & 1 \end{pmatrix} \mid A \in GL_m(R),\, a \in R^m \right\} \subset GL_{m+1}(R).$$

It acts faithfully on R^m by affine automorphisms $x \mapsto Ax + a$ ($x \in R^m$). The group of translations $x \mapsto x + a$ is a normal subgroup $\begin{pmatrix} 1 & R \\ 0 & 1 \end{pmatrix} \lhd GA_m(R)$, with the quotient group naturally isomorphic to $GL_m(R)$ via the map $GA_m(R) \longrightarrow GL_m(R)$ given by $(x \mapsto Ax + a) \mapsto A$.

(16.4) In particular, the affine group

$$GA_1(\mathbf{Z}/n\mathbf{Z}) = \begin{pmatrix} (\mathbf{Z}/n\mathbf{Z})^* & \mathbf{Z}/n\mathbf{Z} \\ 0 & 1 \end{pmatrix}$$

occurring in (16.2.2) acts faithfully on $\mathbf{Z}/n\mathbf{Z}$ by

$$\begin{pmatrix} u & a \\ 0 & 1 \end{pmatrix} : x \mapsto ux + a;$$

this gives a natural injection $GA_1(\mathbf{Z}/n\mathbf{Z}) \hookrightarrow S_n$. Note that $|GA_1(\mathbf{Z}/n\mathbf{Z})| = n\varphi(n)$.

If $n \geq 3$, then $\{\pm 1\} \subset (\mathbf{Z}/n\mathbf{Z})^*$ and the (normal) subgroups

$$GA_1(\mathbf{Z}/n\mathbf{Z}) \supset \begin{pmatrix} \pm 1 & \mathbf{Z}/n\mathbf{Z} \\ 0 & 1 \end{pmatrix} \supset \begin{pmatrix} 1 & \mathbf{Z}/n\mathbf{Z} \\ 0 & 1 \end{pmatrix} \tag{16.4.1}$$

are isomorphic to $GA_1(\mathbf{Z}/n\mathbf{Z}) \supset D_{2n} \supset C_n$, with

$$GA_1(\mathbf{Z}/n\mathbf{Z})/C_n \xrightarrow{\sim} (\mathbf{Z}/n\mathbf{Z})^*, \qquad GA_1(\mathbf{Z}/n\mathbf{Z})/D_{2n} \xrightarrow{\sim} (\mathbf{Z}/n\mathbf{Z})^*/\{\pm 1\}. \tag{16.4.2}$$

For $n = 3, 4$ and 6 we have $(\mathbf{Z}/n\mathbf{Z})^* = \{\pm 1\}$, hence $GA_1(\mathbf{Z}/n\mathbf{Z}) = D_{2n}$.

(16.5) Example. If p is a prime number and $a \in \mathbf{Q}^*$, $a \notin \mathbf{Q}^{*p}$, then the polynomial $f(X) = X^p - a$ is irreducible in $\mathbf{Q}[X]$ (exercise!), hence $[\mathbf{Q}(\sqrt[p]{a}) : \mathbf{Q}] = p$, for any fixed p-th root $\sqrt[p]{a} \in \mathbf{C}$. The splitting field of f inside $\mathbf{C}$ is equal to $L = \mathbf{Q}(\zeta_p, \sqrt[p]{a})$, where $\zeta_p = e^{2\pi i/p}$. As the degrees $[\mathbf{Q}(\zeta_p) : \mathbf{Q}] = p - 1$ and $[\mathbf{Q}(\sqrt[p]{a}) : \mathbf{Q}] = p$ are relatively prime, the full degree $[L : \mathbf{Q}]$ is equal to $[L : \mathbf{Q}] = p(p-1) = |GA_1(\mathbf{F}_p)|$ (cf. 3.30). It follows that the injective group homomorphism (16.2.2) is an isomorphism in this case:

$$\rho : \mathrm{Gal}(f) = \mathrm{Gal}(\mathbf{Q}(\zeta_p, \sqrt[p]{a})/\mathbf{Q}) \xrightarrow{\sim} GA_1(\mathbf{F}_p).$$

(16.6) Exercise. *Let $a \in \mathbf{Q}^*$, $a \notin \mathbf{Q}^{*2}$, $-4a \notin \mathbf{Q}^{*4}$.*
(1) Show that the polynomial $X^4 - a$ is irreducible in $\mathbf{Q}[X]$.
(2) Assume, in addition, that $-a \notin \mathbf{Q}^{*2}$. Show that $X^4 - a$ is irreducible in $\mathbf{Q}(i)[X]$. Deduce that (16.2.2) induces an isomorphism

$$\rho : \mathrm{Gal}(\mathbf{Q}(i, \sqrt[4]{a})/\mathbf{Q}) \xrightarrow{\sim} GA_1(\mathbf{Z}/4\mathbf{Z}) = D_8.$$

Give a list of all subgroups of D_8 and of all subfields of $\mathbf{Q}(i, \sqrt[4]{a})$.

(16.7) Exercise. *(1) Determine all integers $n \geq 1$ for which the field $\mathbf{Q}(\mu_n) \subset \mathbf{C}$ (resp. $\mathbf{Q}(\mu_n) \cap \mathbf{R} \subset \mathbf{R}$) is of the form $\mathbf{Q}(\sqrt{a_1}, \ldots, \sqrt{a_r})$, for some $a_j \in \mathbf{Q}$.*
(2) Given $n \geq 1$, determine the number of subfields $F \subset \mathbf{Q}(\mu_n)$ satisfying $[F : \mathbf{Q}] = 2$.

(16.8) Proposition. *Let K be a field, $n \geq 1$ an integer such that $char(K) \nmid n$ and L/K a Galois extension. The Galois group $\mathrm{Gal}(L(\mu_n)/K(\mu_n))$ is then canonically isomorphic to a subgroup of $\mathrm{Gal}(L/K)$.*

Proof. This is a special case of Exercise 10.13(1), but we give a full argument here. According to Proposition 9.3, L is a splitting field over K of a separable polynomial $f \in K[X]$. The field $L(\mu_n)$ is then

a splitting field of f over $K(\mu_n)$. In particular, $L(\mu_n)/K(\mu_n)$ is a Galois extension. The restriction of any $g \in \mathrm{Gal}(L(\mu_n)/K(\mu_n))$ to L is an element of $\mathrm{Gal}(L/K)$; this defines a group homomorphism $\mathrm{res} : \mathrm{Gal}(L(\mu_n)/K(\mu_n)) \longrightarrow \mathrm{Gal}(L/K)$. Let $\alpha_1, \ldots, \alpha_r \in L(\mu_n)$ be the roots of f in $L(\mu_n)$; then $L = K(\alpha_1, \ldots, \alpha_r)$ and $L(\mu_n) = K(\mu_n)(\alpha_1, \ldots, \alpha_r)$. If $g \in \mathrm{Ker}(\mathrm{res})$, then $g|_{K(\mu_n)} = \mathrm{id}$ and $g(\alpha_j) = \alpha_j$ for all $j = 1, \ldots, r$; thus $g = \mathrm{id}$. As a result, the restriction homomorphism res is injective, as claimed.

17. Solvability by radicals

In this section we prove a celebrated result of Galois characterising those polynomial equations which can be solved by taking iterated roots $\sqrt[n]{a}$.

(17.1) Definition. *A group G is* **solvable** *if there exists a finite chain of subgroups $G = G_0 \supset G_1 \supset \cdots \supset G_k = \{e\}$ $(k < \infty)$ such that $G_{i+1} \lhd G_i$ and G_i/G_{i+1} is an abelian group for all $i = 0, \ldots, k-1$. If G is finite, this is equivalent to an apparently stronger condition requiring each quotient $G_i/G_{i+1} \xrightarrow{\sim} \mathbf{Z}/p_i\mathbf{Z}$ to be a cyclic group of prime order.*

(17.2) Examples. (1) An abelian group is solvable: $G \supset \{e\}$.
(2) For any commutative ring R, the affine group $G = GA_1(R)$ is solvable: $G \supset G_1 = (R, +) \supset \{0\}$, $G/G_1 \xrightarrow{\sim} R^*$.
(3) $G = S_3$ is solvable: $S_3 \supset A_3 \supset \{e\}$.
(4) $G = S_4$ is solvable: the action of S_4 on the polynomials y_1, y_2, y_3 from 1.4.6 defines a surjective group homomorphism $\pi : S_4 \longrightarrow S_3$, whose kernel is isomorphic to $C_2 \times C_2$. This yields a chain of subgroups $S_4 \supset A_4 \supset C_2 \times C_2 \supset \{e\}$ (see [Es], 10.8).
(5) A simple non abelian group is not solvable (recall that a group G is **simple** if G has no normal subgroup $H \neq G, \{e\}$).
(6) Any subgroup $H \subset G$ of a solvable group G is solvable (take $H_i = H \cap G_i$).
(7) The image of a solvable group G by any group homomorphism $f : G \longrightarrow K$ is solvable (take $f(G)_i = f(G_i)$).
(8) For $n \geq 5$ the group A_n is simple ([De 1], Thm. I.5.1). It follows from (5) (resp. from (6)) that A_n (resp. S_n) is not solvable for $n \geq 5$.
(9) Any group of order p^n (where p is a prime number and $n \geq 0$) is solvable (see Corollary 14.4).

(17.3) Definition. *Let K be a field.*
(1) A finite extension L/K is a **radical extension** *if there exists a tower of extensions*

$$K = K_0 \subset \cdots \subset K_j = K(\alpha_1, \ldots, \alpha_j) = K_{j-1}(\alpha_j) \subset \cdots \subset K_m = L$$

such that, for each $j = 1, \ldots, m$, $K_j = K_{j-1}(\alpha_j)$ with $\alpha_j^{n_j} = a_j \in K_{j-1}^*$ for some integer $n_j \geq 1$ satisfying $\mathrm{car}(K) \nmid n_j$. [Note that L/K is a separable extension, since each layer K_j/K_{j-1} is separable.]
(2) Let $f \in K[X]$ be a non-constant separable polynomial. The polynomial equation $f(X) = 0$ is **solvable by radicals over** *K if there exists a radical extension L of K containing a splitting field of f ($\Longleftrightarrow$ containing all roots of f).*

(17.4) Example. $K_0 = \mathbf{Q} \subset K_1 = \mathbf{Q}(\sqrt[4]{5}) \subset K_2 = \mathbf{Q}(\sqrt[3]{2 + \sqrt[4]{5}})$ (for a fixed choice of cubic and quartic roots).

(17.5) Proposition. *Each radical extension of K is contained in a radical extension L/K with the following properties.*
(i) L/K is a Galois extension.
(ii) $K_1 = K(\mu_n)$, where $\mathrm{char}(K) \nmid n$.
(iii) For each $j = 1, \ldots, m$, $K_j = K_{j-1}(\alpha_j)$ with $\alpha_j^n = a_j \in K_{j-1}^*$.

Proof. For example, the extension from Example 17.4 is contained in the following radical extension: $L = \mathbf{Q}(\mu_{12}, \rho^j \sqrt[3]{2 + i^k \sqrt[4]{5}};\ j = 0, 1, 2,\ k = 0, 1, 2, 3)$. In general, one proceeds recursively in the same manner; the details are left to the reader.

(17.6) Theorem (Galois). *Let K be a field, let $f \in K[X]$ be a non-constant separable polynomial.*
(1) If the equation $f = 0$ is solvable by radicals over K, then the Galois group $\mathrm{Gal}(f)$ (over K) is solvable.
(2) Conversely, if the group $\mathrm{Gal}(f)$ is solvable and its order is not divisible by $char(K)$, then the equation $f = 0$ is solvable by radicals over K.

Proof. (1) Let F be a splitting field of f over K. According to Proposition 17.5, there exists a radical extension $K = K_0 \subset \cdots \subset K_m = L$ satisfying (i)–(iii) for which $L \supset F$. Denote the corresponding Galois groups by $G = \mathrm{Gal}(L/K)$ and $G_j = \mathrm{Gal}(L/K_j)$ $(j = 0, \ldots, m)$. For $j = 0$ (resp. for $j = 1, \ldots, m-1$) the extension $K_j \hookrightarrow K_{j+1} = K_j(\sqrt[r]{a_{j+1}})$ is a Galois extension (hence $G_{j+1} \lhd G_j$) and its Galois group $\mathrm{Gal}(K_{j+1}/K_j) = G_j/G_{j+1}$ is abelian, by Proposition 12.5(5) (resp. by Corollary 15.5(4)). In particular, G is a solvable group, hence its quotient $\mathrm{Gal}(f) = \mathrm{Gal}(F/K) = G/\mathrm{Gal}(L/F)$ is also solvable, thanks to 17.2(7).
(ii) Let $n = |\mathrm{Gal}(f)|$, let F be a splitting field of f over K; then $F(\mu_n)$ is a splitting field of f over $K(\mu_n)$ and $\mathrm{Gal}(f(\mu_n)/K(\mu_n))$ is a solvable group, being a subgroup of $\mathrm{Gal}(f) = \mathrm{Gal}(F/K)$ (by Proposition 16.8). The extension $K(\mu_n)/K$ is radical, which means that we can replace K by $K(\mu_n)$ and assume that $\mu_n \subset K$. There exists a chain of subgroups $\mathrm{Gal}(f) = G = H_0 \supset \cdots \supset H_l = \{e\}$ such that $\forall j = 0, \ldots, l-1$ $H_{j+1} \lhd H_j$ and H_j/H_{j+1} is an abelian group of order dividing n. The fixed fields $K_j = F^{H_j}$ form a tower of extensions $K_0 = K \subset K_1 \subset \cdots \subset K_l = F$ with $\mathrm{Gal}(K_{j+1}/K_j) = H_j/H_{j+1}$. Theorem 15.9(2) implies that each layer K_{j+1}/K_j is of the form $K_{j+1} = K_j(\sqrt[n]{b_1}, \ldots \sqrt[n]{b_k})$ for some $b_i \in K_j$, hence F/K is a radical extension.

(17.7) Example. The equation $f(X) = X^5 - X + 1 = 0$ is not solvable by radicals over $\mathbf{Q}$, since the group $\mathrm{Gal}(f) = S_5$ (see 11.13) is not solvable.

(17.8) General polynomial equation of degree n. Let F be a field, $L = F(x_1, \ldots, x_n)$ the field of rational functions in n variables $x_1, \ldots, x_n$ and $K = L^{S_n} = F(\sigma_1, \ldots, \sigma_n)$ the subfield of symmetric rational functions. We know (see 10.3) that L/K is a Galois extension and $\mathrm{Gal}(L/K) = S_n$. More precisely, $L = K(x_1, \ldots, x_n)$ is a splitting field over K of the separable polynomial

$$f(X) = (X - x_1) \cdots (X - x_n) = X^n - \sigma_1 X^{n-1} + \sigma_2 X^{n-2} - \cdots + (-1)^n \sigma_n \in K[X].$$

As S_n is not solvable for $n \geq 5$, it follows from Theorem 17.6(1) that the "general polynomial equation of degree n"

$$X^n - \sigma_1 X^{n-1} + \sigma_2 X^{n-2} - \cdots + (-1)^n \sigma_n = 0$$

over $K = F(\sigma_1, \ldots, \sigma_n)$ is not solvable by radicals for any $n \geq 5$.

(17.9) Proposition. *(1) A subgroup of S_5 is transitive $\iff$ it is conjugate to $S_5, A_5, GA_1(\mathbf{F}_5), GA_1(\mathbf{F}_5) \cap A_5 = D_{10}$ or C_5.*
(2) A transitive subgroup of S_5 is solvable $\iff$ it is conjugate to a subgroup of $GA_1(\mathbf{F}_5)$ (namely, to $GA_1(\mathbf{F}_5), D_{10}$ or C_5).

Proof. The statment (2) follows from (1) and the fact that S_5, A_5 are not solvable, but $GA_1(\mathbf{F}_5)$ is.
(1) It is sufficient to prove "$\Longrightarrow$". If $G \subset S_5$ is a transitive subgroup, then it contains a 5-cycle. After conjugation we can assume that G contains C_5 (a 5-Sylow subgroup of G) in the form (16.4.1). Let $N = N_G(C_5)$ be the normaliser of C_5 in G.

If $N = G$, then $G \subset N_{S_5}(C_5) = GA_1(\mathbf{F}_5)$, which implies that $G = GA_1(\mathbf{F}_5), D_{10}$ or C_5.

If $N \subsetneq G$, then the number of subgroups of G conjugate to C_5 is an integer $d > 1$. As $d \mid |G|$ and $d \equiv 1 \pmod 5$ (by Theorem 14.9), $d = 6$. If $C \neq C' \subset G$ are distinct subgroups conjugate to C_5, then $C \cap C' = \{e\}$ and all elements of $C \setminus \{e\}$ are 5-cycles. It follows that G contains at least $4d = 24$ 5-cycles, hence all 5-cycles in S_5. As $(ijklm)(ijmlk) = (ikj)$, G contains all 3-cycles; but the 3-cycles generate A_5, so $G \supset A_5$, which implies that $G = S_5$ or A_5.

(17.10) A suitable resolvent $u(x_1, \ldots, x_5)$ for which $H = GA_1(\mathbf{F}_5)$ can be used to decide whether a given irreducible separable polynomial of degree $\deg(f) = 5$ has a solvable Galois group, at least if $char(K) \neq 2$. See [Co, §13.2] for details.

(17.11) Theorem (Galois). *Let p be a prime number, let K be a field such that $char(K) \nmid p!$, let $f \in K[X]$ be an irreducible polynomial of degree $\deg(f) = p$ (it is automatically separable). Let L be a splitting field of f over K. The following conditions are equivalent:*
(1) The equation $f = 0$ is solvable by radicals over K.
(2) The Galois group $\mathrm{Gal}(f)$ is solvable.
(3) The Galois group $\mathrm{Gal}(f) \subset S_p$ is conjugate to a subgroup of $GA_1(\mathbf{F}_p)$.
(4) For any pair of distinct roots $\alpha \neq \beta$ of f we have $L = K(\alpha, \beta)$.
(5) There exist distinct roots $\alpha \neq \beta$ of f such that $L = K(\alpha, \beta)$.

Proof. The equivalence (1) $\iff$ (2) follows from Theorem 17.6. The implications (3) $\implies$ (2) and (4) $\implies$ (5) are automatic.
(2) $\implies$ (3): see [De 1], Ex. I.5.18. This implication was proved by Galois.
(3) $\implies$ (4): we can assume that $\mathrm{Gal}(f) \subset GA_1(\mathbf{F}_p)$. If

$$\sigma = \begin{pmatrix} u & a \\ 0 & 1 \end{pmatrix} \subset \mathrm{Gal}(L/K(\alpha,\beta)) \subset \mathrm{Gal}(f) \subset GA_1(\mathbf{F}_p),$$

then the affine transformation $x \mapsto ux + a$ fixes both $\alpha, \beta \in \mathbf{F}_p$:

$$u\alpha + a = \alpha, \qquad u\beta + a = \beta \implies (u-1)(\alpha-\beta) = 0 \implies u = 1 \implies a = 0 \implies \sigma = \mathrm{id}.$$

As a result, $\mathrm{Gal}(L/K(\alpha,\beta)) = \{\mathrm{id}\}$, hence $L = K(\alpha, \beta)$.
(5) $\implies$ (3): in the tower of fields $K \subset K(\alpha) \subset K(\alpha,\beta) = L$ we have $[K(\alpha) : K] = p$ and $[L : K(\beta)] = m < p$. In particular, $G = \mathrm{Gal}(f) \subset S_p$ is a group of order $|G| = pm$ with $m < p$. According to Corollary 14.2 there exists a subgroup $H \subset G$ of order $|H| = p$. The number of subgroups of G conjugate to H is an integer $d \equiv 1 \pmod p$ dividing $|G| = pm$ (by Theorem 14.9), which implies that $d = 1$ and $H = C_p \lhd G \subset S_p$. It follows that $G \subset N_{S_p}(C_p) = GA_1(\mathbf{F}_p)$.

18. Complements

(18.1) Theorem (Algebraic independence of field embeddings). *Let $L \supset K \subset M$ be fields. If $|K| = \infty$, then any finite set of (distinct) homomorphisms of K-algebras $\sigma_1, \ldots, \sigma_r \in \mathrm{Hom}_{K-\mathrm{Alg}}(L, M)$ is algebraically independent over M. In other words, if $I \subset \mathbf{N}^r$ is a finite subset and if $a_n \in M$ $(n \in I)$ satisfy*

$$\forall y \in L \quad \sum_{n \in I} a_n\, y(n) = 0 \qquad\qquad (y(n) = \sigma_1(y)^{n_1} \cdots \sigma_r(y)^{n_r},\ n = (n_1, \ldots, n_r)), \tag{$\star$}$$

then $a_n = 0$ for all $n \in I$.

Proof. Assume that $(\star)$ holds, with some $a_n \neq 0$. After throwing away zero terms, we can assume that $a_n \neq 0$ for all $n \in I \neq \emptyset$. Furthermore, we can assume that I is chosen in such a way that its cardinality $|I| > 1$ is minimal, among all non-zero relations $(\star)$.

Fix $z \in L$ and $m \in I$ and consider, for each $y \in L$, the following linear combination of $(\star)$ for yz and y:

$$\forall y \in L \quad 0 = \sum_{n \in I} a_n\, (yz)(n) - \big(\sum_{n \in I} a_n\, y(n)\big) z(m) = \sum_{n \in I \setminus \{m\}} a_n\, (z(n) - z(m))\, y(n), \tag{18.1.1}$$

which is a relation of the type $(\star)$ with fewer terms than the cardinality of I. It follows that each coefficient in (18.1.1) must be zero:

$$\forall n \in I\ \forall z \in L \quad z(n) = z(m). \tag{18.1.2}$$

Fix $n \in I$, $n \neq m$ and set $I_\pm = \{i \mid \mathrm{sgn}(m_i - n_i) = \pm 1\} \subset \{1, \dots, r\}$. The relation (18.1.2) can be rewritten as

$$\forall z \in L \quad \prod_{i \in I_+} \sigma_i(z)^{c_i} = \prod_{j \in I_-} \sigma_j(z)^{-c_j}, \tag{18.1.3}$$

where $c_i = m_i - n_i$. For $z = x \in K$ we obtain

$$\forall x \in K \quad x^{c_+} = x^{-c_-} \qquad\qquad (c_\pm = \sum_{I_\pm} c_i),$$

hence $c_+ = -c_- > 0$ (since $|K| = \infty$). The relation (18.1.3) for $z = x + y$ with $x \in K$ and $y \in L$ reads as follows:

$$\forall x \in K\ \forall y \in L \quad \prod_{i \in I_+} (x + \sigma_i(y))^{c_i} = \prod_{j \in I_-} (x + \sigma_j(y))^{-c_j}.$$

After expanding both sides as polynomials in x, we obtain, for each $y \in L$,

$$\forall x \in K \quad x^{c_+} + x^{c_+ - 1} \sum_{i \in I_+} c_i\, \sigma_i(y) + \cdots = x^{-c_-} + x^{-c_- - 1} \sum_{j \in I_-} -c_j\, \sigma_j(y) + \cdots.$$

Again, the assumption $|K| = \infty$ implies that the coefficients must match, hence

$$\forall y \in L \quad \sum_{i \in I} c_i\, \sigma_i(y) = 0.$$

Corollary 15.11 tells us that $c_i = 0$ for all $i \in I$, which is a contradiction. Theorem is proved.

(18.2) Corollary. *For any Galois extension $K \hookrightarrow L$ of infinite fields the elements of* $\mathrm{Gal}(L/K)$ *are algebraically independent over any field $M \supset L$. [See [La, Thm. 12.1-2] for Artin's original proof.]*

(18.3) Dedekind's determinant. Let G be a finite group. Let X_g be variables indexed by elements $g \in G$. Consider the matrix $M = (X_{g^{-1}h})$ (its rows and columns again indexed by elements of G). A study of the determinant of M was at the origin of the theory of representations of finite groups.

For example, for $G = \mathbf{Z}/2\mathbf{Z}$ and $G = \mathbf{Z}/3\mathbf{Z}$ we have, respectively,

$$\begin{vmatrix} X_0 & X_1 \\ X_1 & X_0 \end{vmatrix} = (X_0 + X_1)(X_0 - X_1),$$

and

$$\begin{vmatrix} X_0 & X_1 & X_2 \\ X_2 & X_0 & X_1 \\ X_1 & X_2 & X_0 \end{vmatrix} = (X_0 + X_1 + X_2)(X_0 + \zeta_3 X_1 + \zeta_3^2 X_2)(X_0 + \zeta_3^2 X_1 + \zeta_3 X_2) \qquad (\zeta_3 = e^{2\pi i/3}).$$

In general, $\det(M) = X_e^{|G|} + \cdots \in \mathbf{Z}[X_g]$ is a non-zero polynomial with integral coefficients which factors in $\mathbf{C}[X_g]$ as a product

$$\prod_\rho \det(M[\rho])^{\dim(\rho)}, \tag{18.3.1}$$

where ρ runs through all irreducible complex representations of G and

$$M[\rho] = \sum_{g \in G} X_g \rho(g).$$

This follows from the decomposition (15.12.3) $\mathbf{C}[G] \xrightarrow{\sim} \oplus_\rho \rho^{\oplus \dim(\rho)}$ and the fact that M is the matrix of left multiplication $\mathbf{C}[G] \longrightarrow \mathbf{C}[G]$ by $\sum_{g \in G} X_g\, g \in \mathbf{C}[G]$.

(18.4) Theorem (Normal basis theorem). *Let $K \hookrightarrow L$ be a Galois extension with Galois group $G = \mathrm{Gal}(L/K)$. There exists $\alpha \in L$ with the following equivalent properties.*
(1) The elements $g(\alpha) \in L$ $(g \in G)$ are linearly independent over K.
(2) The elements $g(\alpha) \in L$ $(g \in G)$ form a basis of L/K.
(3) The map $K[G] \longrightarrow L$ sending $\sum a_g g$ to $\sum a_g\, g(\alpha)$ is a surjective homomorphism of (left) $K[G]$-modules (i.e., L is a cyclic $K[G]$-module).
(4) The map from (3) is an isomorphism of $K[G]$-modules (i.e., L is a free $K[G]$-module of rank one).

Proof. The equivalences (1) $\iff$ (3) and (2) $\iff$ (4) hold by definition and (1) $\iff$ (2) holds since $[L : K] < \infty$.

Assume first that $|K| = \infty$. For $\alpha \in L$ denote by $M_\alpha = ((g^{-1}h(\alpha))_{g,h\in G}) \in M_{|G|}(L)$ the image of the matrix M from 18.3 under the ring homomorphism $Z[X_g]_{g\in G} \longrightarrow L$ sending each X_g to $g(\alpha)$.

The polynomial $\det(M) \in \mathbf{Z}[X_g]_{g\in G}$ from 18.3 is non-zero in $K[X_g]_{g\in G}$. Theorem 18.1 implies that there exists $\alpha \in L$ such that $\det(M_\alpha) \neq 0$. We claim that such an α has property (1). Indeed, if

$$\sum_{h\in G} a_h\, h(\alpha) = 0$$

with $a_h \in K$, then we have a system of linear equations for a_h

$$\forall g \in G \quad 0 = \sum_{h\in G} g^{-1}(a_h\, h(\alpha)) = \sum_{h\in G} a_h\, (g^{-1}h)(\alpha).$$

As the determinant of this system $\det(M_\alpha) \neq 0$, all values $a_h = 0$ must vanish. Therefore α satisfies (1).

Assume now that $K = \mathbf{F}_q$ is a finite field; then $L = \mathbf{F}_{q^n}$, where $n = [L : K]$. The group G is cyclic of order n, generated by the Frobenius map $\varphi_q : a \mapsto a^q$. The group algebra $K[G]$ is isomorphic to $K[X]/(X^n - 1)$, with φ_q corresponding to the image $\overline{X}$ of X in $K[X]/(X^n - 1)$.

Consider L as a $K[X]$-module, with X acting as φ_q. We have $X^n - 1 = \varphi_q^n - 1 = 0$ on L. On the other hand, if $P(X) = X^m + a_1X^{m-1} + \cdots + a_m \in K[X]$ is a monic polynomial of degree $m < n$, then

$$|\{a \in L \mid P(\varphi_q)a = 0\}| = |\{a \in L \mid a^{q^m} + \cdots + a_m = 0\}| \leq q^m < q^n,$$

which means that $X^n - 1$ is the minimal polynomial of $\varphi_q \in \mathrm{End}_K(L)$. If we write

$$L \xrightarrow{\sim} K[X]/(P_1) \oplus \cdots K[X]/(P_r)$$

as in II.4.7, with non-constant monic polynomials $P_i \in K[X]$ satisfying $P_1 \mid P_2 \mid \cdots \mid P_r$, then $\deg(P_1) + \cdots + \deg(P_r) = [L : K] = n$. On the other hand, P_r is the minimal polynomial of φ_q; thus $P_r = X^n - 1$. This implies that $r = 1$ and L is a cyclic $K[X]$-module, hence a cyclic $K[X]/(X^n - 1)$-module, proving (3).

(18.5) Theorem (Hilbert's theorem 90). *Let $K \hookrightarrow L$ be a Galois extension with cyclic Galois group $G = \mathrm{Gal}(L/K)$ of order n; let $\sigma \in G$ be a generator. Then: $\beta \in L^*$ satisfies $N_{L/K}(\beta) = 1$ $\iff$ there exists $\alpha \in L^*$ such that $\beta = \alpha/\sigma(\alpha)$.*

Proof. We have $N_{L/K}(\alpha/\sigma(\alpha)) = 1$, since $N_{L/K}(\beta) = \beta\,\sigma(\beta)\cdots\sigma^{n-1}(\beta)$. Conversely, if $N_{L/K}(\beta) = 1$, fix $\gamma \in L$ and consider the following expression:

$$\alpha = \sum_{j=0}^{n-1} (\beta\,\sigma(\beta)\cdots\sigma^j(\beta))\,\sigma^j(\gamma) = \beta\gamma + \beta\,\sigma(\beta)\sigma(\gamma) + \cdots + N_{L/K}(\beta)\sigma^{n-1}(\gamma) \in L.$$

We have

$$\sigma(\alpha) = \gamma + \sigma(\beta)\sigma(\gamma) + \cdots + \sigma(\beta)\cdots\sigma^{n-1}(\beta)\sigma^{n-1}(\gamma),$$

hence $\alpha = \beta\,\sigma(\alpha)$. Linear independence of field embeddings $\sigma^j \in \mathrm{Hom}_{K-\mathrm{Alg}}(L, L)$ proved in Corollary 15.11 implies that there exists $\gamma \in L$ for which $\alpha \neq 0$. Therefore $\beta = \alpha/\sigma(\alpha)$.

(18.6) Rational points on the unit circle. Hilbert's Theorem 90 for the quadratic extension $\mathbf{Q}(i)/\mathbf{Q}$ states that $x+iy \in \mathbf{Q}(i)$ satisfies $(x+iy)(x-iy) = 1 \iff$ there exists $u+iv \in \mathbf{Q}(i)$ such that $x+iy = (u+iv)/(u-iv) = (u+iv)^2/(u^2+v^2)$. We can assume that $u+iv \in \mathbf{Z}[i]$; this leads to an explicit description of all Pythagorean triples, i.e., of positive solutions $a, b, c \in \mathbf{Z}$ of $a^2+b^2=c^2$:

$$\begin{aligned} (2+i)^2 &= 3+4i, \qquad (3+2i)^2 = 5+12i, \qquad (4+i)^2 = 15+8, \qquad \dots \\ 3^2+4^2 &= 5^2, \qquad 5^2+12^2 = 13^2, \qquad 15^2+8^2 = 17^2, \qquad \dots \end{aligned}$$

(18.7) Theorem (additive version of Hilbert's theorem 90). *Under the assumptions of Theorem 18.5, $\beta \in L$ satisfies $\mathrm{Tr}_{L/K}(\beta) = 0 \iff$ there exists $\alpha \in L$ such that $\beta = \alpha - \sigma(\alpha)$.*

Proof. Again, $\mathrm{Tr}(\alpha - \sigma(\alpha)) = 0$, since $\mathrm{Tr}_{L/K}(\beta) = \beta + \sigma(\beta) + \cdots + \sigma^{n-1}(\beta)$. Conversely, if $\mathrm{Tr}_{L/K}(\beta) = 0$, fix $\gamma \in L$ and consider

$$\delta = \sum_{j=0}^{n-1} (\beta + \sigma(\beta) + \cdots + \sigma^j(\beta))\, \sigma^j(\gamma) = \beta\gamma + (\beta + \sigma(\beta))\, \sigma(\gamma) + \cdots + \mathrm{Tr}_{L/K}(\beta)\sigma^{n-1}(\gamma) \in L.$$

We have

$$\sigma(\delta) = \sigma(\beta)\sigma(\gamma) + \cdots + (\sigma(\beta) + \cdots + \sigma^{n-1}(\beta))\, \sigma^{n-1}(\gamma),$$

hence

$$\delta - \sigma(\delta) = \beta(\gamma + \sigma(\gamma) + \cdots + \cdots \sigma^{n-1}(\gamma)) = \beta\, \mathrm{Tr}_{L/K}(\gamma).$$

According to Theorem 7.7(2) there exists $\gamma \in L$ such that $a := \mathrm{Tr}_{L/K}(\gamma) \neq 0$. The element $\alpha = a^{-1}\delta$ then satisfies $\alpha - \sigma(\alpha) = \beta$.

In the special case when $char(K) \mid n$ (which is sufficient for 18.8 below) one can construct α directly as

$$\alpha = \sum_{j=1}^{n-1} j\, \sigma^j(\beta).$$

Here is another proof of the general case of Theorem 18.7 based on Theorem 18.4(4). The group algebra $K[G]$ is isomorphic to $K[X]/(X^n-1)$ and $\mathrm{Tr}_{L/K}$ corresponds to multiplication by $(1+X+\cdots+X^{n-1}) = (X^n-1)/(X-1)$. The statement we wish to prove is equivalent to saying that a polynomial $g \in K[X]$ satisfies

$$(X^n-1) \mid (1+X+\cdots+X^{n-1})g(X)$$

if and only if $g(X)$ is divisible by $(X-1)$, which is immediate.

(18.8) Artin-Schreier extensions. Artin-Schreier theory describes explicitly Galois extensions L/K of fields of characteristic $p>0$ whose Galois groups $G(L/K)$ are isomorphic to $(\mathbf{Z}/p\mathbf{Z})^r$. Such extensions are of the form $K(\alpha_1, \dots, \alpha_r)/K$, where α_i is a root of $X^p - X - a_i = 0$, for suitable $a_i \in K$. The algebraic formalism is analogous to that of Kummer theory, as explained in the following table. Proofs are the same as in §15.

If we modify Definition 17.3 by allowing certain layers K_j/K_{j-1} to be Artin-Schreier extensions $K(\alpha)/K$, $\alpha^p - \alpha \in K$ (if $char(K) = p$), then Theorem 17.6 still holds for these generalised radical extensions, with no restriction on $char(K)$.

Kummer theory	Artin-Schreier theory
$char(K) \nmid n$	$char(K) = p$
$\overline{K}^* \xrightarrow{n} \overline{K}^*,\ z \mapsto z^n$	$\varphi - 1 : \overline{K} \longrightarrow \overline{K},\ z \mapsto z^p - z$
$\mu_n = \mathrm{Ker}(\overline{K}^* \xrightarrow{n} \overline{K}^*) \subset K^*$	$\mathbf{F}_p = \mathrm{Ker}(\varphi - 1 : \overline{K} \longrightarrow \overline{K}) \subset K$ (automatic)
$K^* \longrightarrow K^*/K^{*n},\ a \mapsto \overline{a}$	$K \longrightarrow K/(\varphi - 1)K,\ a \mapsto \overline{a}$
$a_1, \ldots, a_r \in K^*$	$a_1, \ldots, a_r \in K$
$f_i(X) = X^n - a_i = \prod_{\zeta \in \mu_n}(X - \zeta\alpha_i)$	$f_i(X) = X^p - X - a_i = \prod_{b \in \mathbf{F}_p}(X - \alpha_i - b)$
$f_i(\alpha_i) = 0$	$f_i(\alpha_i) = 0$
$L = K(\alpha_1, \ldots, \alpha_r)$	$L = K(\alpha_1, \ldots, \alpha_r)$
$G = \mathrm{Gal}(L/K)$	$G = \mathrm{Gal}(L/K)$
$\Delta' = \mathrm{Ker}(K^*/K^{*n} \longrightarrow L^*/L^{*n})$	$\Delta' = \mathrm{Ker}(K/(\varphi - 1)K \longrightarrow L/(\varphi - 1)L)$
$\Delta = \langle \overline{a}_1, \ldots, \overline{a}_r \rangle \subset \Delta'$	$\Delta = \langle \overline{a}_1, \ldots, \overline{a}_r \rangle \subset \Delta'$
$(\ ,\ ) : G \times \Delta' \longrightarrow \mu_n$	$(\ ,\ ] : G \times \Delta' \longrightarrow \mathbf{F}_p$
$(\sigma, \overline{a}) = \sigma(\alpha)/\alpha,\ a = \alpha^n,\ \alpha \in L^*$	$(\sigma, \overline{a}] = \sigma(\alpha) - \alpha,\ a = (\varphi - 1)(\alpha),\ \alpha \in L$
$(\sigma\tau, \overline{a}) = (\sigma, \overline{a})(\tau, \overline{a})$	$(\sigma\tau, \overline{a}] = (\sigma, \overline{a}] + (\tau, \overline{a}]$
$(\sigma, \overline{a}\overline{b}) = (\sigma, \overline{a})(\sigma, \overline{b})$	$(\sigma, \overline{a}\overline{b}] = (\sigma, \overline{a}] + (\sigma, \overline{b}]$
$G \hookrightarrow \mathrm{Hom}_{\mathbf{Z}}(\Delta, \mu_n)$	$G \hookrightarrow \mathrm{Hom}_{\mathbf{Z}}(\Delta, \mathbf{F}_p)$
$G = G[n]$ finite abelian group	$G = G[p]$ finite abelian group
$\Delta' \hookrightarrow \mathrm{Hom}_{\mathbf{Z}}(G, \mu_n)$	$\Delta' \hookrightarrow \mathrm{Hom}_{\mathbf{Z}}(G, \mathbf{F}_p)$
$\Delta = \Delta' \xrightarrow{\sim} \mathrm{Hom}_{\mathbf{Z}}(G, \mu_n)$	$\Delta = \Delta' \xrightarrow{\sim} \mathrm{Hom}_{\mathbf{Z}}(G, \mathbf{F}_p)$
$\chi : G' = \mathrm{Gal}(L'/K) \longrightarrow \mu_n$	$\chi : G' = \mathrm{Gal}(L'/K) \longrightarrow \mathbf{F}_p$
$\alpha = \sum_{\tau \in G'} \chi(\tau^{-1})\tau(\beta),\ \beta \in L'$	$\alpha = \sum_{\tau \in G'} \chi(\tau^{-1})\tau(\beta),\ \beta \in L'$
$\chi : \sigma \mapsto \sigma(\alpha)/\alpha$	$\chi : \sigma \mapsto \sigma(\alpha) - \alpha$
$G' = G'[n] \Longrightarrow L' = K(\alpha_1, \ldots, \alpha_r),$	$G' = G'[p] \Longrightarrow L' = K(\alpha_1, \ldots, \alpha_r),$
$\alpha_i^n \in K^*$	$(\varphi - 1)(\alpha_i) = \alpha_i^p - \alpha_i \in K$

More generally, Galois extensions of fields of characteristic p with abelian Galois groups of p-power order can be described in terms of the operator $\varphi - 1$ acting on Witt vectors (see [La, Ex. VI.50]).

(18.9) Exercise. *Let K be a field, $char(K) = p > 0$, $a \in K$, $a \notin (\varphi - 1)K$. Artin-Schreier theory tells us that $L = K(\alpha)$ with $a = (\varphi - 1)(\alpha) = \alpha^p - \alpha$ is a Galois extension of K of degree $[L : K] = p$. Show that $\alpha^{p-1} \notin (\varphi - 1)L$. As a result, $M = L(\beta)$ with $\alpha = (\varphi - 1)(\beta)$ is again a Galois extension of L of degree p.*

(18.10) Exercise. *Let K be a field, $a \in K$, $r \geq 1$ an integer and p a prime number.*
(1) Assume that $p \neq 2$ or $char(K) = p$ or $p^r = 2$. Show that the polynomial $X^{p^r} - a$ is irreducible in $K[X]$ $\iff$ $a \notin K^p$.
(2) Assume that $p = 2$, $char(K) \neq 2$ and $r \geq 2$. Show that the polynomial $X^{p^r} - a$ is irreducible in $K[X]$ $\iff$ $a \notin K^2$ and $-4a \notin K^4$. [Note that $X^4 + 4b^4 = (X^2 + 2bX^2 + 2b^2)(X^2 - 2bX^2 + 2b^2)$.]

(18.11) Exercise. *Assume that $K \subsetneq L$ is a non-trivial finite extension, with $L = \overline{L}$ algebraically closed.*
(1) L/K is a separable ($\Longrightarrow$ Galois) extension. [Hint: if not, consider $K(\sqrt[p^r]{a})$.]
(2) If $[L : K] = p$ is a prime number, then $char(K) \neq p$. [Hint: use 18.9.]
(3) If $[L : K] = p$ is a prime number, then $L = K(\sqrt[p]{a})$ for some $a \in K \setminus K^p$.
(4) If $[L : K] = p$ is a prime number, then $p = 2$ and $L = K(i)$, where $i^2 = -1$. [Hint: consider $K(\sqrt[p^2]{a})$.]
(5) $L = K(i)$, where $i^2 = -1$.
(6) $char(K) = 0$. [Hint: if $char(K) = \ell$, consider $\mathbf{F}_\ell(\mu_{2^n}) \subset L$ and its intersection with K, for large n.]

(18.12) Galois (and non-Galois) descent. Let $K \hookrightarrow L$ be a Galois extension, with Galois group $G = \mathrm{Gal}(L/K)$. Galois descent is a general principle which states that an object defined over K is the same thing as an object defined over L which is invariant under the action of G.

For example, if Z is an algebraic variety defined by a system of equations with coefficients in K, then an L-valued point of Z is defined over $K \iff$ it is fixed by G:

$$Z(K) = Z(L^G) = Z(L)^G.$$

The isomorphism $K = L^G$ can be reformulated in more abstract terms as follows. The formula (7.10.3) defines an isomorphism of K-algebras

$$L \otimes_K L \xrightarrow{\sim} \prod_{g\in G} L, \qquad a \otimes b \mapsto (g(a)b)_{g\in G}. \tag{18.12.1}$$

Composing (18.12.1) with the maps

$$d_0 : L \longrightarrow L \otimes_K L, \quad d_0(a) = a \otimes 1, \qquad d_1 : L \longrightarrow L \otimes_K L, \quad d_1(a) = 1 \otimes a$$

we obtain

$$a \mapsto (g(a))_{g\in G}, \qquad a \mapsto (a)_{g\in G},$$

which means that the sequence

$$0 \longrightarrow L^G \longrightarrow L \xrightarrow{d_0-d_1} L \otimes_K L$$

is an exact sequence of K-vector spaces.

It turns out that

$$0 \longrightarrow K \longrightarrow L \xrightarrow{d_0-d_1} L \otimes_K L$$

is an exact sequence of K-vector spaces, for an arbitrary finite extension L/K. This is a beginning of non-Galois descent, which works for suitable ("faithfully flat") extensions of rings.

(18.13) Infinite Galois extensions. An infinite algebraic extension $K \hookrightarrow L$ is a **Galois extension** if the following equivalent conditions hold:

$$K = L^{\mathrm{Aut}(L/K)} \iff L = \bigcup_\alpha K_\alpha, \quad K_\alpha/K \text{ finite Galois extension.}$$

Every element $g \in G$ of the Galois group $G = \mathrm{Gal}(L/K) := \mathrm{Aut}(L/K)$ defines, by restriction, a compatible system of elements of the finite Galois groups $G_\alpha = \mathrm{Gal}(K_\alpha/K)$: there are (surjective) restriction homomorphisms $\mathrm{res}_{\alpha\beta} : G_\beta \longrightarrow G_\alpha$ whenever $K_\alpha \hookrightarrow K_\beta$ and the automorphisms $g|_{K_\alpha} \in G_\alpha$ satisfy $\mathrm{res}_{\alpha\beta}(g|_{K_\beta}) = g|_{K_\alpha}$. As a result, the collection of all restriction homomorphisms

$$\mathrm{Gal}(L/K) \longrightarrow \prod_\alpha \mathrm{Gal}(K_\alpha/K) = \prod_\alpha G_\alpha, \qquad g \mapsto g|_{K_\alpha} \tag{18.13.1}$$

factors through the subgroup

$$\varprojlim_\alpha G_\alpha = \{(g_\alpha)_\alpha \mid g_\alpha \in G_\alpha, \mathrm{res}_{\alpha\beta}(g_\beta) = g_\alpha\} \subset \prod_\alpha G_\alpha, \tag{18.13.2}$$

called the **projective limit** of the finite groups G_α.

It turns out that this recipe defines a group isomorphism

$$G = \mathrm{Gal}(L/K) \xrightarrow{\sim} \varprojlim_\alpha G_\alpha.$$

Moreover, the projective limit (18.13.2) has a natural ("pro-finite") topology, as a closed subgroup of the product group (18.13.1) equipped with the product topology (each finite group G_α having the discrete topology). This makes G into a compact Hausdorff group, with basis of neighbourhoods of the neutral element given by the open normal subgroups $\mathrm{Gal}(L/K_\alpha) = \mathrm{Ker}(G \longrightarrow G_\alpha)$.

The Galois correspondence in this case (due to Krull) is a bijection

$$\{F \text{ field} \,|\, K \hookrightarrow F \hookrightarrow L\} \longleftrightarrow \{\text{closed subgroups } H \subset G\}, \qquad F = L^H, \qquad H = \mathrm{Aut}(L/F).$$

Note that open subgroups of G are precisely closed subgroups of finite index; they correspond to fields F which are finite over K.

(18.14) Examples of infinite Galois extensions. (i) Let $p_1 < p_2 < p_3 < \cdots$ be an increasing sequence of prime numbers. Consider the fields

$$\mathbf{Q} = K \subset \cdots \subset K_n = \mathbf{Q}(\sqrt{p_1}, \ldots \sqrt{p_n}) \subset L = \bigcup_{n=1}^{\infty} K_n = \mathbf{Q}(\sqrt{p_n} \mid n \geq 1).$$

Theorem 15.7 gives, for each $n \geq 1$, an isomorphism $G_n = \mathrm{Gal}(K_n/\mathbf{Q}) \xrightarrow{\sim} \prod_{j=1}^{n} \{\pm 1\}$. When put together, these isomorphisms induce a group isomorphism

$$G = \mathrm{Aut}(L/\mathbf{Q}) \xrightarrow{\sim} \prod_{j=1}^{\infty} \{\pm 1\}, \qquad g \mapsto (g(\sqrt{p_j})/\sqrt{p_j})_{j \geq 1}.$$

Which subgroups of G correspond to subfields $F \subset L$ of degree $[F : \mathbf{Q}] = 2$? Such a field is necessarily contained in K_n, for suitable $n \geq 1$, which means that $H = \mathrm{Aut}(L/F) \supset U_n = \mathrm{Gal}(L/K_n) \xrightarrow{\sim} \prod_{j=n}^{\infty}$ is an open subgroup and $G/H = G_n/\mathrm{Im}(H) \xrightarrow{\sim} \mathbf{Z}/2\mathbf{Z}$.

Note that the group G is a vector space over the field $\mathbf{F}_2$. The kernel of any non-zero linear map $G \longrightarrow \mathbf{F}_2$ is a subgroup $H' \subset G$ satisfying $G/H' \xrightarrow{\sim} \mathbf{Z}/2\mathbf{Z}$, and vice versa. However, H' is a **closed** subgroup of $G \iff$ it is open, which is equivalent to H' containing U_n for some $n \geq 1$. In this case $L^{H'} = F$ is a field of degree $[F : \mathbf{Q}] = 2$. If H' is not closed, then $L^{H'} = \mathbf{Q}$.

(ii) Let $K = \mathbf{F}_q \subset L = \overline{\mathbf{F}}_q = \bigcup_{n \geq 1} \mathbf{F}_{q^n}$. In this case $\mathbf{F}_{q^m} \subset \mathbf{F}_{q^n} \iff m \mid n$ and

$$\mathrm{Gal}(\overline{\mathbf{F}}_q/\mathbf{F}_q) = \varprojlim_n (\mathbf{Z}/n\mathbf{Z}, +) = (\widehat{\mathbf{Z}}, +) \xrightarrow{\sim} \prod_{\ell \text{ prime}} \mathbf{Z}_\ell.$$

(iii) Let $K = \mathbf{Q} \subset L = \bigcup_{n \geq 1} \mathbf{Q}(\mu_n) = \mathbf{Q}(\mu_\infty)$. We have

$$\mathrm{Gal}(\mathbf{Q}(\mu_\infty)/\mathbf{Q}) = \varprojlim_n (\mathbf{Z}/n\mathbf{Z})^* = \widehat{\mathbf{Z}}^* \xrightarrow{\sim} \prod_{\ell \text{ prime}} \mathbf{Z}_\ell^*.$$

For a fixed prime number ℓ, consider $\mathbf{Q}(\mu_{\ell^\infty}) = \bigcup_{r \geq 1} \mathbf{Q}(\mu_{\ell^r})$; then

$$\mathrm{Gal}(\mathbf{Q}(\mu_{\ell^\infty})/\mathbf{Q}) = \varprojlim_r (\mathbf{Z}/\ell^r\mathbf{Z})^* = \mathbf{Z}_\ell^*.$$

IV. Commutative algebra: a geometric perspective

This chapter treats more advanced topics in commutative ring theory, especially those related to algebraic geometry. We continue to follow the conventions of I.1.1 and I.1.4.

1. Finiteness conditions for rings

(1.1) The finiteness conditions for field extensions defined in III.3.2 and III.3.4 have obvious analogues for rings.

(1.2) Definition. *Let $i : A \longrightarrow B$ be a ring homomorphism (in other words, B is an A-algebra via i). We say that B is a* **finite** *A-algebra if it is finitely generated as an A-module, i.e., if there exists an integer $n \in \mathbf{N}$ and $b_1, \ldots, b_n \in B$ such that $B = i(A)b_1 + \cdots + i(A)b_n$ ($\Longleftrightarrow$ the A-module B is isomorphic to a quotient of A^n by some submodule). Similarly, B is an A-algebra* **of finite type** *if there exists an integer $n \in \mathbf{N}$ and $b_1, \ldots, b_n \in B$ such that $B = i(A)[b_1, \ldots, b_n]$ ($\Longleftrightarrow$ the A-algebra B is isomorphic to a quotient of the polynomial ring $A[X_1, \ldots, X_n]$ by some ideal).*

(1.3) Examples. (i) For field extensions $K \hookrightarrow L$ these notions coincide with those introduced in III.3.4 and III.3.2, respectively.
(ii) If K is a field, then any intermediate ring $K \subset B \subset K[x]$ is a K-algebra of finite type (exercise!).
(iii) This is no longer true if we consider subrings of the polynomial ring in several variables. For example, $B = K[y, xy, x^2y, x^3y, \ldots] \subset K[x, y]$ is not a K-algebra of finite type (exercise!).
(iv) On the other hand, $K[x_1, \ldots, x_n]^{S_n} = K[\sigma_1, \ldots, \sigma_n]$ is a K-algebra of finite type. A more general result of this kind (due to Hilbert and E. Noether) will be proved in Corollary 2.11 below.
(v) Corollary II.3.9 can be reformulated by saying that any algebra of finite type over a noetherian ring is noetherian.

(1.4). In the next section we are going to generalise the concept of an algebraic element and an algebraic field extension defined in III.3.8.

2. Integral and finite ring extensions

(2.1) Historically, integrality was first studied in number theory. It was observed that various finite extensions of $\mathbf{Q}$ contain natural subrings generalising $\mathbf{Z} \subset \mathbf{Q}$. Typical examples included

$$\mathbf{Z}[\sqrt{d}] \subset \mathbf{Q}(\sqrt{d}), \qquad \mathbf{Z}[\zeta_n] \subset \mathbf{Q}(\zeta_n),$$

where $d \in \mathbf{Z} \setminus \{0, 1\}$ is square-free and $\zeta_n = e^{2\pi i/n}$. However, $\zeta_3 = (-1 + i\sqrt{3})/2$, which implies that

$$\mathbf{Q}(\sqrt{-3}) = \mathbf{Q}(\zeta_3), \qquad \mathbf{Z}[\sqrt{-3}] = \mathbf{Z} \cdot 1 + \mathbf{Z} \cdot \sqrt{-3} \subsetneq \mathbf{Z} \cdot 1 + \mathbf{Z} \cdot \zeta_3 = \mathbf{Z}[\zeta_3].$$

Which of the two subrings $\mathbf{Z}[\sqrt{-3}]$ and $\mathbf{Z}[\zeta_3]$ of $\mathbf{Q}(\sqrt{-3}) = \mathbf{Q}(\zeta_3)$ is the right analogue of $\mathbf{Z}$?

(2.2) More generally, given a field $K \supset \mathbf{Q}$ of finite degree $[K : \mathbf{Q}] < \infty$, what can we say about subrings $B \subset K$ of the form $B = \mathbf{Z}b_1 + \cdots \mathbf{Z}b_m$, for some $m \geq 1$ and $b_j \in B$?

The key observation is the following. If $b \in B$, then all products bb_i are integral linear combinations of $b_1, \ldots, b_m$, which means that there exists a matrix with integral coefficients $U \in M_m(\mathbf{Z})$ such that

$$b \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} = U \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \subset K^m.$$

As a result, b is an eigenvalue of U, hence

$$f(b) = 0, \qquad f(X) = \det(X \cdot I - U) \in \mathbf{Z}[X].$$

Note that f is a **monic** polynomial with integral coefficients, which justifies the following definition.

(2.3) Definition. *Let $i : A \longrightarrow B$ be a ring homomorphism. An element $b \in B$ is* **integral over** *A if there exists a* **monic** *polynomial $f \in A[X]$ such that $f(b) = 0 \in B$ (this is a slight abuse of language; purists would have written $(i(f))(b) = 0$, where $i(f) \in i(A)[X] \subset B[X]$ is the image of f in $B[X]$). We say that B is an* **integral** *A-***algebra** *if each $b \in B$ is integral over A.*

(2.4) Definition. *A* **ring extension** *is an injective ring homomorphism $A \hookrightarrow B$.*

(2.5) Remarks. (i) A field extension $K \hookrightarrow L$ is algebraic $\iff$ $K \hookrightarrow L$ is an integral ring extension.
(ii) We know (Prop. III.3.15(3)) that a field extension is finite $\iff$ it is algebraic and of finite type. An analogous result for ring extensions is proved in Proposition 2.8(6) below.
(iii) $\beta \in \mathbf{C}$ is integral over $\mathbf{Z}$ $\iff$ β is an algebraic integer. The discussion in 2.2 suggests that one should consider, for any field K algebraic over $\mathbf{Q}$, the set $O_K = \{\beta \in K \mid \beta \text{ integral over } \mathbf{Z}\}$. According to Proposition 2.8(4) below, O_K is a subring of K, called **the ring of integers of** K. See Example 3.7 and Corollary 4.4 for more results on O_K.

(2.6) Proposition. *Let $A \hookrightarrow B$ be a ring extension, let $b \in B$. The following properties are equivalent:*
(1) b is integral over A.
(2) $A[b]$ is a finite A-algebra ($\iff$ $A[b]$ is a finitely generated A-module).
(3) There exists a **faithful** *$A[b]$-module M which is finitely generated as an A-module. [Recall that an R-module M is faithful if for each $r \in R \setminus \{0\}$ there exists $m \in M$ such that $rm \neq 0$.]*

Proof. (1) $\implies$ (2): if $b^n + a_1 b^{n-1} + \cdots + a_n = 0$ for some $a_i \in A$, then an easy induction shows that $\forall k \geq 0 \quad b^{n+k} \in Ab^{n-1} + \cdots + Ab + A$; thus $A[b] = Ab^{n-1} + \cdots + Ab + A$ is a finitely generated A-module. The implication (2) $\implies$ (3) is automatic (take $M = A[b]$), so it remains to prove that (3) $\implies$ (1). If $M = Am_1 + \cdots + Am_r$, then the inclusion $bM \subset M$ implies that

$$b \begin{pmatrix} m_1 \\ \vdots \\ m_r \end{pmatrix} = U \begin{pmatrix} m_1 \\ \vdots \\ m_r \end{pmatrix} \subset M^r,$$

for a suitable matrix $U \in M_r(A)$ with coefficients in A. Its characteristic polynomial $f(X) = \det(X \cdot I - U) \in A[X]$ is monic; the goal is to show that $f(b) = 0$. We have

$$(b \cdot I - U) \begin{pmatrix} m_1 \\ \vdots \\ m_r \end{pmatrix} = 0 \quad \implies \quad f(b) \begin{pmatrix} m_1 \\ \vdots \\ m_r \end{pmatrix} = \mathrm{adj}(b \cdot I - U)(b \cdot I - U) \begin{pmatrix} m_1 \\ \vdots \\ m_r \end{pmatrix} = 0 \in M^r,$$

hence $f(b)m_i = 0 \in M$ for each $i = 1, \ldots, r$; thus $f(b)m = 0$ for all $m \in M$. As M is a faithful $A[b]$-module and $f(b) \in A[b]$, it follows that $f(b) = 0$, hence b is integral over A.

(2.7) Corollary. *A finite ring extension $A \hookrightarrow B$ is integral.*

Proof. Take $M = B$ in Proposition 2.6(3).

(2.8) Proposition-Definition. *Let $A \hookrightarrow B \hookrightarrow C$ be ring extensions.*
(1) If the ring extensions $A \hookrightarrow B$ and $B \hookrightarrow C$ are finite, so is $A \hookrightarrow C$.
(2) If $b_1, \ldots, b_n \in B$ are integral over A, then $A[b_1, \ldots, b_n] \subset B$ is a finite ($\implies$ integral) A-algebra.
(3) The ring extensions $A \hookrightarrow B$ and $B \hookrightarrow C$ are integral $\iff$ $A \hookrightarrow C$ is integral.
(4) The set $B' = \{b \in B \mid b \text{ integral over } A\}$ is a subring of B containing A, called the **integral closure** *(or* **normalisation***) of A in B.*
(5) The normalisation C' of A in C coincides with the normalisation C'' in C of any ring B'' satisfying $A \hookrightarrow B'' \hookrightarrow B'$.
(6) The ring extension $A \hookrightarrow B$ is finite $\iff$ B is an integral A-algebra of finite type.

Proof. (1) If $B = Ab_1 + \cdots + Ab_m$ and $C = Bc_1 + \cdots + Bc_n$, then $C = \sum_{i,j} Ab_i c_j$. (2) Thanks to (1), this follows by induction from the case $n = 1$, which was proved in Proposition 2.6(2). The implication "$\Longleftarrow$"

in (3) is automatic. Conversely, if both ring extensions $A \hookrightarrow B$ and $B \hookrightarrow C$ are integral, then each $c \in C$ satisfies $c^n + b_1 c^{n-1} + \cdots + b_n = 0$ for some $b_i \in B$. In particular, c is integral over $A' = A[b_1, \ldots, b_n]$, which is a finite A-algebra by (2). It follows that $M = A'[c] = A'c^{n-1} + \cdots + A'$ is a finitely generated A-module, hence c is integral over A, by Proposition 2.6(3). In (4), for any $b, b' \in B'$, the subring $A[b, b'] \subset B$ is a finite A-algebra by (2), which means that $b \pm b', bb' \in A[b, b']$ are integral over A, by Proposition 2.6(3) applied to $M = A[b, b']$. The inclusion $C' \supseteq C''$ in (5) is a consequence of (3); the opposite inclusion is automatic. The implication "$\Longrightarrow$" in (6) is immediate; the converse "$\Longleftarrow$" follows from (2).

(2.9) Theorem (E. Artin-Tate). *Let $A \hookrightarrow B \hookrightarrow C$ be ring extension. Assume that A is a noetherian ring, C is an A-algebra of finite type and C is an integral B-algebra. Then B is an A-algebra of finite type.*

Proof. By assumption, there exist integers $m, n \geq 1$ and elements $c_i \in C$, $b_{i,j} \in B$ $(1 \leq i \leq m,\ 1 \leq j \leq n)$ such that $C = A[c_1, \ldots, c_m]$ and $c_i^n + b_{i,1} c_i^{n-1} + \cdots + b_{i,n} = 0$. In particular, each c_i is integral over the subring $B' = A[\{b_{i,j}\}_{i,j}] \subset B$. Note that B' is a noetherian ring, by Corollary II.3.9. Moreover, $C = B'[c_1, \ldots, c_m]$ is a finitely generated B'-module, by Proposition 2.8(2). The ring B is a submodule of a finitely generated B'-module C, which implies that B is also a finitely generated B'-module, thanks to Proposition II.3.6: $B = B'b_1 + \cdots + B'b_r$. However, B' is a ring, which means that $B = B'[b_1, \ldots, b_r] = A[\{b_{i,j}, b_k\}_{i,j,k}]$ is an A-algebra of finite type.

(2.10) Corollary (E. Noether). *If C is an algebra of finite type over a noetherian ring A and G a finite group acting on C by homomorphisms of A-algebras, then C^G is an A-algebra of finite type.*

Proof. We need to check that each $c \in C$ is integral over $B = C^G$, which follows from the fact that c is a root of the monic polynomial $f(X) = \prod_{g \in G}(X - g(c)) \in C^G[X]$.

(2.11) Corollary (Hilbert, E. Noether). *If K is a field and G a finite group acting on the polynomial ring $K[x_1, \ldots, x_n]$ by homomorphisms of K-algebras, then $K[x_1, \ldots, x_n]^G$ is a K-algebra of finite type.*

(2.12) The abstract argument in the proof of Theorem 2.9 does not give any concrete information about the set of generating elements of B (as an A-algebra). In the situation of Corollary 2.11, the ring of invariants $K[x_1, \ldots, x_n]^G$ is generated by invariant polynomials of degree $\leq |G|$, provided $char(K) = 0$ (E. Noether) or, more generally, if $char(K) \nmid |G|$ (Fleischmann, Fogarty).

(2.13) Exercise. *For each $n \geq 1$ (resp. $n \geq 3$) let the cyclic group C_n (resp. the dihedral group D_{2n}) act on $\mathbf{C}[x, y]$ as follows: a fixed generator r of C_n acts by $x \mapsto \zeta_n x$, $y \mapsto \zeta_n^{-1} y$ (and a fixed element s of $D_{2n} \setminus C_n$ interchanges x and y). Determine the structure of the $\mathbf{C}$-algebra $\mathbf{C}[x, y]^{C_n}$ (resp. $\mathbf{C}[x, y]^{D_{2n}}$).*
[There are other interesting finite subgroups $G \subset SU(2) \subset GL_2(\mathbf{C})$, namely, two-fold coverings (via the surjective homomorphism $\pi : SU(2) = Spin(3) \longrightarrow SO(3)$) of symmetry groups $\pi(G) \subset SO(3)$ of regular polyhedra. The corresponding algebras of invariants $\mathbf{C}[x, y]^G$ are discussed in [Kl]. The case of the icosahedron group $\pi(G) \xrightarrow{\sim} A_5$ is particularly interesting, as it is related to the problem of solving a general polynomial equation of degree 5.]

(2.14) Hilbert's 14-th problem asked whether $\mathbf{C}[x_1, \ldots, x_n]^G$ is a $\mathbf{C}$-algebra of finite type in the case when $G \subset GL_n(\mathbf{C})$ is a matrix group (such as $SL_n(\mathbf{C})$ or $SO(n)$). It turns out that the answer is "yes" if G is a reductive group, but "no" in general (a counterexample was found by Nagata).

3. Integrally closed domains

(3.1) Definition. *A domain A is* **integrally closed** *(or* **normal***) if it is equal to its integral closure (= normalisation) in* $\mathrm{Frac}(A)$. *[As we shall see, normal domains are "nicer" than non-normal ones.]*

(3.2) Proposition. *Any UFD is integrally closed.*

Proof. Assume that A is a UFD and $ab^{-1} \in \mathrm{Frac}(A)$ is a root of $f(X) = X^n + a_1 X^{n-1} + \cdots + a_n \in A[X]$. We can also assume that $\gcd(a, b) = 1$; the equality

$$a^n + a_1 a^{n-1} b + \cdots + a_n b^n = 0$$

implies that every irreducible element $x \mid b$ divides a^n, hence $x \mid a$, which contradicts the assumption $\gcd(a,b)=1$. It follows that $b \in A^*$, hence $ab^{-1} \in A$.

(3.3) (i) Assume that A is a domain with fraction field $K = \mathrm{Frac}(A)$ and $K \hookrightarrow L$ is a field extension. If $\alpha \in L$ is algebraic over K, then there exist $a_i \in A$ with $a_0 \neq 0$ such that $a_0\alpha^n + a_1\alpha^{n-1} + \cdots + a_n = 0$. After multiplying this equation by a_0^{n-1}, we obtain $(a_0\alpha)^n + a_1(a_0\alpha)^{n-1} + \cdots + a_n a_0^{n-1} = 0$. Therefore $a_0\alpha \in L$ is integral over K.

Applying this argument to each element of a basis of L/K we see that there exists a basis of L/K consisting of elements integral over A.

(ii) If $A \hookrightarrow B$ is a ring extension and B is a domain, then (i) implies that any set of generators of B as an A-module generates $\mathrm{Frac}(B)$ as a vector space over $\mathrm{Frac}(A)$. In particular, if B is a finite A-algebra, then $[\mathrm{Frac}(B) : \mathrm{Frac}(A)] < \infty$.

(3.4) Proposition. *Let A be an integrally closed domain and $K = \mathrm{Frac}(A) \hookrightarrow L$ be a finite field extension. For $\alpha \in L$ the following properties are equivalent.*
(1) α is integral over A.
(2) The characteristic polynomial $P_{L/K,\alpha}(X) \in K[X]$ lies in $A[X]$.
(3) The minimal polynomial $f(X) \in K[X]$ of α over K lies in $A[X]$.

Proof. The implication (2) $\Longrightarrow$ (1) (resp. (3) $\Longrightarrow$ (2)) follows from the fact that $P_{L/K,\alpha}(\alpha) = 0$ (resp. from $P_{L/K,\alpha} = f^{[L:K(\alpha)]}$). It remains to prove that (1) implies (3). If α is integral over A, write $f(X) = \prod(X - \alpha_i) \in M[X]$, where M is a splitting field of f over K and $\alpha_1 = \alpha$. By assumption, there exists a monic polynomial $g \in A[X]$ such that $g(\alpha) = 0$; then f divides g in $K[X]$, which implies that $g(\alpha_i) = 0$ for all i; thus each α_i is integral over A, and so are all coefficients of f, since they lie in $\mathbf{Z}[\alpha_1, \ldots, \alpha_n]$. In particular, each coefficient of f lies in K and is integral over A, hence is contained in the integrally closed ring A.

(3.5) Note that the implication (1) $\Longrightarrow$ (3) does not hold if A is not integrally closed: it is enough to consider an element $\alpha \in K$ which is integral over A but does not belong to A; then $f(X) = X - \alpha \notin A[X]$ (example: $A = \mathbf{Z}[\sqrt{-3}]$, $\alpha = \zeta_3$).

(3.6) Corollary. *Under the assumptions of Proposition 3.4, $\mathrm{Tr}_{L/K}(\alpha) \in A$ and $N_{L/K}(\alpha) \in A$ for any $\alpha \in L$ integral over A.*

Proof. Up to a sign, the trace (resp. the norm) of α is equal to one of the coefficients of $P_{L/K,\alpha}(X)$.

(3.7) Example. If $A = \mathbf{Z}$, then $K = \mathbf{Q}$ and L is a field of finite degree $[L : \mathbf{Q}] < \infty$ over $\mathbf{Q}$. Proposition 3.4 tells us that $\alpha \in L$ lies in the ring of integers O_L (= the normalisation of $\mathbf{Z}$ in L) $\iff$ $P_{L/\mathbf{Q},\alpha}(X) \in \mathbf{Z}[X]$.

In the simplest non-trivial case when L is a quadratic field (i.e., $[L : \mathbf{Q}] = 2$), then $L = \mathbf{Q}(\sqrt{d})$ for a square-free integer $d \in \mathbf{Z} \setminus \{0, 1\}$. Writing $\alpha \in L$ as $\alpha = a + b\sqrt{d}$ $(a, b \in \mathbf{Q})$, we have $P_{L/\mathbf{Q},\alpha}(X) = X^2 - 2aX + (a^2 - db^2)$; thus $\alpha \in O_L \iff 2a,\ a^2 - db^2 \in \mathbf{Z}$. An easy calculation (exercise!) shows that

$$O_K = \mathbf{Z}[\beta] = \mathbf{Z} \cdot 1 + \mathbf{Z} \cdot \beta, \qquad \beta = \begin{cases} \sqrt{d}, & \text{if } d \equiv 2, 3 \pmod 4 \\ \frac{1+\sqrt{d}}{2}, & \text{if } d \equiv 1 \pmod 4. \end{cases}$$

If $d < 0$, then the arithmetic of O_L becomes more and more complicated as $|d| \longrightarrow \infty$. In particular, O_L is a euclidean ring $\iff$ $d = -1, -2, -3, -7, -11$; it is a UFD $\iff$ it is a PID $\iff$ $d = -1, -2, -3, -7, -11, -19, -43, -67, -163$. The latter property is related to the fact that, for each $d = -7, -11, -19, -43, -67, -163$, the quadratic polynomial $n^2 + n + (1 - d)/4$ with discriminant d takes prime values for all $n = 0, 1, \ldots, (1 - d)/4 - 2$.

(3.8) One can show that, if A is an integrally closed domain, so is $A[X]$ ([De 2, Thm. 8.23]). However, we are not going to use this fact in the sequel.

4. Finiteness of normalisation and Noether's Normalisation Lemma

(4.1) In §4 we are going to investigate the following question. Given a domain A with fraction field $\mathrm{Frac}(A) = K$ and a finite field extension $K \hookrightarrow L$, under what conditions is the integral closure B of A in L a finite A-algebra (i.e., a finitely generated A-module)?

(4.2) One can often use the following reduction steps. Assume that

$$A_1 \hookrightarrow A_2 \hookrightarrow B \hookrightarrow B'$$

are extensions of domains with respective fraction fields

$$K_1 \hookrightarrow K_2 \hookrightarrow L \hookrightarrow L'$$

satisfying the following conditions:

- A_2 is a finite A_1-algebra;
- B is the integral closure of A_1 (hence of A_2, by Proposition 2.8(5)) in L;
- B' is the integral closure of A_1 (hence of A_2) in L'.

In this case the following implications hold.

(i) B is finite over $A_1 \iff B$ is finite over A_2 (by Proposition 2.8(1));

(ii) if A_1 is noetherian and B' is finite over A_1, then B is finite over A_1 (by Proposition II.3.6).

(4.3) Theorem. *If A is* **integrally closed** *and the extension $K \hookrightarrow L$ is* **separable**, *then:*
(1) B is contained in a finitely generated A-module.
(2) If A is a noetherian ring, then B is a finitely generated A-module.
(3) If A is a PID, then B is a free A-module of finite rank (equal to $[L:K]$).

Proof. (1) Let $n = [L:K]$. As explained in 3.3(i), there exist elements $\beta_1, \ldots, \beta_n \in B$ forming a basis of L/K. Let $b \in B$; write $b = \sum_{j=1}^n \lambda_j \beta_j$ ($\lambda_j \in K$). For each β_i we have $b\beta_i \in B$, hence $\mathrm{Tr}_{L/K}(b\beta_i) = \sum_{j=1}^n \mathrm{Tr}_{L/K}(\beta_i\beta_j)\lambda_j \in A$, by Corollary 3.6. This condition can be written in terms of the matrix $M = (\mathrm{Tr}_{L/K}(\beta_i\beta_j)_{1\le i,j\le n}) \in M_n(A)$ as follows:

$$M \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in A^n \subset K^n.$$

After multiplying this relation by the adjoint matrix $\mathrm{adj}(M) \in M_n(A)$, we obtain that $\det(M)\lambda_i \in A$ for each $i = 1, \ldots, n$. The separability assumption implies that $\det(M) \in A \cap K^* = A \setminus \{0\}$, thanks to Theorem 7.7(2). Therefore $\lambda_i \in \det(M)^{-1}A$, hence $B \subset A\beta_1' + \cdots + A\beta_n'$, where $\beta_i' = \det(M)^{-1}\beta_i$.
(2) This follows from (1) and Proposition II.3.6.
(3) By (2), B is a finitely generated A-module. It is torsion-free (since it is contained in a $\mathrm{Frac}(A)$-vector space), hence free of finite rank r, by Theorem II.4.6. The argument from 3.3(i) implies that a basis of B as a free A-module is a basis of the field extension L/K; thus $r = n$.

(4.4) Corollary. *If $[L:\mathbf{Q}] = n < \infty$, then there exists a basis $\alpha_1, \ldots, \alpha_n$ of $L/\mathbf{Q}$ such that $O_L = \mathbf{Z}\alpha_1 + \cdots + \mathbf{Z}\alpha_n$.*

Proof. This is a special case of Theorem 4.3(3) for $A = \mathbf{Z}$.

(4.5) Example (F.K. Schmidt). Let k be a field of characteristic $p > 0$; fix a power series $f \in k[[T]]$ which is transcendental over $k(T)$ (for example, consider the universal case when $k = \mathbf{F}_p(t_0, t_1, \ldots)$ and $f = t_0 + t_1T + t_2T^2 + \cdots$). The homomorphism of k-algebras

$$k[X,Y] \longrightarrow k[[T]], \qquad g(X,Y) \mapsto g(T, f(T))$$

is then injective; it extends to a field embedding

$$\alpha_f : k(X,Y) \hookrightarrow k((T)) = \mathrm{Frac}(k[[T]]).$$

The subrings

$$B = \alpha_f^{-1}(k[[T]]) \subset k(X,Y) = L, \qquad A = B \cap k(X, Y^p) \subset k(X, Y^p) = K$$

have the following properties: both A and B are integrally closed noetherian domains (in fact, discrete valuation rings; see IV.14 below), $\mathrm{Frac}(A) = K$, $\mathrm{Frac}(B) = L$, L/K is a purely inseparable extension of degree $[L : K] = p$, B is the normalisation of A in L and B is **not** a finitely generated A-module (see [BGR, 1.6.2] and an extended discussion in [Re, 9.4]).

This implies that $A' = A[Y] = AY^{p-1} + AY^{p-2} + \cdots + A \subset B$ is a noetherian domain for which $\mathrm{Frac}(A') = \mathrm{Frac}(B)$, but whose integral closure in $\mathrm{Frac}(A')$ is equal to B, hence is not a finitely generated A'-module.

(4.6) The previous example is somewhat discourageing (there are even more sophisticated examples, in which A is noetherian, but B is not). However, it turns out that all rings naturally encountered in algebraic or arithmetic geometry are "excellent" (see [Gr, §7], [ILO, Exp. I]). In particular, Question 4.1 has a positive answer for them. We are going to prove an important special case of this in Theorem 4.10, which is indispensable for constructing normalisations in algebraic geometry. The main tool will be Noether's Normalisation Lemma 4.8 (proved by E. Noether under the assumption $|k| = \infty$ and by Nagata in general). See 5.12 below for a geometric interpretation of this statement.

(4.7) Definition. *Let k be a field, let A be a k-algebra. Elements $a_1, \ldots, a_n \in A$ are* **algebraically independent over** *k if $f(a_1, \ldots, a_n) \neq 0$ for each non-zero polynomial $f \in k[X_1, \ldots, X_n]$ (which is equivalent to saying that the evaluation homomorphism*

$$k[X_1, \ldots, X_n] \longrightarrow k[a_1, \ldots, a_n], \qquad f \mapsto f(a_1, \ldots, a_n)$$

is an isomorphism of k-algebras).

(4.8) Noether's Normalisation Lemma. *Let k be a field, let A be a k-algebra of finite type. There exist elements $a_1, \ldots, a_d \in A$ which are algebraically independent over k and for which $k[a_1, \ldots, a_d] \hookrightarrow A$ is a finite ring extension.*

Proof. We argue by induction on the number of generators of $A = k[b_1, \ldots, b_m]$ as a k-algebra. If the elements $b_1, \ldots, b_m \in A$ are algebraically independent over k, then we take $a_i = b_i$. If not, we apply Lemma 4.9 below and the induction hypothesis to obtain finite ring extensions

$$k[a_1, \ldots, a_d] \hookrightarrow A' \hookrightarrow A = A'[b_m],$$

with $a_1, \ldots, a_d$ algebraically independent over k. The extension $k[a_1, \ldots, a_d] \hookrightarrow A$ is then finite, thanks to Proposition 2.8(1).

(4.9) Lemma. *Let k be a field, let $A = k[b_1, \ldots, b_m]$ a k-algebra of finite type. If the elements $b_1, \ldots, b_m \in A$ are not algebraically independent over k, then there exist $u_1, \ldots, u_{m-1} \in A$ such that b_m is integral over $A' = k[u_1, \ldots, u_{m-1}]$ and $A = A'[b_m]$. Moreover, we can take:*
$u_i = b_i - b_m^{N^i}$ for any sufficiently large integer $N >> 0$ (Nagata);
$u_i = b_i - \lambda_i b_m$ for suitable $\lambda_i \in k$ if $|k| = \infty$ (E. Noether).

Proof. There exists a non-zero polynomial $P \in k[X_1, \ldots, X_m]$ such that $P(b_1, \ldots, b_m) = 0$. Fix an integer $N > 1$ and write P in terms of new variables $Y_i = X_i - X_m^{N^i}$ $(1 \leq i \leq m-1)$ and X_m: $P(X_1, \ldots, X_m) = Q(Y_1, \ldots, Y_{m-1}, X_m)$. We have $Q(u_1, \ldots, u_{m-1}, b_m) = 0$. Each monomial $cX_1^{r_1} \cdots X_m^{r_m}$ is equal to

$$cX_m^{r_m} \prod_{i=1}^{m-1} (Y_i + X_m^{N^i})^{r_i} = cY_1^{r_1} \cdots Y_{m-1}^{r_{m-1}} X_m^{r_m} + \cdots + cX_m^{r_m + r_1 N + \cdots + r_{m-1} N^{m-1}}.$$

The highest power of X_m in this expression is equal to $r_m + r_1 N + \cdots + r_{m-1} N^{m-1}$.

If N is bigger than all exponents r_i appearing in P, then the values of $r_1, \ldots, r_m$ for each monomial occurring in P are determined by $r_m + r_1 N + \cdots + r_{m-1} N^{m-1}$. In particular, the coefficient $g(Y_1, \ldots, Y_{m-1})$ at the maximal power of X_m appearing in $Q(Y_1, \ldots, Y_{m-1}, X_m)$ lies in $k \setminus \{0\} \subset k[Y_1, \ldots, Y_{m-1}] \setminus \{0\}$. The identity $g^{-1}Q(u_1, \ldots, u_{m-1}, b_m) = 0$ then gives an integral equation for b_m over $k[u_1, \ldots, u_{m-1}]$.

If $|k| = \infty$ we use new variables $Y_i = X_i - \lambda_i b_m$ $(1 \leq i \leq m-1)$ and X_m: $P(X_1, \ldots, X_m) = Q(Y_1, \ldots, Y_{m-1}, X_m)$. Denote by $P_d \in k[X_1, \ldots, X_m]$ the sum of all monomials of degree $d = \deg(P)$

occurring in P; then $\deg(Q) = d$ and X_m^d occurs in $Q(Y_1, \ldots, Y_{m-1}, X_m)$ with coefficient equal to $c = P_d(\lambda_1, \ldots, \lambda_{m-1}, 1)$, which is non-zero for suitable $\lambda_i \in k$. After dividing Q by c we obtain an integral equation of degree d for b_m over $k[u_1, \ldots, u_{m-1}]$.

(4.10) Theorem (E. Noether). *Let k be a field, let A be a k-algebra of finite type which is a domain, let $K = \mathrm{Frac}(A) \hookrightarrow L$ be a finite field extension. Then the normalisation B of A in L is a finitely generated A-module (in particular, it is again a k-algebra of finite type which is a domain).*
[Note that we **do not** *assume that A is integrally closed.]*

Proof. **First proof:** there exist elements $\beta_1, \ldots, \beta_n \in B$ which form a basis of L/K, by 3.3(i). An application of 4.2(i) with $A_1 = A$ and $A_2 = A[\beta_1, \ldots, \beta_n]$ implies that we can replace A by $A[\beta_1, \ldots, \beta_n]$, hence assume that $K = L$.

According to Noether's Normalisation Lemma, A is finite over a polynomial ring $A_0 = k[a_1, \ldots, a_d]$. Another application of 4.2, this time to $A_1 = A_0$ and $A_2 = A$ shows that we can replace A by A_0 (and K by $\mathrm{Frac}(A_0) = k(a_1, \ldots, a_d)$), hence assume that $A = k[a_1, \ldots, a_d]$ is a polynomial ring over k (but L is no longer equal to K).

If it was possible to find A_0 for which the field extension $\mathrm{Frac}(A)/\mathrm{Frac}(A_0)$ was **separable** (which is automatic if $char(k) = 0$), then we could conclude by Theorem 4.3(2) applied to the integrally closed domain A_0 and the field extension $L/K = \mathrm{Frac}(A)/\mathrm{Frac}(A_0)$. If $char(k) = p > 0$, such A_0 exists under the assumption that the field k is perfect, but not in general. The argument requires an understanding of (in-)separability for non-algebraic field extensions ([ZS1], II.13 Thm. 31; V.4 Thm. 8). In particular, this works if the field $k = \overline{k}$ is algebraically closed ($|k| = \infty$ in this case, which means that the original version of Noether's Normalisation Lemma, not Nagata's refinement, is sufficient). The statement for a general field k can be reduced to the case $k = \overline{k}$ rather easily ([ZS1], V.4 Thm. 9).

Second proof: we reproduce the argument given in [Ei, Cor. 13.13]. As above, we can assume that $A = k[a_1, \ldots, a_d]$ is a polynomial ring over k. According to 4.2(ii), we are free to replace L by any finite extension; we can assume, therefore, that L/K is a normal extension (by Proposition III.8.6). In the tower of fields $K \hookrightarrow L_1 = L^G \hookrightarrow L$, where $G = \mathrm{Aut}(L/K)$, the extension $K \hookrightarrow L_1$ is purely inseparable (by Proposition 10.12(6)) and $L_1 \hookrightarrow L$ is a Galois (hence separable) extension, by Theorem 10.1.

It is enough to show that the integral closure B_1 of A in L_1 is finite over A, since Theorem 4.3(2) will then apply to B_1 (and the extension $L_1 \hookrightarrow L$) and show that B is finite over B_1, hence over A. We can replace, therefore, L by L_1 and assume that the extension L/K is purely inseparable. If $K = L$, then there is nothing to prove. If not, then $char(k) = p > 0$ and there exists a power $q = p^r$ such that $L = K(y_1, \ldots, y_m)$ and $y_i^q = f_i/g_i$ for some polynomials $f_i, g_i \in A = k[a_1, \ldots, a_d]$. If $c_1, \ldots, c_t \in k$ are the non-zero coefficients of these polynomials, then the field $k' = k(\sqrt[q]{c_1}, \ldots, \sqrt[q]{c_t}) \hookrightarrow \overline{L}$ is a finite extension of k and there is a tower of finite field extensions

$$L = K(\{\sqrt[q]{f_i/g_i}\}_i) \hookrightarrow K(\{\sqrt[q]{f_i}, \sqrt[q]{g_i}\}_i) \hookrightarrow K(\{\sqrt[q]{c_i}, \sqrt[q]{a_j}\}_{i,j}) = k'(\sqrt[q]{a_1}, \ldots, \sqrt[q]{a_d}) = L'.$$

According to 4.2(ii), we can replace L by L'. However, the integral closure B' of A in L' is equal to the polynomial ring $B' = k'[\sqrt[q]{a_1}, \ldots, \sqrt[q]{a_d}]$, which is finite over $A = k[a_1, \ldots, a_d]$.

5. Algebra and geometry

We are now ready to develop a dictionary between algebra and geometry based on the duality between points and functions which was alluded to in the Introduction (and elaborated on in I.2.6 and I.4.4). The reader is encouraged to consult [Re], which gives a very readable account of the geometric intuition behind various topics in commutative algebra.

In §5 we are going to work over an arbitrary base field K.

(5.1) For any point $a = (a_1, \ldots, a_n) \in K^n$ the evaluation map

$$\mathrm{ev}_a : K[X_1, \ldots, X_n] \longrightarrow K, \qquad f \mapsto f(a) = f(a_1, \ldots, a_n) \tag{5.1.1}$$

is a surjective homomorphism of K-algebras with kernel

$$\mathfrak{m}_a = \mathrm{Ker}(\mathrm{ev}_a) = (X_1 - a_1, \ldots, X_n - a_n) \in \mathrm{Max}(K[X_1, \ldots, X_n]).$$

In particular,

$$f(a) = 0 \iff f \in \mathfrak{m}_a$$

and ev_a induces an isomorphism of K-algebras

$$\overline{\mathrm{ev}}_a : K[X_1, \ldots, X_n]/\mathfrak{m}_a \xrightarrow{\sim} K, \qquad f \pmod{\mathfrak{m}_a} \mapsto f(a).$$

This example is a principal motivation for the following general definition.

(5.2) Definition. *Let A be a ring, let $P \in \mathrm{Spec}(A)$ be a prime ideal. The* **residue field of** *P is the field $k(P) = \mathrm{Frac}(A/P)$. We use the functional notation and think of any element $f \in A$ as a function and of P as a point. The* **value of** *f* **at** *P is then defined to be $f(P) = f \pmod{P} \in A/P \subset k(P)$. In particular, $f(P) = 0 \iff f \in P$.*

(5.3) Note that, in this generality, a non-zero element $f \neq 0$ of A can have values $f(P) = 0$ at all $P \in \mathrm{Spec}(A)$. Example: $A = \mathbf{Z}/4\mathbf{Z}$, $\mathrm{Spec}(A) = \{(2)\}$, $f = 2 \in A$. This example is quite representative, as the following general statement shows.

(5.4) Proposition. *Let A be a ring, let $I \subset A$ be an ideal. Then*

$$\bigcap_{P \in \mathrm{Spec}(A)} P = \sqrt{(0)}, \qquad \bigcap_{\substack{P \in \mathrm{Spec}(A) \\ P \supset I}} P = \sqrt{I}.$$

In other words, $f \in A$ vanishes at each $P \in \mathrm{Spec}(A) \iff f$ belongs to the nilradical of A.

Proof. It is enough to consider $I = (0)$ (the general case follows from this special case for A/I). The inclusion $P \supset \sqrt{(0)}$ holds for any $P \in \mathrm{Spec}(A)$, by definition of a prime ideal. Conversely, if $f \in A$ and $f \notin \sqrt{(0)}$, then the ring $A[1/f] = A[Y]/(Yf - 1)$ is non-zero, by Lemma 5.5 below. The inverse image under the canonical morphism $i : A \longrightarrow A[Y]/(Yf - 1)$ of any $\mathfrak{m} \in \mathrm{Max}(A[1/f])$ (which exists, by Theorem I.6.6) is a prime ideal $P = i^{-1}(\mathfrak{m}) \in \mathrm{Spec}(A)$. As $i(f)$ is invertible in $A[1/f]$, it is not contained in $\mathfrak{m}$, hence $f \notin P$.

(5.5) Lemma (inverting f). *Let A be a ring, let $f \in A$. The ring $A[1/f] = A[Y]/(Yf - 1)$ is the zero ring $\iff f \in \sqrt{(0)}$.*

Proof. $A[1/f] = 0 \iff$ there exists $P(Y) = a_0 + a_1 Y + \cdots + a_n Y^n \in A[Y]$ such that $(1 - Yf)P(Y) = 1$. The latter equality is equivalent to

$$\forall i = 0, \ldots, n \quad a_i = f^i, \qquad f^{n+1} = 0;$$

therefore: $A[1/f] = 0 \iff f$ is nilpotent.

(5.6) Returning back to the situation of 5.1, fix an ideal $I \subset K[X_1, \ldots, X_n]$. As in I.4.4, we can attach to I an algebro-geometric object Z defined by the system of polynomial equations

$$Z : \forall f \in I \quad f = 0 \tag{5.6.1}$$

and the K-algebra $O(Z) = K[X_1, \ldots, X_n]/I$ of regular functions on Z. Conversely, any K-algebra of finite type is obtained in this way. According to Hilbert's basis theorem (Theorem II.3.8), the ideal I is finitely generated $I = (f_1, \ldots, f_r)$, which means that

$$Z : f_1 = \cdots = f_r = 0$$

is given by a finite system of polynomial equations. Similarly, for each K-algebra B, the set of B-valued points of Z is equal to

$$Z(B) = \{a \in B^n \mid \forall f \in I \;\; f(a) = 0\} = \{a \in B^n \mid f_1(a) = \cdots = f_r(a) = 0\}.$$

The discussion in 5.1 implies that a point $a \in K^n$ satisfies

$$a \in Z(K) \iff \forall f \in I \;\; f(a) = 0 \iff \forall f \in I \;\; f \in \mathfrak{m}_a \iff I \subset \mathfrak{m}_a.$$

Furthermore, for each $a \in Z(K)$ the evaluation map (5.1.1) induces a surjective homomorphism of K-algebras

$$\overline{\mathrm{ev}}_a : O(Z) \longrightarrow K$$

with kernel

$$\overline{\mathfrak{m}}_a = \mathfrak{m}_a/I = \mathrm{Ker}(\overline{\mathrm{ev}}_a) = (\overline{X}_1 - a_1, \ldots, \overline{X}_n - a_n) \in \mathrm{Max}(O(Z)).$$

(5.7) To sum up, we have

$$\mathrm{Max}(O(Z)) = \{\mathfrak{m}/I \mid \mathfrak{m} \in \mathrm{Max}(K[X_1, \ldots, X_n]),\, \mathfrak{m} \supset I\} \tag{5.7.1}$$

and the discussion in 5.6 defines an injective map

$$Z(K) \longrightarrow \mathrm{Max}(O(Z)), \qquad a \mapsto \mathrm{Ker}(\overline{\mathrm{ev}}_a) = \mathfrak{m}_a/I. \tag{5.7.2}$$

If the field $K = \overline{K}$ is **algebraically closed**, a version of Hilbert's theorem on zeroes (see Theorem 6.5 below) states that

$$\mathrm{Max}(K[X_1, \ldots, X_n]) = \{\mathfrak{m}_a \mid a \in K^n\},$$

which implies that the map (5.7.2) is **bijective**. In other words, classical points of Z correspond to maximal ideals of $O(Z)$.

(5.8) Morphisms. Given two K-algebras of finite type $O(Z_1) = K[X_1, \ldots, X_m]/I_1$ and $O(Z_2) = K[Y_1, \ldots, Y_n]/I_2$ corresponding to

$$Z_1 : \forall f \in I_1 \quad f = 0, \qquad Z_2 : \forall g \in I_2 \quad g = 0,$$

any homomorphism of K-algebras

$$\alpha : O(Z_1) \longrightarrow O(Z_2)$$

induces maps

$$\alpha_B^* : Z_2(B) = \mathrm{Hom}_{A-\mathrm{Alg}}(O(Z_2), B) \longrightarrow \mathrm{Hom}_{A-\mathrm{Alg}}(O(Z_1), B) = Z_1(B), \qquad b \mapsto b \circ \alpha$$

for all K-algebras B, by Proposition I.4.4(iii). In concrete terms, write

$$\alpha(X_i \pmod{I_1}) = h_i(Y_1, \ldots, Y_n) \pmod{I_2} \qquad (1 \le i \le m).$$

The polynomials $h_i \in K[Y_1, \ldots, Y_n]$ satisfy

$$\forall f \in I_1 \quad f(h_1(Y), \ldots, h_m(Y)) \in I_2$$

and the map α_B^* is given by the formula

$$\alpha_B^* : (b = (b_1, \ldots, b_n) \in Z_2(B)) \mapsto (h_1(b), \ldots, h_m(b)) \in Z_1(B).$$

We consider the system of maps α_B^*, which are compatible with respect to homomorphisms of K-algebras $B \longrightarrow B'$, as being induced by a "geometric morphism"

$$\alpha^* : Z_2 \longrightarrow Z_1.$$

Note that α is determined by the collection of maps α_B^*, since $\alpha \in Z_1(O(Z_2))$ is the image of the **tautological point** $\mathrm{id} \in \mathrm{Hom}_{A-\mathrm{Alg}}(O(Z_2), O(Z_2)) = Z_2((O(Z_2))$ by $\alpha^*_{O(Z_2)}$ (this is a special case of Yoneda's lemma, an elementary but useful statement from category theory).

Let us give a few examples of such morphisms.

(5.9) Vertical projection onto a horizontal line. The polynomial ring $K[T]$ (resp. $K[X,Y]$) is the ring of regular functions on an affine line $\mathbf{A}^1_K$ (resp. on an affine plane $\mathbf{A}^2_K$) over K. The inclusion

$$\alpha : K[T] \hookrightarrow K[X,Y], \qquad g(T) \mapsto g(X) \tag{5.9.1}$$

corresponds to

$$\alpha_B^* : B^2 \longrightarrow B, \qquad (b_1, b_2) \mapsto b_1,$$

hence to the projection on the first factor

$$\alpha^* = \mathrm{pr}_1 : \mathbf{A}^2_K \longrightarrow \mathbf{A}^1_K.$$

(5.10) Inclusion $Z \hookrightarrow \mathbf{A}^n_K$. The projection

$$\mathrm{pr} : K[X_1, \ldots, X_n] = O(\mathbf{A}^n_K) \longrightarrow K[X_1, \ldots, X_n]/I = O(Z) \tag{5.10.1}$$

corresponds to the tautological inclusion of Z to the n-dimensional affine space $\mathbf{A}^n_K$, since

$$\mathrm{pr}_B^* : Z(B) \hookrightarrow B^n, \qquad (b_1, \ldots, b_n) \mapsto (b_1, \ldots, b_n).$$

(5.11) Combination of inclusion and projection. Any non-constant polynomial $f \in K[X,Y]$ defines a plane curve over K

$$C : f = 0, \qquad C \hookrightarrow \mathbf{A}^2_K.$$

Combining (5.9.1) with (5.10.1) we obtain morphisms

$$\beta : K[T] \hookrightarrow K[X,Y] \xrightarrow{\mathrm{pr}} K[X,Y]/(f) = O(C), \qquad g(T) \mapsto \overline{g(X)} = g(\overline{X}) = g(X) \pmod f \tag{5.11.1}$$

and

$$\beta^* : C \hookrightarrow \mathbf{A}^2_K \xrightarrow{\mathrm{pr}_1} \mathbf{A}^1_K, \qquad (b_1, b_2) \mapsto b_1,$$

which is given by projecting the curve C vertically onto a horizontal line.

(5.12) Relation to finiteness. In the special case of 5.11 when $f(X,Y) = XY - 1$, then

$$C : XY - 1 = 0, \qquad C \hookrightarrow \mathbf{A}^2_K$$

is a hyperbola and the morphism (5.11.1) can be rewritten using the isomorphism

$$O(C) = K[X,Y]/(XY-1) \xrightarrow{\sim} K[X, 1/X] \subset K(X), \qquad \overline{X} \mapsto X, \qquad \overline{Y} \mapsto 1/X$$

as

$$\beta : K[T] \hookrightarrow K[X, 1/X], \qquad g(T) \mapsto g(X). \tag{5.12.1}$$

Note that the ring extension (5.12.1) **is not finite**. Geometrically, this corresponds to the fact that the induced map on points

$$\beta^* : C \longrightarrow \mathbf{A}^1_K, \qquad (b_1, b_2) \mapsto b_1$$

has finite fibres (in fact, $\beta_B^* : C(B) \xrightarrow{\sim} B^* \hookrightarrow B = \mathbf{A}^1_K(B)$ is injective, for any K-algebra B), but **is not proper** (say, for $K = B = \mathbf{C}$): the inverse image of a compact subset of $A^1(\mathbf{C})$ need not be compact in

$C(\mathbf{C})$. Indeed, if $(\beta^*_{\mathbf{C}})(b_1, b_2) = b_1$ remains bounded ("finite") but tends to 0, then $b_2 = 1/b_1$ is not bounded ("goes to infinity"). This is the origin of the terminology "finite".

In order to remedy this situation we can fix $c \in K^*$ and consider a slightly skewered projection

$$\gamma^* : C \longrightarrow \mathbf{A}^1_K$$

corresponding to

$$\gamma : K[T] \longrightarrow K[X,Y]/(XY-1) \xrightarrow{\sim} K[X, 1/X], \qquad g(T) \mapsto \overline{g(X+cY)} = g(X + c/X)$$

On the level of points,

$$\gamma^*_B : (b_1, b_2) \mapsto b_1 + c/b_2.$$

In this case γ **is a finite ring extension**, since

$$K[X, 1/X] = \mathrm{Im}(\gamma) \cdot 1 + \mathrm{Im}(\gamma) \cdot X$$

and the map

$$\gamma^*_{\mathbf{C}} : C(\mathbf{C}) = \mathbf{C}^* \longrightarrow \mathbf{A}^1(\mathbf{C}) = \mathbf{C}, \qquad b_1 \mapsto b_1 + c/b_1$$

is proper.

Attentive reader will have noticed that we have been discussing here a very special case of Lemma 4.9. Noether's Normalisation Lemma can be reformulated by saying that for any Z in (5.6.1) there exists a morphism $Z \longrightarrow \mathbf{A}^d_K$ which is "nice" in the sense that it has finite fibres (is "quasi-finite") and is proper.

The ring extension (5.12.1) has various arithmetic analogues, the simplest one being the inclusion $\mathbf{Z} \hookrightarrow \mathbf{Z}[1/2]$.

As we shall see in Proposition 11.4(2) below, **finite** K-algebra homomorphisms $O(Z_1) \longrightarrow O(Z_2)$ induce **surjective** maps $Z_2(\overline{K}) \longrightarrow Z_1(\overline{K})$. The lack of surjectivity of $\beta^*_{\overline{K}} = \overline{K}^* \hookrightarrow \overline{K} = \mathbf{A}^1_K(\overline{K})$ is related, therefore, to the fact that (5.12.1) is not a finite ring extension.

(5.13) Normalisation in geometry (example). The plane curve

$$C : Y^2 - X^3 = 0, \qquad C \subset \mathbf{A}^2_K$$

has a "singular point" $O = (0,0) \in C(K)$ at the origin. Intersecting C with the system of all non-vertical lines passing through O

$$L_t : Y - tX = 0 \qquad\qquad (t \in K)$$

we obtain points $(t^2, t^3) \in C(K)$. The assignment

$$t \mapsto (t^2, t^3), \qquad B \longrightarrow C(B)$$

makes sense for any K-algebra B and is compatible with K-algebra homomorphisms $B \longrightarrow B'$, which means that it comes from a geometric morphism

$$\alpha^* : \mathbf{A}^1_K \longrightarrow C. \tag{5.13.1}$$

The corresponding K-algebra homomorphism $\alpha : O(C) \longrightarrow O(\mathbf{A}^1_K) = K[T]$ must send $\overline{X}$ (resp. $\overline{Y}$) to T^2 (resp. to T^3), hence is given by

$$\alpha : O(C) = K[X,Y]/(Y^2 - X^3) \longrightarrow O(\mathbf{A}^1_K) = K[T], \qquad g(X,Y) \pmod{(Y^2 - X^3)} \mapsto g(T^2, T^3).$$

Note that

$$O(C) = K[\overline{X}] + \overline{Y}K[\overline{X}], \qquad \overline{Y}^2 = \overline{X}^3, \qquad \alpha(K[\overline{X}]) = K[T^2], \qquad \alpha(\overline{Y}K[\overline{X}]) = T^3K[T^2],$$

which implies that α is injective and

$$\mathrm{Im}(\alpha) = K + T^2K[T] = K[T^2, T^3] \subsetneq K[T].$$

We use α to identify $O(C)$ with $K[T^2, T^3]$; the map (5.13.1) is then induced by the inclusion $K[T^2, T^3] \hookrightarrow K[T]$.

What is the main difference between the two rings $K[T^2, T^3]$ and $K[T]$? They have a common fraction field $\mathrm{Frac}(K[T^2, T^3]) = \mathrm{Frac}(K[T])$, but the element $T \in K[T]$ does not belong to $K[T^2, T^3]$. This implies that $O(C) = K[T^2, T^3]$ **is not integrally closed** and $O(\mathbf{A}^1_K) = K[T]$ is its integral closure. Geometrically, this corresponds to the fact that C has a singular point $O = \alpha^*(0)$ and the map $\alpha^* : \mathbf{A}^1_K \longrightarrow C$ is a **desingularisation of** C.

This is not an accident. As we shall see in IV.15, any curve C for which $O(C)$ is a domain can be desingularised $\alpha : \widetilde{C} \longrightarrow C$ by replacing $O(C)$ by its integral closure $O(\widetilde{C})$ in $\mathrm{Frac}(O(C))$ (Theorem 4.10 will come handy at this point).

(5.14) Exercise. *In the situation of 5.13, show that the map $\alpha^*_K : K \longrightarrow C(K)$ is bijective, but there is no homomorphism of K-algebras $K[T] = O(\mathbf{A}^1_K) \longrightarrow O(C)$ for which $\beta^*_K : C(K) \xrightarrow{\sim} K$ is the inverse of α^*_K.*

(5.15) Exercise. *What happens if we replace C in 5.13 by $C' : Y^2 - X^2(X+1) = 0$?*

(5.16) Arithmetic analogue. The rings appearing in 5.13 have the following arithmetic analogues. In the table below, $\widetilde{A}$ is the integral closure of A in $\mathrm{Frac}(A)$.

Geometry	Arithmetic
$K[X]$	$\mathbf{Z}$
$Y^2 - X^3 = 0$	$Y^2 + 4 = 0$
$Y/X = T$	$Y/2 = T$
$T^2 - X = 0$	$T^2 + 1 = 0$
$A = K[X,Y]/(Y^2 - X^3) = K[T^2, T^3]$	$A = \mathbf{Z}[Y]/(Y^2 + 4) = \mathbf{Z}[2i]$
$\widetilde{A} = K[X,T]/(T^2 - X) = K[T]$	$\widetilde{A} = \mathbf{Z}[T]/(T^2 + 1) = \mathbf{Z}[i]$
$\mathrm{Frac}(A) = K(T)$	$\mathrm{Frac}(A) = \mathbf{Q}(i)$

6. Hilbert's Theorem on Zeroes ("Nullstellensatz")

There are several versions of this fundamental result. We collect most of them in Theorems 6.5 and 6.8 below. As before, K is an arbitrary field.

(6.1) Example: the affine line. We know that

$$\mathrm{Max}(K[T]) = \{(f) \mid f \in K[T] \setminus K \text{ monic irreducible}\}.$$

In particular, the map

$$\mathbf{C} \xrightarrow{\sim} \mathrm{Max}(\mathbf{C}[T]), \qquad a \mapsto (T - a)$$

is bijective and the map

$$\mathbf{C} \xrightarrow{\sim} \mathrm{Max}(\mathbf{C}[T]) \longrightarrow \mathrm{Max}(\mathbf{R}[T]), \qquad a \mapsto (T - a) \cap \mathbf{R}[T]$$

is surjective, with fibres given by $\mathrm{Gal}(\mathbf{C}/\mathbf{R})$-orbits in $\mathbf{C}$, since

$$(T-a)\cap \mathbf{R}[T] = \begin{cases} (T-a) & \text{if } a\in\mathbf{R} \\ (T-a)(T-\overline{a}) & \text{if } a\notin \mathbf{R}. \end{cases}$$

(6.2) More generally, for each $a=(a_1,\dots,a_n)\in \overline{K}^n$ the image of the evaluation morphism

$$\mathrm{ev}_a : K[X_1,\dots,X_n] \longrightarrow \overline{K}, \qquad f\mapsto f(a)$$

is the finite field extension $K(a_1,\dots,a_n)$ of K and its kernel $\mathrm{Ker}(\mathrm{ev}_a)\subset K[X_1,\dots,X_n]$ is a maximal ideal of $K[X_1,\dots,X_n]$. As we shall see in Theorem 6.5(2) below, all maximal ideals are obtained in this way. Moreover, if $f(a)=0$, then $f(\sigma(a))=0$, for all $\sigma\in \mathrm{Aut}(\overline{K}/K)$. Therefore maximal ideals $\mathfrak{m}\subset K[X_1,\dots,X_n]$ correspond to $\mathrm{Aut}(\overline{K}/K)$-orbits in $\overline{K}^n$.

(6.3) Exercise. *Describe explicitly all maximal ideals $\mathfrak{m}\subset \mathbf{R}[X,Y]$.*

(6.4) Lemma. *Let $A\hookrightarrow B$ be an integral ring extension, with B a domain. Then A is a field $\iff$ B is a field.*

Proof. If A is a field, then $A[b]$ is finite-dimensional A-vector space, for any $b\in B$ (since b is integral over A). If $b\neq 0$, then multiplication by b is an injective (since B is a domain) A-linear endomorphism of $A[b]$, hence it is surjective. In particular, 1 lies in its image, which means that b is invertible.

Conversely, if B is a field, then each non-zero element $a\in A$ has an inverse $b\in B$. As b is integral over A, we have $b^n+a_1b^{n-1}+\cdots+a_n=0$ for some $a_i\in A$. Multiplying this equation by a^{n-1} we obtain $b=-(a_1+a_2a+\cdots+a_na^{n-1})\in A$.

(6.5) Theorem (Nullstellensatz). *Let K be a field, let $O(Z)=K[X_1,\dots,X_n]/I$ (with $I=(f_1,\dots,f_r)$) be a K-algebra of finite type (corresponding to $Z:\forall f\in I\ \ f=0,\ \ Z\hookrightarrow \mathbf{A}^n_K$).*
(1) If $O(Z)$ is a field, then $[O(Z):K]<\infty$.
(2) The map

$$Z(\overline{K})=\mathrm{Hom}_{K-\mathrm{Alg}}(O(Z),\overline{K})\longrightarrow \mathrm{Max}(O(Z)), \qquad \lambda\mapsto \mathrm{Ker}(\lambda)$$

is well-defined and surjective.
(3) The fibre of the map (2) above $\overline{\mathfrak{m}}\in\mathrm{Max}(O(Z))$ consists of all homomorphisms $\lambda=\sigma\circ\mathrm{pr}: O(Z)\longrightarrow O(Z)/\overline{\mathfrak{m}}=k(\overline{\mathfrak{m}})\xrightarrow{\sigma}\overline{K}$, where $\sigma\in\mathrm{Hom}_{K-\mathrm{Alg}}(k(\overline{\mathfrak{m}}),\overline{K})$. In particular, the fibres of the map (2) are precisely the $\mathrm{Aut}(\overline{K}/K)$-orbits in $Z(\overline{K})$.
(4) If $K=\overline{K}$ is algebraically closed, then the map (2) is bijective.
(5) The following properties are equivalent:

$$Z(\overline{K})=\emptyset \iff O(Z)=0 \iff 1\in I \iff \exists\, g_1,\dots,g_r\in K[X_1,\dots,X_n]\quad \sum_{i=1}^{r} g_if_i=1.$$

(6) $f\in K[X_1,\dots,X_n]$ satisfies $f|_{Z(\overline{K})}=0 \iff f\in\sqrt{I}$.

Proof. (1) Noether's Normalisation Lemma 4.8 tells us that $O(Z)$ is finite over a polynomial algebra $K[a_1,\dots,a_d]$, for some $d\geq 0$. If $O(Z)$ is a field, so is $K[a_1,\dots,a_d]$, by Lemma 6.4; thus $d=0$ and $O(Z)$ is finite over K.
(2) The image of any morphism of K-algebras $\lambda: O(Z)\longrightarrow\overline{K}$ (which is isomorphic to $O(Z)/\mathrm{Ker}(\lambda)$) is equal to $K[\lambda(\overline{X}_1),\dots,\lambda(\overline{X}_n)]=K(\lambda(\overline{X}_1),\dots,\lambda(\overline{X}_n))\subset\overline{K}$, hence is a finite extension of K, by Proposition III.3.15(2). Therefore $\mathrm{Ker}(\lambda)$ is a maximal ideal of $O(Z)$. Conversely, for any $\overline{\mathfrak{m}}\in\mathrm{Max}(O(Z))$ the residue field $k(\overline{\mathfrak{m}})=O(Z)/\overline{\mathfrak{m}}$ is a K-algebra of finite type, hence $[k(\overline{\mathfrak{m}}):K]<\infty$, by (1). According to Theorem III.5.6(2) there exists a homomorphism of K-algebras $k(\overline{\mathfrak{m}})\hookrightarrow\overline{K}$; the kernel of the composite morphism $\lambda: O(Z)\longrightarrow k(\overline{\mathfrak{m}})\hookrightarrow\overline{K}$ then coincides with $\overline{\mathfrak{m}}$.
(3), (4) The factorisation $\lambda=\sigma\circ\mathrm{pr}$ of each λ mapping to $\overline{\mathfrak{m}}$ by the map (2) is automatic. The homomorphisms $\sigma\in\mathrm{Hom}_{K-\mathrm{Alg}}(k(\overline{\mathfrak{m}}),\overline{K})$ form one $\mathrm{Aut}(\overline{K}/K)$-orbit, by Proposition III.9.7.
(5) The only non-trivial implication is "$O(Z)\neq 0\Longrightarrow Z(\overline{K})\neq\emptyset$", which follows from (2), since $\mathrm{Max}(O(Z))$ is non-empty if $O(Z)\neq 0$.

(6) If $f^n \in I$ for some $n \geq 1$, then $f(a)^n = 0 \in \overline{K}$ for each $a \in Z(\overline{K})$, hence $f(a) = 0$. Conversely, if $f \notin \sqrt{I}$, we must show that there exists $a \in Z(\overline{K})$ such that $f(a) \neq 0$. The idea is to invert f by introducing a new equation $Yf - 1 = 0$, which removes all zeroes of f and reduces (6) to the statement of (5). In concrete terms, consider the ideal $J \subset K[X_1, \ldots, X_n, Y]$ generated by I and $Yf-1$ (cf. Lemma 5.5) and the K-algebra of finite type $O(Z') = K[X_1, \ldots, X_n, Y]/J$ corresponding to $Z' \hookrightarrow \mathbf{A}_K^{n+1}$ defined by the equations $\forall g \in J \;\; g = 0$.

We have $O(Z') = O(Z)[Y]/(Y\overline{f} - 1)$, where $\overline{f} = f \pmod{I} \in O(Z)$. The assumption $f \notin \sqrt{I}$ implies that $\overline{f}$ is not nilpotent in $O(Z)$, hence $O(Z') \neq 0$, by Lemma 5.5. The statement (5) for Z' shows that there exists a point $(a_1, \ldots, a_n, b) \in Z'(\overline{K})$. By definition of Z',

$$f(a_1, \ldots, a_n)b = 1, \qquad \forall g \in I \quad g(a_1, \ldots, a_n) = 0,$$

which means that $a = (a_1, \ldots, a_n) \in Z(\overline{K})$ and $f(a) \neq 0$, as required.

(6.6) Corollary. *For any ideal $I \subset A = K[X_1, \ldots, X_n]$,*

$$\bigcap_{\substack{\mathfrak{m} \in \mathrm{Max}(A) \\ \mathfrak{m} \supset I}} \mathfrak{m} = \sqrt{I}, \qquad \bigcap_{\overline{\mathfrak{m}} \in \mathrm{Max}(A/I)} \overline{\mathfrak{m}} = \sqrt{(0)}.$$

[The intersection of all maximal ideals of a ring B is called the **Jacobson radical of** *B. The statement above says that the nilradical and the Jacobson radical coincide if B is an algebra of finite type over a field.]*

Proof. For $f \in K[X_1, \ldots, X_n]$, the condition $f|_{Z(\overline{K})} = 0$ in (6) is equivalent to $f \in \mathfrak{m}$ for all $\mathfrak{m} \in \mathrm{Max}(K[X_1, \ldots, X_n])$ satisfying $\mathfrak{m} \supset I$, by (2) combined with (5.7.1).

(6.7) Reduced and non-reduced K-algebras. As observed in I.6.5, the set of maximal ideals does not change if a ring is replaced by the corresponding reduced ring. In the situation of Theorem 6.5,

$$O(Z)^{\mathrm{red}} = K[X_1, \ldots, X_n]/\sqrt{I} = O(Z_{\mathrm{red}}),$$

where

$$Z_{\mathrm{red}} : \forall g \in \sqrt{I} \quad g = 0$$

is the reduced system of polynomial equations attached to I (for example, if $I = (XY, Y^2) \subset K[X, Y]$, then $\sqrt{I} = (Y)$). In this case $Z_{\mathrm{red}}(B) = Z(B)$ for any reduced K-algebra B (for example, a field) and the canonical projection $O(Z) \longrightarrow O(Z_{\mathrm{red}})$ induces a bijection

$$\mathrm{Max}(O(Z_{\mathrm{red}})) \xrightarrow{\sim} \mathrm{Max}(O(Z)).$$

(6.8) Theorem-Definition (geometric version of the Nullstellensatz). *Let $K = \overline{K}$ be an algebraically closed field. An* **algebraic set** *in K^n is a subset of K^n of the form*

$$V_K(I) = \{a \in K^n \mid \forall f \in I \quad f(a) = 0\} = V_K(\sqrt{I}),$$

for some ideal $I \subset K[X_1, \ldots, X_n]$. Conversely, for any algebraic set $V \subset K^n$ the set

$$I(V) = \{f \in K[X_1, \ldots, X_n] \mid \forall a \in V \quad f(a) = 0\} = \sqrt{I(V)}$$

is an ideal of $K[X_1, \ldots, X_n]$. The maps

$$V_K : \{\text{ideals of } K[X_1, \ldots, X_n]\} \longrightarrow \{\text{algebraic sets in } K^n\}$$

and

$$I : \{\text{algebraic sets in } K^n\} \longrightarrow \{\text{ideals of } K[X_1, \ldots, X_n]\}$$

satisfy $I(V_K(I)) = \sqrt{I}$ *and* $V_K(I(V)) = V$. *They induce mutually inverse bijections*

$$\{\text{algebraic sets in } K^n\} \overset{\sim}{\longleftrightarrow} \{\text{ideals } I \subset K[X_1, \ldots, X_n] \text{ satisfying } I = \sqrt{I}\}$$

(such ideals are called **radical ideals***).*

Proof. The equality $I(V_K(I)) = \sqrt{I}$ is a reformulation of Theorem 6.5(6). If $V = V_K(I)$, then $I(V) = I(V_K(I)) = \sqrt{I}$ and $V_K(I(V)) = V_K(\sqrt{I}) = V_K(I) = V$. Finally, if $I = \sqrt{I}$, then $I(V_K(I)) = \sqrt{I} = I$.

(6.9) If the field $K = \overline{K}$ is algebraically closed and **sufficiently large** (i.e., containing arbitrarily large finite sets of elements algebraically independent over $\mathbf{Q}$ or $\mathbf{F}_p$; for example, $K = \mathbf{C}$), then it is easy to prove that $I(V_K(P)) = P$ for any **prime ideal** $P \in \mathrm{Spec}(K[X_1, \ldots, X_n])$ ([Mu 2, Thm. 1.5]). As any radical ideal I of $K[X_1, \ldots, X_n]$ is an intersection of finitely many prime ideals $I = P_1 \cap \cdots \cap P_r$, by Cor. 8.11(2) below, it follows that $I(V_K(I)) = I(V_K(P_1) \cup \cdots \cup V_K(P_r)) = I(V_K(P_1)) \cap \cdots \cap I(V_K(P_r)) = P_1 \cap \cdots \cap P_r = I$.

(6.10) Does $Z(K)$ determine Z_{red} if $K = \overline{K}$? At first glance, Theorem 6.8 tells us that the answer is "yes", since $O(Z_{\mathrm{red}}) = K[X_1, \ldots, X_n]/\sqrt{I}$ and $\sqrt{I} = I(Z(K))$. However, this description is not intrinsic, since it depends not only on Z_{red}, but on its inclusion into the affine space $\mathbf{A}^n_K$.

A more refined question is the following. If $K = \overline{K}$ and if $\alpha : O(Z_1) \longrightarrow O(Z_2)$ is a homomorphism of K-algebras of finite type for which $\alpha^*_K : Z_2(K) \overset{\sim}{\longrightarrow} Z_1(K)$ is bijective, is α an isomorphism?

The answer to this question is "no" – we have already seen a counterexample in 5.13, in which Z_2 is the normalisation of a singular curve Z_1. Another example is given by the relative Frobenius morphism if $char(K) = p > 0$: take $Z_1 = Z_2 = \mathbf{A}^1_K$ and

$$\alpha : K[X] \longrightarrow K[X], \qquad g(X) \mapsto g(X^p).$$

In this case $\mathrm{Im}(\alpha) = K[X^p] \subsetneq K[X]$, but $\alpha^*_K : K \longrightarrow K$, $a \mapsto a^p$ is bijective.

It may come as a surprise that these are, essentially, the only two sources of possible counterexamples to α being an isomorphism. Zariski's Main Theorem [Mu 1, III.9] implies that, if $K = \overline{K} \supset \mathbf{Q}$, $O(Z_1)$ and $O(Z_2)$ are domains, $O(Z_1)$ is integrally closed and $\alpha^*_K : Z_2(K) \overset{\sim}{\longrightarrow} Z_1(K)$ is bijective, then α is an isomorphism.

(6.11) Exercise. *If A is a $\mathbf{Z}$-algebra of finite type, then the residue field $k(m) = A/\mathfrak{m}$ of each maximal ideal $\mathfrak{m} \in \mathrm{Max}(A)$ is finite. [Hint: apply Theorem 6.5(1) to $K = \mathbf{Z}/\mathbf{Z} \cap \mathfrak{m}$ and $O(Z) = k(\mathfrak{m})$, and then show that $\mathbf{Z} \cap \mathfrak{m} \neq 0$.]*

(6.12) Zeta-functions and counting points over finite fields. Let A be a $\mathbf{Z}$-algebra of finite type. The **zeta-function of** A is defined as

$$\zeta(A, s) = \prod_{\mathfrak{m} \in \mathrm{Max}(A)} \left(1 - \frac{1}{N(\mathfrak{m})^s}\right)^{-1},$$

where $N(\mathfrak{m}) = |A/\mathfrak{m}|$ is the cardinality of the finite field $k(\mathfrak{m}) = A/\mathfrak{m}$. For example, for $A = \mathbf{Z}$ we obtain the Riemann zeta-function

$$\zeta(\mathbf{Z}, s) = \prod_{p \text{ prime}} \left(1 - \frac{1}{p^s}\right)^{-1} = \sum_{n=1}^{\infty} \frac{1}{n^s} = \zeta(s).$$

In general, if $char(k(\mathfrak{m})) = p$, then $\mathfrak{m} \supset pA$ and $k(\mathfrak{m}) = k(\overline{\mathfrak{m}})$, where $\overline{\mathfrak{m}} = \mathfrak{m}/pA \in \mathrm{Max}(A/pA)$, which implies that

$$\zeta(A, s) = \prod_{p \text{ prime}} \zeta(A/pA, s).$$

Fix a prime number p and replace A by A/pA (still to be denoted by A), which will be an $\mathbf{F}_p$-algebra of finite type: $A = \mathbf{F}_p[X_1, \ldots, X_n]/I = O(Z)$ (for $K = \mathbf{F}_p$). The zeta-function of A can be expressed in terms of the action of the Frobenius map

$$\varphi : (a_1, \ldots, a_n) \mapsto (a_1^p, \ldots, a_n^p)$$

on the set

$$Z(\overline{\mathbf{F}}_p) = \{a = (a_1, \ldots, a_n) \in \overline{\mathbf{F}}_p^n \mid \forall f \in I \;\; f(a) = 0\},$$

as follows: for each $n \geq 1$,

$$Z(\mathbf{F}_{p^n}) = Z(\overline{\mathbf{F}}_p)^{\varphi^n = 1}$$

is a disjoint union of orbits of length d, for various divisors $d \mid n$, under the action of φ. According to Theorem 6.5(3), the set of orbits of length d is in bijection with $\{\mathfrak{m} \in \mathrm{Max}(A) \mid \deg(\mathfrak{m}) = d\}$, where $N(\mathfrak{m}) = p^{\deg(\mathfrak{m})}$. As a result,

$$\log \zeta(A, s) = \sum_{\mathfrak{m} \in \mathrm{Max}(A)} -\log(1 - p^{-s \deg(\mathfrak{m})}) = \sum_{\mathfrak{m} \in \mathrm{Max}(A)} \sum_{k=1}^{\infty} \frac{(p^{-s})^{k \deg(\mathfrak{m})}}{k} =$$
$$\sum_{n=1}^{\infty} \left(\sum_{\substack{\mathfrak{m} \in \mathrm{Max}(A) \\ \deg(\mathfrak{m}) \mid n}} \deg(\mathfrak{m}) \right) \frac{(p^{-s})^n}{n} = \sum_{n=1}^{\infty} |Z(\mathbf{F}_{p^n})| \, \frac{(p^{-s})^n}{n}.$$

In particular,

$$\zeta(\mathbf{F}_p[X_1, \ldots, X_d], s) = \exp\left(\sum_{n=1}^{\infty} \frac{p^{(d-s)n}}{n} \right) = \left(1 - \frac{1}{p^{s-d}}\right)^{-1}$$

and

$$\zeta(\mathbf{Z}[X_1, \ldots, X_d], s) = \prod_{p \text{ prime}} \left(1 - \frac{1}{p^{s-d}}\right)^{-1} = \zeta(s - d).$$

(6.13) Exercise (back to circle one). *Let K be a field of characteristic $char(K) \neq 2$. The aim of the first two parts of this exercise is to describe the K-valued points of the circle $C : X^2 + Y^2 - 1 = 0$.*
(1) If there exists $i \in K$ such that $i^2 + 1 = 0$, then the map

$$C(K) \longrightarrow K^*, \qquad (x, y) \mapsto x + iy$$

is bijective.
(2) If no such $i \in K$ exists, then $L = K[T]/(T^2 + 1) = K(i)$ is a field, $[L : K] = 2$ and there is an exact sequence

$$0 \longrightarrow K^* \longrightarrow L^* \xrightarrow{f} C(K) \longrightarrow 0,$$

where $f(u + iv) = (x, y) \iff (u + iv)/(u - iv) = x + iy$.
(3) If p is a prime number and $n \geq 1$, then

$$|C(\mathbf{F}_{p^n})| = p^n - \left(\frac{-1}{p}\right)^n, \qquad \left(\frac{-1}{p}\right) = \begin{cases} \pm 1 & p \equiv \pm 1 \pmod 4 \\ 0 & p = 2. \end{cases}$$

(4) The zeta-function of the $\mathbf{Z}$-algebra $A = \mathbf{Z}[X, Y]/(X^2 + Y^2 - 1)$ corresponding to the circle is equal to

$$\zeta(A, s) = \zeta(s - 1) L\left(\left(\frac{-1}{\cdot}\right), s\right), \qquad L\left(\left(\frac{-1}{\cdot}\right), s\right) = \prod_{p \text{ prime}} \left(1 - \left(\frac{-1}{p}\right) \frac{1}{p^s}\right)^{-1} = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{(2n-1)^s}.$$

(6.14) Exercise. *Deduce Theorem 6.5(1) from Theorem 2.9, without using Noether's Normalisation Lemma.* ∎
[Hint: if a field L is a K-algebra of finite type, Theorem 2.9 implies that $K(X_1, \dots, X_n)$, where $n = \mathrm{tr.deg}_K L$, is also a K-algebra of finite type.]

7. Zariski topology on $\mathrm{Spec}(A)$

For any algebraically closed field K, the collection of all algebraic sets in $K^n \xrightarrow{\sim} \mathrm{Max}(K[X_1, \dots, X_n])$ satisfies the axioms of the system of all closed sets of a suitable topology – the **Zariski topology** – on K^n (it is much coarser than the usual topology of $\mathbf{C}^n$ if $K = \mathbf{C}$).

The definition of this topology makes sense for $\mathrm{Max}(A)$, for any ring A. In fact, it is much better to consider this topology on $\mathrm{Spec}(A)$, not just on its subset $\mathrm{Max}(A)$.

We should think of $\mathrm{Spec}(A)$ as being the geometric object on which A is the ring of regular functions. In this section we consider $\mathrm{Spec}(A)$ only as a topological space, but it has an additional structure (see 10.16(ii) below).

(7.1) Let A be any ring. It is useful to think in terms of the functional notation introduced in 5.2: $f \in A$ is a "function", $P \in \mathrm{Spec}(A)$ is a "point" and the image $f \pmod P \in A/P \subset k(P) = \mathrm{Frac}(A/P)$ of f is the "value of f at P". In particular, $f(P) = 0 \iff f \in P$.

Using this dictionary, we can translate the definitions from 6.8 as follows: there are natural maps

$$V : \{\text{subsets of } A\} \longrightarrow \{\text{subsets of } \mathrm{Spec}(A)\}, \qquad V(S) = \{P \in \mathrm{Spec}(A) \mid S \subset P\} = V(I) = V(\sqrt{I}),$$

where $I = (S)$ is the ideal generated by S, and

$$I : \{\text{subsets of } \mathrm{Spec}(A)\} \longrightarrow \{\text{subsets of } A\}, \qquad I(E) = \{f \in A \mid \forall P \in E \ \ f \in P\} = \bigcap_{P \in E} P.$$

Note that $I(E)$ is an ideal of A satisfying $I(E) = \sqrt{I(E)}$.

(7.2) Proposition-Definition. *The map $I \mapsto V(I)$ (where I is an ideal of A) has the following properties.*
(1) $I \subset J \Longrightarrow V(I) \supset V(J)$.
(2) $\sqrt{I} \subset \sqrt{J} \iff V(I) \supset V(J)$. In particular, $V(I) = V(J) \iff \sqrt{I} = \sqrt{J}$.
(3) $V((0)) = \mathrm{Spec}(A)$.
(4) $V(I) = \emptyset \iff I = (1)$.
(5) For any collection of ideals $I_\alpha \subset A$, the intersection $\bigcap_\alpha V(I_\alpha) = V(\sum_\alpha I_\alpha)$.
(6) $V(I_1) \cup V(I_2) = V(I_1 \cap I_2) = V(I_1 I_2)$.
(7) The sets $V(I)$ form the system of closed sets of a topology on $\mathrm{Spec}(A)$, called the **Zariski topology**.
(8) The sets $D(f) = \mathrm{Spec}(A) \setminus V((f)) = \{P \in \mathrm{Spec}(A) \mid f \notin P\}$ $(f \in A)$ form a basis of open sets of this topology.
(9) The closure of a point $P \in \mathrm{Spec}(A)$ is the set $\overline{\{P\}} = \bigcap_{I \subset P} V(I) = V(P) = \{Q \in \mathrm{Spec}(A) \mid Q \supset P\}$. In particular, the point P is closed $\iff P \in \mathrm{Max}(A)$.

Proof. (1) and (3) are immediate. (2) If $\sqrt{I} \subset \sqrt{J}$, then $V(I) = V(\sqrt{I}) \supset V(\sqrt{J}) = V(J)$. Conversely, the formula $\sqrt{I} = \bigcap_{P \in V(I)} P$ proved in Proposition 5.4 shows that $V(I) \supset V(J)$ implies $\sqrt{I} \subset \sqrt{J}$. (4) is a special case of (2). The statement (5) is also clear, since $P \in \bigcap_\alpha V(I_\alpha) \iff \forall \alpha \ \ I_\alpha \subset P \iff \sum_\alpha I_\alpha \subset P$. (6) The inclusions $I_1 I_2 \subset I_1 \cap I_2 \subset I_i$ imply that $V(I_1 I_2) \supset V(I_1 \cap I_2) \supset V(I_1) \cup V(I_2)$. If $P \in V(I_1 I_2) \setminus V(I_1)$, then there exists $f_1 \in I_1$ such that $f_1 \notin P$. For each $f_2 \in I_2$ we have $f_1 f_2 \in I_1 I_2 \subset P$, hence $f_2 \in P$; thus $I_2 \subset P$ and $P \in V(I_2)$, proving that $V(I_1 I_2) = V(I_1) \cup V(I_2)$. The statement (7) follows from (1)–(6). For any ideal I we have $V(I) = \bigcap_{f \in I} V((f))$, by (5); thus $\mathrm{Spec}(A) \setminus V(I) = \bigcup_{f \in I} D(f)$. Finally, (9) follows from the definition of the topology.

(7.3) Examples. (0) $\mathrm{Spec}(A) = \emptyset \iff A = 0$.
(1) If K is a field, then $\mathrm{Spec}(K)$ consists of one point (0).

(2) If A is a domain, then $\eta = (0) \in \mathrm{Spec}(A)$ and $\overline{\{\eta\}} = \mathrm{Spec}(A)$ (η is dense in $\mathrm{Spec}(A)$; we say that η is a (in fact, "the") **generic point** of $\mathrm{Spec}(A)$).
(3) $\mathrm{Spec}(\mathbf{C}[T]) = \{\eta\} \cup \mathrm{Max}(\mathbf{C}[T])$, where $\eta = (0)$ and $\mathrm{Max}(\mathbf{C}[T]) \xrightarrow{\sim} \mathbf{C}$ ($(T-a) \mapsto a$). The closed sets in $\mathrm{Spec}(\mathbf{C}[T])$ are $\mathrm{Spec}(\mathbf{C}[T])$ itself and arbitrary finite subsets of $\mathbf{C}$. This is a one-dimensional object (an affine line over $\mathbf{C}$).
(4) $\mathrm{Spec}(\mathbf{Z}) = \{\eta\} \cup \mathrm{Max}(\mathbf{Z})$, where $\eta = (0)$ and $\mathrm{Max}(\mathbf{Z}) = \{(p) \mid p \text{ prime number}\}$. Again, closed sets are $\mathrm{Spec}(\mathbf{Z})$ itself and all finite subsets of $\mathrm{Max}(\mathbf{Z})$. This is also a one-dimensional object, but of arithmetic nature.
(5) $\mathrm{Spec}(\mathbf{C}[X,Y]) = \{(0)\} \cup \{(f) \mid f \in \mathbf{C}[X,Y] \setminus \mathbf{C} \text{ irreducible}\} \cup \mathrm{Max}(\mathbf{C}[X,Y])$, where $\mathrm{Max}(\mathbf{C}[X,Y]) = \{(X-a, Y-b) \mid (a,b) \in \mathbf{C}^2\}$, (0) is the generic point of $\mathrm{Spec}(\mathbf{C}[X,Y])$ and (f) is the generic point of the curve $f(X,Y) = 0$: the closure of (f) consists of (f) together with $\{(X-a, Y-b) \mid (a,b) \in \mathbf{C}^2,\ f(a,b) = 0\}$ (the set of all closed points on this curve). This is a surface – an affine plane – over $\mathbf{C}$.
(6) $\mathrm{Spec}(\mathbf{Z}[Y]) = \{(0)\} \cup \{(p) \mid \text{prime number}\} \cup \{(f) \mid f \in \mathbf{Z}[Y] \setminus \mathbf{Z},\ ct(f) = 1,\ f \text{ irreducible in } \mathbf{Q}[Y]\} \cup \mathrm{Max}(\mathbf{Z}[Y])$, where $(\mathbf{Z}[Y]) = \{(p, f) \mid p \text{ a prime number},\ f \in \mathbf{Z}[Y] \text{ monic},\ f \pmod p \text{ irreducible in } \mathbf{F}_p[Y]\}$. See [Mu 1, II.1] for a picture of this object (an "arithmetic surface" – it has one arithmetic and one geometric dimension).

(7.4) Corollary. *The maps from 7.1 have the following properties.*
(1) For each subset $E \subset A$, $V(I(E))$ is the closure of E.
(2) For each ideal J of A, $I(V(J)) = \sqrt{J}$.
(3) The maps V and I define mutually inverse bijections

$$\{\text{ideals } I \subset A \text{ satisfying } I = \sqrt{I}\} \xleftrightarrow{\sim} \{\text{closed subsets of } \mathrm{Spec}(A)\}.$$

(7.5) Exercise. *(1) For any ring homomorphism $\alpha : A \longrightarrow B$, the map $\alpha^* : \mathrm{Spec}(B) \longrightarrow \mathrm{Spec}(A)$ defined in I.6.4 ($\alpha^*(Q) = \alpha^{-1}(Q)$) is continuous. More precisely, $(\alpha^*)^{-1}(V(I)) = V(\alpha(I)B)$.*
(2) The projection $\mathrm{pr} : A \longrightarrow A^{\mathrm{red}}$ induces a homeomorphism $\mathrm{pr}^* : \mathrm{Spec}(A^{\mathrm{red}}) \xrightarrow{\sim} \mathrm{Spec}(A)$.
(3) For any ideal I of A, the map $\mathrm{pr}^* : \mathrm{Spec}(A/I) \longrightarrow \mathrm{Spec}(A)$ induced by the projection $\mathrm{pr} : A \longrightarrow A/I$ is injective, with image equal to $V(I)$. It induces a homeomorphism $\mathrm{pr}^* : \mathrm{Spec}(A/I) \xrightarrow{\sim} V(I)$.
(4) In general, $\mathrm{Im}(\alpha^*) \subset V(I) \iff \alpha(I) \subset \sqrt{(0)} \iff I/\sqrt{I} \subset \mathrm{Ker}(\alpha_{\mathrm{red}} : A^{\mathrm{red}} \longrightarrow B^{\mathrm{red}})$. In particular, $\mathrm{Im}(\alpha^*)$ is dense in $\mathrm{Spec}(A) \iff \alpha_{\mathrm{red}}$ is injective.

(7.6) Examples. (1) If $\alpha : \mathbf{Z} \hookrightarrow \mathbf{Q}$, then the image of $\alpha^* : \mathrm{Spec}(\mathbf{Q}) \hookrightarrow \mathrm{Spec}(\mathbf{Z})$ is the generic point (0) of $\mathrm{Spec}(\mathbf{Z})$.
(2) Similarly for $\alpha : \mathbf{C}[X] \hookrightarrow \mathbf{C}(X)$.
(3) Let K be a field. The ring $A = K[X,Y]/(XY)$ is generated as a K-algebra by the elements $\overline{X}, \overline{Y}$ satisfying $\overline{XY} = 0$. In particular, each $P \in \mathrm{Spec}(A)$ contains $\overline{X}$ or $\overline{Y}$, hence $\mathrm{Spec}(A) = V((\overline{X})) \cup V((\overline{Y}))$. The projections $A \longrightarrow A/(\overline{X}) = K[Y]$ and $A \longrightarrow A/(\overline{Y}) = K[X]$ induce $\mathrm{Spec}(K[Y]) \xrightarrow{\sim} V((\overline{X}))$ and $\mathrm{Spec}(K[X]) \xrightarrow{\sim} V((\overline{Y}))$. In particular, if $K = \overline{K}$ is algebraically closed, then $V((\overline{X})) = \{(\overline{X})\} \cup \{(\overline{X}, \overline{Y} - b) \mid b \in K\}$ and $V((\overline{Y})) = \{(\overline{Y})\} \cup \{(\overline{Y}, \overline{X} - a) \mid a \in K\}$, with $V((\overline{X})) \cap V((\overline{Y})) = \{(\overline{X}, \overline{Y})\}$. This is in line with geometric intuition: the equation $XY = 0$ represents a union of the horizontal $Y = 0$ and the vertical $X = 0$ axes in the affine plane $\mathbf{A}^2_K$.

(7.7) Proposition. $\mathrm{Spec}(A)$ *is disconnected* $\iff$ $A = A_1 \times A_2$ *is a product of two non-zero rings* $A_i \neq 0$.

Proof. If $A = A_1 \times A_2$ and $P \in \mathrm{Spec}(A)$, then $P = I_1 \times I_2$ for some ideals $I_i \subset A_i$, by Proposition I.3.3. The product ring $A_1/I_1 \times A_2/I_2 = A/P$ is a domain, which is equivalent to either $I_1 = A_1$ and $I_2 \in \mathrm{Spec}(A_2)$, or $I_2 = A_2$ and $I_1 \in \mathrm{Spec}(A_1)$. Therefore the projections $p_i : A \longrightarrow A_i$ induce a decomposition $\mathrm{Spec}(A) = \mathrm{Spec}(A_1) \cup \mathrm{Spec}(A_2)$ (disjoint union).

Conversely, if $\mathrm{Spec}(A) = V(I_1) \cup V(I_2)$ is a disjoint union of two non-empty closed subsets $V(I_i)$, then $I_i \neq (1)$, $I_1 + I_2 = (1)$ and $\sqrt{I_1 I_2} = \sqrt{(0)}$, by Proposition 7.2(2),(3),(5). In particular, there exist $f_i \in I_i$ such that $f_1 + f_2 = 1$. As $f_1 f_2$ is nilpotent, we have $(f_1 f_2)^n = 0$ for some $n \geq 1$. We can write $1 = (f_1 + f_2)^{2n} = f_1^n g_1 + f_2^n g_2 = e_1 + e_2$, where $e_i \in I_i$, $e_1 + e_2 = 1$ and $e_1 e_2 = (f_1 f_2)^n g_1 g_2 = 0$. These two idempotents define a decomposition $A = A_1 \times A_2$ with $A_i = e_i A$. Finally, $A_i \neq 0$, since $I_i \neq (1)$ ($\Longrightarrow e_i \neq 0$).

(7.8) Proposition. $\mathrm{Spec}(A)$ *is (quasi-)compact (i.e., every open covering has a finite subcovering), for any ring* A. *[Note that, if* $\mathrm{Spec}(A) \neq \mathrm{Max}(A)$*, then* $\mathrm{Spec}(A)$ *is not a Hausdorff topological space.]*

Proof. Any open covering $\bigcup U_\alpha = \mathrm{Spec}(A)$ is of the form $U_\alpha = \mathrm{Spec}(A) \setminus V(I_\alpha)$, where $V(\sum I_\alpha) = (1)$, by Proposition 7.2(4),(5). The element $1 \in A$ lies in a finite sum $I_{\alpha_1} + \cdots + I_{\alpha_n}$; thus $U_{\alpha_1} \cup \cdots \cup U_{\alpha_n} = \mathrm{Spec}(A)$, by reversing the argument.

(7.9) Proposition. *If* A *is an algebra of finite type over a field, then* $V \cap \mathrm{Max}(A)$ *is dense in* V*, for any closed subset* $V \subset \mathrm{Spec}(A)$. *[Rings* A *having this property – which is equivalent to the fact that every prime ideal of* A *is an intersection of maximal ideals – are called* **Jacobson rings***.]*

Proof. Write $V = V(J)$, where $J = \sqrt{J}$ is an ideal of A. The closure of $V \cap \mathrm{Max}(A) = \{\mathfrak{m} \in \mathrm{Max}(A) \mid \mathfrak{m} \supset J\}$ in $\mathrm{Spec}(A)$ is equal to $V(I)$, where

$$I = \bigcap_{\substack{\mathfrak{m} \in \mathrm{Max}(A) \\ \mathfrak{m} \supset J}} \mathfrak{m}.$$

However, Corollary 6.6 tells us that

$$I/J = \bigcap_{\overline{\mathfrak{m}} \in \mathrm{Max}(A/J)} \overline{\mathfrak{m}} = \sqrt{J}/J = (0),$$

hence $I = J$.

8. Irreducible components and minimal prime ideals

(8.1) The ring $A = K[X,Y]/(XY)$ from Example 7.6(3) is the ring of functions on the plane curve $Z \hookrightarrow \mathbf{A}^2_K$, $Z : XY = 0$, which is a union of two lines $Z_1 : X = 0$ and $Z_2 : Y = 0$ (the **irreducible components** of Z). This is reflected in the decomposition

$$\mathrm{Spec}(A) = \mathrm{Spec}(A/(\overline{X})) \cup \mathrm{Spec}(A/(\overline{Y})) = \mathrm{Spec}(K[Y]) \cup \mathrm{Spec}(K[X]) = V((\overline{X})) \cup V((\overline{Y})),$$

where $(\overline{X}), (\overline{Y}) \in \mathrm{Spec}(A)$ are the **minimal prime ideals of** A (with respect to inclusion).

The existence of a similar decomposition of $\mathrm{Spec}(A)$ for an arbitrary noetherian ring A will be established by purely topological considerations in Proposition 8.10 below.

(8.2) Proposition-Definition. *A non-empty topological space* X *is* **irreducible** *if the following equivalent conditions hold.*
(1) If $X_1, X_2 \subsetneq X$ *are closed subsets, then* $X_1 \cup X_2 \neq X$.
(2) If $U_1, U_2 \subset X$ *are non-empty open subsets, then* $U_1 \cap U_2$ *is non-empty.*
(3) Every non-empty open subset of X *is dense in* X.
[Note that a Hausdorff space is irreducible $\iff$ *it consists of one point.]*

Proof. (3) (resp. (1)) is equivalent to (2) by definition (resp. by taking U_i to be the complement of X_i, and vice versa).

(8.3) Exercise. *(1) An irreducible space is connected.*
(2) The image of an irreducible space by a continuous map is irreducible.
(3) The closure of an irreducible subspace (in particular, of a point) is irreducible.

(8.4) What does this mean for the topological space $\mathrm{Spec}(A) \xrightarrow{\sim} \mathrm{Spec}(A^{\mathrm{red}})$, where A is an arbitrary ring? Recall from Proposition 7.2(9) that, for any $P, Q \in \mathrm{Spec}(A)$,

$$\overline{\{P\}} \supset \overline{\{Q\}} \iff Q \in \overline{\{P\}} \iff P \subset Q; \quad \text{thus} \quad \overline{\{P\}} = \overline{\{Q\}} \iff P = Q. \tag{8.4.1}$$

(8.5) Proposition. *(1)* $\mathrm{Spec}(A) \xrightarrow{\sim} \mathrm{Spec}(A^{\mathrm{red}})$ *is irreducible* $\iff$ A^{red} *is a domain* $\iff$ $\sqrt{(0)} \in \mathrm{Spec}(A)$.

(2) If this is the case, then $\sqrt{(0)}$ *is the unique dense point (=* **the generic point***) of* $\mathrm{Spec}(A)$.
(3) The maps V *and* I *from 7.1 induce mutually inverse bijections*

$$\{\text{prime ideals } P \subset A\} \overset{\sim}{\longleftrightarrow} \{\text{irreducible closed subsets of } \mathrm{Spec}(A)\}.$$

For each $P \in \mathrm{Spec}(A)$*, the point* P *is the unique dense point (=* **the generic point***) of* $V(P) = \overline{\{P\}} \xrightarrow{\sim} \mathrm{Spec}(A/P)$.

Proof. (1) We can assume that $A = A^{\mathrm{red}}$ is reduced. Thanks to Exercise 8.3(3), the only non-trivial implication to prove is that $\mathrm{Spec}(A)$ is reducible whenever A is not a domain. If $a, b \in A$ are non-zero elements with $ab = 0$, then $\mathrm{Spec}(A) = V((ab)) = V((a)) \cup V((b))$. We claim that $V((a)), V((b)) \neq \mathrm{Spec}(A)$, hence $\mathrm{Spec}(A)$ is reducible. Indeed, if $V((a)) = \mathrm{Spec}(A)$, then $\sqrt{(a)} = \sqrt{(0)} = (0)$ (by Proposition 7.2(2)), which contradicts the assumption $a \neq 0$ (and similarly for $V((b))$).
(2) Combine 7.3(2) with (8.4.1).
(3) According to (1), a closed subset $V(I) = V(\sqrt{I}) \xrightarrow{\sim} \mathrm{Spec}(A/I) \xrightarrow{\sim} \mathrm{Spec}(A/\sqrt{I})$ of $\mathrm{Spec}(A)$ is irreducible $\iff$ $(A/I)^{\mathrm{red}} = A/\sqrt{I}$ is a domain $\iff$ $\sqrt{I} = P \in \mathrm{Spec}(A)$. The remaining statements follow from (2) applied to A/P.

(8.6) Proposition-Definition. *An* **irreducible component** *of a non-empty topological space* X *is a* **maximal irreducible subset** *of* X *with respect to inclusion (it is closed, by Exercise 8.3(3)). Each point* $x \in X$ *is contained in an irreducible component of* X *(hence* X *is the union of its irreducible components).*

Proof. The set $\{A \subset X \mid A \text{ is irreducible}, x \in A\}$ contains $\{x\}$ and is inductive (exercise), hence contains a maximal element, by Zorn's Lemma.

(8.7) Proposition. *Let* $A \neq 0$ *be a ring.*
(1) The irreducible components of $\mathrm{Spec}(A)$ *are the subsets* $\overline{\{P\}} = V(P) \xrightarrow{\sim} \mathrm{Spec}(A/P)$*, where* $P \in \mathrm{Spec}(A)$ *is a* **minimal prime ideal of** A *(with respect to inclusion).*
(2) For every $Q \in \mathrm{Spec}(A)$ *there exists a minimal prime ideal* $P \subset Q$.
(3) The intersection of all minimal prime ideals of A *is equal to the nilradical* $\sqrt{(0)}$.

Proof. (1), (2) This is a translation of Proposition 8.6, using the dictionary of Proposition 8.5(3). The statement (3) follows from (2) and Proposition 5.4.

(8.8) Proposition-Definition. *A topological space* X *is* **noetherian** *if the following equivalent conditions hold.*
(1) Every descending chain of closed subsets $Z_1 \supset Z_2 \supset \cdots$ *of* X *stabilises (there exists* j *such that* $Z_k = Z_j$ *for all* $k \geq j$*).*
(2) Every non-empty set of closed subsets of X *has a minimal element.*
(3) Every ascending chain of open subsets $U_1 \subset U_2 \subset \cdots$ *of* X *stabilises (there exists* j *such that* $U_k = U_j$ *for all* $k \geq j$*).*
(4) Every non-empty set of open subsets of X *has a maximal element.*

Proof. The equivalences (1) $\iff$ (3) and (2) $\iff$ (4) are immediate. The equivalence (3) $\iff$ (4) is proved as in II.3.1.

(8.9) Proposition. $\mathrm{Spec}(A)$ *is a noetherian topological space* $\iff$ *every ascending chain* $I_1 \subset I_2 \subset \cdots$ *of ideals of* A *satisfying* $I_i = \sqrt{I_i}$ *stabilises. In particular,* $\mathrm{Spec}(A)$ *is a noetherian space if* A *is a noetherian ring.*

Proof. This is a translation of the condition 8.8(1) using Corollary 7.4(3).

(8.10) Proposition. *A noetherian topological space* X *has only finitely many irreducible components* $X_1, \ldots, X_r$*. These components satisfy* $X_i \not\subset X_j$ *if* $i \neq j$ *and* $X = X_1 \cup \cdots \cup X_r$*. Conversely, if* $X_1, \ldots, X_r$ *are irreducible closed subsets of* X *with these properties, then they are the irreducible components of* X.

Proof. We can assume that X is non-empty. Consider the set S of all closed subsets of X which cannot be written as a union of finitely many irreducible closed subsets. If S is non-empty, then it contains a minimal

element Y, which is not irreducible; thus $Y = Y_1 \cup Y_2$ for some closed subsets $Y_i \subsetneq Y$. By minimality of Y, each Y_i is a union of finitely many irreducible closed subsets, hence so is Y. This contradiction implies that S is empty; in particular, $X = X_1 \cup \cdots \cup X_r$, where each X_i is an irreducible closed subset. After removing several X_i if necessary, we can assume that this decomposition is irredundant, i.e., that $X_i \not\subset X_j$ if $i \neq j$.

If Y is an irreducible component of X, then $Y = (Y \cap X_1) \cup \cdots \cup (Y \cap X_r)$. Irreducibility of Y implies that $Y \cap X_i = Y$ for some i; thus $Y \subset X_i$, hence $Y = X_i$, by maximality of Y.

Conversely, we must show that each X_i is an irreducible component of X. If $X_i \subset Z$ with Z irreducible, then $Z = (Z \cap X_1) \cup \cdots \cup (Z \cap X_r)$ implies again that $Z \subset X_j$ for some j; thus $X_i \subset Z \subset X_j$, hence $i = j$ and $Z = X_i$. Therefore X_i is a maximal irreducible subset.

(8.11) Corollary. *(1) A noetherian ring $A \neq 0$ has only finitely many minimal prime ideals $P_1, \ldots, P_r \in \mathrm{Spec}(A)$. These prime ideals satisfy $\sqrt{(0)} = P_1 \cap \cdots \cap P_r$ and $P_i \not\supset P_j$ if $i \neq j$, and are characterised by these two properties. The irreducible components of $\mathrm{Spec}(A)$ are the subsets $\overline{\{P_i\}} = V(P_i) \xrightarrow{\sim} \mathrm{Spec}(A/P_i)$.*
(2) Let $I \neq A$ be an ideal of a noetherian ring A. There are only finitely many prime ideals $P_1, \ldots, P_r \in \mathrm{Spec}(A)$ which are minimal among those prime ideals of A which contain I. These prime ideals satisfy $\sqrt{I} = P_1 \cap \cdots \cap P_r$ and $P_i \not\supset P_j$ if $i \neq j$, and are characterised by these two properties. The irreducible components of $V(I) \xrightarrow{\sim} \mathrm{Spec}(A/I) \subset \mathrm{Spec}(A)$ are the subsets $\overline{\{P_i\}} = V(P_i) \xrightarrow{\sim} \mathrm{Spec}(A/P_i)$.

Proof. (1) Combine Proposition 8.7 with Proposition 8.10 for $X = \mathrm{Spec}(A)$. The statement (2) is equivalent to (1) applied to A/I.

(8.12) Examples. (i) If A is a UFD and $I = (f)$ is a non-zero principal ideal, then $f = u \prod_{i=1}^{r} f_i^{n_i}$, where $u \in A^*$, $r \geq 0$, $n_i \geq 1$ and $f_i \in A$ are irreducible elements such that $f_i \nmid f_j$ for $i \neq j$. Each principal ideal $P_i = (f_i)$ is a prime ideal and

$$\sqrt{(f)} = (f_1 \cdots f_r) = (f_1) \cap \cdots \cap (f_r).$$

(ii) If each ideal $P_i = \mathfrak{m}_i \in \mathrm{Max}(A)$ in Corollary 8.11 is maximal, then $\mathfrak{m}_i + \mathfrak{m}_j = (1)$ for $i \neq j$, hence $\mathfrak{m}_1 \cap \cdots \cap \mathfrak{m}_r = \mathfrak{m}_1 \cdots \mathfrak{m}_r$ and

$$(A/I)^{\mathrm{red}} = A/\sqrt{I} = A/(\mathfrak{m}_1 \cap \cdots \cap \mathfrak{m}_r) \xrightarrow{\sim} A/\mathfrak{m}_1 \times \cdots \times A/\mathfrak{m}_r$$

is a finite product of fields, by the Chinese remainder theorem I.3.5.

(iii) If A is a finite algebra over a field K (i.e., if $\dim_K(A) < \infty$), then A is noetherian and each prime ideal $P \in \mathrm{Spec}(A)$ is maximal, by Lemma 6.4 applied to $K \hookrightarrow A/P$. As a result, $\mathrm{Spec}(A) = \mathrm{Max}(A) = \{P_1, \ldots, P_r\}$, where P_i are the minimal prime ideals of A. As in (ii), we obtain $A^{\mathrm{red}} = k(P_1) \times \cdots \times k(P_r)$, where each $k(P_i) = A/P_i$ is a finite field extension of K. In fact, the following exercise shows that $(P_1 \cdots P_r)^n = (0)$ for some $n \geq 1$, which implies that $A = A/P_1^n \times \cdots \times A/P_r^n$, again by Corollary I.3.5.

(8.13) Exercise. *Let A be an* **artinian ring**, *i.e., a ring satisfying the descending chain condition for ideals: every descending chain of ideals $I_1 \supset I_2 \supset \cdots$ of A stabilises.*
(1) If A is a domain, then it is a field.
(2) A has only finitely many prime ideals $\mathfrak{m}_1, \ldots, \mathfrak{m}_r$, all of which are maximal.
(3) The ideal $\mathfrak{m}_1 \cdots \mathfrak{m}_r$ is nilpotent: there exists $n \geq 1$ such that $(\mathfrak{m}_1 \cdots \mathfrak{m}_r)^n = (0)$.
(4) $A = A/\mathfrak{m}_1^n \times \cdots \times A/\mathfrak{m}_r^n$.
(5) $A^{\mathrm{red}} = A/\mathfrak{m}_1 \times \cdots \times A/\mathfrak{m}_r$ is a finite product of fields.
(6) A is a noetherian ring.
(7) Conversely, a noetherian ring in which all prime ideals are maximal (e.g., the ring A/I in Example 8.12(ii)) is artinian.

(8.14) Exercise. *Let A be an algebra of finite type over a field. Consider $\mathrm{Max}(A) \subset \mathrm{Spec}(A)$ with the induced topology. Show that the maps "intersection with $\mathrm{Max}(A)$" and "closure in $\mathrm{Spec}(A)$" define mutually inverse bijections*

$$\{\text{closed subsets of } \mathrm{Spec}(A)\} \xleftrightarrow{\sim} \{\text{closed subsets of } \mathrm{Max}(A)\}$$

and

$$\{\text{irreducible closed subsets of } \mathrm{Spec}(A)\} \xleftrightarrow{\sim} \{\text{irreducible closed subsets of } \mathrm{Max}(A)\}.$$

(8.15) Irreducibility and extension of scalars. Let K be a field. An **affine K-variety** $Z \hookrightarrow \mathbf{A}^n_K$ is given by a system of polynomial equations

$$Z : \forall f \in I \quad f = 0,$$

where $I \in \mathrm{Spec}(K[X_1, \ldots, X_n])$ is a **prime** ideal. In other words, the ring $A = O(Z) = K[X_1, \ldots, X_n]/I$ of regular functions on Z is a domain ($\iff$ A is reduced and $\mathrm{Spec}(A)$ is irreducible). The fraction field $\mathrm{Frac}(A)$ is the **field of rational functions on** Z.

For any field extension $K \hookrightarrow L$ we can consider the extension of scalars of Z to L, i.e., the object $Z_L \hookrightarrow \mathbf{A}^n_L$ given by the same system of equations, but this time over L. If we write $I = (f_1, \ldots, f_r)$, then

$$A_L = O(Z_L) = L[X_1, \ldots, X_n]/(f_1, \ldots, f_r) = A \otimes_K L.$$

When is Z_L an L-variety? The following examples are instructive.

(i) $A = \mathbf{R}[X, Y]/(X^2 + Y^2)$, $K = \mathbf{R}$, $L = \mathbf{C}$. In this case $A_{\mathbf{C}} = \mathbf{C}[X, Y]/((X + iY)(X - iY))$ is reduced, but $\mathrm{Spec}(A_{\mathbf{C}})$ is reducible: it corresponds to a pair of complex conjugate lines intersecting transversally.

Note that the element $t = \overline{X}/\overline{Y} \in \mathrm{Frac}(A)$ satisfies $t^2 + 1 = 0$. In other words, **the field of constants of** Z, namely, $\overline{K} \cap \mathrm{Frac}(A) = \mathbf{R}(t) \xrightarrow{\sim} \mathbf{C}$, is a non-trivial separable extension of K.

(ii) Let K be a non-perfect field, $char(K) = p$, $a \in K \setminus K^p$. The K-algebra $A = K[X, Y]/(Y^p - a)$ is a domain. If L is a field extension of K containing α such that $\alpha^p = a$, then $A_L = L[X, Y]/(Y^p - a) = L[X, Y]/((Y - \alpha)^p)$ is not reduced; it represents a line over L with multiplicity p. The corresponding reduced ring $(A_L)^{\mathrm{red}} = L[X, Y]/(Y - \alpha) \xrightarrow{\sim} L[X]$ is a domain, so Z_L is not an L-variety, but $(Z_L)^{\mathrm{red}}$ is.

The field of constants of Z is $\overline{K} \cap \mathrm{Frac}(A) = K(\alpha)$, a non-trivial purely inseparable extension of K.

(iii) Let $K = \mathbf{F}_p(a, b)$, where a, b are variables. Again, the K-algebra $A = K[X, Y]/(Y^p + aX^p + b)$ is a domain, but its extension of scalars to any field $L \supset K$ containing elements α, β such that $\alpha^p = a$ and $\beta^p = b$ is a non-reduced ring $A_L = L[X, Y]/((Y + \alpha X + \beta)^p)$ corresponding again to a line with multiplicity p ($(A_L)^{\mathrm{red}} = L[X, Y]/(Y + \alpha X + \beta) \xrightarrow{\sim} L[X]$ is again a domain). Unlike in (ii), $\overline{K} \cap \mathrm{Frac}(A) = K$. However, the elements $\overline{X}, \overline{Y}, 1 \in \mathrm{Frac}(A)$ which are linearly independent over K become linearly dependent in any field extension of $\mathrm{Frac}(A)$ containing $K(\alpha, \beta) = K^{1/p}$, since $\overline{Y} + \alpha\overline{X} + \beta = 0$ in such a field ("$\mathrm{Frac}(A)$ and $K^{1/p}$ are not linearly disjoint over K").

(8.16) What is going on? There exists a theory of (in-)separability for arbitrary field extensions ([Ei, A.1.2-1.3], [ZS 1, II.15]). In examples 8.15(ii) and (iii) the field extension $\mathrm{Frac}(A)/K$ is not separable (for each element $z \in \mathrm{Frac}(A)$ transcendental over K the extension $\mathrm{Frac}(A)/K(z)$ is a finite inseparable extension; cf. the comments in the first proof of Theorem 4.10).

The general result is the following ([ZS 1, III.15], [Mu 1, II.4, Prop. 4]). If A is a domain of finite type over K and L is an algebraically closed field extension of K, then:

$$A \otimes_K L \text{ is reduced} \iff \mathrm{Frac}(A) \otimes_K L \text{ is reduced} \iff \mathrm{Frac}(A)/K \text{ is separable} \tag{8.16.1}$$

and

$$(A \otimes_K L)^{\mathrm{red}} \text{ is a domain} \iff \sqrt{(0)} \in \mathrm{Spec}(\mathrm{Frac}(A) \otimes_K L) \iff K^{sep} \cap \mathrm{Frac}(A) = K. \tag{8.16.2}$$

It follows that

$$A \otimes_K L \text{ is a domain} \iff (0) \in \mathrm{Spec}(\mathrm{Frac}(A) \otimes_K L) \iff \mathrm{Frac}(A)/K \text{ is separable and } \overline{K} \cap \mathrm{Frac}(A) = K. \tag{8.16.3}$$

Note that $A \otimes_K K'$ is contained in $A \otimes_K L$, for any algebraically closed field L containing K'; thus the condition (8.16.3) is equivalent to $A \otimes_K K'$ being a domain **for every field extension** $K \hookrightarrow K'$.

Note that the field of constants $\overline{K} \cap \mathrm{Frac}(A)$ is always a finite extension of K, by Exercise 9.6(2).

9. Dimension theory for affine algebras over a field

There are at least four different definitions of dimension in commutative algebra; the fact that they all agree is a non-trivial (and very useful) theorem. This theory applies both in algebraic geometry and in arithmetic. We are going to discuss two global definitions of dimension; a local definition will be mentioned in passing in 13.8 below. Throughout §9, K is an arbitrary field.

(9.1) Our first goal is to define the dimension of classical algebro-geometric objects, namely

$$Z : \forall f \in I \quad f = 0, \qquad Z \hookrightarrow \mathbf{A}^n_K,$$

where $I \subset K[X_1, \ldots, X_n]$ is the ideal of polynomial functions vanishing on Z.

We can visualise Z in at least three different ways: either as the set of classical points $Z(\overline{K}) \xrightarrow{\sim} \mathrm{Max}(A \otimes_K \overline{K})$, where $A = O(Z) = K[X_1, \ldots, X_n]/I$ is the ring of regular functions on Z, or as $\mathrm{Spec}(A)$ or as $\mathrm{Max}(A) \xrightarrow{\sim} Z(\overline{K})/\operatorname{Aut}(\overline{K}/K)$. Each of these three sets is equipped with the Zariski topology.

What is the dimension of Z equal to? Geometric intuition suggests that we should first decompose $Z = Z_1 \cup \cdots \cup Z_r$ into a union of its irreducible components and then define

$$\dim(Z) = \max_{1 \le i \le r} \dim(Z_i), \tag{9.1.1}$$

where $\dim(Z_i)$ is the number of independent parameters defining Z_i. We need to translate this intuition into a purely algebraic language involving only the ring of regular functions $A = O(Z)$. We use the decomposition into irreducible components of $\mathrm{Spec}(A)$, which is equivalent to an analogous decomposition of $\mathrm{Max}(A)$, by Exercise 8.14 (examples 8.15(i)-(iii) tell us that such a decomposition need not be preserved under extension of scalars $K \hookrightarrow L$, but it turns out that $\dim(Z_L) = \dim(Z)$; cf. Theorem 11.6 below).

(9.2) The first step is easy. Any K-algebra of finite type $A \neq 0$ has finitely many minimal prime ideals $P_1, \ldots, P_r$ $(r \ge 1)$, which correspond to the irreducible components $\overline{\{P_i\}} = V(P_i) \xrightarrow{\sim} \mathrm{Spec}(A/P_i)$ of $\mathrm{Spec}(A)$, by Corollary 8.11(1). In line with (9.1.1), we impose

$$\dim_1(A) = \max_{1 \le i \le r} \dim_1(A/P_i) \tag{9.2.1}$$

(we use the subscript "1" to distinguish this geometrically intuitive dimension from the abstract dimension defined in 9.14 below).

In the second step we consider the domain $A_i = A/P_i$ (which depends only on the reduced ring A^{red} and which is again a K-algebra of finite type). According to Noether's Normalisation Lemma 4.8, A_i is finite over a suitable polynomial ring $K[a_1, \ldots, a_{d_i}] \hookrightarrow A_i$ (this inclusion corresponds to a "finite" geometric morphism $Z_i \longrightarrow \mathbf{A}^{d_i}_K$, where Z_i is the irreducible component of Z with the ring of regular functions $O(Z_i) = A_i$). The corresponding extension of the fields of rational functions $K(a_1, \ldots, a_{d_i}) \hookrightarrow \mathrm{Frac}(A_i)$ is again finite, by 3.3(ii).

Proposition 9.5 below shows that d_i depends only on the field extension $\mathrm{Frac}(A_i)/K$: it is the maximal number of elements of $\mathrm{Frac}(A_i)$ that are algebraically independent over K (the **transcendence degree** of $\mathrm{Frac}(A_i)$ over K). We define, therefore,

$$\dim_1(A/P_i) = \mathrm{tr.deg}_K \mathrm{Frac}(A/P_i) \tag{9.2.2}$$

and

$$\dim_1(A) = \max_{1 \le i \le r} \mathrm{tr.deg}_K \mathrm{Frac}(A/P_i). \tag{9.2.3}$$

As observed above, the collection of domains A/P_i depends only on the reduced ring A^{red}, hence

$$\dim_1(A) = \dim_1(A^{\mathrm{red}}). \tag{9.2.4}$$

(9.3) Definition. *A* **transcendence basis** *of a field extension $K \hookrightarrow L$ is a subset $B \subset L$ such that L is algebraic over $K(B)$ and B is algebraically independent over K (i.e., every finite subset of B is algebraically independent over K).*

(9.4) We can take, for example, $B = \{X_1, \dots, X_n\}$ (resp. $B = \emptyset$) if $L = K(X_1, \dots, X_n)$ (resp. if L is algebraic over K). We are going to show that the cardinality of B depends only on the field extension L/K. For our purposes, it will be enough to consider finitely generated field extensions (see [ZS 1, II.12] for the general case).

(9.5) Proposition-Definition. *Let $K \hookrightarrow L$ be a field extension of finite type.*
(1) If B (resp. S) is a finite subset of L such that B is algebraically independent over K (resp. L is algebraic over $K(S)$), then $|B| \leq |S|$. In particular, $|B|$ is bounded above by the number of generators of the field extension L/K.
(2) If $|B|$ is maximal, then B is a transcendence basis of L/K and all transcendence bases of L/K have cardinality $|B|$. We say that L/K has **transcendence degree** *(notation: $\mathrm{tr.deg}_K L$) equal to $|B|$.*
(3) Any finite set S of generators of the field extension L/K contains a transcendence basis of L/K.

Proof. (1) We can assume that $B = \{b_1, \dots, b_n\}$ with $n > 0$. We claim that there exists $s_1 \in S$ which is transcendental over $K(b_2, \dots, b_n)$. Indeed, if all $s \in S$ were algebraic over $K(b_2, \dots, b_n)$, then the extension $K(S)/K(b_2, \dots, b_n)$, hence also $L/K(b_2, \dots, b_n)$, would be algebraic, but $b_1 \in L$ is not algebraic over $K(b_2, \dots, b_n)$. Therefore $s_1, b_2, \dots, b_n$ are algebraically independent over K. Repeating the procedure, we obtain $s_1, \dots, s_n \in S$, which are algebraically independent over K, hence distinct; thus $|B| = n \leq |S|$.
(2) If B has maximal cardinality among algebraically independent (over K) subsets of L, then B is a transcendence basis of L/K, by definition. If B' is another transcendence basis of L/K, then (1) for $S = B'$ yields $|B| \leq |B'|$, hence $|B| = |B'|$, by maximality of $|B|$.
(3) The proof of (1) applied to S and any transcendence basis B implies that S contains a subset $S' \subset S$ consisting of algebraically independent elements (over K) which has cardinality $|S'| = |B|$; it is a transcendence basis of L/K, by (2).

(9.6) Exercise. *A field extension $K \hookrightarrow L$ is* **purely transcendental** *if $L = K(B)$ for some transcendence basis B of L/K.*
(1) If $K \hookrightarrow L \hookrightarrow M$ are field extensions with $K \hookrightarrow L$ purely transcendental, then the minimal polynomials over K and over L of any element $\alpha \in M$ which is algebraic over K coincide.
(2) If $K \hookrightarrow M$ is a field extension of finite type, then $\{\alpha \in M \mid \alpha$ is algebraic over $K\}$ is a finite extension of K.
(3) If $K \hookrightarrow M$ is a field extension of finite type, then for any intermediate field $K \hookrightarrow L \hookrightarrow M$, the field extensions $K \hookrightarrow L$ and $L \hookrightarrow M$ are of finite type and $\mathrm{tr.deg}_K M = \mathrm{tr.deg}_K L + \mathrm{tr.deg}_L M$.

(9.7) Proposition. *If $A \neq 0$ is a K-algebra of finite type, then:*

$$\dim_1(A) = 0 \iff \text{each minimal prime ideal of } A \text{ is maximal} \iff \mathrm{Spec}(A) = \mathrm{Max}(A) \iff$$
$$\iff A^{\mathrm{red}} \text{ is a finite product of fields}$$

Proof. Let $P_1, \dots, P_r$ be the minimal prime ideals of A. If $\dim_1(A) = 0$, then each $A/P_i \subset \mathrm{Frac}(A/P_i)$ is a K-algebra of finite type consisting of elements algebraic over K, hence is a finite field extension of K, by Proposition III.3.10 and III.3.15. It follows that $P_i \in \mathrm{Max}(A)$, hence $\mathrm{Spec}(A) = \mathrm{Max}(A) = \{P_1, \dots, P_r\}$. As in Example 8.12(ii), $A^{\mathrm{red}} \xrightarrow{\sim} A/P_1 \times \cdots \times A/P_r$. Conversely, if A^{red} is a finite product of fields, then each of these fields must be finite over K, by Theorem 6.5(1), hence A^{red} is as in Example 8.12(iii), which implies that $\dim_1(A^{\mathrm{red}}) = 0$.

If $\dim_1(A) = d > 0$, then at least one of the rings A/P_i is finite over $K[X_1, \dots, X_d]$, hence A/P_i is not a field and $\mathrm{Spec}(A) \neq \mathrm{Max}(A)$.

(9.8) The key property of dimension is the following special case of Krull's Principal Ideal Theorem (see 10.21 below for the general statement). It makes precise the geometric intuition according to which dimension decreases by one if we impose one non-trivial condition $f = 0$.

(9.9) Theorem (A special case of Krull's Principal Ideal Theorem (Hauptidealsatz)). *Let A be a domain of finite type over K, let $f \in A$, $f \neq 0$, let $P \in \operatorname{Spec}(A)$ be minimal among the prime ideals containing f (in other words, $\overline{\{P\}} = V(P) \xrightarrow{\sim} \operatorname{Spec}(A/P)$ is an irreducible component of $V((f)) \xrightarrow{\sim} \operatorname{Spec}(A/(f))$, the locus of the points where $f = 0$). Then*

$$\operatorname{tr.deg}_K \operatorname{Frac}(A/P) = \operatorname{tr.deg}_K \operatorname{Frac}(A) - 1.$$

[Note that the existence of P, which we assume, implies that $f \notin A^*$.]

Proof. The following argument is due to Tate (see [Mu 1, I.7. Thm.2]). It consists of three simple steps.
Step 1: reduction to the case when $V(P)$ is the only irreducible component. This is done by localising, i.e., by finding a function $g \in A$ which identically vanishes on the remaining irreducible components (but not on $V(P)$) and then removing all zeroes of g by imposing a new equation $Yg - 1 = 0$ (cf. the proof of Theorem 6.5(6)).

In concrete terms, let $P_1, \dots, P_r$ (with $P_1 = P$) be as in Corollary 8.11(2), for $I = (f)$; then $\sqrt{(f)} = P_1 \cap \dots \cap P_r$. For each $i \neq 1$ choose $g_i \in P_i \setminus P$; then $g = g_2 \cdots g_r \in P_2 \cdots P_r \setminus P$, which means that $V(P_2) \cup \dots \cup V(P_r) \subset V((g))$, but $V(P) \not\subset V((g))$. We remove the set of all zeroes $V((g))$ of g by replacing A by the K-algebra of finite type

$$B = A[Y]/(Yg - 1) = A[1/g] \subset \operatorname{Frac}(A) = \operatorname{Frac}(B).$$

Note that $PB \in \operatorname{Spec}(B)$, since $B/PB = (A/P)[1/\overline{g}]$, $\overline{g} = g \pmod P \in (A/P) \setminus \{0\}$. On the other hand, $P_iB = B$ for $i \neq 1$, since $1 = g \cdot \overline{Y} \in P_iB$; thus

$$\sqrt{(f)}B = (P_1 \cap \dots \cap P_r)B \subset P_1B \cap \dots \cap P_rB = PB = PP_2 \cdots P_rB \subset (P_1 \cap \dots \cap P_r)B = \sqrt{(f)}B,$$

which implies that $\sqrt{(f)}B = PB$ is a prime ideal of B and $\operatorname{Frac}(B/PB) = \operatorname{Frac}(A/P)$. As a result, we can replace the triple (A, f, P) by (B, f, PB) and assume that $\sqrt{(f)} = P$ is a prime ideal.
Step 2: reduction to the case when A is a polynomial ring. This is done by combining Noether's Normalisation Lemma 4.8 with a norm argument. More precisely, there exists a polynomial subalgebra $A_0 = K[X_1, \dots, X_d] \subset A$ such that A is finite over A_0. We want to replace all objects

$$\sqrt{(f)} = P \subset A \subset \operatorname{Frac}(A) = L$$

by

$$P_0 = P \cap A_0 \subset A_0 = K[X_1, \dots, X_d] \subset \operatorname{Frac}(A_0) = L_0 = K(X_1, \dots, X_d).$$

As A is finite over A_0, so is L over L_0 and A/P and A_0/P_0; thus

$$\operatorname{tr.deg}_K L = \operatorname{tr.deg}_K L_0 = d, \qquad \operatorname{tr.deg}_K \operatorname{Frac}(A/P) = \operatorname{tr.deg}_K \operatorname{Frac}(A_0/P_0).$$

(9.10) Lemma. *$P_0 = \sqrt{(f_0)}$, where $f_0 = N_{L/L_0}(f)$.*

Proof of Lemma. According to Proposition 3.4, the minimal polynomial equation for f over the integrally closed ring A_0 is of the form $f^n + a_{n-1}f^{n-1} + \dots + a_0$, with all $a_i \in A_0$. In particular, $f_0 = \pm a_n^{[L:L_0(f)]} \in A_0$. Moreover, $a_0 \in A_0 \cap fA \subset A_0 \cap P = P_0$, which implies that $f_0 \in P_0$, hence $\sqrt{(f_0)} \subset P_0$. Conversely, if $h \in P_0$, then $h^m \in fA$ for some $m \geq 1$, hence $h^{m[L:L_0]} = N_{L/L_0}(h) \in N_{L/L_0}(f)N_{L/L_0}(A) \subset f_0A_0 = (f_0)$ (using Corollary 3.6), which proves that $P_0 \subset \sqrt{(f_0)}$. Lemma is proved.
Step 3: elementary verification of Theorem 9.9 in the case when A is a polynomial ring. Thanks to Step 2 and Lemma 9.10 we can replace the triple $(A, f, P = \sqrt{(f)})$ by $(A_0 = K[X_1, \dots, X_d], f_0, P_0 = \sqrt{(f_0)})$. The condition $\sqrt{(f_0)}) = P_0 \in \operatorname{Spec}(A_0)$ implies, by Example 8.12(i), that $P_0 = (f_{00})$, where $f_{00} \in K[X_1, \dots, X_d]$ is an irreducible element (non-constant). Theorem 9.9 is then a consequence of the following elementary statement.

(9.11) Lemma. *Let $f_{00} \in K[X_1, \ldots, X_d]$ be an irreducible non-constant polynomial. The ring $R = K[X_1, \ldots, X_d]/(f_{00})$ is a domain and* $\mathrm{tr.deg}_K \operatorname{Frac}(R) = d - 1$.

Proof of Lemma. The ring R is a domain, by Proposition I.5.7. We can assume that the variable X_d occurs in f_{00}; its image $\overline{X_d} \in R$ is then algebraic over $K[\overline{X}_1, \ldots, \overline{X}_{d-1}] \subset R$, which implies that $\mathrm{tr.deg}_K \operatorname{Frac}(R) \leq d - 1$. Conversely, if $\overline{X}_1, \ldots, \overline{X}_{d-1}$ were algebraically dependent over K, then there would be a non-zero polynomial $h \in K[X_1, \ldots, X_{d-1}]$ such that $\overline{h} = 0$, hence $h \in f \cdot K[X_1, \ldots, X_d]$, which is impossible; therefore $\mathrm{tr.deg}_K \operatorname{Frac}(R) \geq d - 1$.

(9.12) Corollary. *If A is a K-algebra of finite type and $P_0 \subsetneq P_1$ are prime ideals of A, then*

$$\dim_1(A/P_1) \leq \dim_1(A/P_0) - 1,$$

with equality if P_1 is minimal among the prime ideals $\neq P_0$ containing P_0.

Proof. We can replace A by A/P_0, hence assume that $P_0 = (0)$ and A is a domain. Let $f \in P_1$ be any non-zero element and let $Q \subset P_1/(f)$ be a minimal prime ideal of $A/(f)$; then $Q = P/(f)$ for a prime ideal $P \subset P_1$ of A which is minimal among prime ideals containing f. As A/P_1 is a quotient of A/P, we have $\dim_1(A/P_1) \leq \dim_1(A/P)$, with equality if P_1 is minimal among non-zero prime ideals (since $P = P_1$ then). Applying Theorem 9.9 to the triple (A, f, P), we obtain

$$\dim_1(A/P) = \dim_1(A/P_0) - 1,$$

which concludes the proof.

(9.13) In view of Corollary 9.12, it is natural to study the lengths of chains of prime ideals of A

$$P_0 \subsetneq P_1 \subsetneq \cdots \subsetneq P_r, \tag{9.13.1}$$

which correspond to chains of **irreducible closed subsets** of $\operatorname{Spec}(A)$

$$V(P_0) \supsetneq V(P_1) \supsetneq \cdots \supsetneq V(P_r), \qquad V(P_i) = \overline{\{P_i\}} \xrightarrow{\sim} \operatorname{Spec}(A/P_i).$$

This makes sense for arbitrary rings A. Corollary 9.12 then serves as a motivation for the following general definition.

(9.14) Definition. *(1) Let $X \neq \emptyset$ be a topological space. The* **Krull dimension of** X *is*

$$\dim(X) = \sup\{r \geq 0 \mid \exists\, Z_0 \supsetneq Z_1 \supsetneq \cdots \supsetneq Z_r,\ Z_i \text{ irreducible closed subset of } X\} \in \mathbf{N} \cup \{+\infty\}.$$

(2) Let $A \neq 0$ be a ring. The **Krull dimension of** A *is*

$$\dim(A) = \dim(\operatorname{Spec}(A)) = \sup\{r \geq 0 \mid \exists\, P_0 \subsetneq P_1 \subsetneq \cdots \subsetneq P_r,\ P_i \in \operatorname{Spec}(A)\} \in \mathbf{N} \cup \{+\infty\}.$$

(9.15) Examples. (i) $\dim(A) = 0 \iff \operatorname{Spec}(A) = \operatorname{Max}(A)$. In particular, a domain of $\dim(A) = 0$ is a field.
(ii) For $A = K[X_1, \ldots, X_n]$ the chain

$$(0) \subsetneq (X_1) \subsetneq \cdots \subsetneq (X_1, \ldots, X_n)$$

has length n, hence $\dim(K[X_1, \ldots, X_n]) \geq n = \dim_1(K[X_1, \ldots, X_n])$. Theorem 9.16 below implies that this is, in fact, an equality.
(iii) If A is a noetherian ring, then all chains (9.13.1) have finite length. However, this does not necessarily imply that $\dim(A) < \infty$. In fact, Nagata constructed an example of a noetherian ring of infinite dimension (see [De 2, Ex. 6.7], [Re, 9.4(3), Ex. 9.2]).
(iv) Chains of prime ideals (9.13.1) with fixed P_0 are in bijection with the corresponding chains

$$(0) \subsetneq P_1/P_0 \subsetneq \cdots \subsetneq P_r/P_0$$

of prime ideals of the domain A/P_0.
(v) Proposition 8.7 tells us that every $P_0 \in \mathrm{Spec}(A)$ contains a minimal prime ideal Q. Chains (9.13.1) with P_0 containing a fixed minimal ideal Q then correspond bijectively to chains

$$P_0/Q \subsetneq P_1/Q \subsetneq \cdots \subsetneq P_r/Q$$

in the domain A/Q. As a result,

$$\dim(A) = \sup\{\dim(A/Q) \mid Q \text{ minimal prime ideal}\},$$

which is analogous to (9.2.1).

(9.16) Theorem. *Let $A \neq 0$ be a K-algebra of finite type.*
(1) $\dim(A) = \dim_1(A)$.
(2) Let $P \subset Q$ be prime ideals of A. If $P = P_0 \subsetneq P_1 \subsetneq \cdots \subsetneq P_r = Q$ is a chain of prime ideals of A between P and Q, then

$$r \leq \dim_1(A/P) - \dim_1(A/Q) = \mathrm{tr.deg}_{\mathrm{Frac}(A/Q)} \mathrm{Frac}(A/P),$$

with equality if the chain is **saturated** *(i.e., if it cannot be refined). [Rings for which all saturated chains between any fixed pair $P \subset Q$ of prime ideals have the same length are called* **catenary***.]*
(3) If $P \in \mathrm{Spec}(A)$ and if If $P = P_0 \subsetneq P_1 \subsetneq \cdots \subsetneq P_t$ is an increasing chain of prime ideals starting at P, then

$$t \leq \dim_1(A/P) = \mathrm{tr.deg}_K \mathrm{Frac}(A/P),$$

with equality if the chain is saturated (i.e., if it cannot be refined nor extended beyond P_t).

Proof. The statement (2) for $P = (0)$ and $Q \in \mathrm{Max}(A)$ implies (1) in the case when A is a domain (since A/Q is a finite field extension of K, by the Nullstellensatz, hence $\dim_1(A/Q) = 0$). The general case of (1) follows by combining (9.2.1) with Example 9.15(v).

Any chain in (2) has finite length, since A is noetherian. Corollary 9.12 tells us that

$$\forall\, i = 0, \ldots, r-1 \quad \dim_1(A/P_{i+1}) \leq \dim_1(A/P_i) - 1,$$

with equalities everywhere $\iff$ the chain is saturated. Taking the sum of these (in)equalities yields (2).
(3) There exists $Q \in \mathrm{Max}(A)$ containing P_t. Replacing P_t with Q and applying (2) to this modified chain, we obtain, as in the proof of (1),

$$t \leq \dim_1(A/P) - \dim_1(A/Q) = \dim_1(A/P),$$

with equality if the original chain was saturated.

(9.17) Nagata constructed a non-catenary local noetherian domain, which has saturated chains of prime ideals (between (0) and the unique maximal ideal) of different lengths (2 and 3). See [Re, 9.4(2), Ex. 9.4].

(9.18) Theorem. *Let A be a domain of finite type over K, let $f_1, \ldots, f_r \in A$. If the ring $A' = A/(f_1, \ldots, f_r)$ is non-zero, then $\dim(A'/P') \geq \dim(A) - r$, for every minimal prime ideal P' of A'.*

Proof. This follows from Theorem 9.9 by induction on r.

(9.19) Example. The union of the three coordinate axes in $\mathbf{A}^3_K$ is given by three equations

$$X_1X_2 = X_1X_3 = X_2X_3 = 0.$$

In this example $A = K[X_1, X_2, X_3]$, $f_1 = X_2X_3$, $f_2 = X_1X_3$, $f_3 = X_1X_2$, $r = 3$, $\dim(A) = 3$, $\dim(A'/P') = 1$ for every minimal prime ideal P' of A'. In fact, the ideal $(f_1, f_2, f_3) \subset A$ **cannot** be generated by two elements (exercise!).

10. Localisation

We saw in 5.4-5.5 how to formally invert an arbitrary element f of a ring A, by passing to a new ring $A[1/f] = A[Y]/(Yf - 1)$, in which (the image of) f has an inverse, namely, $\overline{Y}$. However, the ring $A[1/f]$

can sometimes be zero. In this section we are going to study rings obtained by inverting an arbitrary subset of A.

(10.1) Let A be any ring. A subset $S \subset A$ is **multiplicative** if $1 \in S$ and $st \in S$ for all $s, t \in S$.

Our goal is to construct a "universal" ring $S^{-1}A$ containing fractions $a/s = \frac{a}{s}$ ($a \in A$, $s \in S$) satisfying the usual relations. We want to imitate the classical procedure for constructing the fraction field of a domain A (for $S = A \setminus \{0\}$), when a/s is the class of a pair $(a, s) \in A \times S$ with respect to the equivalence relation

$$(a, s) \sim (a', s') \iff as' - a's = 0. \tag{10.1.1}$$

The problem is that the relation (10.1.1) need not be transitive if A is not a domain. The correct general definition is the following:

$$(a, s) \sim (a', s') \iff \exists\, t \in S \quad t(as' - a's) = 0. \tag{10.1.2}$$

One checks that (10.1.2) is, indeed, an equivalence relation on $A \times S$ and that the set $S^{-1}A$ of equivalence classes forms a ring with respect to the usual operations

$$\frac{a}{s} + \frac{a'}{s'} = \frac{as' + sa'}{ss'}, \qquad \frac{a}{s}\,\frac{a'}{s'} = \frac{aa'}{ss'},$$

where $a/s = \frac{a}{s}$ denotes the equivalence class of (a, s). The zero (resp. the unit) of $S^{-1}A$ is $0/1$ (resp. $1/1$).

(10.2) Proposition. *(1) $S^{-1}A = 0 \iff 0 \in S$.*
(2) The canonical map $\alpha : A \longrightarrow S^{-1}A$, $\alpha(a) = a/1$, is a ring homomorphism with kernel $\mathrm{Ker}(\alpha) = \{a \in A \mid \exists s \in S \;\; sa = 0\}$.
(3) (Universal property) If $\beta : A \longrightarrow B$ is a ring homomorphism such that $\beta(S) \subset B^*$, then there is a unique ring homomorphism $\overline{\beta} : S^{-1}A \longrightarrow B$ such that $\overline{\beta} \circ \alpha = \beta$, namely, $\overline{\beta}(a/s) = \beta(a)\beta(s)^{-1}$.
(4) If $S = \{f^n \mid n \geq 0\}$ for some $f \in A$, then the map $\overline{\beta} : S^{-1}A \xrightarrow{\sim} A[Y]/(Yf - 1) = A[1/f]$ corresponding to the canonical morphism $\beta : A \longrightarrow A[Y]/(Yf - 1)$ is an isomorphism of A-algebras.
(5) If A is a domain and $0 \notin S$, then $S^{-1}A = \{a/s \in \mathrm{Frac}(A) \mid a \in A,\, s \in S\} \subset \mathrm{Frac}(A)$.

Proof. Easy exercise ([Re, 6.2]).

(10.3) $S^{-1}A$ is a limit of $A[1/f]$ ($f \in S$). Note that, if A is a domain and $0 \notin S$, then

$$S^{-1}A = \bigcup_{f \in S} A[1/f] \subset \mathrm{Frac}(A).$$

A similar relation holds in general, since $S^{-1}A$ requires inverting all elements of S. However, the rings $A[1/f]$ are no longer contained in a common big ring. Nevertheless, $S^{-1}A$ is the **direct limit** of $A[1/f]$ ($f \in S$):

$$S^{-1}A = \varinjlim_{f \in S} A[1/f]. \tag{10.3.1}$$

(10.4) What is a direct limit? Assume that I is a non-empty ordered set with the property that for all $i, j \in I$ there exists $k \in I$ such that $i < k$ and $j < k$. In addition, assume that we are given sets X_i ($i \in I$) and maps $\alpha_{ij} : X_i \longrightarrow X_j$ ($i < j$) which are transitive: $\alpha_{ik} = \alpha_{jk} \circ \alpha_{ij}$. The direct (or inductive) limit of the X_i is the set

$$X = \varinjlim_{i \in I} X_i = \coprod_{i \in I} X_i / \sim$$

defined as the quotient of the disjoint union of the sets X_i by the equivalence relation

$$x_i \sim x_j \iff \exists\, k \quad i < k,\, j < k,\, \alpha_{ik}(x_i) = \alpha_{jk}(x_j).$$

For each $i \in I$ there is a natural map $\beta_i : X_i \longrightarrow X$ sending x_i to the class of x_i in X; these maps satisfy $\beta_j \circ \alpha_{ij} = \beta_i$. The direct limit X is the universal object among sets having this property. If all X_i are contained in some set Y (which means that the maps α_{ij} are the corresponding inclusions), then

$X = \bigcup X_i \subset Y$. If all X_i are groups (rings, modules ...) and all α_{ij} are homomorphisms of groups (rings, modules ...), then X is again a group (ring, module ...).

In (10.3.1), if $f \mid f'$, then $f' = fg$ for some $g \in A$ and the ring homomorphism

$$A[1/f] \longrightarrow A[1/f'], \qquad \frac{a}{f^n} \mapsto \frac{ag^n}{(fg)^n}$$

depends only on f and f', not on g. Moreover, these maps are transitive for $f \mid f' \mid f''$, hence give rise to the direct limit $\varinjlim A[1/f]$ (indexed by $I = S$ with order $f < f' \iff f \mid f'$).

(10.5) A construction similar to that in 10.1 works for an arbitrary A-module M. One checks that the relation

$$(m, s) \sim (m', s') \iff \exists\, t \in S \quad t(ms' - m's) = 0$$

on $M \times S$ is an equivalence relation and that the set $S^{-1}M$ of equivalence classes forms an $S^{-1}A$-module with respect to the operations

$$\frac{m}{s} + \frac{m'}{s'} = \frac{ms' + sm'}{ss'}, \qquad \frac{a}{s}\,\frac{m'}{s'} = \frac{am'}{ss'},$$

where $m/s = \frac{m}{s} \in S^{-1}M$ denotes the equivalence class of (m, s). Again,

$$S^{-1}M = \varinjlim_{f \in S} M[1/f],$$

where $M[1/f] = \{1, f, f^2, \ldots\}^{-1}M$. We leave it as an exercise to formulate an analogue of Proposition 10.2 for $S^{-1}M$. One can show that $S^{-1}M = S^{-1}A \otimes_A M$, but we are not going to use this fact. The most important property of the localisation for modules is the following exactness property.

(10.6) Proposition ($M \mapsto S^{-1}M$ is an exact functor). *If $M \xrightarrow{f} N \xrightarrow{g} P$ is an exact sequence of A-modules, then*

$$S^{-1}M \xrightarrow{f'} S^{-1}N \xrightarrow{g'} S^{-1}P,$$

where $f'(m/s) = f(m)/s$ and $g'(n/s) = g(n)/s$, is an exact sequence of $S^{-1}A$-modules.

Proof. One checks that f' (and g') are well-defined homomorphisms of $S^{-1}A$-modules. The condition $\mathrm{Im}(f) = \mathrm{Ker}(g)$ implies that $g \circ f = 0$, hence $g' \circ f' = 0$, which means that $\mathrm{Im}(f') \subset \mathrm{Ker}(g')$. Conversely, if $g'(n/s) = g(n)/s = 0 \in S^{-1}P$, then there exists $t \in S$ such that $tg(n) = g(tn) = 0 \in P$, hence $tn \in \mathrm{Ker}(g) = \mathrm{Im}(f)$, $tn = f(m)$ for some $m \in M$. It follows that $f'(m/st) = f(m)/st = tn/st = n/s$; therefore $\mathrm{Ker}(g') \subset \mathrm{Im}(f')$.

(10.7) Corollary. *If $L \subset M$ is an A-submodule, then $S^{-1}L \subset S^{-1}M$ and $S^{-1}(M/L) = S^{-1}M/S^{-1}L$. In particular, if I is an ideal of A, then $S^{-1}I$ is an ideal of $S^{-1}A$.*

(10.8) Proposition (Localisation commutes with quotients). *For any ideal I of A the image $\overline{S}$ of S in $\overline{A} = A/I$ is a multiplicative subset of $\overline{A}$ and the canonical map $S^{-1}A/S^{-1}I \longrightarrow \overline{S}^{-1}\overline{A}$ sending $a/s \pmod{S^{-1}I}$ to $\overline{a}/\overline{s}$ is a ring isomorphism.*

Proof. If we view $\overline{A}$ as an A-module, then $\overline{S}^{-1}\overline{A} = S^{-1}\overline{A}$ is canonically isomorphic to $S^{-1}A/S^{-1}I$ as an $S^{-1}A$-module, by Corollary 10.7. It is easy to check that the map in question is a ring homomorphism.

(10.9) Definition. *If $P \in \mathrm{Spec}(A)$, then $S = A \setminus P$ is a multiplicative set. The corresponding localisation A_P (resp. M_P) is called the* **localisation of A (resp. of M) at P.**

(10.10) Examples. (i) If A is a domain and $P = (0)$, then $A_P = \mathrm{Frac}(A)$.
(ii) If $A = \mathbf{Z}$ and $P = (p)$, where p is a prime number, then

$$\mathbf{Z}_{(p)} = \{a/b \mid a, b \in \mathbf{Z},\ p \nmid b\} \subset \mathbf{Q}.$$

(iii) If $A = K[X]$ for a field K and $P = (X - c)$ $(c \in K)$, then

$$K[X]_{(X-c)} = \{f/g \mid f, g \in K[X],\, g(c) \neq 0\} \subset K(X).$$

This is the ring of those rational functions on the affine line $\mathbf{A}^1_K$ which are defined at the point $c \in K$.
(iv) Both rings in (ii) and (iii) are examples of **discrete valuation rings** (see §14 below), since

$$\mathbf{Z}_{(p)} \setminus \{0\} = \bigcup_{n\geq 0} p^n\,(\mathbf{Z}_{(p)})^*, \qquad K[X]_{(X-c)} \setminus \{0\} = \bigcup_{n\geq 0} (X-c)^n\,(K[X]_{(X-c)})^*.$$

As we shall see, discrete valuation rings are non-singular local objects of dimension one.
(v) If $Q \subset P$ are prime ideals of A, then $QA_P \in \mathrm{Spec}(A_P)$ and one can check that the localisation of A_P at QA_P coincides with A_Q. For example, if $A = K[X,Y]$ (K a field), $Q = (X)$ and $P = (X,Y)$, then $A_P = \{f/g \mid f, g \in K[X,Y],\, g(0,0) \neq 0\}$ and $A_Q = \{f/g \mid f, g \in K[X,Y],\, X \nmid g\}$. Geometrically, A_P (resp. A_Q) consists of those rational functions on the plane $\mathbf{A}^2_K$ which are defined at the origin $(0,0)$ (resp. which are defined at sufficiently "generic" points of the line $X = 0$).
(vi) The ring $A = K[X,Y]/(XY^2)$ is generated as a K-algebra by $\overline{X}, \overline{Y}$, where $\overline{XY}^2 = 0$. We are going to determine the localisations of A at the two minimal prime ideals $(\overline{X}), (\overline{Y}) \in \mathrm{Spec}(A)$ of A. The localisation $A_{(\overline{X})}$ contains $1/\overline{Y}$, hence $\overline{X}/1 = 0 \in A_{(\overline{X})}$. It follows that

$$A_{(\overline{X})} = (A/(\overline{X}))_{(0)} = \mathrm{Frac}(K[Y]) = K(Y).$$

Similarly, $1/\overline{X} \in A_{(\overline{Y})}$, hence $\overline{Y}^2/1 = 0 \in A_{(\overline{Y})}$; thus

$$A_{(\overline{Y})} = (A/(\overline{Y}^2))_{(\overline{Y})} = (K[X,Y]/(Y^2))_{(\overline{Y})} = K(X)[Y]/(Y^2)$$

(the dual numbers over $K(X)$). Geometrically, $A = O(Z)$, where $Z : XY^2 = 0$ ($Z \hookrightarrow \mathbf{A}^2_K$) is a union of the vertical axis $X = 0$ with the double horizontal axis $Y^2 = 0$, since $(XY^2) = (X) \cap (Y^2)$.

(10.11) Geometric interpretation of A_P. For each $f \in A$, the ring $A[1/f]$ can be interpreted as the ring of functions on $D(f) = \mathrm{Spec}(A) \setminus V((f)) = \{P \in \mathrm{Spec}(A) \mid f \notin P\}$. Indeed, the value of any element $a/f^n \in A[1/f]$ ($a \in A$) at $P \in D(f)$ is equal to

$$\frac{a}{f^n}(P) = \frac{a \pmod P}{(f \pmod P)^n} \in \mathrm{Frac}(A/P) = k(P),$$

which makes sense, since $f(P) = f \pmod P \neq 0 \in A/P$.

For fixed $P \in \mathrm{Spec}(A)$, the sets $D(f)$ for $f \notin P$ form a basis of open neighbourhoods of P in $Spec(A)$, which means that (10.3.1) for $S = A \setminus P$

$$A_P = \varinjlim_{f \notin P} A[1/f] = \varinjlim_{P \in D(f)} A[1/f] \tag{10.11.1}$$

identifies A_P with the ring of **germs of functions on** $\mathrm{Spec}(A)$ **at** P. In general, a germ at a point x is represented by a function defined on some open set $U \ni x$; two functions define the same germ if they become equal on some open neighbourhood of x contained in their common domain of definition. This is a very general and very useful concept.

For example, we can consider continuous or differentiable or holomorphic functions. In the latter case, the ring of germs at $0 \in \mathbf{C}$ is the ring $\mathbf{C}\{X\}$ of power series $f \in \mathbf{C}[[X]]$ with positive radius of convergence.

(10.12) Ideals of $S^{-1}A$. The canonical morphism $\alpha : A \longrightarrow S^{-1}A$ induces the extension and restriction maps

$$e : \{\text{ideals of } A\} \longrightarrow \{\text{ideals of } S^{-1}A\}, \qquad e(I) = S^{-1}A \cdot \alpha(I) = S^{-1}I$$

$$r : \{\text{ideals of } S^{-1}A\} \longrightarrow \{\text{ideals of } A\}, \qquad r(J) = \alpha^{-1}(J).$$

(10.13) Proposition. *(1) $e(r(J)) = J$ holds for all ideals J of $S^{-1}A$.*
(2) $r(e(I)) = \{a \in A \mid \exists s \in S \;\; sa \in I\}$ (= the "S-saturation of I") holds for all ideals I of A.

(3) The maps r and e define mutually inverse bijections between the set of all ideals of $S^{-1}A$ and the set of all S-saturated ideals

$$\{\text{ideals } I \text{ of } A \text{ such that } [s \in S,\ sa \in I \Longrightarrow a \in I]\}$$

of A.
(4) If A is a noetherian ring, so is $S^{-1}A$.
(5) An ideal I of A satisfies $e(I) = S^{-1}A \iff r(e(I)) = A \iff I \cap S \neq \emptyset$.
(6) For a prime ideal $Q \in \mathrm{Spec}(A)$ the extension $S^{-1}Q = e(Q)$ is a prime ideal of $S^{-1}A$ (resp. is equal to $S^{-1}A$) if $Q \cap S = \emptyset$ (resp. if not).
(7) The canonical homomorphism $\alpha : A \longrightarrow S^{-1}A$ induces an injective map $\alpha^* : \mathrm{Spec}(S^{-1}A) \longrightarrow \mathrm{Spec}(A)$ with image $\mathrm{Im}(\alpha^*) = \{Q \in \mathrm{Spec}(A) \mid Q \cap S = \emptyset\}$.

Proof. (1), (2) This is an easy calculation ([Re, 6.3]). (3) is a consequence of (1) and (2). If $I = (f_1, \ldots, f_r)$, then $e(I) = (f_1, \ldots, f_r)$ in $S^{-1}A$. As every ideal of $S^{-1}A$ is of the form $e(I)$, this proves (4). In (5), $S^{-1}I = S^{-1}A \iff \overline{S}^{-1}(A/I) = 0$ (by Proposition 10.8), which is equivalent to $0 \in \overline{S} \iff I \cap S \neq \emptyset$. Similarly, in (6) we have $S^{-1}A/S^{-1}Q = \overline{S}^{-1}(A/Q)$, which is a domain (resp. is the zero ring) if $0 \in \overline{S}$ (resp. if not), where $\overline{S}$ is the image of S in A/Q. The map α^* in (7) is given by r; as $e \circ r = \mathrm{id}$, α^* is injective. By (3), $Q \in \mathrm{Spec}(A)$ lies in $\mathrm{Im}(\alpha^*) \iff Q$ is S-saturated, which is equivalent to $Q \cap S = \emptyset$ (by (6)).

(10.14) Proposition-Definition. *(1) Let $P \in \mathrm{Spec}(A)$. The map $Q \mapsto QA_P = (A \setminus P)^{-1}Q$ defines a bijection $\{Q \in \mathrm{Spec}(A) \mid Q \subset P\} \xrightarrow{\sim} \mathrm{Spec}(A_P)$. In particular, the ring A_P has a* **unique maximal ideal** *(rings having this property are called* **local rings***), namely PA_P. The residue field of A_P is equal to $A_P/PA_P = \mathrm{Frac}(A/P) = k(P)$.*
(2) The **semi-localisation of A at a finite set of prime ideals** *$P_1, \ldots, P_r \in \mathrm{Spec}(A)$ is the ring $S^{-1}A$ for $S = A \setminus (P_1 \cup \cdots \cup P_r)$. The map $Q \mapsto Q \cdot S^{-1}A = S^{-1}Q$ defines a bijection $\{Q \in \mathrm{Spec}(A) \mid \exists i \ \ Q \subset P_i\} \xrightarrow{\sim} \mathrm{Spec}(S^{-1}A)$. In particular, $\mathrm{Max}(S^{-1}A) = \{S^{-1}P_1, \ldots, S^{-1}P_r\}$.*

Proof. (1) Combine Proposition 10.13(7) with 10.8, which gives $A_P/PA_P = \{A/P \setminus \{0\}\}^{-1}(A/P) = \mathrm{Frac}(A/P)$. In (2) we must check that a prime ideal $Q \in \mathrm{Spec}(A)$ satisfying $Q \subset P_1 \cup \cdots \cup P_r$ is contained in one of the P_i, which is a special case of the following lemma.

(10.15) Lemma (Prime avoidance). *Let $J, I_1, \ldots, I_n$ be ideals of A such that $J \subset I_1 \cup \cdots \cup I_n$ and that at most two among $I_1, \ldots, I_n$ are not prime ideals. Then $J \subset I_i$ for some i.*

Proof. By induction on n, we can assume that $n \geq 2$ and that J is not contained in $\bigcup_{i \neq j} I_i$ for any j; we have, therefore, $x_j \in J$ such that $x_j \notin \bigcup_{i \neq j} I_i$, for all $j = 1, \ldots, n$. We can also assume that I_1 is a prime ideal if $n > 2$. The element $y = x_1 + x_2 \cdots x_n \in J$ then satisfies $y \notin I_1 \cup \cdots \cup I_n$, which is a contradiction.

(10.16) Examples. (i) According to Proposition 10.14(1), chains of prime ideals $P_0 \subsetneq P_1 \subsetneq \cdots \subsetneq P_r$ of $A \neq 0$ for which P_r is contained in a fixed prime ideal $P \in \mathrm{Spec}(A)$ correspond bijectively to chains $P_0A_P \subsetneq P_1A_P \subsetneq \cdots \subsetneq P_rA_P$ of prime ideals of A_P. In particular,

$$\dim(A_P) = \sup\{r \geq 0 \mid \exists P_0 \subsetneq P_1 \subsetneq \cdots \subsetneq P_r \subset P,\ P_i \in \mathrm{Spec}(A)\},$$

$$\dim(A) = \sup_{\mathfrak{m} \in \mathrm{Max}(A)} \dim(A_{\mathfrak{m}})$$

and, thanks to 9.15(iv),

$$\dim(A) \geq \dim(A_P) + \dim(A/P).$$

(ii) For every $f \in A$, the canonical morphism $\alpha : A \longrightarrow A[1/f]$ induces a bijection

$$\alpha^* : \mathrm{Spec}(A[1/f]) \xrightarrow{\sim} \{P \in \mathrm{Spec}(A) \mid f \notin P\} = D(f) = \mathrm{Spec}(A) \setminus V((f)) = \mathrm{Spec}(A) \setminus \mathrm{Spec}(A/(f)),$$

by Proposition 10.13(7) for $S = \{f^n \mid n \geq 0\}$. In particular, each open set $D(f) \subset \mathrm{Spec}(A)$ is naturally of the form $\mathrm{Spec}(-)$.

This is not true for more general open subsets. For example, the plane with one point removed, $\mathbf{A}^2_K \setminus \{(0,0)\} = \mathrm{Spec}(K[X,Y]) \setminus \{(X,Y)\}$ is not of the form $\mathrm{Spec}(-)$. However, making this statement precise requires some non-trivial foundational work. Here is a hint: the "true" $\mathrm{Spec}(A)$ is not just the topological space we have been studying here, but it is equipped with rings of regular functions $O(U)$ on each open subset $U \subset \mathrm{Spec}(A)$. For example, $O(D(f)) = A[1/f]$, and $O(U)$ for general U is determined by a glueing property.

In the special case $U = \mathrm{Spec}(K[X,Y]) \setminus \{(X,Y)\}$ we can write $U = D(X) \cup D(Y)$, where $D(X) \cap D(Y) = D(XY)$ ($D(X)$ is the plane with the Y axis removed, $D(Y)$ the plane with the X-axis removed). The ring $O(U)$ of functions regular on $\mathbf{A}^2_K \setminus \{(0,0)\}$ should be the intersection of $O(D(X)) = K[X,Y,1/X]$ and $O(D(Y)) = K[X,Y,1/Y]$ inside $O(D(XY)) = K[X,Y,1/X,1/Y]$, which is equal to $K[X,Y] = O(\mathbf{A}^2_K)$.

(10.17) Exercise. *Assume that $0 \notin S$.*
(1) If A is an integrally closed domain, so is $S^{-1}A$.
(2) If A is a PID, so is $S^{-1}A$.
(3) If A is a UFD, so is $S^{-1}A$.

(10.18) Exercise. *The* **support** *of an A-module M is $\mathrm{supp}(M) = \{P \in \mathrm{Spec}(A) \mid M_P \neq 0\}$ (geometrically, one should think of the collection $\{M_P\}$ as a family of spaces living over $\mathrm{Spec}(A)$).*
(1) If $P \in \mathrm{supp}(M)$, then $V(P) \subset \mathrm{supp}(M)$.
(2) $M = 0 \iff \forall P \in \mathrm{Spec}(A) \quad M_P = 0 \iff \forall \mathfrak{m} \in \mathrm{Max}(A) \quad M_\mathfrak{m} = 0$.
(3) If $0 \longrightarrow M \longrightarrow N \longrightarrow P \longrightarrow 0$ is an exact sequence of A-modules, then $\mathrm{supp}(N) = \mathrm{supp}(M) \cup \mathrm{supp}(P)$.
(4) If I is an ideal of A, then $\mathrm{supp}(A/I) = V(I)$.
(5) If M is a finitely generated module over a noetherian ring M, then there exists a finite filtration $0 = M_0 \subset M_1 \subset \cdots \subset M_r = M$ ($r \geq 0$) by submodules such that $M_{i+1}/M_i \xrightarrow{\sim} A/Q_i$ for some $Q_i \in \mathrm{Spec}(A)$.
(6) In the situation of (5), $\mathrm{supp}(M) = \bigcup_{i=1}^r V(Q_i)$. [This is the beginning of the theory of primary decomposition; see [Re, §7] and [Ei, ch. 3] for more details.]

(10.19) Proposition. *Let A be a domain with fraction field K.*
(1) $\bigcap_{\mathfrak{m} \in \mathrm{Max}(A)} A_\mathfrak{m} = A$ (the intersection taken in K).
(2) If A_m is integrally closed for each $\mathfrak{m} \in \mathrm{Max}(A)$, then A is integrally closed.

Proof. (1) For $x \in K$ the set $I = \{a \in A \mid ax \in A\}$ is an ideal of A. Moreover, $x \in A_\mathfrak{m} \iff I \not\subset \mathfrak{m}$. In particular, if $x \in \bigcap_{\mathfrak{m} \in \mathrm{Max}(A)} A_\mathfrak{m}$, then $I = A$, which means that $x \in A$.
(2) If $x \in K$ is integral over A, then it is integral over each $A_\mathfrak{m}$, hence is contained in $\bigcap_{\mathfrak{m} \in \mathrm{Max}(A)} A_\mathfrak{m} = A$.

(10.20) Theorem. *If A is a domain of finite type over a field, then*

$$\forall P \in \mathrm{Spec}(A) \quad \dim(A) = \dim(A_P) + \dim(A/P).$$

Proof. Let $\mathfrak{m} \in \mathrm{Max}(A)$ be any maximal ideal containing P. There exists a saturated chain $(0) \subsetneq P_1 \subsetneq \cdots \subsetneq P_k = P$ of prime ideals of A between (0) and P of length $k = \dim(A_P)$ and a saturated chain of prime ideals $P = Q_0 \subsetneq \cdots \subsetneq Q_l = \mathfrak{m}$ between P and $\mathfrak{m}$ of length $l = \dim(A/P)$. Their concatenation $(0) \subsetneq P_1 \subsetneq \cdots \subsetneq P \subsetneq \cdots \subsetneq Q_l$ is a saturated chain of prime ideals between (0) and $\mathfrak{m}$ of length $k + l = \dim(A_P) + \dim(A/P)$. According to Theorem 9.16(3), the length of this chain is equal to $\dim(A)$.

(10.21) The general form of Krull's Principal Ideal Theorem (Hauptidealsatz) is the case $n = 1$ of the following statement ([Ei, Thm. 10.2], [M, Thm. 13.5], [De 2, Thm. 7.2]).

If A is a noetherian ring and $f_1, \ldots, f_n \in A$, then any $P \in \mathrm{Spec}(A)$ which is minimal among prime ideals containing $(f_1, \ldots, f_n)$ satisfies $\dim(A_P) \leq n$.

We have proved this result in the case when A is an algebra of finite type over a field, by Theorem 9.18 and Theorem 10.20 (one can replace A by its quotient by a minimal prime ideal, hence suppose that A is a domain).

(10.22) Theorem. *If A is a domain of finite type over a field, then A is a UFD $\iff$ every prime ideal $P \in \mathrm{Spec}(A)$ which is minimal among non-zero prime ideals ($\iff \dim(A_P) = 1$) is principal. [In fact, this result holds for arbitrary noetherian domains, but the proof requires Krull's Principal Ideal Theorem.]*

Proof. We know that a noetherian domain A is a UFD $\iff$ the principal ideal (f) generated by any irreducible element f of A is a prime ideal, by Corollary II.3.13. Assume that this condition holds. If $P \in \mathrm{Spec}(A)$ is minimal among non-zero prime ideals, take any $a \in P$, $a \neq 0$. The element a is not invertible in A, hence can be written as a non-trivial product $a = f_1 \cdots f_r$ of irreducible elements f_i ($r \geq 1$). The ideal P is prime, hence $f_i \in P$ for some i; then $(f_i) \subset P$ is a non-zero prime ideal, hence $(f_i) = P$, by minimality of P.

Conversely, assume that the condition in the statement of the theorem holds. If $f \in A$ is irreducible, let $P \in \mathrm{Spec}(A)$ be minimal among prime ideals containing f; then $\dim(A_P) = 1$, by 10.21 for $n = 1$ ($\dim(A_P) \neq 0$, since $P \neq (0)$). Our assumtion then says that $P = (g)$ is principal. The inclusion $(f) \subset (g)$ implies that $g \mid f$, hence $f = gu$ for some $u \in A^*$ (g is not invertible); therefore $(f) = (g)$ is a prime ideal.

11. Fibres of a morphism and theorems of Cohen-Seidenberg

(11.1) Given a ring homomorphism $\alpha : A \longrightarrow B$, what can be said about the fibres of the induced map $\alpha^* : \mathrm{Spec}(B) \longrightarrow \mathrm{Spec}(A)$?

For example, if A and B are algebras of finite type over a field K, then $\alpha^* : Q \mapsto \alpha^{-1}(Q)$ maps $\mathrm{Max}(B)$ to $\mathrm{Max}(A)$ (since $A/\alpha^{-1}(Q) \subset B/Q$ and B/Q is a finite field extension of K if $Q \in \mathrm{Max}(B)$, by the Nullstellensatz). If $K = \overline{K}$ is algebrically closed, then the map $\alpha^* : \mathrm{Max}(B) \longrightarrow \mathrm{Max}(A)$ is just the map between the classical points of the algebraic sets corresponding to B and A.

(11.2) Proposition. *Let $\alpha : A \longrightarrow B$ be a ring homomorphism. For any $P \in \mathrm{Spec}(A)$ there is a natural bijection between $(\alpha^*)^{-1}(P) \subset \mathrm{Spec}(B)$ and $\mathrm{Spec}(B_P/PB_P)$, where $B_P = S^{-1}B$ for $S = \alpha(A \setminus P)$ and $B_P/PB_P = B_P/\alpha(P)B_P$. [Exercise: this bijection is, in fact, a homeomorphism.]*

Proof. A prime ideal $Q \in \mathrm{Spec}(B)$ satisfies

$$\alpha^*(Q) \supset P \iff Q \supset \alpha(P)B \iff Q \in V(\alpha(P)B), \qquad \alpha^*(Q) \subset P \iff Q \subset \alpha(P) \iff Q \cap \alpha(A \setminus P) = \emptyset.$$

The statement then follows by applying 7.5(3) and 10.13.(7).

(11.3) From now on we concentrate on the case when $\alpha : A \longrightarrow B$ is injective, i.e., when $A \hookrightarrow B$ is a ring extension. In this case $\alpha^*(Q) = Q \cap A$ and $\mathrm{Im}(\alpha^*)$ is dense in $\mathrm{Spec}(A)$, by 7.5(4).

(11.4) Proposition. *Let $\alpha : A \hookrightarrow B$ be an integral ring extension.*
(1) $Q \in \mathrm{Spec}(B)$ lies in $\mathrm{Max}(B) \iff P = Q \cap A \in \mathrm{Max}(A)$.
(2) The map $\alpha^* : \mathrm{Spec}(B) \longrightarrow \mathrm{Spec}(A)$ is surjective.
(3) The map $\alpha^* : \mathrm{Spec}(B) \longrightarrow \mathrm{Spec}(A)$ is closed (i.e., the image of a closed set is closed).
(4) If $Q \subset Q'$ are prime ideals of B such that $Q \cap A = Q' \cap A = P$, then $Q = Q'$.
(5) Moreover, if $A \hookrightarrow B$ is a finite ring extension, then the map α^* has finite fibres.

Proof. (1) Apply Lemma 6.4 to the integral ring extension $A/P \hookrightarrow B/Q$.
(2) We can assume that $A \neq 0$. Fix $P \in \mathrm{Spec}(A)$ and consider the localised integral ring extension $\alpha_P : A_P \hookrightarrow B_P$, where $B_P = S^{-1}B$ for $S = A \setminus P$. There exists $\mathfrak{n} \in \mathrm{Max}(B_P)$ (since $A_P \neq 0$); then $\alpha_P^{-1}(\mathfrak{n}) = \mathfrak{n} \cap A_P \in \mathrm{Max}(A_P) = \{PA_P\}$, by (1); thus $\alpha_P^{-1}(\mathfrak{n}) = PA_P$. If we denote by $i : A \longrightarrow A_P$ (resp. $j : B \longrightarrow B_P$) the canonical maps, then the prime ideal $Q = j^{-1}(\mathfrak{n}) \in \mathrm{Spec}(B)$ satisfies $Q \cap A = \alpha^{-1}(j^{-1}(\mathfrak{n})) = (j \circ \alpha)^{-1}(\mathfrak{n}) = (\alpha_P \circ i)^{-1}(\mathfrak{n}) = i^{-1}(\alpha_P^{-1}(\mathfrak{n})) = i^{-1}(PA_P) = P$.
(3) Let $V(J) \xrightarrow{\sim} \mathrm{Spec}(B/J) \subset \mathrm{Spec}(B)$ be a closed subset (where J is an ideal of B). If $Q \supset J$ is a prime ideal of B, then $P = \alpha^*(Q) = Q \cap A \supset J \cap A = I$; thus $\alpha^*(V(J)) \subset V(I) \xrightarrow{\sim} \mathrm{Spec}(A/I)$. On the other hand, the statment (2) applied to the integral ring extension $A/I \hookrightarrow B/J$ tells us that the restriction of α^* to $\mathrm{Spec}(B/J) \longrightarrow \mathrm{Spec}(A/I)$ is surjective, hence $\alpha^*(V(J)) = V(I)$ is closed in $\mathrm{Spec}(A)$.

(4) As in the proof of (2) we localise at $S = A \setminus P$. We obtain an integral ring extension $A_P \hookrightarrow B_P$ and two prime ideals $QB_P \subset Q'B_P$ of B_P whose intersection with A_P is the *maximal* ideal PA_P of A_P. Applying (1), we deduce that both QB_P and $Q'B_P$ are maximal ideals of B_P. They must coincide, since one is contained in the other. The equality $QB_P = Q'B_P$ implies that $(Q'/Q)_P = 0$ (by Corollary 10.7), hence for each $b \in Q'$ there exists $s \in S$ (hence $s \notin Q$) such that $sb \in Q$; thus $b \in Q$, since Q is a prime ideal. This shows that $Q = Q'$.
(5) If $P \in \mathrm{Spec}(A)$ and B is finite over A, then B_P/PB_P is finite over the field $A_P/PA_P = \mathrm{Frac}(A/P)$, hence $(\alpha^*)^{-1}(P) \xrightarrow{\sim} \mathrm{Spec}(B_P/PB_P)$ is finite, by 8.12(iii).

(11.5) Theorem ("Going up"). *Let $A \hookrightarrow B$ be an integral ring extension. Given prime ideals $P_0 \subset P_1$ of A and $Q_0 \in \mathrm{Spec}(B)$ such that $Q_0 \cap A = P_0$, then there exists $Q_1 \in \mathrm{Spec}(B)$ such that $Q_0 \subset Q_1$ and $Q_1 \cap A = P_1$.*

Proof. Proposition 11.4(2) applied to the integral ring extension $A/P_0 \hookrightarrow B/Q_0$ tells us that there exists a prime ideal of B/Q_0, necessarily of the form Q_1/Q_0 for some $Q_1 \in \mathrm{Spec}(B)$ containing Q_0, such that $Q_1/Q_0 \cap A/P_0 = P_1/P_0$; thus $Q_1 \cap A = P_1$.

(11.6) Theorem. *If $0 \neq A \hookrightarrow B$ is an integral ring extension, then* $\dim(A) = \dim(B)$.

Proof. If $Q_0 \subsetneq Q_1 \subsetneq \cdots \subsetneq Q_r$ are distinct prime ideals of B, then $Q_0 \cap A \subsetneq Q_1 \cap A \subsetneq \cdots \subsetneq Q_r \cap A$ are distinct prime ideals of A, by Proposition 11.4(4); thus $\dim(B) \leq \dim(A)$. Conversely, given distinct prime ideals $P_0 \subsetneq P_1 \subsetneq \cdots \subsetneq P_r$ of A, we deduce from Proposition 11.4(2) and a repeated application of "going up" that there exist prime ideals $Q_0 \subset Q_1 \subset \cdots \subset Q_r$ of B satisfying $Q_i \cap A = P_i$ for all i. These prime ideals are necessarily distinct, hence $\dim(B) \geq \dim(A)$.

(11.7) Exercise. *Use Theorem 11.6 together with Noether's Normalisation Lemma to prove, inductively, that* $\dim(K[X_1, \ldots, X_n] \leq n$. *Deduce from this (and Noether's Normalisation Lemma) the statement 9.16(1). [However, this method is insufficient for proving 9.16(2). To do so, one must combine Noether's Normalisation Lemma with the following result.]*

(11.8) Theorem ("Going down"). *Let $A \hookrightarrow B$ be an integral extension of domains, with A integrally closed. Given prime ideals $P_0 \subset P_1$ of A and $Q_1 \in \mathrm{Spec}(B)$ such that $Q_1 \cap A = P_1$, then there exists $Q_0 \in \mathrm{Spec}(B)$ such that $Q_0 \subset Q_1$ and $Q_0 \cap A = P_0$.*

Proof. The proof is more involved that that of "going up". As we are not going to use this result, we refer the interested reader to [De 2, III.13] or [Ei, Thm. 13.9].

12. Algebraic and analytic local rings

(12.1) Let K be a field, let $Z \hookrightarrow \mathbf{A}^n_K$ be the zero locus of polynomials $f_1, \ldots, f_r \in K[X_1, \ldots, X_n]$:

$$Z : f_1 = \cdots = f_r = 0.$$

The ring of regular functions on Z

$$O(Z) = K[X_1, \ldots, X_n]/(f_1, \ldots, f_r)$$

contains information about global geometry of Z. What can we say about local geometry of Z around a given point?

In order to simplify the notation, write $B = K[X_1, \ldots, X_n]$, $J = (f_1, \ldots, f_r)$ and $A = O(Z) = B/J$. If we visualise Z as $\mathrm{Spec}(A)$, then a "point" means a prime ideal $P \in \mathrm{Spec}(A)$; it is of the form $P = Q/J$, where $Q \in \mathrm{Spec}(B)$ and $Q \supset J$.

According to (10.11.1) we can interpret the localisation A_P as the **local ring of Z at P** – its elements are represented by functions regular on a suitable open neighbourhood of P in the Zariski topology.

As localisation commutes with quotients, by Proposition 10.8, we have

$$A_P = B_Q/JB_Q.$$

In the special case when $P = \mathfrak{m} \in \mathrm{Max}(A)$ corresponds to a K-rational point $a = (a_1, \ldots, a_n) \in Z(K) \subset K^n$ of Z, then $Q = (X_1 - a_1, \ldots, X_n - a_n)$, $\mathfrak{m} = (\overline{X}_1 - a_1, \ldots, \overline{X}_n - a_n)$ and

$$A_{\mathfrak{m}} = K[X_1, \ldots, X_n]_{(X_1 - a_1, \ldots, X_n - a_n)}/(f_1, \ldots, f_r).$$

(12.2) Unfortunately, it turns out that the local rings $A_{\mathfrak{m}}$ are not local enough. The problem is that non-empty Zariski open sets are quite large (for example, on an irreducible curve they are just complements of finite sets of closed points). They do not allow us to "zoom in" on truly small neighbourhoods of $\mathfrak{m}$. Here is a typical example.

Let $char(K) \neq 2$,

$$Z : Y^2 - X^2(X+1) = 0, \qquad Z \hookrightarrow \mathbf{A}^2_K,$$

$a = O = (0,0) \in Z(K)$, $Q = (X, Y)$. The polynomial $Y^2 - X^2(X+1)$ is irreducible in both $B = K[X,Y]$ and $B_Q = K[X,Y]_{(X,Y)}$, which means that both rings

$$A = B/(f), \qquad A_m = B_Q/(f)$$

are integral domains. However, the curve Z has two branches at the singular point O. This can be seen by expanding everything in terms of power series in X and Y (the "local coordinates of $\mathbf{A}^2_K$ at O"). In other words, we replace the local ring B_Q of the plane at the origin by the power series ring

$$\widehat{B} = \varprojlim_k B/Q^k = \varprojlim_k K[X,Y]/(X,Y)^k = K[[X,Y]]. \tag{12.2.1}$$

Note that $B_Q = K[X,Y]_{(X,Y)}$ is a subring of $\widehat{B}$, since any power series (in particular, a polynomial) of the form $1 + h$ with $h \in Q$ has an inverse $(1+h)^{-1} = 1 - h + h^2 - h^3 + \cdots \in \widehat{B}$.

As observed in I.4.11, there exists $g \in 1 + XK[[X]]$ such that $g^2 = 1 + X$, which implies that

$$Y^2 - X^2(X+1) = (Y - Xg)(Y + Xg) \in K[[X,Y]]$$

is reducible in the power series ring $K[[X,Y]]$. The two local branches are given by $Y \pm Xg = 0$.

It is natural to interpret the ring

$$K[[X,Y]]/(f) = K[[X,Y]]/((Y - Xg)(Y + Xg)) \xrightarrow{\sim} K[[X,Y_1]]/((Y_1 - X)(Y_1 + X)) \qquad (Y = Y_1 g) \tag{12.2.2}$$

as the **analytic local ring of Z at $\mathfrak{m}$**. Note that, after a non-algebraic change of coordinates $Y = Y_1 g$, the two branches become linearised: the equation $(Y_1 - X)(Y_1 + X) = 0$ represents two lines intersecting transversally.

However, we must be careful in dealing with the ring (12.2.2). It is of the form $\widehat{B}/J\widehat{B}$, but it is not clear that it depends only on $A_{\mathfrak{m}}$. It is true that completions such as (12.2.1) commute with quotients (cf. 12.10 below), but the proof requires some work ([Ei, Thm. 7.2]). If we admit this result (which will be proved in a special case in Theorem 13.11(1) below), then we can, indeed, identify the ring $\widehat{B}/J\widehat{B}$ from (12.2.2) with the $\mathfrak{m}$-adic completion of A (or of $A_{\mathfrak{m}}$):

$$\widehat{B}/J\widehat{B} \xrightarrow{\sim} \widehat{A} = \varprojlim_k A/\mathfrak{m}^k = \varprojlim_k A_{\mathfrak{m}}/(\mathfrak{m}A_{\mathfrak{m}})^k. \tag{12.2.3}$$

These are special cases of I-adic completions, which generalise power series rings. We now prove a few basic results about such rings.

(12.3) Definition. *Let I be an ideal in a ring A, let M be an A-module. The* **I-adic completion of M** *is the projective limit (see Definition I.3.7)*

$$\widehat{M} = \varprojlim_k M/I^k M.$$

It is a module over the I-adic completion of A

$$\widehat{A} = \varprojlim_k A/I^k.$$

There are canonical homomorphisms $M \longrightarrow \widehat{M}$ and $A \longrightarrow \widehat{A}$.

(12.4) Examples. (i) $A = K[X_1, \ldots, X_n]$, $I = (X_1 - a_1, \ldots, X_n - a_n) = \mathfrak{m} \in \mathrm{Max}(K[X_1, \ldots, X_n])$, $\widehat{A} = K[[X_1 - a_1, \ldots, X_n - a_n]]$. Any power series $g \in A$ with non-zero constant term (in particular, any polynomial that does not vanish at a) is invertible in $\widehat{A}$, which implies that $A_{\mathfrak{m}}$ is naturally a subring of $\widehat{A}$.
(ii) $A = \mathbf{Z}$, $I = (p) = \mathfrak{m} \in \mathrm{Max}(\mathbf{Z})$ (p a prime number), $\widehat{A} = \mathbf{Z}_p$. Again, $A_{\mathfrak{m}} = \mathbf{Z}_{(p)}$ is a subring of $\widehat{A} = \mathbf{Z}_p$.
(iii) $A = K[X, Y]$, $I = (X) = P \in \mathrm{Spec}(A) \setminus \mathrm{Max}(A)$, $\widehat{A} = (K[Y])[[X]]$. However, A_P is not a subring of $\widehat{A}$ (for example, $1/Y \in A_P$).
(iv) $A = K[X, Y]_{(X)}$, $I = (X) = \mathfrak{m} \in \mathrm{Max}(A)$, $\widehat{A} = (K(Y))[[X]]$. In this case $A_{\mathfrak{m}} = A$ is a subring of $\widehat{A}$.
(v) $A = \mathbf{R}[X]$, $I = (X^2 + 1) = \mathfrak{m} \in \mathrm{Max}(A)$, $\widehat{A} \xrightarrow{\sim} \mathbf{C}[[Y]]$ (this example appeared in the midterm).

(12.5) Exercise. *If $\mathfrak{m} \in \mathrm{Max}(A)$, then $A/\mathfrak{m}^k = A_{\mathfrak{m}}/\mathfrak{m}^k A_{\mathfrak{m}} = A_{\mathfrak{m}}/(\mathfrak{m}A_{\mathfrak{m}})^k$ for all $k \geq 1$; therefore the $\mathfrak{m}$-adic completion $\widehat{A}$ of A coincides with the $\mathfrak{m}A_{\mathfrak{m}}$-adic completion of the local ring $A_{\mathfrak{m}}$. In particular, there is a canonical homomorphism $A_{\mathfrak{m}} \longrightarrow \widehat{A}$. [Hint: use exact sequences $0 \longrightarrow \mathfrak{m}^k/\mathfrak{m}^{k+1} \longrightarrow A/\mathfrak{m}^{k+1} \longrightarrow A/\mathfrak{m}^k \longrightarrow 0$ and induction on k.]*

(12.6) In examples 12.4(i),(ii),(iv),(v), the canonical homomorphism $A_{\mathfrak{m}} \longrightarrow \varprojlim A_{\mathfrak{m}}/(\mathfrak{m}A_{\mathfrak{m}})^k$ was injective. In general, we should think of the canonical homomorphism $A \longrightarrow \widehat{A}$ as a power series expansion. What can we say about functions with trivial expansion, namely, about

$$\mathrm{Ker}(A \longrightarrow \widehat{A}) = \bigcap_{k \geq 1} I^k \subset A \ ?$$

(12.7) Theorem (Krull). *If I is an ideal in a* **noetherian** *ring A, then*

$$\bigcap_{k \geq 1} I^k = (0) \iff \text{no element of } 1 + I \text{ is a divisor of zero in } A.$$

Proof. We use the notation $I^\infty := \bigcap_{k \geq 1} I^k$. If there exists $a \in I$ for which $1 - a$ is a divisor of zero, then $x = xa$ for some $x \neq 0$, which implies that $x = xa = xa^2 = \cdots \in I^\infty$.

Conversely, assume that no element of $1 + I$ is a divisor of zero. It is enough to show that each element $x \in I^\infty$ satisfies $x \in xI$, since an equality $x = xa$ with $a \in I$ yields $(1 - a)x = 0$, hence $x = 0$.

The ideal $I = (f_1, \ldots, f_r)$ is finitely generated, since A is noetherian. For each $k \geq 1$ there exists a homogeneous polynomial $P_k \in A[X_1, \ldots, X_r]$ of degree k such that $x = P_k(f_1, \ldots, f_r)$. The chain of ideals $(P_1) \subset (P_1, P_2) \cdots$ in the noetherian ring $A[X_1, \ldots, X_r]$ stabilises, which means that $P_{n+1} \in (P_1, \ldots, P_n)$ for some $n \geq 1$. Writing $P_{n+1} = Q_n P_1 + \cdots + Q_1 P_n$ with $Q_i \in A[X_1, \ldots, X_r]$ homogeneous of degree i, we obtain $x = x((Q_1 + \cdots + Q_n)(f_1, \ldots, f_r)) \in xI$.

(12.8) Corollary. *If A is a noetherian ring, then $\bigcap_{k \geq 1} I^k = (0)$ in each of the following cases.*
(1) $I \neq A$ and A is a domain.
(2) $I \subset \bigcap_{\mathfrak{m} \in \mathrm{Max}(A)} \mathfrak{m}$ (since $1 + I \subset A \setminus \bigcup_{\mathfrak{m}} \mathfrak{m} = A^*$).
(3) $I \neq A$ and A is a local ring (since $I \subset \mathfrak{m}$, $\mathrm{Max}(A) = \{\mathfrak{m}\}$).

(12.9) Proposition. *If $(A, \mathfrak{m})$ is a noetherian local ring and $J = (f_1, \ldots, f_r) \subset A$ is any ideal, then $\bigcap_{k \geq 1}(J + \mathfrak{m}^k) = J$ and the $\mathfrak{m}$-adic completion $\widehat{A} = \varprojlim A/\mathfrak{m}^k$ satisfies $J = A \cap J\widehat{A}$. In particular, A/J is naturally a subring of $\widehat{A}/J\widehat{A}$.*

Proof. There is nothing to prove if $J = A$, so we can assume that $J \subset \mathfrak{m}$. Corollary 12.8(3) applied to the local ring $(A/J, \mathfrak{m}/J)$ tells us that $\bigcap_{k \geq 1}((J + \mathfrak{m}^k)/J) = J/J$, but the L.H.S. contains $(\bigcap_{k \geq 1}(J + \mathfrak{m}^k))/J$,

which proves the non-trivial inclusion $\bigcap_{k\geq 1}(J+\mathfrak{m}^k)\subset J$. This implies that $A\cap J\widehat{A}=\{a\in A\mid a=\sum f_ig_i,\ g_i\in\widehat{A}\}\subset\bigcap_{k\geq 1}(J+\mathfrak{m}^k)=J$ (writing each $g_i\pmod{\mathfrak{m}^k}$ as $h_{i,k}\pmod{\mathfrak{m}^k}$ for some $h_{i,k}\in A$).

(12.10) If $J\neq A$ in 12.9, then $J\subset\mathfrak{m}$ and there is a canonical ring homomorphism

$$\widehat{A}/J\widehat{A}\longrightarrow(A/J)\widehat{\ }, \tag{12.10.1}$$

where $(A/J)\widehat{\ }$ denotes the $\mathfrak{m}$-adic (or $\mathfrak{m}/J$-adic, which amounts to the same) completion of A/J.

Another general result of Krull ([Ei, Thm. 7.2], [M, Thm. 8.11]) implies that (12.10.1) is an isomorphism, of which (12.2.3) (with A in 12.9 being equal to $K[X_1,\dots,X_n]_{(X_1-a_1,\dots,X_n-a_n)}$) is a special case. In Theorem 13.11(1) below we prove that (12.10.1) is an isomorphism in the "smooth" case.

13. Local rings in the smooth case and regularity

(13.1) Implicit function theorem in real (resp. complex) analysis. Recall that, if $f_1,\dots,f_r$ are C^∞ (resp. holomorphic) real-valued (resp. complex-valued) functions on some open set $U\subset\mathbf{R}^n$ (resp. $U\subset\mathbf{C}^n$) containing $a=(a_1,\dots,a_n)\in\mathbf{R}^n$ (resp. $a\in\mathbf{C}^n$) such that

$$\det\Big(\Big(\frac{\partial f_i}{\partial X_j}\Big)(a)\Big)_{1\leq i,j\leq r}\neq 0, \tag{13.1.1}$$

then there exists an open subset $U'\subset U$ containing a such that the projection

$$\mathrm{pr}: Z=\{x\in U\mid f_1(x)=\cdots=f_r(x)=0\}\longrightarrow\mathbf{R}^{n-r}\ \ (\text{resp. }\mathbf{C}^{n-r}),\qquad (x_1,\dots,x_n)\mapsto(x_{r+1},\dots,x_n)$$

is a diffeomorphism (resp. a biholomorphic map) between $Z\cap U'$ and the open neighbourhood $\mathrm{pr}(Z\cap U')$ of $(a_{r+1},\dots,a_n)\in\mathbf{R}^{n-r}$ (resp. in $\mathbf{C}^{n-r}$). In other words, $x_{r+1}-a_{r+1},\dots,x_n-a_n$ are "local coordinates on Z at a".

(13.2) We are going to study the condition (13.1.1) in the algebraic situation of 12.1, when K is a field,

$$f_1,\dots,f_r\in K[X_1,\dots,X_n]=B,\qquad Z: f_1=\cdots=f_r=0,\qquad Z\hookrightarrow\mathbf{A}^n_K,$$
$$A=O(Z)=B/J,\qquad J=(f_1,\dots,f_r).$$

According to Theorem 9.18, the dimension of each irreducible component $V(P_i)$ (where P_i is a minimal prime ideal of A) of $\mathrm{Spec}(A)$ (= of Z) satisfies

$$\dim(A/P_i)\geq n-r. \tag{13.2.1}$$

(13.3) "Definition". *We say that $P\in\mathrm{Spec}(A)$* **is a smooth point of Z over K** *if, after renumbering of the coordinates $X_1,\dots,X_n$,*

$$\det\Big(\Big(\frac{\partial f_i}{\partial X_j}\Big)(Q)\Big)_{1\leq i,j\leq r}\neq 0\in\mathrm{Frac}(B/Q)=\mathrm{Frac}(A/P), \tag{13.3.1}$$

where $P=Q/J$, $Q\in\mathrm{Spec}(B)$. In particular, if $P'\subset P$ and P is a smooth point of Z over K, so is P'. [The attentive reader will have noticed that this "definition" is somewhat horrible, as it depends on the presentation of A as $A=B/J$ and on the choice of a set of generators $f_1,\dots,f_r$ of the ideal J. The true definition is suggested by the statement of Proposition 13.11 below.]

(13.4) Example. The simplest case is that of a plane curve

$$Z: f(X,Y)=0,\qquad Z\hookrightarrow\mathbf{A}^2_K$$

smooth at the origin $O=(0,0)$. In other words,

$$Q = (X, Y), \qquad f \in K[X, Y], \qquad f(0,0) = 0, \qquad \frac{\partial f}{\partial X}(0,0) \neq 0.$$

After multiplying f by a non-zero constant in K^* we can assume that

$$f = X + cY + \sum_{i+j \geq 2} c_{i,j} X^i Y^j.$$

An easy division algorithm for power series shows that

$$\forall g \in K[[X, Y]] \; \exists h \in K[[X, Y]] \quad g - fh \in K[[Y]];$$

moreover, h is unique. In algebraic terms, the composite map

$$K[[Y]] \hookrightarrow K[[X, Y]] \longrightarrow K[[X, Y]]/(f)$$

is an isomorphism of K-algebras. This a formal power series counterpart of 13.1: from an analytic point of view, Z is isomorphic around O to a line with the coordinate Y. A general statement of this kind will be proved in Theorem 13.11(1) below.

(13.5) Proposition (Smoothness at K-rational points and tangent spaces). *Let $Q = (X_1 - a_1, \ldots, X_n - a_n) \in \operatorname{Max}(B)$, where $a = (a_1, \ldots, a_n) \in Z(K) \subset K^n$, $P = Q/J = \mathfrak{m} = (\overline{X}_1 - a_1, \ldots, \overline{X}_n - a_n) \in \operatorname{Max}(A)$ (then $A/\mathfrak{m} = B/Q = K$).*
(1) The dual of the tangent space $T_a Z$ to Z at the point a (see I.4.6) is naturally isomorphic to the K-vector space

$$Q/(Q^2 + J) = \mathfrak{m}/\mathfrak{m}^2 = \mathfrak{m}A_{\mathfrak{m}}/\mathfrak{m}^2 A_{\mathfrak{m}} = \mathfrak{m}A_{\mathfrak{m}}/(\mathfrak{m}A_{\mathfrak{m}})^2.$$

(2) $\dim_K(T_a Z) \geq n - r$, with equality $\iff$ $\mathfrak{m}$ is a smooth point of Z over K.
(3) If (13.3.1) (= (13.1.1)) holds, then the elements $\overline{X}_{r+1} - a_{r+1}, \ldots, \overline{X}_n - a_n \pmod{\mathfrak{m}^2}$ form a basis of the K-vector space $\mathfrak{m}/\mathfrak{m}^2$ and $\mathfrak{m}A_{\mathfrak{m}} = (\overline{X}_{r+1} - a_{r+1}, \ldots, \overline{X}_n - a_n)$. [This is an algebraic version of the statement that $X_{r+1} - a_{r+1}, \ldots, X_n - a_n$ are "local coordinates" on Z at a.]

Proof. (1) By definition,

$$T_a Z = \{y \in K^n \mid My = 0\}, \qquad M = \left(\frac{\partial f_i}{\partial X_j}(a)\right) \in M_{r \times n}(K)$$

(we consider y as a column vector). The entries of the matrix M can be expressed purely in terms of $Q \supset J$, as follows.

$$Q/Q^2 = \bigoplus_{j=1}^{n} K \cdot (X_j - a_j) \pmod{Q^2}, \qquad f_i \pmod{Q^2} = \sum_{j=1}^{n} \frac{\partial f_i}{\partial X_j}(a) \; (X_j - a_j) \pmod{Q^2}.$$

If we identify $\{(X_j - a_j) \pmod{Q^2}\}$ with the dual basis of the standard basis of K^n, then we obtain that $T_a Z \subset T_a \mathbf{A}^n_K = K^n = \operatorname{Hom}_K(Q/Q^2, K)$ is equal to

$$T_a Z = \operatorname{Ker}(\operatorname{Hom}_K(Q/Q^2, K) \longrightarrow \operatorname{Hom}_K(J/Q^2, K)) = \operatorname{Hom}_K(Q/(Q^2 + J), K).$$

Alternatively, one can use the abstract description of $T_a Z$ given in I.4.6. The point $a \in Z(K)$ corresponds to $\lambda \in \operatorname{Hom}_{K-\mathrm{Alg}}(A, K)$ and the tangent space $T_a Z$ to the set

$$\{\widetilde{\lambda} \in \operatorname{Hom}_{K-\mathrm{Alg}}(A, K[\varepsilon]) \mid \mathrm{pr} \circ \widetilde{\lambda} = \lambda\},$$

where $\mathrm{pr} : K[\varepsilon] = K + K\varepsilon \longrightarrow K$ is the canonical projection. In other words, $\widetilde{\lambda}(c + x) = c + \mu(x)\varepsilon$ for all $c \in K$ and $x \in \mathfrak{m}$, for a certain map $\mu : \mathfrak{m} \longrightarrow K$. The fact that $\widetilde{\lambda}$ commutes with products is equivalent to

$$\forall\, c_i \in K \;\; \forall\, x_i \in \mathfrak{m} \quad \mu(c_1 x_1 + c_2 x_2 + x_1 x_2) = c_1 \mu(x_1) + c_2 \mu(x_2). \tag{13.5.1}$$

This condition implies that $\mu(c_1x_1) = c_1\mu(x_1)$, $\mu(x_1x_2) = 0$ and $\mu(x_1(1+x_2)) = \mu(x_1)$. The universal property 10.2(3) implies that $\widetilde{\lambda}$ canonically extends to a homomorphism of K-algebras $A_{\mathfrak{m}} \longrightarrow K[\varepsilon]$. We can replace, therefore, A by $A_{\mathfrak{m}}$. In this case $1 + \mathfrak{m}A_{\mathfrak{m}} \subset A_{\mathfrak{m}}^*$ and the condition (13.5.1) implies that, for all $x_1, x_2 \in \mathfrak{m}$,

$$\mu(x_1) + \mu(x_2) = \mu(x_1) + \mu(x_2(1+y)) = \mu(x_1 + x_2(1+y) + x_1x_2(1+y)) = \mu(x_1 + x_2),$$

if we let $y = (1+x_1)^{-1} - 1 \in \mathfrak{m}$. It follows that μ is of the form $\mu = \widetilde{\mu} \circ \mathrm{pr}$, where $\mathrm{pr} : \mathfrak{m}A_{\mathfrak{m}} \longrightarrow \mathfrak{m}A_{\mathfrak{m}}/(\mathfrak{m}A_{\mathfrak{m}})^2$ and $\widetilde{\mu} : \mathfrak{m}A_{\mathfrak{m}}/(\mathfrak{m}A_{\mathfrak{m}})^2 = \mathfrak{m}/\mathfrak{m}^2 \longrightarrow K$ is K-linear. Conversely, any such μ satisfies (13.5.1).
(2) $\dim_K(T_aZ) = n - \mathrm{rk}(M) \geq n - r$, with equality $\iff$ $\mathfrak{m}$ is a smooth point of Z over K.
(3) By assumption, $M = (M_1|M_2)$, where $M_1 \in GL_r(K)$ and $M_2 = M_{r\times(n-r)}(K)$, which implies that the projection

$$T_aZ \subset K^r \times K^{n-r} \longrightarrow K^{n-r}$$

on the last $n - r$ coordinates is an isomorphism. This is equivalent, by (1), to the fact that $\overline{X}_{r+1} - a_{r+1}, \ldots, \overline{X}_n - a_n \pmod{\mathfrak{m}^2}$ form a basis of the K-vector space $\mathfrak{m}/\mathfrak{m}^2 = \mathfrak{m}A_{\mathfrak{m}}/(\mathfrak{m}A_{\mathfrak{m}})^2$. The remaining statement follows from Corollary 13.7 below applied to $A_{\mathfrak{m}}$ and $N = \mathfrak{m}A_{\mathfrak{m}}$.

(13.6) Nakayama's Lemma. *Let A be a ring, let $J \subset A$ be an ideal contained in all maximal ideals of A. If M is a finitely generated A-module satisfying $M = JM$, then $M = 0$.*

Proof. Let $M = Am_1 + \cdots + Am_n$. By assumption, there exists a matrix $U \in M_n(J)$ such that

$$\begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix} = U \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix} \in M^n,$$

which implies that

$$\det(I-U) \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix} = \mathrm{adj}(I-U)(I-U) \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix} = 0 \in M^n,$$

hence $\det(I-U) \cdot m = 0$ for all $m \in M$. However,

$$\det(I-U) \in 1 + J \subset A \setminus \bigcup_{\mathfrak{m} \in \mathrm{Max}(A)} \mathfrak{m} = A^*;$$

therefore $m = 0$ for all $m \in M$.

(13.7) Corollary. *If $(A, \mathfrak{m})$ is a local ring, N is a finitely generated A-module and $n_1, \ldots, n_r \in N$ are elements whose images $\overline{n}_1, \ldots, \overline{n}_r \in N/\mathfrak{m}N$ generate $N/\mathfrak{m}N$ as a vector space over $A/\mathfrak{m}$, then $n_1, \ldots, n_r$ generate N as an A-module.*

Proof. $M = N/(An_1 + \cdots + An_r)$ is a finitely generated A-module satisfying $M/\mathfrak{m}M = 0$; thus $M = 0$ and $N = An_1 + \cdots + An_r$.

(13.8) Dimension of local noetherian rings. The first definition of dimension $\dim_1$ (= the number of independent parameters) has the following local variant.

Let $(A, \mathfrak{m})$ be a local noetherian ring. For every $N \geq 1$, the quotient ring $A/\mathfrak{m}^N$ is artinian, hence of dimension zero. It is natural to consider, therefore, the following integer:

$$\dim_3(A) = \min\{k \geq 0 \mid \exists\, x_1, \ldots, x_k \in \mathfrak{m} \;\; \exists\, N \geq 1 \quad (x_1, \ldots, x_k) \supset \mathfrak{m}^N\} \in \mathbf{N}$$

(one should think of $x_1, \ldots, x_k$ as a system of "weak local coordinates"). Note that $\dim_3(A)$ is bounded above by the minimal number of generators of the ideal $\mathfrak{m}$, which is equal to $\dim_{A/\mathfrak{m}}(\mathfrak{m}/\mathfrak{m}^2)$, by Corollary 13.7. Therefore

$$\dim_3(A) \leq \dim_{A/\mathfrak{m}}(\mathfrak{m}/\mathfrak{m}^2). \tag{13.8.1}$$

One of the main results of dimension theory states that

$$\dim_3(A) = \dim(A), \tag{13.8.2}$$

which is, essentially, a reformulation of Krull's Principal Ideal Theorem in its general form 10.21; see [Ei, Cor. 10.7] (in particular, all local noetherian rings have finite dimension!). As a result,

$$\dim(A) \leq \dim_{A/\mathfrak{m}}(\mathfrak{m}/\mathfrak{m}^2). \tag{13.8.3}$$

In the situation considered in Proposition 13.5(1), (13.8.3) reads as

$$\dim(A_{\mathfrak{m}}) \leq \dim_K(T_a Z), \tag{13.8.4}$$

with $\dim(A_{\mathfrak{m}})$ being equal to the maximum of dimensions of all irreducible components of $\mathrm{Spec}(A)$ (= of Z) containing $\mathfrak{m}$ (= a). Recall that (13.2.1) tells us that

$$n - r \leq \dim(A_{\mathfrak{m}}). \tag{13.8.5}$$

In particular, if we admit (13.8.2), then (13.8.4) becomes an equality $\iff$ $\mathfrak{m}$ is a smooth point of $\mathrm{Spec}(A)$. This observation leads naturally to the following definition.

(13.9) Definition. *A noetherian local ring $(A, \mathfrak{m})$ is* **regular** *if* $\dim(A) = \dim_{A/\mathfrak{m}}(\mathfrak{m}/\mathfrak{m}^2)$.

(13.10) As observed in 13.8, smoothness of a K-rational point is equivalent to regularity of the corresponding local ring, provided we admit the fundamental equality $\dim_3 = \dim$. We are now going to relate smoothness to regularity using power series expansions and completions. This methods yields directly the equality $\dim_3(A_{\mathfrak{m}}) = \dim(A_{\mathfrak{m}})$ and the isomorphism (12.10.1) at smooth K-rational points.

(13.11) Theorem. *Let $Q = (X_1 - a_1, \ldots, X_n - a_n) \in \mathrm{Max}(B)$, where $a = (a_1, \ldots, a_n) \in Z(K) \subset K^n$, $P = Q/J = \mathfrak{m} = (\overline{X}_1 - a_1, \ldots, \overline{X}_n - a_n) \in \mathrm{Max}(A)$. Assume that (13.3.1) (= (13.1.1)) holds.*
(1) (Formal implicit function theorem) The canonical maps $\alpha : K[[\overline{X}_{r+1} - a_{r+1}, \ldots, \overline{X}_n - a_n]] \longrightarrow \widehat{B}/J\widehat{B}$, where $\widehat{B} = K[[X_1 - a_1, \ldots, X_n - a_n]] = \varprojlim_k B/Q^k = \varprojlim_k B_Q/Q^k B_Q$, and $\beta : \widehat{B}/J\widehat{B} \longrightarrow \widehat{A}$, where $\widehat{A} = \varprojlim_k A/\mathfrak{m}^k = \varprojlim_k A_{\mathfrak{m}}/\mathfrak{m}^k A_{\mathfrak{m}}$, are isomorphisms of K-algebras.
(2) The ring $A_{\mathfrak{m}}$ is a domain.
(3) $\mathfrak{m}A_{\mathfrak{m}} = (\overline{X}_{r+1} - a_{r+1}, \ldots, \overline{X}_n - a_n)$.
(4) $\mathrm{Spec}(A)$ (= Z) has a unique irreducible component containing $\mathfrak{m}$ (= a); its dimension is $\dim(A_{\mathfrak{m}}) = n - r$.
(5) $A_{\mathfrak{m}}$ is a regular local ring.
(6) $A_{\mathfrak{m}}$ is a UFD. [In fact, one can show that every regular local ring is a UFD.]

Proof. (1) After a linear change of variables $X_1, \ldots, X_r$ we can assume that $\partial f_i/\partial X_j(a) = \delta_{ij}$ $(1 \leq i, j \leq r)$. This means that the power series expansion of $f_i - (X_i - a_i)$ in $K[[X_1 - a_1, \ldots, X_n - a_n]]$ contains only linear terms involving $X_j - a_j$ for $j > r$ and terms of degree ≥ 2. As in 13.4, an easy division algorithm shows that every $g \in K[[X_1 - a_1, \ldots, X_n - a_n]]$ can be written in a unique way as $g = f_1 h + g_1$, where $g_1 \in K[[X_2 - a_2, \ldots, X_n - a_n]]$. By induction, this implies that α is an isomorphism. The statement about β follows from the fact that $A/\mathfrak{m}^k = B/(Q^k + J) = \widehat{B}/(Q^k\widehat{B} + J\widehat{B})$, which is isomorphic (via α) to $K[[\overline{X}_{r+1} - a_{r+1}, \ldots, \overline{X}_n - a_n]]/(\overline{X}_{r+1} - a_{r+1}, \ldots, \overline{X}_n - a_n)^k$.
(2) Proposition 12.9 applied to the local ring (B_Q, Q) and the ideal JB_Q tells us that $A_{\mathfrak{m}} = B_Q/JB_Q$ is a subring of $\widehat{B}/J\widehat{B}$, which is a domain, by (1).

(3) The two ideals $I = (\overline{X}_{r+1} - a_{r+1}, \ldots, \overline{X}_n - a_n) \subset (\overline{X}_1 - a_1, \ldots, \overline{X}_n - a_n) = \mathfrak{m}A_{\mathfrak{m}} \subset A_{\mathfrak{m}}$ generate the same ideals of $\widehat{B}/J\widehat{B}$, since α is an isomorphism. Applying β, we obtain that $I\widehat{A} = (\mathfrak{m}A_{\mathfrak{m}})\widehat{A}$, hence $I = \mathfrak{m}A_{\mathfrak{m}}$, by Proposition 12.9 applied to $A_{\mathfrak{m}}$.
(4) Uniqueness follows from (2). The arguments in (1) and (2) show that, for each $i \leq r$,

$$B_i = B_Q/(f_1, \ldots, f_i)B_Q \hookrightarrow \widehat{B}/(f_1, \ldots, f_i)\widehat{B} \xrightarrow{\sim} K[[\overline{X}_{i+1} - a_{i+1}, \ldots, \overline{X}_n - a_n]]$$

is a domain and the class of f_{i+1} in B_i (if $i < r$) maps to $\overline{X}_{i+1} - a_{i+1}$, which implies that it is neither zero nor invertible in B_i. Therefore $\dim(B_{i+1}) = \dim(B_i/(f_{i+1})) = \dim(B_i) - 1$, by Theorem 9.9 (strictly speaking, we should work with the rings $B_i' = B[1/h]/(f_1, \ldots, f_r)B[1/h]$ of finite type over K, where $h \in B$, $h(a) \neq 0$ is chosen in such a way that $V(h)$ contains all irreducible components of Z not containing a; then B_i' injects into B_i), hence $\dim(A_m) = \dim(B_r) = \dim(B_0) - r = n - r$.
(5) Combine (4) with Proposition 13.5(2).
(6) The $\mathfrak{m}A_{\mathfrak{m}}$-adic completion $\widehat{A}$ of $A_{\mathfrak{m}}$ is isomorphic to a power series ring over K, which is a UFD [ZS 2, VII.1, Thm. 6]. To show that $A_{\mathfrak{m}}$ itself is a UFD one must check that the condition I.5.7(ii') for $\widehat{A}$ implies the same condition for $A_{\mathfrak{m}}$; see the arguments in ([Mu 1, III.7, Lemma 2] and [Mu 2, Thm. 1.28]).

(13.12) The UFD property. (i) The statement of Theorem 13.11(6) can be reformulated geometrically, using Theorem 10.20, as follows: any closed irreducible subset $V(\widetilde{P}) \xrightarrow{\sim} \mathrm{Spec}(A/\widetilde{P}) \subset \mathrm{Spec}(A)$ of dimension $\dim(A/\widetilde{P}) = \dim(A_{\mathfrak{m}}) - 1$ containing $\mathfrak{m}$ is given by a single equation $g = 0$, on a suitable open neighbourhood of $\mathfrak{m} \in \mathrm{Spec}(A)$ in the Zariski topology. Indeed, $\dim(A_{\widetilde{P}}) = 1$, hence $\widetilde{P}A_{\mathfrak{m}} = gA_{\mathfrak{m}}$ for some $g \in A_{\mathfrak{m}}$, which gives already $\widetilde{P}A[1/h] = gA[1/h]$ for some $h \in A \setminus \mathfrak{m}$ and $g \in A[1/h]$.
(ii) This property can sometimes hold even in the non-smooth case. The following examples are discussed in detail in [Mu 1, III.7, Ex. J]. Let K be a field of characteristic $char(K) \neq 2$, let $f(X_1, \ldots, X_n)$ be a non-degenerate quadratic form over K in $n \geq 3$ variables. Then f is irreducible and the quadratic cone

$$Z : f = 0, \qquad Z \subset \mathbf{A}^n_K$$

has ring of functions $A = O(Z) = K[X_1, \ldots, X_n]/(f)$, which is a domain. The origin $O = (0, \ldots, 0) \in Z(K)$ (which corresponds to $Q = (X_1, \ldots, X_n) \in \mathrm{Max}(K[X_1, \ldots, X_n])$ and $\mathfrak{m} = Q/(f) \in \mathrm{Max}(A)$) is not a smooth point of Z over K.
(iii) If $n = 3$ and $f = X_1^2 + X_2^2 - X_3^2$, then the factorisations $\overline{X}_1 \cdot \overline{X}_1 = (\overline{X}_3 + \overline{X}_2)(\overline{X}_3 - \overline{X}_2) \in A$ show that A (in fact, even $A_{\mathfrak{m}}$) is not a UFD. Geometrically, this means that the prime ideal $\widetilde{P} = (\overline{X}_1, \overline{X}_3 - \overline{X}_2)$ defining a line L on Z passing through O cannot be defined by a single equation (the equation $\overline{X}_3 - \overline{X}_2 = 0$ defines a double line, since $A/(\overline{X}_3 - \overline{X}_2) \xrightarrow{\sim} K[X_1, X_2]/(X_1^2)$; geometrically, the plane $H : X_3 - X_2 = 0$ is tangent to Z and the intersection $Z \cap H$ is the line L taken with multiplicity two).
(iv) Similarly, if $n = 4$ and $f = X_1^2 + X_2^2 - X_3^2 - X_4^2$, then the factorisations $(\overline{X}_1 + \overline{X}_4)(\overline{X}_1 - \overline{X}_4) = (\overline{X}_3 + \overline{X}_2)(\overline{X}_3 - \overline{X}_2) \in A$ show that A (in fact, even $A_{\mathfrak{m}}$) is not a UFD. The ideal $(\overline{X}_1 - \overline{X}_4, \overline{X}_3 - \overline{X}_2)$ defines a plane lying on Z which cannot be defined (locally around O) by a single equation.
(v) By contrast, A is a UFD for $n \geq 5$ ([Mu 1, III.7, Ex. J]).

(13.13) Proposition. *If $\mathfrak{m} \in \mathrm{Max}(A)$ corresponds to a K-rational point $a \in Z(K)$ and if $A_{\mathfrak{m}}$ is regular, then $\mathfrak{m}$ is a smooth point of Z over K in the following sense: there exist $g_1, \ldots, g_s \in J$ and $h \in B$ such that $h(a) \neq 0$ ($\iff h \notin \mathfrak{m}$), $JB[1/h] = (g_1, \ldots, g_s)B[1/h]$ and, after renumbering of the coordinates $X_1, \ldots, X_n$,*

$$\det\Big(\Big(\frac{\partial g_i}{\partial X_j}\Big)(a)\Big)_{1 \leq i,j \leq s} \neq 0. \tag{13.13.1}$$

[Geometrically, the equality $JB[1/h] = (g_1, \ldots, g_s)B[1/h]$ means that Z is given on the Zariski open neighbourhood $D(h) \subset \mathrm{Spec}(B) = \mathbf{A}^n_K$ of Q by the set of equations $g_1 = \cdots = g_s = 0$.]

Proof. Choose a minimal set $g_1, \ldots, g_s \in J \subset Q$ whose linear terms define the tangent space T_aZ; then

$$n - s = \dim_K(T_aZ) = \dim(A_{\mathfrak{m}}),$$

since $A_{\mathfrak{m}}$ is regular. Consider the ideal $J' = (g_1, \ldots, g_s) \subset J$ and the ring $A' = B/J' \supset \mathfrak{m}' = Q/J'$. By construction, $\mathfrak{m}'$ is a smooth point of $Z' : g_1 = \cdots = g_s = 0$, which yields (13.13.1), after renumbering of the coordinates.

In Spec(B), we have $V(J) \subset V(J')$. Geometrically, $V(J')$ contains a unique irreducible component containing Q, by Theorem 13.11(2); it has dimension $n - s$, by Theorem 13.11(4). As $\dim(A_{\mathfrak{m}}) = n - s$, $V(J)$ also contains an irreducible component containing Q of dimension $n-s$. This suggests that $V(J) \cap U = V(J') \cap U$ for some open neighbourhood of $\mathfrak{m} \in \mathrm{Spec}(A)$.

We can prove a more precise property algebraically. The previous statements can be rephrased by saying that $A'_{\mathfrak{m}'}$ is a domain and that there exists a minimal prime ideal $\widetilde{P} \subset A_{\mathfrak{m}}$ such that

$$\dim(A_{\mathfrak{m}}/\widetilde{P}) = n - s = \dim(A'_{\mathfrak{m}'}). \tag{13.13.2}$$

Consider the surjection $\alpha : A'_{\mathfrak{m}'} \longrightarrow A_{\mathfrak{m}}/\widetilde{P}$ between two domains of the same dimension. If α is not an isomorphism, then any chain of prime ideals $P_0 \subsetneq \cdots \subsetneq P_r$ in $A_{\mathfrak{m}}$ gives rise to a longer chain $(0) \subsetneq \mathrm{Ker}(\alpha) \subset \alpha^{-1}(P_0) \subsetneq \cdots \subsetneq \alpha^{-1}(P_r)$ in $A'_{\mathfrak{m}'}$, which contradicts (13.2.2). It follows that α is an isomorphism, which implies that $\alpha' : A'_{\mathfrak{m}'} \longrightarrow A_{\mathfrak{m}}$ and $A_{\mathfrak{m}} \longrightarrow A_{\mathfrak{m}}/\widetilde{P}$ are also isomorphisms (in particular, $A_{\mathfrak{m}}$ is a domain, which means that Z has a unique irreducible component containing $\mathfrak{m}$). The fact that α' is an isomorphism is equivalent to $JB_Q = (g_1, \ldots, g_s)B_Q$, which implies that $JB[1/h] = (g_1, \ldots, g_s)B[1/h]$ for some $h \notin \mathfrak{m}$, since both ideals J and $(g_1, \ldots, g_s)$ are finitely generated.

(13.14) Smoothness = regularity. We can sum up Theorem 13.11(5) and Proposition 13.13 by saying that

$$\forall \mathfrak{m} \in \mathrm{Max}(A) \text{ such that } A/\mathfrak{m} = K \qquad \mathfrak{m} \text{ is a smooth point of } Z \text{ over } K \iff A_{\mathfrak{m}} \text{ is a regular local ring.}$$

This equivalence still holds if $A/\mathfrak{m}$ is a separable extension of K (in particular, for all $\mathfrak{m} \in \mathrm{Max}(A)$ if K is a perfect field).

If $A/\mathfrak{m}$ is an inseparable extension of K, then smoothness still implies regularity, but the converse need not hold. Here is a simple example.

(13.15) Smoothness $\neq$ regularity. Let K be a non-perfect field of $char(K) = 3$, let $c \in K \setminus K^3$. Consider the curve

$$Z : Y^2 - (X^3 - c) = 0, \qquad Z \hookrightarrow \mathbf{A}^2_K$$

with $A = O(Z) = K[X, Y]/(Y^2 - (X^3 - c))$. The extension $L = K(\gamma)$, where $\gamma^3 = c$, is a purely inseparable cubic extension of K. The point $a = (\gamma, 0) \in Z(L)$ gives rise to maximal ideals $Q' = (X - \gamma, Y) \in \mathrm{Max}(L[X, Y])$, $Q = Q' \cap K[X, Y] = (X^3 - c, Y) \in \mathrm{Max}(K[X, Y])$ and $\mathfrak{m} = Q/(Y^2 - (X^3 - c)) = (\overline{X}^3 - c, \overline{Y}) = (\overline{Y}^2, \overline{Y}) = (\overline{Y}) \in \mathrm{Max}(A)$. In particular, $\mathfrak{m}A_{\mathfrak{m}} = (\overline{Y})$ is generated by one element, which means that $A_{\mathfrak{m}}$ is a regular local ring of dimension one. Its residue field is $A/\mathfrak{m} = K[X, Y]/(Y, Y^2 - (X^3 - c)) = K[X]/(X^3 - c) = L$.

On the other hand, the polynomial $f = Y^2 - (X^3 - c)$ satisfies

$$\frac{\partial f}{\partial X}(a) = \frac{\partial f}{\partial Y}(a) = 0,$$

which means that $\mathfrak{m}$ is not a smooth point of Z over K. In fact, the same curve over L becomes

$$Z_L : Y^2 - (X - \gamma)^3 = 0,$$

as in 5.13. The localisation of $O(Z_L) = L[X, Y]/(Y^2 - (X - \gamma)^3)$ at $(\overline{X} - \alpha, \overline{Y})$ is not regular.

(13.16) Smoothness and dimension of intersections. The existence of the "local coordinates" $\overline{X}_i - a_i$ $(r < i \leq n)$ on Z around a smooth K-rational point $a \in Z(K)$ established in Proposition 13.5 and Theorem 13.11 has the following important consequence ([Mu 1, III.6, Prop. 4], [Mu 2, Prop. 3.28]).

If we embed Z into $Z \times Z \subset \mathbf{A}^n_K \times \mathbf{A}^n_K$ diagonally ($\Delta : b \mapsto (b,b)$) and if we denote by X_i (resp. by X'_i) the coordinates on the first (resp. on the second) copy of $\mathbf{A}^n_K$, then $\Delta(Z) \subset Z \times Z$ is given, in a suitable Zariski open neighbourhood of $\Delta(\mathfrak{m})$, by $n-r$ equations

$$\overline{X}_i - \overline{X}'_i = 0, \qquad r < i \leq n.$$

If $J_1, J_2 \supset J$ are ideals of $K[X_1, \ldots, X_n]$ containing Q, they define $Z_1 \subset Z$ and $Z_2 \subset Z$ containing $\mathfrak{m}$. Their intersection $Z_1 \cap Z_2 \subset Z$ is equal to $(Z_1 \times Z_2) \cap \Delta(Z)$. As a result, any irreducible component Y of $Z_1 \cap Z_2$ containing $\mathfrak{m}$ is obtained (in a Zariski open neighbourhood U of $\mathfrak{m}$) from $Z_1 \times Z_2$ by imposing $n-r$ equations. Theorem 9.18 applied to $A = O(Z_1 \times Z_2)/P_0$, for any minimal prime ideal P_0, then yields

$$\dim(Y) \geq \dim(Z_1) + \dim(Z_2) - (n-r) = \dim(Z_1) + \dim(Z_2) - \dim(Z) \tag{13.16.1}$$

(we can shrink U so that Z becomes irreducible).

The inequality need not hold if $\mathfrak{m}$ is not smooth. The following standard example ([Mu 1, III.6, Ex. I]) coincides with the quadratic cone from 13.12(iv), after an obvious linear change of coordinates. Let $Z_1, Z_2 \hookrightarrow Z \hookrightarrow \mathbf{A}^4_K$ be given by

$$Z : X_0X_3 - X_1X_2 = 0, \qquad Z_1 : X_0 = X_1 = 0, \qquad Z_2 : X_2 = X_3 = 0.$$

Then $Y = Z_1 \cap Z_2$ is the non-smooth point $O = \{0,0,0,0\}$ of Z and

$$\dim(Z_i) = 2, \qquad \dim(Z) = 3, \qquad \dim(Y) = 0. \tag{13.16.2}$$

A proper understanding of this example involves projective geometry. If we consider $(X_0 : X_1 : X_2 : X_3)$ as homogeneous coordinates in the projective space $\mathbf{P}^3_K$, then $\widetilde{Z}$ (= the image of $Z \setminus \{O\}$ in $\mathbf{P}^3_K$) is a projective quadratic surface

$$\widetilde{Z} : X_0X_3 - X_1X_2 = 0, \qquad \widetilde{Z} \hookrightarrow \mathbf{P}^3_K$$

and each $\widetilde{Z}_i$ (= the image of $Z_i \setminus \{O\}$ in $\mathbf{P}^3_K$) is a projective line on the quadric $\widetilde{Z}$. Their intersection $\widetilde{Z}_1 \cap \widetilde{Z}_2 = \emptyset$ is empty.

In fact $\widetilde{Z}$ is isomorphic to the product $\mathbf{P}^1_K \times \mathbf{P}^1_K$ of two projective lines via the Segre embedding

$$\mathbf{P}^1_K \times \mathbf{P}^1_K \hookrightarrow \mathbf{P}^3_K, \qquad (a:b), (c:d) \mapsto (ab : ac : bd : cd)$$

and $\widetilde{Z}_1 = \{(0:1)\} \times \mathbf{P}^1_K = \{0\} \times \mathbf{P}^1_K$, $\widetilde{Z}_2 = \{(1:0)\} \times \mathbf{P}^1_K = \{\infty\} \times \mathbf{P}^1_K$. In other words, this example is a projective analogue of the fact that two parallel affine lines $\{0\} \times \mathbf{A}^1_K$ and $\{1\} \times \mathbf{A}^1_K$ do not intersect in $\mathbf{A}^1_K \times \mathbf{A}^1_K = \mathbf{A}^2_K$.

14. Discrete valuation rings

(14.1) Discrete valuation rings are the simplest one-dimensional rings, namely, regular one-dimensional local rings. An archetypal example is provided by the local ring $A = K[X]_{(X)}$ of a line $\mathbf{A}^1_K$ over a field K at the origin and by its completion $\widehat{A} = K[[X]]$.

Both A and $\widehat{A}$ are principal local domains. In particular, their maximal ideals are principal, generated by X (a "local parameter"). Every non-zero element a of A resp. $\widehat{A}$ can be written in a unique way as $a = X^n u$, where u is invertible and $n = v(a) \geq 0$ is the **valuation of** a (= the order of vanishing of a, considered as a function on $\mathbf{A}^1_K$, at the origin).

(14.2) Definition. *A* **discrete valuation** *on a field K is a surjective map $v : K \longrightarrow \mathbf{Z} \cup \{+\infty\}$ such that*
(i) $v(x) = +\infty \iff x = 0$.
(ii) $\forall x, y \in A \quad v(xy) = v(x) + v(y)$.
(iii) $\forall x, y \in A \quad v(x+y) \geq \min(v(x), v(y))$.
The **valuation ring** *of v is the subring $A = \{x \in K \mid v(x) \geq 0\}$. Its multiplicative group is equal to $A^* = \{x \in A \mid v(x) = 0\}$, which implies that A is a local ring with maximal ideal $\mathfrak{m} = \{x \in A \mid v(x) > 0\} =$*

(t), *where* $t \in A$ *is any element of valuation* $v(t) = 1$ *(such an element is called a* **uniformiser** *or a* **local parameter** *of* A*).*

(14.3) Definition. *A domain* A *is a* **discrete valuation ring** *(often abbreviated as DVR) if it is the valuation ring of a suitable discrete valuation* v *on its fraction field* $K = \mathrm{Frac}(A)$*. In concrete terms,* A *is a local domain whose maximal ideal* $\mathfrak{m} = (t)$ *is principal and for which* $A \setminus \{0\} = \bigcup_{n\in\mathbf{N}} t^n A^*$ *(a disjoint union). Indeed, this decomposition gives a disjoint union* $K \setminus \{0\} = \bigcup_{n\in\mathbf{Z}} t^n A^*$ *and the corresponding discrete valuation* $v(t^n A^*) = n$*.*

(14.4) Exercise. *If* A *is a discrete valuation ring with uniformiser* t*, so is its completion*

$$\widehat{A} = \varprojlim_n A/t^n A.$$

(14.5) Examples. (i) $K = k(X)$, where k is a field. As in I.5.10, every irreducible non-constant polynomial $\pi \in k[X]$ defines a discrete valuation on $K = \mathrm{Frac}(k[X])$ given by

$$v_\pi(f/g) = v_\pi(f) - v_\pi(g),$$

where $v_\pi(f)$ is the maximum exponent $n \geq 0$ for which π^n divides f in $k[X]$. The corresponding valuation ring is $A = k[X]_{(\pi)}$ and π is its uniformiser. If $\pi = X - c$ for some $c \in k$, then $\widehat{A} = k[[X - c]]$ and $\mathrm{Frac}(\widehat{A}) = k((X - c))$.
(ii) $K = k(X)$ has another discrete valuation, namely

$$v_\infty(f/g) = \deg(g) - \deg(f),$$

whose uniformiser is $1/X$ (a "local parameter at $\infty \in \mathbf{P}^1(k)$"). In this case $\widehat{A} = k[[1/X]]$ and $\mathrm{Frac}(\widehat{A}) = k((1/X))$.
(iii) $K = \mathbf{Q}$ has p-adic discrete valuations (where p is a prime number) given by

$$v_p(p^n \frac{a}{b}) = n, \qquad a, b \in \mathbf{Z},\ p \nmid ab.$$

The valuation ring of v_p is equal to $\mathbf{Z}_{(p)}$, its uniformiser is p and its completion is the ring of p-adic integers $\mathbf{Z}_p$ (whose fraction field is the field of p-adic numbers $\mathbf{Q}_p$).

(14.6) Before we pass to an abstract characterisation of discrete valuation rings we need a few definitions.

(14.7) Definition. *A* **fractional ideal** *of a domain* A *is an* A*-submodule* $I \subset K = \mathrm{Frac}(A)$ *of the form* $I = \alpha^{-1} I_0$*, where* $\alpha \in A \setminus \{0\}$ *and* I_0 *is a non-zero ideal of* A*.*

(14.8) Exercise. *If* I *and* J *are fractional ideals of* A*, so are* $I + J = \{x + y \mid x \in I,\ y \in J\}$*,* $IJ = \{\sum_{i=1}^r x_i y_i \mid r \geq 0,\ x_i \in I,\ y_i \in J\}$ *and* $I^{-1} = \{x \in K \mid xI \subset A\}$*.*

(14.9) Definition. *A fractional ideal* I *of a domain* A *is* **invertible** *if there exists a fractional ideal* J *such that* $IJ = A$ *(*$\iff$ $II^{-1} = A$*). For a non-zero ideal* $I \subset A$*, this is equivalent to the existence of a non-zero ideal* $I' \subset A$ *such that* $II' = (a)$ *is a principal ideal.*

(14.10) Proposition. *An invertible fractional ideal* I *is finitely generated (as an* A*-module).*

Proof. If $II^{-1} = A$, then there exist $x_i \in I$ and $y_i \in I^{-1}$ such that $\sum_{i=1}^r x_i y_i = 1$. Multiplying this identity by an arbitrary $z \in I$, we obtain that $z = \sum_{i=1}^r (zy_i)x_i \in Ax_1 + \cdots + Ax_r$, since $zy_i \in A$; thus $I = Ax_1 + \cdots + Ax_r$.

(14.11) Theorem. *Let* A *be a local ring with non-zero maximal ideal* $\mathfrak{m} \neq (0)$ *(in other words,* A *is not a field). The following properties are equivalent.*
(1) A *is a DVR.*
(2) A *is a PID.*
(3) A *is a regular noetherian local ring of dimension* $\dim(A) = 1$*.*
(4) A *is noetherian,* $\mathfrak{m} = (t)$ *is principal and* $\dim(A) \neq 0$*.*

(5) A is a noetherian normal domain of dimension $\dim(A) = 1$.
(6) A is a domain and every fractional ideal of A is invertible.

Proof. The implications (1) $\Longrightarrow$ (2) $\Longrightarrow$ (3) $\Longrightarrow$ (4) and (1) $\Longrightarrow$ (6) are automatic (note that the only non-zero ideals of a DVR are (t^n), $n \geq 0$). The implication (2) $\Longrightarrow$ (5) follows from Proposition 3.2 and Lemma III.3.11.
(4) $\Longrightarrow$ (1): for each $a \in A \setminus \{0\}$ we have either $a \in A \setminus \mathfrak{m} = A^*$, or $a = ta_1$ for some $a_1 \in A \setminus \{0\}$. In the latter case we can apply the same argument to a_1 and continue. This process must stop after finitely many steps, since $\bigcap_{n\geq 1}(t^n) = (0)$, by Krull's Theorem 12.7. Therefore $A \setminus \{0\} = \bigcup_{n\geq 1} t^n A^*$. By assumption, there exists a prime ideal $P \subsetneq \mathfrak{m}$. If $P \neq (0)$, then there exists $a \in P$, $a \neq 0$. As above, $a = t^n u$ with $u \in A^*$ and $n \geq 1$ ($n \neq 0$, since $1 \notin P$). However, $t^n \in P$ implies that $t \in P$ (since $P \in \mathrm{Spec}(A)$), which yields a contradiction $\mathfrak{m} = P$. It follows that $P = (0)$, hence A is a domain (and $\mathrm{Spec}(A) = \{(0), \mathfrak{m}\}$), which implies that $A \setminus \{0\} = \bigcup_{n\geq 1} t^n A^*$ is a disjoint union; thus A is a DVR.
(5) $\Longrightarrow$ (3): the assumptions imply that $\mathrm{Spec}(A) = \{(0), \mathfrak{m}\}$. As $\mathfrak{m} \neq (0)$, the quotient $\mathfrak{m}/\mathfrak{m}^2$ is non-zero, by Nakayama's Lemma 13.6. Fix $x \in \mathfrak{m}$ such that $x \notin \mathfrak{m}^2$. We must show that $\mathfrak{m} = (x)$. Corollary 8.11(2) tells us that $\sqrt{(x)} = \mathfrak{m}$. This ideal is finitely generated, which implies that $\mathfrak{m}^n \subset (x) \subset \mathfrak{m}$ for some $n \geq 2$. For each $y \in \mathfrak{m}^{n-1}$ the ideal $(y/x)\mathfrak{m} \subset A$ cannot be equal to A, since $x \notin \mathfrak{m}^n$; thus $(y/x)\mathfrak{m} \subset \mathfrak{m}$. Proposition 2.6(3) implies that $y/x \in \mathrm{Frac}(A)$ is integral over A, hence $y/x \in A$, which means that $\mathfrak{m}^{n-1} \subset (x)$. Decreasing induction shows that $\mathfrak{m} \subset (x) \subset \mathfrak{m}$, hence $\mathfrak{m} = (x)$.
(6) $\Longrightarrow$ (4): the assumption $(0) \in \mathrm{Spec}(A) \setminus \{\mathfrak{m}\}$ implies that $\dim(A) \neq 0$. According to Proposition 14.10, A is noetherian. In particular, $\mathfrak{m} = (x_1, \ldots, x_r)$ is finitely generated. If $x_i \mid x_j$, then we can omit x_j. As a result, if $\mathfrak{m}$ is not principal, then there exist $x, y \in \mathfrak{m} \setminus \{0\}$ such that $x \nmid y$ and $y \nmid x$. Neither of the ideals $\{a \in A \mid ax \subset (y)\}$ and $\{a \in A \mid ay \subset (x)\}$ contains 1, which means that they are both contained in $\mathfrak{m}$, hence $(x)^{-1} \subset y^{-1}\mathfrak{m}$ and $(y)^{-1} \subset x^{-1}\mathfrak{m}$. It follows that the ideal $I = (x, y)$ satisfies $I^{-1} = (x)^{-1} \cap (y)^{-1} \subset x^{-1}\mathfrak{m} \cap y^{-1}\mathfrak{m}$, hence $II^{-1} \subset xx^{-1}\mathfrak{m} + yy^{-1}\mathfrak{m} = \mathfrak{m}$, which means that I is not invertible. This contradiction implies that $\mathfrak{m} = (t)$ is principal.

(14.12) Branches of plane curves, Newton polygons, Puiseux expansions. Let k be an algebraically closed field of characteristic zero. The complete DVR $k[[T]]$ is a very simple object. The only finite extensions of $\mathrm{Frac}(k[[T]]) = k((T))$ are the fields $k((T^{1/d}))$ ($d \geq 1$), and the normalisation of $k[[T]]$ in $k((T^{1/d}))$ is $k[[T^{1/d}]]$ ([Ei, Cor. 13.15]).

It follows that every polynomial $f \in K[X, T] = \sum c_{i,j} X^i T^j$ with $\deg_X(f) = n \geq 1$ factors as

$$f(X, T) = h(T) \prod_{i=1}^{n} (X - g_i),$$

for suitable $g_i \in k((T^{1/d_i}))$. The exponents $1/d_i$ and the leading terms of the series g_i can be read off from the **Newton polygon of** f, which is the lower boundary of the convex hull of the finite set $S = \{(i, j) \mid c_{i,j} \neq 0\} \subset \{0, 1, \ldots, n\} \times \mathbf{N}$.

For example,

$$X^2 + aT^2X - T = (X - g_1)(X - g_2), \qquad g_{1,2} = \pm T^{1/2} + \cdots$$

and $S = \{(0, 1), (1, 2), (2, 0)\}$. Another example is given by

$$TX^2 + (1 + bT)X - T^4 = (X - g_1)(X - g_2), \qquad g_1 = -T^{-1} + \cdots, \qquad g_2 = T^3 + \cdots$$

and $S = \{(0, 4), (1, 0), (1, 1), (2, 1)\}$. Can you guess the general rule?

15. Dedekind rings

Discrete valuation rings are non-singular one-dimensional local objects. Dedekind rings are non-singular one-dimensional global objects. This is a geometric characterisation, but they first appeared in arithmetic, thanks to the property 15.2 for rings of integers in finite extensions of $\mathbf{Q}$.

(15.1) Definition. *A* **Dedekind ring** *is a domain A which is not a field (even though some authors allow fields to be Dedekind rings ...) and which satisfies the following equivalent conditions.*
(1) A is noetherian and the localisations of A at all maximal ideals are discrete valuation rings.
(2) A is noetherian, integrally closed and all non-zero prime ideals are maximal.
(3) All fractional ideals of A are invertible (equivalently, the set of all fractional ideals of A is a group with respect to multiplication).
[For example, a PID which is not a field is a Dedekind ring.]

Proof. (1) $\Longrightarrow$ (2): by assumption, $A_{\mathfrak{m}}$ is a DVR, for each $\mathfrak{m} \in \mathrm{Max}(A)$. Proposition 10.19(2) then shows that A is integrally closed. Moreover, $\dim(A) = \sup_{\mathfrak{m}\in\mathrm{Max}(A)} \dim(A_{\mathfrak{m}}) = 1$, which implies that non-zero prime ideals of A are maximal.
(2) $\Longrightarrow$ (1): the assumption (2) implies that $\dim(A_{\mathfrak{m}}) = 1$ for all $\mathfrak{m} \in \mathrm{Max}(A)$ and that each $A_{\mathfrak{m}}$ is a normal domain (by 10.17(1)); thus $A_{\mathfrak{m}}$ is a DVR, thanks to Theorem 14.11(5).
(3) $\Longrightarrow$ (1): the assumption (3) implies that A is noetherian, by Proposition 14.10, and that all fractional ideals of $A_{\mathfrak{m}}$ ($\mathfrak{m} \in \mathrm{Max}(A)$) are invertible. As $A_{\mathfrak{m}}$ is not a field, $A_{\mathfrak{m}}$ is a DVR, by Theorem 14.11(6).
(1) $\Longrightarrow$ (3): see Corollary 15.3 below (the reader is invited to check that our reasoning is not circular).

(15.2) Theorem (Unique factorisation into prime ideals). *Let A be a Dedekind ring in the sense of Definition 15.1(1). The (well-defined) maps*

$$\mathrm{div} : \{\mathrm{non-zero\ ideals\ of\ } A\} \longrightarrow \bigoplus_{\mathfrak{m}\in\mathrm{Max}(A)} \mathbf{N}, \qquad I \mapsto (v_{\mathfrak{m}}(I)), \qquad IA_{\mathfrak{m}} = (\mathfrak{m}A_{\mathfrak{m}})^{v_{\mathfrak{m}}(I)}$$

and

$$I : \bigoplus_{\mathfrak{m}\in\mathrm{Max}(A)} \mathbf{N} \longrightarrow \{\mathrm{non-zero\ ideals\ of\ } A\}, \qquad (n_{\mathfrak{m}}) \mapsto \prod_{\mathfrak{m}} \mathfrak{m}^{n_{\mathfrak{m}}}$$

define mutually inverse bijections satisfying $\mathrm{div}(IJ) = \mathrm{div}(I) + \mathrm{div}(J)$.

Proof. According to Corollary 8.11(2), the radical of any non-zero ideal I of A is of the form $\sqrt{I} = \mathfrak{m}_1 \cap \cdots \cap \mathfrak{m}_r = \mathfrak{m}_1 \cdots \mathfrak{m}_r$, for some $\mathfrak{m}_i \in \mathrm{Max}(A)$. As $\sqrt{I}$ is finitely generated, there exists $k \geq 1$ such that $I \supset (\mathfrak{m}_1 \cdots \mathfrak{m}_r)^k$, which implies that $IA_{\mathfrak{m}} = A_{\mathfrak{m}}$ for all maximal ideals $\mathfrak{m} \neq \mathfrak{m}_1, \ldots, \mathfrak{m}_r$. In particular, the map div is well-defined.

For $\mathfrak{m}' \neq \mathfrak{m}$ we have $\mathfrak{m}' + \mathfrak{m} = A$, which implies that $\mathfrak{m}'A_{\mathfrak{m}} = A_{\mathfrak{m}}$. Consequently, $\mathrm{div} \circ I = \mathrm{id}$.

As $I((n_{\mathfrak{m}}))I((n'_{\mathfrak{m}})) = I((n_{\mathfrak{m}} + n'_{\mathfrak{m}}))$, it remains to check that div is injective. If $IA_{\mathfrak{m}} = JA_{\mathfrak{m}}$ for all $\mathfrak{m}$, then the ideal $I_0 = \{x \in A \mid xI \subset J\}$ satisfies $I_0A_{\mathfrak{m}} = A_{\mathfrak{m}}$, hence $(A/I_0)_{\mathfrak{m}} = 0$, for all $\mathfrak{m}$; thus $I_0 = A$ by 10.18(2). In particular, $I \subset J$; by symmetry, $J \subset I$ and $I = J$.

(15.3) Corollary. *Let A be a Dedekind ring in the sense of Definition 15.1(1). Then every non-zero ideal I of A is invertible (hence A satisfies the condition 15.1(3)).*

Proof. Fix $a \in I$, $a \neq 0$. As $(a) \subset I$, we have $(a)A_{\mathfrak{m}} \subset IA_{\mathfrak{m}}$ and $v_{\mathfrak{m}}((a)) \geq v_{\mathfrak{m}}(I)$ for all $\mathfrak{m} \in \mathrm{Max}(A)$. The ideal $J = \prod_{\mathfrak{m}} \mathfrak{m}^{v_{\mathfrak{m}}((a)) - v_{\mathfrak{m}}(I)}$ (the product is finite) then satisfies $IJ = (a)$, since $\mathrm{div}(I) + \mathrm{div}(J) = \mathrm{div}((a))$.

(15.4) Corollary. *Let A be a Dedekind ring. The bijections from Theorem 15.2 naturally extend to mutually inverse isomorphisms of abelian groups*

$$\{\mathrm{fractional\ ideals\ of\ } A\} \xrightarrow{\sim} \bigoplus_{\mathfrak{m}\in\mathrm{Max}(A)} \mathbf{Z}.$$

(15.5) Theorem. *Let A be a noetherian domain of dimension $\dim(A) = 1$, let L be a finite extension of the fraction field $K = \mathrm{Frac}(A)$, let B be the normalisation of A in L. If either (i) A is a Dedekind ring, or (ii) A is an algebra of finite type over a field k, then B is a Dedekind ring (of finite type over k in (ii)).*
[In fact, B is always a Dedekind ring, but the hard part is to show that it is noetherian; this is the content of the Krull-Akizuki theorem [Ei, Thm. 11.13], [M, Thm. 11.7].]

Proof. In the special case when B is a finitely generated A-module, so is every ideal of B (since A is noetherian), which implies that B is also noetherian. Moreover, $\dim(B) = \dim(A) = 1$, by Theorem 11.6. Consequently, B satisfies 15.1(2).

The above argument applies in the case (ii), thanks to Theorem 4.10, and in the case (i) if the extension L/K is separable, by Theorem 4.3(2). By transitivity of normalisation, it remains to treat the case (i) when L/K is a non-trivial purely inseparable extension. In this case $char(K) = p$ and there exists $q = p^r > 1$ such that $L^q \subset K$. The extensions $K \subset L \subset K^{1/q}$ contain rings $A \subset B \subset A^{1/q}$ and $B = \{b \in L \mid b^q \in A\}$. The Frobenius morphism $\varphi_q : A^{1/q} \xrightarrow{\sim} A$ is bijective, which means that $A^{1/q}$ is a Dedekind ring. In particular, for every non-zero ideal $I \subset B$ there exists a non-zero ideal $J \subset A^{1/q}$ such that $(IA^{1/q})J = (\alpha)$ $(\alpha \in A^{1/q})$ is a principal ideal. Taking the q-th powers we obtain $A(I^q J^q) = \alpha^q A$, hence $I(I^{q-1}J^q B) = \alpha^q B$; therefore I is invertible and B is a Dedekind ring, thanks to 15.1(3).

(15.6) In the situation (ii) of Theorem 15.5 we can interpret A as the ring of regular functions $O(Z)$ on an irreducible reduced affine curve Z over k. Taking $L = K = \mathrm{Frac}(A)$, the normalisation B of A in $\mathrm{Frac}(A)$ corresponds to a curve $\widetilde{Z}$, which is regular everywhere. The inclusion $A \hookrightarrow B$ defines a map $\widetilde{Z} \longrightarrow Z$, (a "desingularisation of Z"), which is an isomorphism outside a finite set of closed points. If the field k is perfect, then all points of $\widetilde{Z}$ are smooth over k.

(15.7) What is the relation between normalisation and desingularisation in the higher-dimensional case? If $A = O(Z) = K[X_1, \ldots, X_n]/I$ is a **normal domain**, so are its localisations A_P at all $P \in \mathrm{Spec}(A_P)$. In particular, if P is minimal among non-zero prime ideals, then $\dim(A_P) = 1$ and A_P is a DVR, by Theorem 14.11(5). Geometrically, $V(P) \xrightarrow{\sim} \mathrm{Spec}(A/P) \subset \mathrm{Spec}(A)$ is an "irreducible subvariety" of Z of dimension $\dim(A/P) = \dim(A) - \dim(A_P) = \dim(A) - 1$. The fact that A_P is a regular local ring means that Z is non-singular along $V(P)$. Therefore all singularities of Z occur in dimension $\leq \dim(A) - 2$ (Z is **regular in codimension 1**). Moreover, the discrete valuation on $\mathrm{Frac}(A_P) = \mathrm{Frac}(A)$ attached to A_P defines the order of vanishing of rational functions $f \in \mathrm{Frac}(A)^*$ along $V(P)$.

The cone $Z : X_1^2 + X_2^2 - X_3^2 = 0$ from 13.12(iii) gives an example when A is normal, $\dim(A) = 2$ and there is a singularity in dimension $2 - 2 = 0$: $A_{\mathfrak{m}}$ is not regular if $\mathfrak{m} = (\overline{X}_1, \overline{X}_2, \overline{X}_3)$ is the maximal ideal of A corresponding to the origin $(0, 0, 0) \in Z(K)$.

In general, there is a whole hierarchy of algebraic and geometric properties of rings, such as

$$\text{regular} \Longrightarrow \text{local complete intersection} \Longrightarrow \text{Gorenstein} \Longrightarrow \text{Cohen} - \text{Macaulay}$$

(see [AK], [BH]). Serre's criterion ([Se, III.C, Prop. 9])

$$\text{normal} \iff \text{regular in codimension 1 and Cohen} - \text{Macaulay in codimension 2}$$

implies that, for any irreducible $f \in K[X_1, \ldots, X_n] \setminus K$, the hypersurface $Z : f = 0$ is normal $\iff$ it is regular in codimension one ([Mu 1, III.8, Prop. 2]). On the other hand, Example K(B) in [Mu 1, III.8] shows that there exists a surface in $\mathbf{A}^4_K$ which is regular in codimension one, but which is not normal.

(15.8) If $A = \mathbf{Z}$, $K = \mathbf{Q}$ and $[L : \mathbf{Q}] < \infty$, then we obtain from Theorem 15.5(i) that the ring of algebraic integers O_L of L is a Dedekind ring. In particular, non-zero ideals of O_L have unique factorisation into maximal ideals (= non-zero prime ideals).

(15.9) Ramification. If A is a Dedekind ring, $K = \mathrm{Frac}(A) \hookrightarrow L$ a finite extension and B the normalisation of A in L, then each $\mathfrak{m} \in \mathrm{Max}(A)$ factors in B as

$$\mathfrak{m}B = \prod_{i=1}^{r} \mathfrak{m}_i^{e_i}, \qquad (\mathfrak{m}_i \in \mathrm{Max}(B)),$$

where $e_i \geq 1$ is the **ramification index** of $\mathfrak{m}_i$ above $\mathfrak{m}$. The phenomenon of ramification (when $e_i > 1$) is related to the discriminants studied in III.7. Each residue field $B/\mathfrak{m}_i$ is a finite extension of $A/\mathfrak{m}$; denote by $f_i = [B/\mathfrak{m}_i : A/\mathfrak{m}]$ its degree. There is a fundamental inequality

$$\sum_{i=1}^{r} e_i f_i \leq [L : K], \tag{15.9.1}$$

which becomes an equality

$$\sum_{i=1}^{r} e_i f_i = [L : K] \tag{15.9.2}$$

in the special case when $B_{\mathfrak{m}}$ is a finitely generated module over $A_{\mathfrak{m}}$ (which is true, for example, if L/K is separable or if A is an algebra of finite type over a field or if A is a complete DVR).

(15.10) Exercise. *Discuss the equality (15.9.2) in the case when $A = \mathbf{Z}$, $B = \mathbf{Z}[i]$ (resp. $A = k[X]$, $B = k[Y]$, $Y^2 = X$).*

References

[AK] A. Altman, S. Kleiman, *Introduction to Grothendieck duality theory*, Lecture Notes in Mathematics **146**, Springer, Berlin, 1970.

[Ar] E. Artin, *Galois Theory*, Dover, 1998.

[AM] M. Atiyah, I. G. Macdonald, *Introduction to commutative algebra*, Addison-Wesley, Reading, 1969.

[BGR] S. Bosch, U. Güntzer, R. Remmert, *Non-Archimedean Analysis*, Grundlehren der Mathematischen Wissenschaften **261**, Springer, Berlin, 1984.

[BH] W. Bruns, J. Herzog, *Cohen-Macaulay rings*, Cambridge Studies in Advanced Mathematics **39**, Cambridge University Press, Cambridge, 1993.

[CL] A. Chambert-Loir, *Algèbre corporelle*, École Polytechnique, 2005.

[Co] D. Cox, *Galois Theory*, Wiley, 2004.

[De 1] O. Debarre, *Algèbre 1*, E.N.S., 2013.

[De 2] O. Debarre, *Algèbre 2*, E.N.S., 2013.

[Ei] D. Eisenbud, *Commutative algebra*, Graduate Texts in Mathematics **150**, Springer, New York, 1995.

[Es] J.-P. Escofier, *Théorie de Galois*, Dunod, 2000.

[Ga] D.J.H. Garling, *Galois Theory*, Cambridge University Press, 1986.

[Gr] A. Grothendieck, *Éléments de Géométrie Algébrique IV.2*, Publications Mathématiques de l'Institut des Hautes Études Scientifiques **24**, 1965.

[ILO] L. Illusie, Y. Laszlo, F. Orgogozo, *Travaux de Gabber sur l'uniformisation locale et la cohomologie étale des schémas quasi-excellents.* `arXiv:1207.3648`

[Kl] F. Klein, *Lectures on the icosahedron and the solution of equations of the fifth degree*, Second and revised edition, Dover Publications, New York, 1956.

[La] S. Lang, *Algebra*, Revised third edition, Graduate Texts in Mathematics **211**, Springer, New York, 2002.

[Ma] H. Matsumura, *Commutative ring theory*, Second edition, Cambridge Studies in Advanced Mathematics **8**, Cambridge University Press, Cambridge, 1989.

[Mu 1] D. Mumford, *The Red Book of Varieties and Schemes*, Lecture Notes in Mathematics **1358**, Springer, Berlin, 1988.

[Mu 2] D. Mumford, *Algebraic Geometry I, Complex Projective Varieties*, Grundlehren der Mathematischen Wissenschaften **221**, Springer, Berlin, 1976.

[Re] M. Reid, *Undergraduate commutative algebra*, London Mathematical Society Student Texts **29**, Cambridge University Press, 1995.

[Sa] P. Samuel, *Théorie algébrique des Nombres*, Hermann, 1967.

[Se] J.-P. Serre, *Algèbre locale. Multiplicités*, Lecture Notes in Mathematics **11**, Springer, Berlin, 1965.

[Ti 1] J.-P. Tignol, *Leçons sur la théorie des équations*, Université Catholique de Louvain, 1980.

[Ti 2] J.-P. Tignol, *Galois Theory of Algebraic Equations*, World Scientific, 2001.

[ZS 1,2] O. Zariski, P. Samuel, *Commutative algebra, Vol. 1 + 2*, Graduate Texts in Mathematics **28, 29**, Springer, New York, 1975.