ALGEBRA 4 (MATH 371)
COURSE NOTES
WINTER 2014
VERSION: October 17, 2020

EYAL Z. GOREN,
MCGILL UNIVERSITY

©All rights reserved to the author.

Contents

Part 1. Modules

The concept of a module resembles very much the concept of a vector space V over a field $\mathbb{F}$, which is indeed a special case of a module. Recall that a vector space V is an abelian group and the multiplication of a vector $v \in V$ by a scalar $\lambda \in \mathbb{F}$ may be viewed as an action of $\mathbb{F}$ on V that respects the additive structure of V and, at the same time, preserves the addition and multiplication in $\mathbb{F}$. That translates into the axioms $\lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2$, $(\lambda_1 + \lambda_2)v = \lambda_1 v + \lambda_2 v$, $(\lambda_1\lambda_2)v = \lambda_1(\lambda_2 v)$ and $1v = v$.

This definition utilizes only the structure of abelian group on V and that $\mathbb{F}$ has a structure of a ring. The fact that $\mathbb{F}$ is a very special ring - it is commutative, an integral domain and every non-zero element is invertible - doesn't feature in the definition. This observation gives immediately the definition of a module, as we shall presently see.

1. First definitions

1.1. Modules, submodules and homomorphisms.

Let R be a ring, always associative with 1, but not necessarily commutative. An abelian group $(M, +)$ is called a **left R-module** if we are given a function

$$R \times M \to M, \qquad (r, m) \mapsto rm,$$

such that the following holds:

(1) $r(m_1 + m_2) = rm_1 + rm_2$, for all $r \in R, m_i \in M$.
(2) $1m = m$, for all $m \in M$.
(3) $(r_1 + r_2)m = r_1 m + r_2 m$, for all $r_i \in R, m \in M$.
(4) $(r_1 r_2)m = r_1(r_2 m)$, for all $r_i \in R, m \in M$.

Sometime we write $r \cdot m$ instead of rm to help distinguish the ring element from the module element. It is an easy consequence of the axioms that $r \cdot 0 = 0$ for all $r \in R$.

Note that every $r \in R$ can be viewed as a group homomorphism $[r] : M \to M$, given by $m \mapsto rm$. The function

$$R \to \mathrm{End}(M), \qquad r \mapsto [r]$$

(where $\mathrm{End}(M)$ denotes the group homomorphisms $M \to M$; it has a natural ring structure, where $(f + g)(m) := f(m) + g(m)$ and $fg = f \circ g$), is a ring homomorphism. In turn, a ring homomorphism $R \to \mathrm{End}(M)$ makes M into an R-module.

A **submodule** N of a left R-module M is a subgroup $N \subseteq M$ such that for all $n \in N, r \in R$ we have $rn \in N$. We note that in this case N is itself an R-module. The intersection of any collection of submodules is a submodule.

A **homomorphism**

$$f \colon M_1 \to M_2$$

of left R-modules is a function f such that $f \colon M_1 \to M_2$ is a group homomorphism and $f(rm) = rf(m)$ for all $m \in M_1, r \in R$. For example, if N is a submodule of M then the inclusion map $N \to M$, $n \mapsto n$, is a homomorphism of modules.

Let $f \colon M_1 \to M_2$ be a homomorphism of modules. The **kernel** of f,

$$\mathrm{Ker}(f) = \{m \in M_1 : f(m) = 0\},$$

is a subgroup of M_1 (note that all subgroups are normal), but it has another property: it is a submodule. Indeed, let $m \in \mathrm{Ker}(f)$ and $r \in R$ then $f(rm) = rf(m) = r \cdot 0 = 0$. We remark that if $B \subseteq M_1$ is a submodule then $f(B)$ is a submodule of M_2.

A homomorphism $f \colon M_1 \to M_2$ which is bijective is called an **isomorphism**. In that case, the inverse function $f^{-1} \colon M_2 \to M_1$ is automatically a module homomorphism. We say that M_1 is **isomorphic** to M_2 if there exists an isomorphism $f \colon M_1 \to M_2$. We denote this by $M_1 \cong M_2$. One checks that being isomorphic is an equivalence relation.

Let $m \in M$. The **annihilator** of m, $\mathrm{Ann}(m)$ (or, $\mathrm{Ann}_R(m)$ if we need to specify the ring) is defined as follows.

$$\mathrm{Ann}(m) = \{r \in R : rm = 0\}.$$

We note that $\mathrm{Ann}(m)$ is a left ideal of R. More generally, for a non-empty subset S of M define $\mathrm{Ann}(S) = \{r \in R : rs = 0, \forall s \in S\}$. As $\mathrm{Ann}(S) = \cap_{s\in S}\mathrm{Ann}(s)$ is the intersection of left ideals it is a left ideal too.

If R is a *commutative integral domain* we make another definition: we define the **torsion** of M as

$$\mathrm{Tors}(M) = \{m \in M : \exists r \in R, r \neq 0, rm = 0\} = \bigcup_{\{m\in M:\ \mathrm{Ann}(m)\neq 0\}} \{m\}.$$

One checks that $\mathrm{Tors}(M)$ is a submodule of M. The verification is easy, but it does use crucially the assumptions on R.

Before giving examples, we remark that one can define analogously a **right R-module** M as an abelian group with a function $M \times R \to M,\ (m,r) \mapsto mr$ with the analogous axioms. One gets a theory completely parallel to the one of left R-modules and for that reason we shall restrict our discussion to left R-modules throughout. Although, when we need to use results about right R-modules we shall do so without hesitation.

1.2. **Examples.** The following examples are *key examples*. We shall often return to them.

Example 1.2.1. Let I be a left ideal of R. Then I is an R-module. This applies in particular to R itself. Conversely, if $M \subseteq R$ is an R-submodule of R then M is a left ideal.

Example 1.2.2. Let M be an R-module and $m \in M$. Then $Rm := \{rm : r \in R\}$ is a submodule of M. In general, a submodule N of M will be called **cyclic** if there is an element $m \in N$ such that $N = Rm$. Let $S = \{m_\alpha\}_{\alpha\in A}$ be a collection of elements of M. Define the **submodule generated by** S to be

$$\langle S\rangle := \left\{\sum r_i n_i : \text{finite sum}, r_i \in R, n_i \in S\right\}.$$

This is a submodule and, in fact, is the minimal submodule of M containing S. If $M_1, \dots, M_n$ are submodules of M and $S = M_1 \cup \cdots \cup M_n$ then the submodule generated by S can also be written as

$$M_1 + \cdots + M_n = \{m_1 + \cdots + m_n : m_i \in M_i, \forall i\}.$$

It is also called the **sum** of the modules $\{M_i : 1 \leq i \leq n\}$. More generally, if $\{M_\alpha : \alpha \in I\}$ are a collection of submodules of M, we define their **sum**, $\sum_{\alpha\in I} M_\alpha$, to be the minimal submodule containing all of them; equivalently, the collection of all finite sums $\sum_{i=1}^n m_i$ where each m_i belongs to some $M_{\alpha(i)}$.

Let I be an ideal of R. Define IM as the collection of finite sums $\sum r_i m_i$, where $r_i \in I$ and m_i in M. This is an R-submodule of M.

Example 1.2.3 (Abelian groups)**.** Every abelian group M can be viewed as a module over the ring of integers $\mathbb{Z}$, where for $n \in \mathbb{Z}, g \in M$ we let $ng = g + g + \cdots + g$ (n-times) for $n > 0$, $ng = 0$ for $n = 0$ and $ng = -((-n)g)$ for $n < 0$. A submodule is nothing else then a subgroup and a module homomorphism amounts to a group homomorphism. A quotient module is just a quotient group and so on.

A group is a cyclic $\mathbb{Z}$-module precisely when it is a cyclic group. The torsion of a group are the elements of finite order.

Example 1.2.4 (Vector spaces)**.** Let $\mathbb{F}$ be a field and V a vector space of $\mathbb{F}$. Then V is an $\mathbb{F}$-module. Maps of $\mathbb{F}$-modules are linear transformations over $\mathbb{F}$ and a submodule is a subspace. Quotient modules are quotient spaces. We always have $\mathrm{Tors}(V) = \{0\}$. V is cyclic if and only if $\dim_{\mathbb{F}}(V) \leq 1$.

Example 1.2.5 (Vector spaces with a linear transformation)**.** Let $\mathbb{F}$ be a field. Consider vector spaces over $\mathbb{F}$ equipped with a linear transformation, say (V, T). This notion is equivalent to saying that V is an $\mathbb{F}[x]$-module. Indeed, given such datum (V, T), define a module structure on V by

$$f(x) \cdot v = f(T)(v), \qquad f(x) \in \mathbb{F}[x], v \in V.$$

As usual, if $f(x) = \lambda_n x^n + \cdots + \lambda_0$ then $f(T)$ is the linear transformation $\lambda_n T^n + \cdots + \lambda_0 \mathrm{Id}$, where T^n means $T \circ T \circ \cdots \circ T$ (n-times). That is, the scalars $\lambda \in \mathbb{F}$ act naturally, x acts as T, x^2 acts as T^2 and so on. One easily verifies the module axioms.

Conversely, if V is an $\mathbb{F}[x]$-module, the given action $(\lambda, v) \mapsto \lambda v, \lambda \in \mathbb{F}, v \in V$, makes V into a vector space over $\mathbb{F}$. Further, define

$$T = T_V \colon V \to V, \qquad T(v) := x \cdot v.$$

Here $x \cdot v$ stands for the product of a ring element x with a module element v. The identity $T(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 T(v_1) + \lambda_2 T(v_2)$ follows from $x(\lambda_1 v_1 + \lambda_2 v_2) = x(\lambda_1 v_1) + x(\lambda_2 v_2) = (x\lambda_1)v_1 + (x\lambda_2)v_2 = (\lambda_1 x)v_1 + (\lambda_2 x)v_2 = \lambda_1(xv_1) + \lambda_2(xv_2)$.

Under this dictionary, a homomorphism of $\mathbb{F}[x]$-modules $f \colon V \to W$, corresponds to a linear map $f \colon V \to W$ such that $f \circ T_V = T_w \circ f$. A submodule of V corresponds to a T_V-invariant subspace of V. We summarize that in a table.

Modules	Vector spaces with a linear map
$\mathbb{F}[x]$-module V	v. sp. + lin. transf. (V, T_V)
submodule	T_V-invariant subspace
homomorphism of $\mathbb{F}[x]$-modules $f \colon V \to W$	linear transf. $f \colon V \to W$ satisfying $f \circ T_v = T_w \circ f$

Suppose that V is finite dimensional and let $m(x)$ be the minimal polynomial of T. Then $m(T)$ is the zero map. It follows that every element of V is killed by $m(x) \in \mathbb{F}[x]$. Thus, if V is finite dimensional V is torsion, $V = \mathrm{Tors}(V)$.

When is V a cyclic $\mathbb{F}[x]$ module? This is precisely when there is a vector $v \in V$ such that $\{v, Tv, T^2v, \dots\}$ spans V. Suppose that V is finite dimensional, say of dimension n. Then V is a cyclic module if and only if for some $v \in V$, $\{v, Tv, \dots, T^{n-1}v\}$ span V, because T^n is already a combination of $\{1, T, T^2, \dots, T^{n-1}\}$ by the Cayley-Hamilton theorem. Therefore, $\{v, Tv, \dots, T^{n-1}v\}$ is a basis of V. For similar reasons then, the minimal polynomial must be of degree n as well, hence equal to the characteristic polynomial. Conversely, if the minimal polynomial is equal to the characteristic polynomial then V is cyclic. Namely, there is a vector $v \in V$ such that $\{v, Tv, \dots, T^{n-1}v\}$ is a basis of V. We leave that as an exercise in linear algebra.

Finally, given a vector $v \in V$, what is its annihilator? A linear transformation kills v iff it kills the linear subspace spanned by $\{v, Tv, T^2v, \dots\}$, namely, the cyclic submodule W generated by v. The annihilator is an ideal of $\mathbb{F}[x]$ generated by some polynomial $f(x)$. Then, $f(T)$ is zero on W and must be the polynomial of minimal degree vanishing on W. Thus, $\mathrm{Ann}(v) = \langle f(x) \rangle$, where $f(x)$ is the minimal polynomial of T restricted to W, where W is the minimal T-invariant subspace containing v.

Example 1.2.6. (Hom module). Let R be a commutative ring. Let M, N be R-modules, then $\mathrm{Hom}_R(M, N)$, the R-module homomorphisms from M to N, is itself an R-module, where for $f, g \in Hom_R(M, N), r \in R$, we let

$$(f + g)(m) := f(m) + g(m), \qquad (r \cdot f)(m) = r \cdot f(m).$$

We leave the verification as an exercise. Where is the commutativity of R used?

1.3. **Quotients and the isomorphism theorems.** Let M be an R-module and N a submodule of M. As a group, N is normal in M and so M/N is naturally an abelian group and $M \to M/N$ a group homomorphism. We claim that, further, M/N is naturally an R-module and $M \to M/N$ is an R-module homomorphism. Indeed, define

$$r \cdot (m + N) = r \cdot m + N.$$

This is well-defined. If $m' \in M$ is such that $m + N = m' + N$ then $m' = m + n$ for some $n \in N$. Then, $r \cdot m' + N = r \cdot (m + n) + N = r \cdot m + r \cdot n + N = r \cdot m + N$, because $r \cdot n \in N$. The module axioms now follow automatically. Similarly, it is immediate that $M \to M/N$ is a module homomorphism. The module M/N is called a **quotient module**.

Remark 1.3.1. Note at this point a subtle point. If R is a ring and I is a left ideal of R, I is a left R-submodule of R and so the quotient R/I is a left R-module. However, unless I is a *two-sided* ideal, the quotient R/I is not a ring.

Theorem 1.3.2 (First isomorphism theorem). *Let $f : M \to L$ be a homomorphism of R-modules and let N be a submodule of M contained in* $\mathrm{Ker}(f)$. *There is a unique homomorphism $F\colon M/N \to L$ rendering the diagram commutative:*

$$\begin{tikzcd}
M \arrow[rr, "f"] \arrow[dr, "can."'] & & L \\
& M/N \arrow[ur, "F"'] &
\end{tikzcd}$$

Furthermore, $\mathrm{Ker}(F) = \mathrm{Ker}(f)/N$.

Proof. The R-module homomorphism F, if it exists, would be in particular a homomorphism of groups. Thus, the definition of F is imposed on us:

$$F(m + N) = f(m).$$

We know F is a well-defined group homomorphism with kernel $\mathrm{Ker}(f)/N$. It only remains to check that is a homomorphism of modules. But $F(r(m + N)) = F(rm + N) = f(rm) = rf(m) = rF(m + N)$ and the proof is complete. □

As with groups, this theorem is the basis for a series of results. The proofs are almost identical to those given for groups. In fact, some aspects are simpler because all subgroups are normal and so one can forego some verifications. On the other hand, one needs to check that every group homomorphism constructed in those proofs is also an R-module homomorphism, but that always follows without difficulty. We omit most details.

Corollary 1.3.3. *Suppose that f is surjective. Then* $M/\mathrm{Ker}(f) \cong L$.

Proof. Indeed, then F is surjective and its kernel is $\{0\} = \mathrm{Ker}(f)/\mathrm{Ker}(f)$. Thus, F is an isomorphism. □

Example 1.3.4. Let M be an R-module and $m \in M$. The map

$$R \to M, \qquad r \mapsto rm,$$

is an R-module homomorphism with kernel $\mathrm{Ann}(m)$. We conclude an isomorphism of R-modules

$$R/\mathrm{Ann}(m) \cong Rm.$$

Thus, every cyclic module is isomorphic to R/I for some left ideal I. Conversely, if I is a left-ideal, the R-module R/I is cyclic. For example, it is generated by $\bar{1} = 1 + I \in R/I$.

Theorem 1.3.5 (Second isomorphism theorem). *Let A, B, be submodules of a module M. Then $A+B=\{a+b : a\in A, b\in B\}$ is again a submodule, as is $A\cap B$, and we have an isomorphism of R-modules.*

$$A/(A\cap B)\cong (A+B)/B.$$

z

Theorem 1.3.6 (Third isomorphism theorem). *Let $A\subset B\subset M$ be modules over R. We have an isomorphism of R-modules*

$$(M/A)/(B/A)\cong M/B.$$

Theorem 1.3.7 (Correspondence theorem). *Let $f\colon M\to N$ be a surjective homomorphism of R-modules. There is a bijection between the submodules of M containing the kernel of f and submodules of N. It is given by $M_1\subset M\mapsto f(M_1)$ and $N_1\subset N\mapsto f^{-1}(N_1)$. This correspondence preserves sums and intersections. Further, for $M_1\supseteq M_2\supseteq \mathrm{Ker}(f)$ we have $M_1/M_2\cong f(M_1)/f(M_2)$.*

Example 1.3.8. Let M be an R-module and N a submodule. Then $\mathrm{Ann}(N)$ is a two-sided ideal of R (but, unless R is commutative, the annihilator $\mathrm{Ann}(m)$ of an *element m* of M is *not* a two-sided ideal, only a left ideal). Indeed, let $a\in\mathrm{Ann}(N)$ and $r\in R$. Given $n\in N$ we have $(ra)n=r(an)=r0=0$ and also $(ar)(n)=a(rn)=0$ because $rn\in N$ too. Thus, the quotient $R/\mathrm{Ann}(N)$ is a ring. The R-module N is naturally an $R/\mathrm{Ann}(N)$ module; given a coset $\bar a=a+\mathrm{Ann}(N)$ of $R/\mathrm{Ann}(N)$ define

$$\bar a\cdot n := a\cdot n.$$

This is well defined. Suppose that $a'+\mathrm{Ann}(N)=a+\mathrm{Ann}(N)$ then $a'=a+u$ for some $u\in\mathrm{Ann}(N)$. We find that $\bar a'\cdot n=(a+u)\cdot n=a\cdot n+u\cdot n=a\cdot n=\bar a\cdot n$, because $u\cdot n=0$. In the same way, if I is any two-sided ideal contained in $\mathrm{Ann}(N)$ then N is an R/I-module.

Now, let I be a two-sided ideal of R and M an R-module. The ideal I kills the R-module M/IM. Thus, M/IM is naturally an R/I module. This is a very useful observation.

1.4. **Constructions: direct sum, direct product, free modules.**

Let R be a ring and M_α, for α ranging over some index set I, a collection of R-modules. The **direct product** of the M_α is defined as

$$\prod_{\alpha\in I} M_\alpha := \{(m_\alpha)_{\alpha\in I} : m_\alpha\in M_\alpha,\forall\alpha\}.$$

About the notation $(m_\alpha)_{\alpha\in I}$. We are being a bit colloquial here and think about that as a vector whose α's coordinate belongs to M_α. More pedantically, we can form the disjoint union of the modules M_α, denoted $\coprod M_\alpha$ (we shall omit the precise definition of this and rely on our intuition here), and then we can think of $\prod_{\alpha\in I} M_\alpha$ as the collection of functions $\{f : I\to\coprod M_\alpha : f(\alpha)\in M_\alpha,\forall\alpha\}$. Given such a function we can write it as the vector $(f(\alpha))_{\alpha\in I}$ and given a vector $(m_\alpha)_{\alpha\in I}$ we define a function f by $f(\alpha)=m_\alpha$. This gives a rigorous interpretation to the vector notation. At any rate, we define addition on $\prod_{\alpha\in I} M_\alpha$ by

$$(m_\alpha)_{\alpha\in I}+(n_\alpha)_{\alpha\in I}=(m_\alpha+n_\alpha)_{\alpha\in I}$$

(which corresponds to addition of functions), and multiplication by a scalar $r\in R$ by

$$r(m_\alpha)_{\alpha\in I}=(rm_\alpha)_{\alpha\in I}$$

(that is, we multiply all the coordinates by r). The verification that this is an R-module is immediate. Further, let $\alpha_0\in I$. The map

$$M_{\alpha_0}\to\prod_{\alpha\in I} M_\alpha,\quad m_{\alpha_0}\mapsto (m_\alpha)_{\alpha\in I},$$

where $m_\alpha = m_{\alpha_0}$ if $\alpha = \alpha_0$ and otherwise $m_\alpha = 0$, is an injective module homomorphism. The projection

$$p_{\alpha_0} : \prod_{\alpha \in I} M_\alpha \to M_{\alpha_0}, \quad (m_\alpha)_{\alpha \in I} \mapsto m_{\alpha_0},$$

is a module homomorphism as well.

We define the **direct sum** of the modules M_α, denoted $\oplus_{\alpha \in I} M_\alpha$, as the submodule of $\prod_{\alpha \in I} M_\alpha$ comprised the vectors all whose coordinates, but finitely many, are zero. If I is a finite set then the direct sum and the direct product are the same, but not when I is infinite. We further note that the kernel of

$$p_i : M_1 \oplus M_2 \oplus \cdots \oplus M_n \to M_i, \qquad p_i((m_j)_{j=1}^n) = m_i,$$

is $M_1 \oplus \cdots \oplus M_{i-1} \oplus \{0\} \oplus M_{i+1} \oplus \cdots \oplus M_n$ and

$$(M_1 \oplus M_2 \oplus \cdots \oplus M_n)/(M_1 \oplus \cdots \oplus M_{i-1} \oplus \{0\} \oplus M_{i+1} \oplus \cdots \oplus M_n)q \cong M_i.$$

A more general isomorphism is the following. Let $N_i \subseteq M_i$ be submodules. Then,

$$(M_1 \oplus M_2 \oplus \cdots \oplus M_n)/(N_1 \oplus N_2 \oplus \cdots \oplus N_n) \cong (M_1/N_1) \oplus (M_2/N_2) \oplus \cdots \oplus (M_n/N_n).$$

A particular case of the direct sum construction is the case where each $M_\alpha = R$. In this case we also denote $\oplus_{\alpha \in I} R$ as $R^{\oplus I}$. It is called a **free** R-module. In general an R-module M is called free if $M \cong R^{\oplus I}$ for some set I (if I is empty we define $R^{\oplus I}$ to be the zero module).

Proposition 1.4.1. *Let M be a free R-module, say $M \cong R^{\oplus I}$ for some I. Then there exist elements $\{m_\alpha : \alpha \in I\} \subseteq M$ such that every element of M can be written uniquely as $\sum_{\alpha \in I} r_\alpha m_\alpha$ where almost all $r_\alpha = 0$.*

Conversely, if M is an R module and if there exist elements $\{m_\alpha : \alpha \in I\} \subseteq M$, such that every element of M can be written uniquely as $\sum_{\alpha \in I} r_\alpha m_\alpha$, where almost all $r_\alpha = 0$, then $M \cong R^{\oplus I}$.

Proof. Suppose $f : R^{\oplus I} \to M$ is an isomorphism. Let $m_\alpha = f(e_\alpha)$, where e_α is the vector whose α-coordinate is 1 and all whose other coordinates are zero. The identity $(r_\alpha)_{\alpha \in I} = \sum_\alpha r_\alpha e_\alpha$ holds in $R^{\oplus I}$, and shows that every element of $R^{\oplus I}$ is uniquely a linear combination of $\{e_\alpha : \alpha \in I\}$. It follows by applying f that every element of M is uniquely a linear combination of $\{m_\alpha : \alpha \in I\}$.

The converse is slightly less formal. Suppose that for some elements $\{m_\alpha : \alpha \in I\}$ of M, every element of M can be expressed uniquely as $\sum_\alpha r_\alpha(m) m_\alpha$, with coefficients $r_\alpha(m) \in R$ that are almost all zero. Define a map

$$M \to R^{\oplus I}, \qquad m \mapsto g(m) = (r_\alpha(m))_{\alpha \in I}.$$

This map is well-defined as almost all r_α are 0. Since $m_1 + m_2 = \sum_\alpha (r_\alpha(m_1) + r_\alpha(m_2)) m_\alpha$ it follows that $r_\alpha(m_1) + r_\alpha(m_2) = r_\alpha(m_1 + m_2)$. Thus, $g(m_1 + m_2) = (r_\alpha(m_1 + m_2))_\alpha = (r_\alpha(m_1) + r_\alpha(m_2))_\alpha = (r_\alpha(m_1))_\alpha + (r_\alpha(m_2))_\alpha = g(m_1) + g(m_2)$. The argument for $g(rm) = rg(m)$ is very similar.

Clearly, $g(m) = 0$ implies $r_\alpha(m) = 0$ for all α. But then $m = \sum_\alpha r_\alpha m_\alpha = 0$, so g is injective. Finally, given $(r_\alpha)_\alpha \in R^{\oplus I}$, let $m = \sum_\alpha r_\alpha m_\alpha$, which is well-defined because almost all $r_\alpha = 0$. Then $g(m) = (r_\alpha)_\alpha$ and so g is also surjective. □

Lemma 1.4.2. *Let R be a commutative ring. Let I and J be sets. Then $R^{\oplus I} \cong R^{\oplus J}$ if and only if I and J have the same cardinality.*

Proof. Suppose that I and J have the same cardinality. By definition that means that there is a bijection $\psi : J \to I$. Define

$$f : R^{\oplus I} \to R^{\oplus J}, \qquad f((r_\alpha)_{\alpha \in I}) = (s_\beta)_{\beta \in J},$$

where,

$$s_\beta = r_{\psi(\beta)}.$$

f is a module homomorphism. Indeed $f((r_\alpha)_\alpha + (r'_\alpha)_\alpha) = f((r_\alpha + r'_\alpha)_\alpha) = (r_{\psi(\beta)} + r'_{\psi(\beta)})_\beta = (r_{\psi(\beta)})_\beta + (r'_{\psi(\beta)})_\beta = f((r_\alpha)_\alpha) + f((r'_\alpha)_\alpha)$. Also, $f(r(r_\alpha)_\alpha) = f((rr_\alpha)_\alpha) = (rr_{\psi(\beta)})_\beta = r(r_{\psi(\beta)})_\beta = rf((r_\alpha)_\alpha)$. In contrast to this argument which entirely formal, the converse direction is much deeper.

Let A be a maximal ideal of R. One checks that for every collection of modules $\{M_\alpha\}$ we have

$$A \cdot \bigoplus_\alpha M_\alpha = \cdot \bigoplus_\alpha AM_\alpha.$$

Therefore,

$$A \cdot R^{\oplus I} = A \cdot \bigoplus_{i \in I} R = \bigoplus_{i \in I} A.$$

Since $R^{\oplus I} \cong R^{\oplus J}$ we have $A \cdot R^{\oplus I} \cong A \cdot R^{\oplus J}$ and so $R^{\oplus I}/A \cdot R^{\oplus I} \cong R^{\oplus J}/A \cdot R^{\oplus J}$. But,

$$R^{\oplus I}/A \cdot R^{\oplus I} \cong \bigoplus_{i \in I}(R/A) = (R/A)^{\oplus I},$$

and consequently

$$(R/A)^{\oplus I} \cong (R/A)^{\oplus J}.$$

However, R/A is a field and $(R/A)^{\oplus I}$ is a vector-space over it of dimension $|I|$. From the theory of vector spaces, we conclude $|I| = |J|$. □

Finally, we discuss the notion of internal direct sum. Let M be a module and $\{M_\alpha : \alpha \in I\}$ be family of submodules of M such that for every α, $M_\alpha \cap \sum_{\beta \neq \alpha} M_\beta = \{0\}$. Then the sum $\sum_\alpha M_\alpha$ is called in **internal direct sum**, or simply "direct sum". This is justified by the following lemma.

Lemma 1.4.3. *Let $\{M_\alpha : \alpha \in I\}$ be family of submodules of M such that for every α, $M_\alpha \cap \sum_{\beta \neq \alpha} M_\beta = \{0\}$, then*

$$\sum_{\alpha \in I} M_\alpha \cong \oplus_{\alpha \in I} M_\alpha.$$

Therefore, we allow the abuse of notation and denote the internal direct sum also by $\oplus_{\alpha \in I} M_\alpha$ and may refer to it as simply "direct sum".

Corollary 1.4.4. *Let M be a free R-module, say $M \cong R^{\oplus I}$. We say that M has* **rank** *$|I|$. This notion is well-defined.*

1.5. **The Chinese Remainder Theorem.**

Theorem 1.5.1. *Let R be a commutative ring and M an R-module. Let $I_1, \dots, I_k$ be relatively prime ideals of R. The homomorphism of R-modules*

$$M \to M/I_1M \oplus \cdots M/I_kM, \qquad m \mapsto (m + I_1M, \dots, m + I_kM),$$

is a surjective homomorphism with kernel $I_1M \cap \cdots \cap I_kM = (I_1 \cdots I_k)M$.

We proved this theorem for the particular case of $M = R$. In that case, all quotient modules are in fact rings and we proved that the map is a ring homomorphism. In our case we of course make no such claim. One checks that essentially the same proof works. We remark that the proof gives directly $I_1M \cap \cdots \cap I_kM = (I_1 \cdots I_k)M$ without needing to show first that $I_1M \cap \cdots \cap I_kM = (I_1 \cap \cdots \cap I_k)M$; the equality $I_1M \cap \cdots \cap I_kM = (I_1 \cap \cdots \cap I_k)M$ holds under the assumptions of CRT, and is a consequence of it, but in the general situation it may fail; in fact, even for ideals $\mathfrak{a}, \mathfrak{b}, \mathfrak{c}$ of a ring R it need not be the case that $(\mathfrak{a} \cap \mathfrak{b})\mathfrak{c} = \mathfrak{a}\mathfrak{c} \cap \mathfrak{b}\mathfrak{c}$.

The Chinese remainder theorem looks formal, but it contains some interesting information as the next example shows.

Example 1.5.2. Let (V, T) be a finite dimensional vector space over a field $\mathbb{F}$ and $T : V \to V$ a linear transformation. Let $m(x)$ be the minimal polynomial of T. We consider V as an $\mathbb{F}[x]$-module. Since $m(T)$ annihilates V we may also consider V as an $\mathbb{F}[x]/(m(x))$-module. Consider the decomposition of $m(x)$ into irreducible polynomials over $\mathbb{F}$:

$$m(x) = f_1(x)^{r_1} \cdots f_k(x)^{r_k},$$

where the $f_i(x)$ are distinct monic irreducible polynomials and the r_i positive. Let

$$I_i = (f_i(x)^{r_i}).$$

The ideals $I_1, \dots, I_k$ satisfy the conditions of the CRT and $I_1 \cdots I_k = (m(x))$. We conclude that

$$V \cong V/(f_1(x)^{r_1})V \oplus \cdots \oplus V/(f_k(x)^{r_k})V.$$

In fact, if e_i is the i-th element used in the proof of the CRT (so that $e_i \equiv 0 \pmod{f_j^{r_j}}$ for $j \neq i$ and $e_i \equiv 1 \pmod{f_i^{r_i}}$) then e_iV, which is a submodule of V maps isomorphically onto $V/(f_i(x)^{r_i})V$. That give us a computable way to view the quotient module $V/(f_i(x)^{r_i})V$ also as a submodule of V. The restriction of T to the submodule $V/(f_i(x)^{r_i})V$ is of course killed by $f_i(x)^{r_i}$. But since $m(x) = f_1(x)^{r_1} \cdots f_k(x)^{r_k}$ we conclude that it cannot be killed by a smaller power of f_i (as the minimal polynomial is the lcm of the minimal polynomials of the subspaces $V/(f_i(x)^{r_i})V$). To summarize, we have deduced the so-called **Primary Decomposition Theorem**.

Suppose that the minimal polynomial of T factors as $m(x) = f_1(x)^{r_1} \cdots f_k(x)^{r_k}$. There are subspaces $V_1, \dots, V_k$ of V that are T-invariants such that $V = V_1 \oplus \cdots \oplus V_k$ and such that the minimal polynomial of T on V_i is $f_i(x)^{r_i}$. Furthermore, if $e_i(x)$ is a polynomial such that $f_j^{r_j} | e_i$ for $j \neq i$ and $f_i^{r_i} | (e_i - 1)$, then $V_i = e_i(T)V$.

2. Modules over PID

The main goal of this section is to give a structure theorem for certain R-modules, when R is a PID. The only requirement the modules M have to satisfy is that they are **finitely generated**. Namely, that there are finitely many elements $m_1, \dots, m_n$ in M such that every element of M is of the form $\sum_{i=1}^n r_i m_i$ for some $r_i \in R$. (We say that M is generated by the elements $m_1, \dots, m_n$). Note that there is no requirement that the r_i be unique. For example, $\mathbb{Z}/4\mathbb{Z}$ is a finitely generated $\mathbb{Z}$-module. Because taking the element 1 as an element of the module $\mathbb{Z}/4\mathbb{Z}$ every element of $\mathbb{Z}/4\mathbb{Z}$ can be written as $0 \cdot 1, 1 \cdot 1, 2 \cdot 1, 3 \cdot 1$. But note, for example, that $0 \cdot 1$ can also be written as say $24 \cdot 1$, and so-on, so this writing is far from unique. Note also that every element of $\mathbb{Z}/4\mathbb{Z}$ can be written as a multiple of the element $3 \in \mathbb{Z}/4\mathbb{Z}$. Indeed $0 = 0 \cdot 3, 1 = 3 \cdot 3, 2 = 2 \cdot 3, 3 = 1 \cdot 3$. So a module has many different sets of generators.

We also note the easily proven fact that M is generated by n elements if and only if there is a surjective R-module homomorphism $R^n \to M$. Indeed, $e_i \mapsto m_i$ and so on.

2.1. **Rank.**

Let R be a commutative integral domain. There is no need to assume in this section that R is a PID. Let M be an R module. A subset $\{m_\alpha : \alpha \in I\}$ of elements of M is called **linearly independent** if the only finite linear combination $\sum r_\alpha m_\alpha$ (almost all r_α are 0) that equals 0 is the one where all the $r_\alpha = 0$. That is, there are no non-trivial linear relations between the m_α. Note that then the submodule $\langle \{m_\alpha : \alpha \in I\} \rangle$ is free and isomorphic to $R^{\oplus I}$. We define the **rank** of M to be the supremum of the cardinalities of linearly independent subsets of M. We shall denote it $\mathrm{rk}(M)$. Some care has to be taken with relying too much on intuition from the theory of vector spaces: (1) A module my be generated by 1 element x and yet $\{x\}$ may be linearly dependent set; (2) A maximal linear independent set need not be a basis.

As we have already defined the notion of rank for free modules, we better check that our definitions agree.

Lemma 2.1.1. *The rank of $R^{\oplus I}$ is I, in the sense that there is an independent set of cardinality I in $R^{\oplus I}$ and there isn't an independent set of larger cardinality.*

Proof. Let e_i be the element of $R^{\oplus I}$ all whose coordinates are 0 except for the i-th coordinate which is 1. As $\sum r_\alpha e_\alpha$ is the vector whose α coordinate is r_α, such a sum is 0 if and only if $r_\alpha = 0$ for all α. Thus, $\{e_\alpha : \alpha \in I\}$ is linearly independent.

To prove this is the maximal possible cardinality we shall use the theory of vector spaces. let F be the field of fractions of R. Recall that

$$F = \left\{ \frac{a}{b} : a, b \in R, b \neq 0 \right\};$$

fractions are identified in the usual way, that is $a/b = c/d$ if $ad - bc = 0$, and we have defined the operations in similarity to rational numbers. F is a field in which R embeds as the fractions $\{r/1 : r \in R\}$ and so we have an embedding $R^{\oplus I} \rightarrow F^{\oplus I}$.

Let $\{v_\alpha : \alpha \in J\}$ be a linearly independent set in $R^{\oplus I}$. We claim that it is a linearly independent set in $F^{\oplus I}$ too. Suppose that $\sum_{\alpha \in J} \frac{r_\alpha}{s_\alpha} v_\alpha = 0$ in $F^{\oplus I}$. As only finitely many of the coefficients $\frac{r_\alpha}{s_\alpha}$ are non-zero, we may find some $S \in R$, $S \neq 0$, such that $s_\alpha | S$ for all α such that $r_\alpha \neq 0$. Then, $\sum (S \cdot \frac{r_\alpha}{s_\alpha}) \cdot v_\alpha = 0$ in $R^{\oplus I}$, which implies that $S\frac{r_\alpha}{s_\alpha} = (S/s_\alpha) r_\alpha = 0$ for all α. Thus, $r_\alpha = 0$ for all α and also $\frac{r_\alpha}{s_\alpha} = 0$ for all α.

Now, as our independent set $\{e_\alpha : \alpha \in I\}$ is clearly a basis for $F^{\oplus I}$, it follows that any other independent set has cardinality at most that of I. Therefore, $|J| \leq |I|$. □

Corollary 2.1.2. *Every two maximal linearly independent subset of a module M have the same cardinality.*

Proof. Let $\{x_\alpha : \alpha \in I\}$ be a maximal linearly independent set. Let N be the submodule of M spanned by $\{x_\alpha : \alpha \in I\}$. Let $m \in M$. Then $\{m\} \cup \{x_\alpha : \alpha \in I\}$ is linearly dependent and so, for some r_α and non-zero r we have $rm + \sum r_\alpha x_\alpha = 0$. Thus, for some non-zero r we have $rm \in N$.

Let now $\{y_\gamma : \gamma \in J\}$ be another maximal linearly independent set. From the argument we gave, there are non-zero elements r_γ such that $\{r_\gamma y_\gamma : \gamma \in J\}$ is a subset of N, which is still independent. Since $N \cong R^{\oplus I}$ we must have $|J| \leq |I|$. But, reversing the role of the two independent sets we get the opposite inequality. □

The following lemma is left as an exercise. (This doesn't mean that the its content is less important than other results we have proven.)

Lemma 2.1.3. *We have the following facts concerning the rank of a module M.*

(1) The rank of M is zero if and only if M is torsion.
(2) The rank of M is equal to the rank of $M/Tors(M)$.
(3) Let N be a submodule of M then $rk(N) \leq rk(M)$.
(4) Let $0 \rightarrow M_1 \rightarrow M \rightarrow M_2 \rightarrow 0$ be an exact sequence of R-modules. Then, $rk(M) = rk(M_1) + rk(M_2)$.

2.2. **The Elementary Divisors Theorem (EDT).**

The following theorem is one of the most useful theorems in the theory of modules, especially in applications.

Theorem 2.2.1. *Let R be a PID and M a free R-module of finite rank n. Let $N \subseteq M$ be a submodule. Then:*

(1) N is a free module of rank m, $m \leq n$.

(2) There exists a basis $\{y_1, \dots, y_n\}$ of M and non-zero scalars $a_1 | a_2 | \cdots | a_m$ in R, such that $\{a_1 y_1, a_2 y_2, \dots, a_m y_m\}$ is a basis of N.

Corollary 2.2.2. *Let L and M be free R modules of finite rank. Let $f \colon L \to M$ be an R-module homomorphism. There are bases $y_1, \dots, y_n$ of M and $z_1, \dots, z_t$ of L such that with respect to these bases f has the form*

$$\begin{pmatrix} a_1 & & & & & \\ & \ddots & & & & \\ & & a_m & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

Proof. (Corollary). Let $N = f(L)$, a submodule of M, and choose $y_1, \dots, y_n$ in M and $a_1 | \cdots | a_m$ as in the EDT. Let $z_i \in L$ such that $f(z_i) = a_i y_i$, $i = 1, \dots, m$. Let $z_{m+1}, \dots, z_t$ be a basis for $K = \operatorname{Ker}(f)$. Let $H = \langle z_1, \dots, z_m \rangle$. We claim that

$$H \cap K = 0, \quad H + K = L.$$

First, let $h \in H \cap K$. Then $h = \sum_{i=1}^m r_i z_i$ and so $f(h) = \sum_{i=1}^m r_i a_i y_i$. Since $h \in K$, $f(h) = 0$ and so, because $\{a_1 y_1, \dots, a_m y_m\}$ are a basis for N, each $r_i = 0$ and it follows that $h = 0$. Next we show that $H + K = L$. Given $\ell \in L$, since $f(\ell) \in N$, there are r_i such that $f(\ell) = \sum_{i=1}^m r_i a_i y_i$ and $f(\ell) = f(\sum_{i=1}^m r_i z_i)$. Now, $\sum_{i=1}^m r_i z_i \in H$ and $f(\ell - \sum_{i=1}^m r_i z_i) = 0$. Thus, $\ell = (\sum_{i=1}^m r_i z_i) + (\ell - \sum_{i=1}^m r_i z_i)$ exhibits ℓ as an element of $H + K$. □

Proof. (EDT). If $N = 0$ then N is free (and $m = 0$). Assume henceforth that $N \neq 0$.

Lemma 2.2.3. *There is a homomorphism $\varphi : M \to R$, a scalar $0 \neq a_1 \in R$ and an element $y \in N$ such that $\varphi(y) = a_1$ and for every $\psi : M \to R$, $a_1 | \psi(y)$. Moreover $\varphi(N) = Ra_1$.*

Proof. Let

$$\Sigma = \{\varphi(N) : \varphi \in \operatorname{Hom}_R(M, R)\}.$$

Σ is a collection of ideals (= sub R-modules of R) of R and $(0) \in \Sigma$ (take $\varphi = 0$). As R is a PID, we may choose for every φ an element $a_\varphi \in R$ such that $\varphi(N) = \langle a_\varphi \rangle$. Assume that Σ has no maximal elements with respect to inclusion. Then we get $\varphi_1, \varphi_2, \dots$ such that $\langle a_{\varphi_1} \rangle \subsetneqq \langle a_{\varphi_2} \rangle \subsetneqq \dots$. But $\cup_{i=1}^\infty \langle a_{\varphi_i} \rangle$ is an ideal and so equal to $\langle a \rangle$ for some $a \in R$. As $\cup_{i=1}^\infty \langle a_{\varphi_i} \rangle = \langle a \rangle$, $a \in \langle a_{\varphi_i} \rangle$ for some i and then $\langle a \rangle \subseteq \langle a_{\varphi_i} \rangle \subsetneqq \langle a_{\varphi_{i+1}} \rangle \subseteq \langle a \rangle$ and that's a contradiction.

Let therefore $\varphi : M \to R$ be a homomorphism such that $\langle a_\varphi \rangle$ is a maximal element of Σ (possibly R itself). Let $a_1 = a_\varphi$ and choose $y \in N$ such that $\varphi(y) = a_1$. We now show that φ, a_1 and y have the desired properties.

First, using an isomorphism $g : M \to R^n$, composed with projection on the i-th coordinate, $p_i \circ g : M \cong R^n \to R$, we see that there is a coordinate i such that $(p_i \circ g)(N) \neq 0$. If $a_1 = 0$ then $\langle a_1 \rangle \subsetneqq (p_i \circ g)(N)$, contradicting the maximality of $\langle a_1 \rangle$. Thus, $a_1 \neq 0$.

Next, let $\psi \in \operatorname{Hom}_R(M, R)$ and let $b = \psi(y)$. Let $d = \gcd(a_1, b) = r_1 a_1 + rb$, for some $r_1, r \in R$. Consider $\alpha := r_1 \varphi + r\psi \in \operatorname{Hom}_R(M, R)$. We calculate $\alpha(y) = (r_1 \varphi + r\psi)(y) = r_1 a_1 + rb = d$. Since $d | a_1$ and it follows that $\alpha(N) \supseteq \langle a_1 \rangle$. Since $\langle a_1 \rangle$ has a maximality property relative to Σ we must have $\langle d \rangle = \langle a_1 \rangle$, which implies $a_1 | b$. □

Still keeping the notation of the lemma, it follows in particular that

$$a_1 | (p_i \circ g)(y), \forall i.$$

That implies that there is an element $y_1 \in M$ such that

$$y = a_1 y_1 \in N.$$

Note that $\varphi(y_1) = 1$ because $a_1(\varphi(y_1) - 1) = 0$.

Lemma 2.2.4. *We have the direct sum decompositions:*

$$M = \langle y_1 \rangle + \mathrm{Ker}(\varphi), \qquad N = \langle a_1 y_1 \rangle \oplus (\mathrm{Ker}(\varphi) \cap N).$$

Proof. (Lemma). The argument is very similar to the one given in the proof of Corollary 2.2.2 and so we shall omit it. □

We now prove part (1) of the theorem, by induction on the rank m of N. The fact that $m \leq n$ is clear from the definition of rank. If $m = 0$ then N is torsion and, since M is torsion-free, $N = \{0\}$.

Consider $N \cap \mathrm{Ker}(\varphi)$. If it has rank ℓ then $N = \langle a_1 y_1 \rangle \oplus (\mathrm{Ker}(\varphi) \cap N)$ has rank at least $\ell + 1$. Thus, $\ell \leq m - 1$. Using induction for the submodule $N \cap \mathrm{Ker}(\varphi)$ of M , we conclude that $N \cap \mathrm{Ker}(\varphi)$ is free of rank ℓ and so clearly N is free of rank $\ell + 1$. (It now follows that $\ell + 1 = m$.)

We now prove part (2) by induction on $n = \mathrm{rk}(M)$. The arguments above and part (1) show that $N \cap \mathrm{Ker}(\varphi)$ is a free submodule of rank $m - 1$ of the free module $\mathrm{Ker}(\varphi)$ of rank $n - 1$. Induction gives that there exists a basis $y_2, \dots, y_n$ of $\mathrm{Ker}(\varphi)$ and non-zero scalars $a_2 | a_3 | \cdots | a_m$ such that $N \cap \mathrm{Ker}(\varphi)$ is free with a basis $a_2 y_2, \dots, a_m y_m$. It follows that N is free with a basis $a_1 y_1, a_2 y_2, \dots, a_m y_m$. The only thing left to show is that $a_1 | a_2$. To show that, apply Lemma 2.2.3 to the R-module homomorphism,

$$\psi : M \to R, \qquad \psi(\sum b_i y_i) = b_1 + b_2.$$

As $a_1 = \psi(a_1 y_1)$ we have $\psi(N) \supseteq \langle a_1 \rangle$ and by maximality $\psi(N) = \langle a_1 \rangle$. In particular $\psi(a_2 y_2) = a_2 \in \langle a_1 \rangle$, which gives $a_1 | a_2$. □

2.3. The structure theorem for finitely generated modules over PID: Existence.

Theorem 2.3.1. *(Existence of decomposition in invariant factors form) Let R be a PID and let M be a finitely generated R-module.*

1. *$M \cong R^r \oplus R/(a_1) \oplus \cdots \oplus R/(a_m)$ for some $r \geq 0$ and some non-zero $a_i \in R$ satisfying $a_1 | a_2 | \cdots | a_m$.*
2. *M is torsion-free if and only if M is free. In fact,*

$$Tors(M) \cong R/(a_1) \oplus \cdots \oplus R/(a_m).$$

M is torsion if and only if $r = 0$ and then $Ann(M) = (a_m)$.

Proof. Let $x_1, \dots, x_n$ be generators for M. The function

$$R^n \longrightarrow M, \qquad (r_1, \dots, r_n) \mapsto \sum_i r_i x_i,$$

is a surjective R-module homomorphism. Let N be its kernel, then $M \cong R^n / N$. Using EDT, there exists a basis $\{y_1, \dots, y_n\}$ for R^n and non-zero scalars $a_1 | a_2 | \cdots | a_m$ in R, such that $\{a_1 y_1, \dots, a_m y_m\}$ is a basis for N. Therefore,

$$\begin{aligned} M &\cong R^n / N \\ &\cong R y_1 \oplus \cdots \oplus R y_n / R a_1 y_1 \oplus \cdots \oplus R a_m y_m \oplus \{0\} \oplus \cdots \oplus \{0\} \\ &\cong (\oplus_{i=1}^m R/(a_i)) \oplus R^{n-m}. \end{aligned}$$

This gives us part (1) of the theorem.

Note that $\mathrm{Ann}(R/(a_i)) = (a_i)$, which is a non-zero ideal. Thus, if M is torsion-free then $m = 0$ and so M is free. Conversely, a free module is always torsion free.

In general for modules over integral domains, $\mathrm{Tors}(M_1 \oplus M_2) = \mathrm{Tors}(M_1) \oplus \mathrm{Tors}(M_2)$. As R is torsion free, if M is torsion then we must have $r = 0$, in which case $\mathrm{Ann}(M) = \cap_{i=1}^m (a_i) = (a_m)$. Clearly, if $r = 0$, M is torsion. □

Remark 2.3.2. If any of the a_i are units then $R/(a_i) \cong \{0\}$ and so we may remove such an a_i altogether. Thus, we may assume that none of the a_i are units. Then, the ideals $(a_i), i = 1, \dots, m$, are uniquely determined by M, as we shall see shortly, and so is r. The elements $a_1, \dots, a_m$ (that are determined up to units) are called the **invariant factors** of M.

Let us see what the decomposition theorem gives in familiar cases.

Corollary 2.3.3. *$R = \mathbb{F}$ a field. Then an $\mathbb{F}$-module V is an $\mathbb{F}$-vector space. The only torsion $\mathbb{F}$-module is the zero vector space $\{0\}$. The theorem thus states that a finitely generated $\mathbb{F}$-vector space is isomorphic to $\mathbb{F}^r$, where r is of course the dimension of V.*

Corollary 2.3.4. *Let $R = \mathbb{Z}$. The theorem tells us that every finitely generated abelian group M is isomorphic to*

$$\mathbb{Z}^r \oplus \mathbb{Z}/a_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/a_m\mathbb{Z},$$

where r - the rank - is uniquely determined by M, and the a_i are unique positive integers such that $a_1|a_2|\cdots|a_m$ and $a_1 > 1$.

Corollary 2.3.5. *Let $\mathbb{F}$ be a field, V a finite dimensional vector space over $\mathbb{F}$ and $T : V \to V$ a linear transformation. We view (V, T) as an $\mathbb{F}[x]$-module. In this case we must have $r = 0$ (because already the dimension of $\mathbb{F}[x]$ over $\mathbb{F}$ is infinite) and so V is torsion. We find that*

$$V \cong \oplus_{i=1}^m \mathbb{F}[x]/(a_i(x)), \tag{1}$$

for unique monic non-constant polynomials $a_i(x)$ satisfying

$$a_1(x)|a_2(x)|\cdots|a_m(x).$$

Furthermore, $Ann(V) = (a_m(x))$, and so $a_m(x)$ is the minimal polynomial of T.

Fix i and, via the isomorphism above, consider $\mathbb{F}[x]/(a_i(x))$ as a subspace V_i of V. Let

$$a_i(x) = b_0 + b_1x + \cdots b_{s-1}x^{s-1} + x^s.$$

Consider the action of T on the T-invariant subspace V_i. It corresponds to the action of x on the $\mathbb{F}[x]$-module $\mathbb{F}[x]/(a_i(x))$. The latter has a basis over $\mathbb{F}$ given by $1, x, \dots, x^{s-1}$. If we let $v \in V_i$ correspond to $1 \in \mathbb{F}[x]/(a_i(x))$, then we find that V_i has a basis

$$v, Tv, \dots, T^{s-1}v$$

and the minimal polynomial of T on V_i (which is just the generator of the annihilator of the $\mathbb{F}[x]$-module $\mathbb{F}[x]/(a_i(x))$, that is the ideal $(a_i(x))$) is $a_i(x)$. In particular, it is also equal to the characteristic polynomial of T on V_i. Furthermore, considering the action of x on the basis $\{1, x, \dots, x^{s-1}\}$, we find that the action of T on the basis $\{v, Tv, \dots, T^{s-1}v\}$ is given by the matrix

$$\begin{pmatrix} 0 & 0 & \dots & 0 & -b_0 \\ 1 & 0 & & 0 & -b_1 \\ & 1 & & \vdots & \vdots \\ & & & 1 & -b_{s-1} \end{pmatrix}$$

We summarize some of our discussion: Under the decomposition in (1), we have

- *The minimal polynomial of T is $a_m(x)$;*
- *The characteristic polynomial of T is the product $a_1(x)a_2(x)\cdots a_m(x)$.*

Corollary 2.3.6. *(Existence of decomposition in elementary divisors form) Let M be a finitely generated module over a PID R. Then,*

$$M \cong R^r \oplus R/(p_1^{\alpha_1}) \oplus \cdots \oplus R/(p_t^{\alpha_t}),$$

where the p_i are irreducible (not necessarily distinct) elements of R and the α_i are positive integers.

Proof. Using the invariant factor decomposition, we reduce to the case $M = R/(a)$. Let $a = up_1^{b_1} \cdots p_d^{b_d}$ be the decomposition of a into powers of distinct irreducible elements, where u is a unit. Then, by CRT

$$R/(a) \cong \bigoplus_{i=1}^{d} R/(p_i^{b_i}).$$

□

Remark 2.3.7. In turn, existence of decomposition in elementary divisors form implies decomposition in invariant factors form. Indeed, to simplify the notation let us assume that there are only three irreducible elements appearing. Suppose that altogether the powers of p_1 appearing in the decomposition are $p_1^{a_1}, p_1^{a_2}, \dots, p_1^{a_\ell}$, where $a_1 \leq a_2, \leq \cdots \leq a_\ell$; that the powers of p_2 appearing in the decomposition are $p_2^{b_1}, p_2^{b_2}, \dots, p_2^{b_m}$, where $b_1 \leq b_2, \leq \cdots \leq b_m$; that the powers of p_3 appearing in the decomposition are $p_3^{c_1}, p_3^{c_2}, \dots, p_3^{c_n}$, where $c_1 \leq c_2, \leq \cdots \leq c_m$. Write a table of the following form

$$\begin{array}{|c|c|c|c|c|c|c|c|}
 & & p_1^{a_1} & p_1^{a_2} & p_1^{a_3} & \dots & \dots & p_1^{a_\ell} \\
 & & & & p_2^{b_1} & p_2^{b_2} & \dots & p_2^{b_m} \\
p_3^{c_1} & p_3^{c_2} & \dots & & & & \dots & p_3^{c_n}
\end{array}$$

We make sure to align the rows to the right. Then, using CRT "in reverse" we get a decomposition in invariant factor forms $d_1(x)|d_2(x)| \cdots |d_n(x)$, where $d_1(x)$ is the product of the elements appearing in the first column, $d_2(x)$ is the product of those in the second column, and so on, $d_n(x)$ is the product $p_1^{a_\ell} p_2^{b_m} p_3^{c_n}$. It is clear how to extend this method to more (or less) than 3 irreducible elements.

2.4. **The structure theorem for finitely generated modules over PID: Uniqueness.** We now prove the uniqueness of decomposition in elementary divisors form. We will state, but not prove, the uniqueness for decomposition in invariant factors form. It can be deduced from uniqueness in elementary divisors form using Remark 2.3.7

Theorem 2.4.1. *(Uniqueness of decomposition in elementary divisors form) Let R be a PID and suppose that*

$$R^{r_1} \oplus \bigoplus_i R/(p_i^{a_i}) \cong R^{r_2} \oplus \bigoplus_j R/(p_j^{b_j}),$$

where p_i, q_j are irreducible, a_i, b_j positive integers. Then, after re-indexing if required, we have that the number of summands in each decomposition is the same and

$$r_1 = r_2, \quad , p_i \sim q_i, \quad a_i = b_i, \ \forall i,$$

where $p \sim q$ means that p and q are associate (differ by a unit).

Proof. Let M_1 denote the l.h.s. and M_2 the r.h.s. Since $M_1 \cong M_2$, we have $\operatorname{Tors}(M_1) \cong \operatorname{Tors}(M_2)$ and therefore $R^{r_1} \cong M_1/\operatorname{Tors}(M_1) \cong M_2/\operatorname{Tors}(M_2) \cong R^{r_2}$. By Lemma 1.4.2, $r_1 = r_2$.

It remains to study the torsion part and so we may now assume that

$$M_1 = \bigoplus_i R/(p_i^{a_i}) \cong M_2 = \bigoplus_j R/(p_j^{b_j}).$$

The method of proof is of interest in itself. It teaches us how to extract those parts in the sum associated with a particular prime.

Let p be a prime of R and M an R-module. Let,

$$M_p = \{m \in M : \exists b > 0, p^b m = 0\}.$$

M_p is called the p-**primary** component of M. We note that M_p is a submodule of M. The following properties are easy to check:

- If $M \cong N$ then $M_p \cong N_p$.
- $(M \oplus N)_p \cong M_p \oplus N_p$.

We also claim that

$$(R/(p_i^{a_i}))_p = \begin{cases} R/(p_i^{a_i}) & p \sim p_i, \\ \{0\} & p \nsim p_i. \end{cases}$$

Indeed, the first case where $p \sim p_i$ is immediate as $(p^{a_i}) = (p_i^{a_i})$ and so p^{a_i} kills every element of $R/(p_i^{a_i})$. Suppose then that $p \nsim p_i$. If $m \in R/(p_i^{a_i})$ and $p^b m = 0$, we have $p^b \in \mathrm{Ann}(m)$ and $p_i^{a_i} \in \mathrm{Ann}(m)$ and so also $\gcd(p^b, p_i^{a_i}) \in \mathrm{Ann}(m)$. But the gcd is 1, and so $1 \cdot m = 0$, which implies $m = 0$; we have shown the second case.

Using these results we conclude that

$$(\bigoplus_i R/(p_i^{a_i}))_p = \bigoplus_{\{i:p\sim p_i\}} R/(p^{a_i}).$$

Therefore, since $M_i \cong M_2$ we conclude that

$$\bigoplus_{\{i:p\sim p_i\}} R/(p^{a_i}) \cong \bigoplus_{\{j:p\sim q_j\}} R/(p^{b_j}).$$

We therefore reduced to proving the following statement: Let p be an irreducible element of R, a_i, b_j positive integers. If

$$\bigoplus_{i=1}^{n_1} R/(p^{a_i}) \cong \bigoplus_{i=1}^{n_2} R/(p^{b_j}), \tag{2}$$

where,

$$0 < a_1 \leq a_2 \leq \cdots \leq a_{n_1}, \qquad 0 < b_1 \leq b_2 \leq \cdots \leq b_{n_2},$$

then $n_1 = n_2$ and $a_i = b_i$ for all i. To show that we will use the following Lemma.

Lemma 2.4.2. *Let $M = R/(p^c)$ and $d \geq 0$ an integer. Then*

$$p^d M/p^{d+1}M \cong \begin{cases} \{0\} & d \geq c \\ R/pR & d < c. \end{cases}$$

Proof. (Lemma) Indeed, if $d \geq c$ then every element of M is killed by p^d and so $p^d M/p^{d+1}M$ is just $\{0\}$. For $d < c$ we have $p^d M/p^{d+1}M = p^d R/(p^{d+1}, p^c)R = (p^d)/(p^{d+1})$. The map $R \to (p^d)/(p^{d+1})$ given by multiplication by p^d is a surjective homomorphism with kernel (p) and so $R/(p) \cong p^d M/p^{d+1}M$ in this case. □

Note that $\mathbb{F} := R/(p)$ is a field. Using the lemma, coming back to the general case as in (2), we conclude that

$$p^d M_1/p^{d+1}M_1 \cong \mathbb{F}^{m_1(d)}, \qquad m_1(d) := \sharp\{a_i : a_i > d\},$$

and

$$p^d M_2/p^{d+1}M_2 \cong \mathbb{F}^{m_2(d)}, \qquad m_2(d) := \sharp\{b_i : b_i > d\}.$$

Since $p^d M_1/p^{d+1}M_1 \cong p^d M_2/p^{d+1}M_2$, we must have $m_1(d) = \dim_{\mathbb{F}}(\mathbb{F}^{m_1(d)}) = \dim_{\mathbb{F}}(\mathbb{F}^{m_2(d)}) = m_2(d)$, and that for every $d \geq 0$. Note for example that this implies that $\sharp\{a_i : a_i = d\} = m_1(d-1) - m_1(d) = m_2(d-1) - m_2(d) = \sharp\{b_i : b_i = d\}$. It follows that there is the same number of a_i and b_j and equalities $a_i = b_i$, for all i. □

Theorem 2.4.3. *(Uniqueness in invariant factors form) Suppose that*

$$R^{r_1} \oplus \bigoplus_{i=1}^{n_1} R/(a_i) \cong R^{r_2} \oplus \bigoplus_{i=1}^{n_2} R/(b_i),$$

where the a_i, b_j are non-zero, non-units and $a_1|a_2|\cdots|a_{n_1}$, $b_1|b_2|\cdots|b_{n_2}$. Then, $r_1 = r_2, n_1 = n_2$ and for every i, $a_i \sim b_i$.

2.5. **Applications for the structure theorem for modules over PID.** We mainly consider two cases in this section, abelian groups and vectors spaces endowed with a linear transformation. In fact, those examples were already discussed above and so some of the discussion is brief.

2.5.1. *Abelian groups.* This is the case where $R = \mathbb{Z}$. The structure theorem applies for finitely generated abelian groups A. Every such group A is isomorphic

$$\mathbb{Z}^r \oplus \mathbb{Z}/a_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/a_m\mathbb{Z},$$

where r - the rank - is uniquely determined by A, and the a_i are unique positive integers such that $a_1|a_2|\cdots|a_m$ and $a_1 > 1$. Note that any two positive integers that are associates are actually equal. This allows us to eliminate the "up to unit" ever-present in the structure theorems for a general PID.

Also, every such group is isomorphic to

$$\mathbb{Z}^r \oplus \mathbb{Z}/p_1^{a_1} \oplus \cdots \oplus \mathbb{Z}/p_s^{a_s},$$

where the p_i are primes, $s \geq 0$, $a_i > 0$ and r and the set $\{p_1^{a_1}, \dots, p_s^{a_s}\}$ are uniquely determined by A.

Proposition 2.5.1. *Let $n = p_1^{a_1} \cdots p_r^{a_r}$ be the decomposition of a positive integer. The number of abelian groups of order n, up to isomorphism, is $p(a_1) \cdots p(a_r)$ where $p(\cdot)$ is the partition function.*

Proof. Exercise. □

2.5.2. $\mathbb{F}[x]$*-modules.* Let $\mathbb{F}$ be a field and V a finite dimensional vector space over $\mathbb{F}$, $T : V \to V$ a linear map. As usual, we view V as an $\mathbb{F}[x]$ module where $x \cdot v := T(v)$. As we have already remarked, the finite dimensionality of V implies that it is a torsion $\mathbb{F}[x]$-module.

The structure theorem in invariant factors form gives

$$V \cong \mathbb{F}[x]/(a_1(x)) \oplus \cdots \oplus \mathbb{F}[x]/(a_n(x)),$$

where $a_1(x)|\cdots|a_n(x)$ are uniquely determined monic polynomials.

On the vector space $\mathbb{F}[x]/(a(x))$ the minimal polynomial of T is $a(x)$. On each $\mathbb{F}[x]/(a(x))$ we can describe the action of T as follows: if $a(x) = x^r + b_{r-1}x^{r-1} + \cdots + b_0$, then $1, x, \dots, x^{r-1}$ is a basis over $\mathbb{F}$ for $\mathbb{F}[x]/(a(x))$ and x (and so T) act via the matrix

$$\begin{pmatrix} 0 & 0 & \dots & 0 & -b_0 \\ 1 & 0 & & 0 & -b_1 \\ & 1 & & \vdots & \vdots \\ & & & 1 & -b_{r-1} \end{pmatrix}.$$

This matrix is called the **companion matrix** of $a(x)$ and we shall denote it $C_{a(x)}$. We note that

$$\Delta(C_{a(x)}) = m(C_{a(x)}) = a(x),$$

where $\Delta(C_{a(x)})$ is the characteristic polynomial and $m(C_{a(x)})$ is the minimal polynomial.

All together we get that every matrix (or a linear transformation) can be put in **rational canonical form**,

$$\begin{pmatrix} C_{a_1(x)} & & \\ & \ddots & \\ & & C_{a_n(x)} \end{pmatrix},$$

for unique monic polynomials $a_1(x)|\cdots|a_n(x)$. We can summarize this discussion as follows.

Theorem 2.5.2. *Let $\mathbb{F}$ be a field. Let $\mathrm{GL}_n(\mathbb{F})$ act on $M_n(\mathbb{F})$ by*

$$m \mapsto gmg^{-1}, \quad g \in \mathrm{GL}_n(\mathbb{F}), m \in M_n(\mathbb{F}).$$

The orbits of $\mathrm{GL}_n(\mathbb{F})$ are in bijection with block matrices

$$\begin{pmatrix} C_1 & & \\ & \ddots & \\ & & C_s \end{pmatrix},$$

where: (i) each C_i is a square matrix of the form

$$C_i = \begin{pmatrix} 0 & 0 & \dots & 0 & * \\ 1 & 0 & & 0 & * \\ & 1 & & \vdots & \vdots \\ & & & 1 & * \end{pmatrix};$$

(ii) the sum of the sizes of the matrices C_i is n; (iii)

$$\Delta(C_1)|\cdots|\Delta(C_s).$$

The following corollary is one explanation as to why the invariant factors are called so.

Corollary 2.5.3. *Let A be a matrix in $M_n(\mathbb{F})$ and $\mathbb{K} \supseteq \mathbb{F}$ a field extension. The rational canonical form of A over $\mathbb{K}$ is the same as over $\mathbb{F}$. Consequently, if A_1, A_2 are matrices in $M_n(\mathbb{F})$ and for some $B \in \mathrm{GL}_n(\mathbb{K})$ we have $BA_1B^{-1} = A_2$, then for some matrix $\tilde{B} \in \mathrm{GL}_n(\mathbb{F})$ we have $\tilde{B}A_1\tilde{B}^{-1} = A_2$.*

2.5.3. $\mathbb{F}[x]$*-modules,* $\mathbb{F}$ *algebraically closed.* We continue the analysis of the previous section in the case where $\mathbb{F}$ is algebraically closed. An example to keep in mind is when $\mathbb{F} = \mathbb{C}$, the complex numbers, but the discussion applies to *every* algebraically closed field. In this situation we would like to consider the structure theorem in elementary divisors form. Note that because $\mathbb{F}$ is algebraically closed, the only irreducible monic polynomials are the linear polynomials $x - \lambda$ for $\lambda \in \mathbb{F}$. We then have,

$$V \cong \bigoplus_{i=1}^{n} \mathbb{F}[x]/((x-\lambda_i)^{a_i}).$$

Consider a module $\mathbb{F}[x]/((x-\lambda)^a)$. Note that $\{1, x-\lambda, (x-\lambda)^2, \dots, (x-\lambda)^{a-1}\}$ is a basis. If we write it in opposite order $\{(x-\lambda)^{a-1}, \dots, x-\lambda, 1\}$ then $x - \lambda$ acts by the matrix,

$$\begin{pmatrix} 0 & 1 & & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & & \vdots \\ & & & & 1 \\ 0 & & \dots & & 0 \end{pmatrix}$$

and so x acts by

$$J(\lambda, a) := \begin{pmatrix} \lambda & 1 & & \dots & 0 \\ & \lambda & 1 & \dots & 0 \\ \vdots & & & & \vdots \\ & & & \lambda & 1 \\ 0 & & \dots & & \lambda \end{pmatrix}$$

We conclude that choosing bases this way, every matrix is equivalent to a unique matrix of the form

$$\begin{pmatrix} J(\lambda_1, a_1) & & \\ & \ddots & \\ & & J(\lambda_n, a_n) \end{pmatrix}.$$

This is of course the **Jordan canonical form**.

2.5.4. *Computational issues.* The rational canonical form $\mathrm{diag}(C_{a_1(x)}, \dots, C_{a_m(x)})$ can be calculated quickly over any field. There is *no need* to factor polynomials, while the Jordan canonical form requires factorization. There is no algorithm for factoring polynomials over a general field and so the rational canonical form is advantageous. We will not prove the following theorem, but we may use it for calculations.

Theorem 2.5.4. *(Smith's normal form) Let $A \in M_n(\mathbb{F})$ and consider the matrix $B = xI_n - A \in M_n(\mathbb{F}[x])$. Using repeatedly one of the elementary operations below, one can arrive from B to a matrix of the form*

$$\mathrm{diag}(1, \dots, 1, a_1(x), \dots, a_m(x)),$$

where the $a_i(x)$ are the invariant factors of A.
Elementary operations:

(1) Exchanging 2 rows, or 2 columns.
(2) Adding an $\mathbb{F}[x]$-multiple of a row to another row, and the same with columns.
(3) Multiplying a row, or a column, by a unit of $\mathbb{F}[x]$ (namely, a non-zero scalar in $\mathbb{F}$).

In fact, keeping track of the process one gets a change of basis matrix taking A to its rational canonical form.

We give an example.

Example 2.5.5. Suppose that

$$A = \begin{pmatrix} 2 & -2 & 14 \\ & 3 & -7 \\ & & 2 \end{pmatrix}.$$

The characteristic polynomial is clearly $(x-2)^2(x-3)$ and a quick check shows that the minimal polynomial is $(x-2)(x-3)$. Note that the invariant divisors then can only be

$$a_1(x) = x - 2, \quad a_2(x) = (x-2)(x-3).$$

Lets check that the algorithm above indeed gives the same answer. We form the matrix

$$xI - A = \begin{pmatrix} x-2 & 2 & -14 \\ & x-3 & 7 \\ & & x-2 \end{pmatrix}.$$

We perform row and column operations in $\mathbb{Q}[x]$. We obtain

$$\begin{pmatrix} x-2 & 2 & -14 \\ & x-3 & 7 \\ & & x-2 \end{pmatrix} \to \begin{pmatrix} 2 & x-2 & -14 \\ x-3 & 0 & 7 \\ 0 & 0 & x-2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ \frac{x-3}{2} & -\frac{(x-2)(x-3)}{2} & 7x-14 \\ 0 & 0 & x-2 \end{pmatrix}.$$

We have in the first step switched the 1st and 2nd columns. In the second step we first divided the first column by 2 and then subtracted $(x-2)$ the first column from the second, and we added 14 times the first column to the third. Multiply now the second row by 2 and subtract $x-3$ times the first row to get,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -(x-2)(x-3) & 14(x-2)) \\ 0 & 0 & x-2 \end{pmatrix}.$$

Subtract now 14 times the third row from the second and multiply the second column by -1. After that switch the second and third column and then the second and third rows to arrive at

$$\begin{pmatrix} 1 & & \\ & x-2 & \\ & & (x-2)(x-3) \end{pmatrix}.$$

While the result is not a surprise, it is still satisfying.

Also quotient modules can be calculated effectively when R is a PID. Suppose that

$$f : \mathbb{R}^n \longrightarrow M$$

is a surjective homomorphism of R-modules. Then $M \cong R^n/\mathrm{Ker}(f)$ and one would like to gain understanding into the nature of M this way. Suppose that $y_1, \dots, y_m$ is a basis of $\mathrm{Ker}(f)$ and $x_1, \dots, x_n$ a basis of R^n. The elementary divisors theorem says that there is a change of basis $x_1', \dots, x_n'$ of $x_1, \dots, x_n$, and another change of basis $y_1', \dots, y_m'$ of the basis $y_1, \dots, y_m$ such that now

$$y_i' = a_i x_i', \quad a_1 | \cdots | a_m.$$

Once these bases are computed, we have that $M \cong R^{n-m} \oplus R/(a_1) \oplus \cdots \oplus R/(a_m)$.

Write

$$y_i = \sum_{j=1}^{n} a_{ij} x_j.$$

This gives us a matrix

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}.$$

Suppose that

$$x_i = \sum_j b_{ij} x_j'.$$

This gives us the matrix

$$B = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{nn} \end{pmatrix}.$$

Then, in terms of the basis $\{x_i'\}$, the y_i are gotten using the matrix AB. Also, the change of basis from $y_1, \dots, y_m$ is encoded by a matrix

$$C = \begin{pmatrix} c_{11} & \dots & c_{1m} \\ \vdots & & \vdots \\ c_{m1} & \dots & c_{mm} \end{pmatrix}.$$

(So that $y_i' = \sum_j c_{ij} y_j$). Thus, the relation between the $\{y_i'\}$ and $\{x_i'\}$ is given by the matrix

$$CAB, \quad C \in \mathrm{GL}_m(R), B \in \mathrm{GL}_n(R).$$

The EDT is equivalent to saying that they are choices of C and B such that CAB is an $m \times n$ matrix of the form

$$\begin{pmatrix} a_1 & 0 & & \dots & & 0 \\ & a_2 & & & & \\ & & \ddots & & & \\ & & & a_m & & \vdots \\ & & & & 0 & \\ & & & & & \ddots & \\ & & & & & 0 & \dots & 0 \end{pmatrix}, \qquad a_1 | a_2 | \cdots | a_m, \tag{3}$$

in fact uniquely determined up to units. We can also express that by saying that for every $A = m \times n$ matrix (a_{ij}) with entries in R, the double coset

$$\mathrm{GL}_m(R)(a_{ij})\mathrm{GL}_n(R),$$

contains a matrix as in (3), which is unique up to modifying each entry by a unit.

In practice, we think about the matrix B as giving columns operations and the matrix A as giving row operations. Thus, to find the diagonal matrix of the elementary divisors, we take the matrix (a_{ij}) and perform on it any number of column and row operations until we find the matrix of elementary divisors. We illustrate this by a simple example:

Example 2.5.6. We wish to calculate the structure of the abelian group $\mathbb{Z}^3/N$ where, N is spanned by $(1,1,1), (1,2,1), (3,1,5)$. The matrix A is thus

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 3 & 1 & 5 \end{pmatrix}.$$

Consider the following changes

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 3 & 1 & 5 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 3 & -2 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

where in the first step we have subtracted from the second column and from the third column the first column; in the second step we have subtracted from the second row the first row and from the third row three times the first row; in the last step we have added twice the second row to the third row. We conclude that

$$\mathbb{Z}^3/N \cong \mathbb{Z}/2\mathbb{Z}.$$

See also exercise (6) in this context.

Part 2. **Fields**

3. Basic notions

3.1. **Characteristic.** We recall (see §***) that every field F contains a unique field among the fields $\mathbb{Q}$ and $\mathbb{F}_p$, where p is a prime number. That field is called the prime field of F. If $F \supseteq \mathbb{Q}$ we say that F has characteristic zero; if $F \supseteq \mathbb{F}_p$ we say that F has characteristic p.

3.2. **Degrees.** Consider now two fields L, F such that $L \supseteq F$. We say that L is a **field extension** of F, and sometimes that L/F (read: L over F) is a field extension. We can then view L as an F-vector space. We define the **degree** of L over F as

$$[L : F] = \dim_F(L).$$

Warning: as L and F are also abelian groups, we also have the notion of the index of F in L that we had also denoted by $[L : F]$. The index is not equal to the degree in general. For example, if L is a field with p^2 elements and $F = \mathbb{F}_p$. Then $[L : F] = 2$, but the index of F in L as abelian groups is $\sharp L/\sharp\mathbb{F}_p = p^2/p = p$.

Proposition 3.2.1. *Let $L \supseteq M \supseteq N$ be field extensions. We have*

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

Proof. Let $\{x_\alpha : \alpha \in I\}$ be a basis for M over N. Let $\{y_\beta : \beta \in J\}$ be a basis for L over M. Note that $|I| = [M : N], |J| = [L : M]$ and, by definition, $|I \times J| = |I| \cdot |J|$. We will prove that

$$\{x_\alpha y_\beta : (\alpha, \beta) \in I \times J\}$$

is a basis for L over N. We first show it is a spanning set.

Let $\ell \in L$. We can write

$$\ell = \sum_{\beta\in J} r_\beta y_\beta, \qquad r_\beta \in M, \text{ a.a. zero.}$$

As each $r_\beta \in M$, we can write

$$r_\beta = \sum_{\alpha\in I} s_{\alpha,\beta} x_\alpha, \qquad s_{\alpha,\beta} \in N, \text{ a.a. zero.}$$

Note that is $r_\beta = 0$ then all $s_{\alpha,\beta} = 0$. Now,

$$\ell = \sum_{\beta\in J} r_\beta y_\beta = \sum_{\beta\in J}(\sum_{\alpha\in I} s_{\alpha,\beta}x_\alpha)y_\beta = \sum_{(\alpha,\beta)\in I\times J} s_{\alpha,\beta}x_\alpha y_\beta.$$

It remains to show that $\{x_\alpha y_\beta\}$ is linearly independent over N. Suppose that

$$\sum_{(\alpha,\beta)\in I\times J} s_{\alpha,\beta}x_\alpha y_\beta = 0,$$

where $s_{\alpha,\beta} \in N$ are almost all zero. Then, as $0 = \sum_{\beta\in J}(\sum_{\alpha\in I} s_{\alpha,\beta}x_\alpha)y_\beta$ is a linear combination of the y_β with coefficients in M, we must have for all β that

$$\sum_{\alpha\in I} s_{\alpha,\beta}x_\alpha = 0.$$

Since this is a linear combination of the x_α with coefficients in N, it follows that all $s_{\alpha,\beta} = 0$. □

Corollary 3.2.2. *If $[L : N] < \infty$ then $[L : M]$ and $[M : N]$ are finite too and divide $[L : N]$.*

Example 3.2.3. Suppose that $[L : N]$ is a prime number, then any subfield $L \supseteq M \supseteq N$ is either L or N.

About notation. We often depict the situation above by the following diagram, where a, b and ab denote the degrees:

$$\begin{tikzcd}
L \arrow[d, dash, "a"] \arrow[dd, dash, bend right=60, "ab"] \\
M \arrow[d, dash, "b"] \\
N
\end{tikzcd}$$

3.3. **Construction.** The most fundamental method is the following. Let F be a field and $f(x) \in F[x]$ an irreducible monic polynomial of degree d. Let

$$L = F[x]/(f(x)),$$

then L is a field in which f has a root (viz. the coset $\bar{x} = x + (f(x))$ of x) and $[L : F] = d$.

Proposition 3.3.1. *Let $M \supseteq F$ be a field extension in which f has a root α. Then, there is a unique ring homomorphism (necessarily injective) $\varphi : L \to M$ such that $\varphi(\bar{x}) = \alpha$.*

Proof. Define a homomorphism

$$\tilde{\varphi} : F[x] \to F(\alpha), \quad \tilde{\varphi}(g(x)) = g(\alpha).$$

This is a well defined homomorphism restricting to the identity map on F, identified with the constant polynomials. As $\tilde{\varphi}(f) = f(\alpha) = 0$, $\tilde{\varphi}$ factors by the first isomorphism theorem for rings:

$$\begin{tikzcd}
F[x] \arrow[rr, "\tilde{\varphi}"] \arrow[dr, "\text{can.}"'] & & F(\alpha) \\
& L \arrow[ur, "\varphi"] &
\end{tikzcd}$$

As any homomorphism of rings from a field to a ring is injective (use that the only ideals a field has are the trivial ideals and a homomorphism takes 1 to 1), we find that φ is injective. Its image is a subfield of $F(\alpha)$ that contains $\alpha = \varphi(\bar{x})$ and F. By minimality of $F(\alpha)$ it follows that $\varphi(L) = F(\alpha)$.

Finally, the uniqueness follows from the fact that every element of L is of the form $\sum a_i \bar{x}^i, a_i \in F$ and thus a ring homomorphism that is the identity on F and takes the value α on $\bar{x}$ is uniquely determined. □

Let $M \supset F$ be a field and let $\alpha_1, \dots, \alpha_r$ be any elements of M (possibly with repetitions, possibly in F). Define

$$F(\alpha_1, \dots, \alpha_r) = \bigcap_{\substack{K \subset M \\ K \supset F \cup \{\alpha_1, \dots, \alpha_r\}}} K.$$

It is the minimal subfield of M that contains F and all the elements $\alpha_1, \dots, \alpha_r$.

Example 3.3.2. Consider the field $\mathbb{Q}(\sqrt{2}, \sqrt{3})$, a subfield of $\mathbb{C}$. We have the following diagram of subfields:

$$\begin{tikzcd}
 & \mathbb{Q}(\sqrt{2}, \sqrt{3}) \arrow[dl, dash] \arrow[d, dash] \arrow[dr, dash] & \\
\mathbb{Q}(\sqrt{2}) \arrow[dr, dash] & \mathbb{Q}(\sqrt{6}) \arrow[d, dash] & \mathbb{Q}(\sqrt{3}) \arrow[dl, dash] \\
 & \mathbb{Q} &
\end{tikzcd}$$

We remark that all these fields are distinct. For example, $\mathbb{Q}(\sqrt{3}) \neq \mathbb{Q}(\sqrt{2})$, else for some rational numbers a, b we would have $3 = (\sqrt{3})^2 = (a + b\sqrt{2})^2 = a^2 + 2b^2 + 2ab\sqrt{2}$, this forces either a or b to be zero, and we either get $3 = a^2$ or $3 = 2b^2$, which cannot hold for rational numbers by unique factorization.

Are there any other subfields? This is not clear at all. For example, we may try the field $\mathbb{Q}(\sqrt{2} + \sqrt{3})$. This field contains $-1/(\sqrt{2} + \sqrt{3}) = \sqrt{2} - \sqrt{3}$ and so contains $\sqrt{2} = ((\sqrt{2} + \sqrt{3}) + (\sqrt{2} - \sqrt{3}))/2$ and therefore also $\sqrt{3}$. Thus,

$$\mathbb{Q}(\sqrt{2} + \sqrt{3}) = \mathbb{Q}(\sqrt{2}, \sqrt{3}).$$

As it turns out, the subfields we have listed above are all the subfields. The diagram resembles the diagram of the subgroups of $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ and we will later see, via Galois theory, that this is no mere accident. The fact that our list of fields is exhaustive, reflects that we have accounted for all subgroups of $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

Corollary 3.3.3. *$F(\alpha) \cong F[x]/(f(x))$ and, in particular, every element of $F(\alpha)$ can be expressed uniquely as a polynomial in α of degree at most $d - 1$ with coefficients in F. As $F(\alpha_1, \dots, \alpha_r) = F(\alpha_1, \dots, \alpha_{r-1})(\alpha_r)$, we find that every element of $F(\alpha_1, \dots, \alpha_r)$ is a polynomial in $\alpha_1, \dots, \alpha_r$ with coefficients in F.*

Example 3.3.4. The polynomial $x^3 - 2$ is irreducible over $\mathbb{Q}$ by Eisenstein's criterion. It has the real root $\sqrt[3]{2}$ in $\mathbb{C}$. Thus, $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$ and any element in $\mathbb{Q}(\sqrt[3]{2})$ can be written uniquely as $a + b\sqrt[3]{2} + c(\sqrt[3]{2})^2$ for some $a, b, c \in \mathbb{Q}$.

The following proposition is a strengthening of Proposition 3.3.1, which will be very important in studying automorphisms of fields later.

Proposition 3.3.5. *Let $\sigma : F_1 \to F_2$ be an isomorphism of fields. Let $f_1(x) = a_n x^n + \cdots + a_0 \in F_1(x)$ be an irreducible polynomial, and let*

$$f_2(x) = {}^{\sigma}f_1(x) = \sigma(a_n)x^n + \cdots + \sigma(a_0) \in F_2(x).$$

Let $M_i \supseteq F_i$ be a field in which f_i has a root α_i, $i = 1, 2$. Then $f_2(x)$ is irreducible as well and there exists a unique isomorphism $\varphi : F_1(\alpha_1) \to F_2(\alpha_2)$, restricting to σ on F_1. We denote this by the following diagram:

$$\begin{tikzcd}
F_1(\alpha_1) \arrow[r, "\varphi"] \arrow[d, dash] \arrow[dr, phantom, "\circlearrowleft"] & F_2(\alpha_2) \arrow[d, dash] \\
F_1 \arrow[r, "\sigma"] & F_2
\end{tikzcd}$$

Proof. The ring isomorphism $\sigma : F_1 \to F_2$ induces a ring isomorphism

$$\sigma : F_1[x] \to F_2[x], \qquad \sigma(b_r x^r + \cdots + b_1 x + b_0) = \sigma(b_r)x^r + \cdots + \sigma(b_1)x + \sigma(b_0).$$

We also denote this map $g(x) \mapsto {}^{\sigma}g(x)$. □

Example 3.3.6. Consider the subfield $\mathbb{Q}(\sqrt[3]{2}, \omega\sqrt[3]{2}, \omega^2\sqrt[3]{2}) = \mathbb{Q}(\sqrt[3]{2}, \omega)$ of $\mathbb{C}$ where $\omega = e^{2\pi i/3}$ is a third root of 1 and $\sqrt[3]{2}$ is real. It solves the quadratic polynomial $\frac{x^3-1}{x-1} = x^2 + x + 1$ and so we may also write $\omega = \frac{-1+\sqrt{-3}}{2}$. The field $\mathbb{Q}(\sqrt[3]{2}, \omega\sqrt[3]{2}, \omega^2\sqrt[3]{2})$ is obtained from $\mathbb{Q}$ by adding all the solutions of the polynomial $x^3 - 2$.

We have the following diagram of fields:

$$\begin{tikzcd}
& \mathbb{Q}(\sqrt[3]{2}, \omega) \arrow[dl, dash, "3"'] \arrow[d, dash, "2"] \arrow[dr, dash, "2"] \arrow[drr, dash, "2"] & & \\
\mathbb{Q}(\omega) \arrow[dr, dash, "2"'] & \mathbb{Q}(\sqrt[3]{2}) \arrow[d, dash, "3"] & \mathbb{Q}(\omega\sqrt[3]{2}) \arrow[dl, dash, "3"] & \mathbb{Q}(\omega^2\sqrt[3]{2}) \arrow[dll, dash, "3"] \\
& \mathbb{Q} & &
\end{tikzcd}$$

We claim that all these fields are distinct. In fact, $[\mathbb{Q}(\omega) : \mathbb{Q}] = 2$, while the irreducibility of $x^3 - 2$ implies that $[\mathbb{Q}(\omega^a\sqrt[3]{2}) : \mathbb{Q}] = 3$ for $a = 0, 1, 2$. This explains the degrees at the bottom. It also shows that $\mathbb{Q}(\omega) \neq \mathbb{Q}(\omega^a\sqrt[3]{2})$. Since $\mathbb{Q}(\omega^a\sqrt[3]{2})$ is real for $a = 0$ and not real for $a = 1, 2$, it follows that all the fields we have written in the middle row are distinct, except possibly $\mathbb{Q}(\omega\sqrt[3]{2})$ and $\mathbb{Q}(\omega^2\sqrt[3]{2})$, which we leave as an exercise.

Note also, that as $\mathbb{Q}(\sqrt[3]{2})$ is real, $\omega \notin \mathbb{Q}(\sqrt[3]{2})$ and so $x^2 + x + 1$ is irreducible over $\mathbb{Q}(\sqrt[3]{2})$. Thus, $[\mathbb{Q}(\sqrt[3]{2}, \omega) : \mathbb{Q}(\sqrt[3]{2})] = 2$ and we find that $[\mathbb{Q}(\sqrt[3]{2}, \omega) : \mathbb{Q}] = 6$. The degrees at the top now follow.

The diagram resembles the diagram of subgroups of S_3. We shall see later that S_3 is the Galois group of $\mathbb{Q}(\sqrt[3]{2}, \omega)/\mathbb{Q}$ and we shall be able to conclude that there are no more subfields of $\mathbb{Q}(\sqrt[3]{2}, \omega)$.

We remark the following. The field $\mathbb{Q}(\sqrt[3]{2}, \omega)$ satisfies:

$$[\mathbb{Q}(\sqrt[3]{2}, \omega) : \mathbb{Q}] = [\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] \cdot [\mathbb{Q}(\omega) : \mathbb{Q}].$$

But, we can also write this field as $\mathbb{Q}(\sqrt[3]{2}, \omega\sqrt[3]{2})$ (check!). Now,

$$[\mathbb{Q}(\sqrt[3]{2}, \omega\sqrt[3]{2}) : \mathbb{Q}] \neq [\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] \cdot [\mathbb{Q}(\omega\sqrt[3]{2}) : \mathbb{Q}]$$

(the right hand side is equal to 6, while the left hand side is equal to 9). Thus, it is not clear how to calculate $[F(\alpha_1, \dots, \alpha_r) : F]$. Nonetheless, we have the following lemma.

Lemma 3.3.7. *We have the inequality,*

$$[F(\alpha_1, \dots, \alpha_r) : F] \leq \prod_{i=1}^{r} [F(\alpha_i) : F].$$

Proof. If $[F(\alpha_i) : F] = \infty$ for some i, there's nothing to prove, as both sides are infinite ($F(\alpha_1, \dots, \alpha_r) \supset F(\alpha_i)$ and so $[F(\alpha_1, \dots, \alpha_r) : F]$ is infinite as well).

Assume thus that $[F(\alpha_i) : F]$ is finite for all i. We prove the results by induction on r. The cases $r = 0, 1$ are obvious. Assume the result for $r - 1$. As $[F(\alpha_r) : F]$ is finite, the set $1, \alpha_r, \dots, \alpha_r^n$ is independent over F for some n. So α_r solves some irreducible polynomial f over F. We have

$$F(\alpha_r) \cong F[x]/(f(x)), \qquad [F(\alpha_r) : F] = \deg(f).$$

Now, there exists an irreducible polynomial $g \in F(\alpha_1, \dots, \alpha_{r-1})$ such that $g|f$ and $g(\alpha_r) = 0$. Therefore,

$$\begin{aligned}
[F(\alpha_1, \dots, \alpha_r) : F] &= [F(\alpha_1, \dots, \alpha_r) : F(\alpha_1, \dots, \alpha_{r-1})] \cdot [F(\alpha_1, \dots, \alpha_{r-1}) : F] \\
&= [F(\alpha_1, \dots, \alpha_{r-1})[x]/(g(x)) : F(\alpha_1, \dots, \alpha_{r-1})] \cdot [F(\alpha_1, \dots, \alpha_{r-1}) : F] \\
&\leq \deg(g) \cdot \prod_{i=1}^{r-1} [F(\alpha_i) : F] \\
&\leq \deg(f) \cdot \prod_{i=1}^{r-1} [F(\alpha_i) : F] \\
&\leq [F(\alpha_r) : F] \cdot \prod_{i=1}^{r-1} [F(\alpha_i) : F] \\
&= \prod_{i=1}^{r} [F(\alpha_i) : F]
\end{aligned}$$

□

Remark 3.3.8. The proof gives a criterion for when equality holds. Namely, if for each i the irreducible polynomial satisfied by α_i over F remains irreducible over $F(\alpha_1, \dots, \alpha_{i-1})$ then we have equality.

4. Straight-edge and Compass constructions

4.1. **The problem and the rules of the game.** The problem is this: given an interval of length one, can one construct an interval of a given length ℓ using only a straight-edge and a compass?

It will be useful to formalize the notions. Given a finite set X of points $p_1, \dots, p_n$ in the plane we are allowed to increase X to a larger set of points as follows: we can construct

- a line passing through two points p_i, p_j of X;
- a circle with center at one of the points and radius equal to the length of an interval whose end points are two points in X.

Given two such lines, or two such circles, or a line and a circle we may increase X by adjoining the points of intersection.

Let us now assume that two points p_1, p_2 in the plane are given.

Definition 4.1.1. A point p in the plane is **constructible** from two points p_1, p_2 if by repeated application of the procedures above we can arrive at a set X such that $p \in X$. More generally, a point p is constructible from a set Y if by repeated application of the procedures above we arrive at a set X such that $p \in X$.

Definition 4.1.2. A length r (i.e., a non-negative real number r) is **constructible** if, assigning the interval $[p_1, p_2]$ the length 1, there are two constructible points x, y whose distance from each other is r.

More generally, a length r is constructible from a set Y if by repeated application of the procedures above we arrive at a set X such that r is the distance between two points in X.

To study whether a point is constructible or not, hence whether a length is constructible or not, it will be useful to introduce coordinates. We choose our coordinates such that $p_1 = (0, 0)$ and $p_2 = (1, 0)$. It will also be useful to introduce the following terminology:

Definition 4.1.3. The **field of definition** of a constructible set of points $X = \{(x_1, y_1), \cdots, (x_n, y_n)\}$ is the field $\mathbb{Q}(x_1, y_1, \dots, x_n, y_n)$. We shall denote this field by $\mathbb{Q}(X)$.

Example 4.1.4. Constructing $\sqrt{2}$. Draw the line through $(0,0)$ and $(1,0)$ and the circle of radius 1 centered at $(1,0)$. We obtain $(2,0)$ as an intersection point. Draw two circles of radius 2, one about $(0,0)$ and the other about $(2,0)$. Draw the line passing through the intersection points $(1, \pm\sqrt{3})$ of these two circles. This is a line perpendicular to the line through (0, 0) and (2, 0) that passes through (1, 0). Draw a circle of radius 1 about (1, 0). It intersects the perpendicular line at the point (1, 1). The distance between $(1,1)$ and $(0,0)$ is $\sqrt{2}$. Incidentally, note that we have also constructed $\sqrt{3}$ as the distance between $(1, \sqrt{3})$ and $(1,0)$.

Example 4.1.5. If we can construct lengths a, b we can construct $a + b, ab$ and $\sqrt{a}$. See Figure 1.

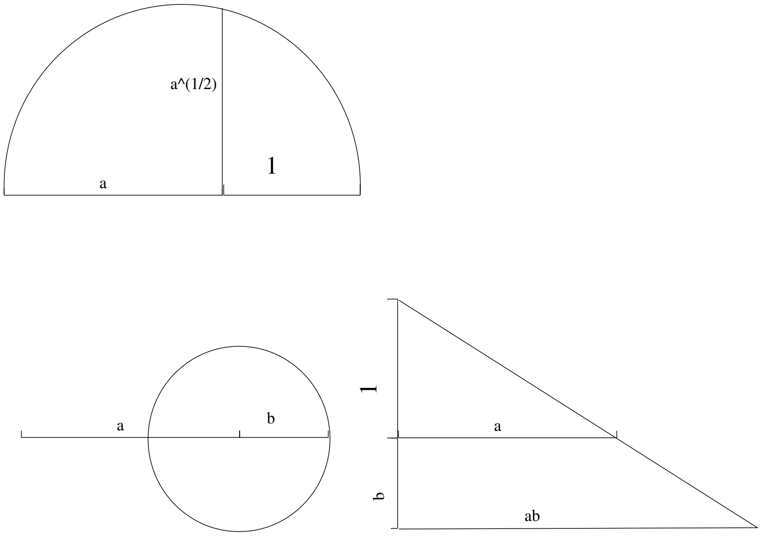


FIGURE 1. Straight-edge and compass constructions

Example 4.1.6. We can construct an angle of $\pi/3$. Given $(0,0), (1,0)$ we can construct $(1/2, 0)$ as well as the unit circle about $(0,0)$. We can construct a line perpendicular to the line through $(0,0)$ and $(1,0)$ and passing through $(1/2,0)$. This line intersects the circle at the point $x = (1/2, \sqrt{3}/2)$. The line through x and $(0,0)$ forms an angle of $\pi/3$ with the line through $(0,0)$ and $(1,0)$.

Theorem 4.1.7. *Let X be a set of points constructible from a set of points Y. Then $[\mathbb{Q}(X) : \mathbb{Q}(Y)] = 2^k$ for some $k \geq 0$. Let r be a length constructible from Y then $[\mathbb{Q}(Y, r) : \mathbb{Q}(Y)] = 2^j$ for some $j \geq 0$.*

In particular, let X be a set of constructible points. The field $\mathbb{Q}(X)$ is of degree 2^k over $\mathbb{Q}$ for some $k \geq 0$. Let r be a constructible length then $[\mathbb{Q}(r) : \mathbb{Q}] = 2^j$ for some $j \geq 0$.

Proof. The set X is obtained by repeated applications of the procedures above from the set Y. We may assume that X contains all the points these procedures yield (i.e., both points of intersection if a circle is involved). Indeed, if X' is this larger set then $[\mathbb{Q}(X) : \mathbb{Q}(Y)]$ divides $[\mathbb{Q}(X') : \mathbb{Q}(Y)]$. The result for $\mathbb{Q}(X)$ therefore follows then from the result for $\mathbb{Q}(X')$. The same holds for r; if r is a distance between points in X, it is a distance between points in X'.

The induction start when $X = Y$ ($X = \{(0,0), (1,0)\}$ in the special case) and then the result holds as $\mathbb{Q}(X) = \mathbb{Q}(Y)$ and if r is a length between points in Y then $r^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2$ for some points $(x_1, y_1), (x_2, y_2)$ in the set Y and thus $r^2 \in \mathbb{Q}(Y)$.

Suppose we proved that $[\mathbb{Q}(X) : \mathbb{Q}(Y)] = 2^k$. Let us consider the set X^+ obtained from X by adding the points of intersections of a line and a circle, two lines, or two circles.

The line through points $(x_1, y_1), (x_2, y_2)$ in X can be written as $t(x_1, y_1) + (1-t)(x_2, y_2)$. The circle of radius r, where $r = \sqrt{(x_3 - x_4)^2 + (y_3 - y_4)^2}$ for some points $(x_3, y_3), (x_4, y_4)$ of X, centered at a point (x_5, y_5) of X, can be written as $(x - x_5)^2 + (y - y_5)^2 = r^2$. Then, the intersection points are the solutions of

$$(tx_1 + (1-t)x_2 - x_5)^2 + (ty_1 + (1-t)y_2 - y_5)^2 = r^2.$$

This is at most a quadratic equation over the field $\mathbb{Q}(X)(r)$. Since $[\mathbb{Q}(X)(r) : \mathbb{Q}(Y)] = [\mathbb{Q}(X)(r) : \mathbb{Q}(X)][\mathbb{Q}(X) : \mathbb{Q}(Y)]$ is a power of 2, the solutions t_1, t_2 lie in the field $\mathbb{Q}(X)(r)(t_1, t_2)$ that has degree a power of two over $\mathbb{Q}(Y)$. It follows that if X^+ is the set obtained from X by adding the new points then $\mathbb{Q}(X^+)$ has degree a power of 2 over $\mathbb{Q}(Y)$. Since the distance between any two points of X^+ lies either in $\mathbb{Q}(X^+)$ or in a quadratic extension of it, any length r constructed from X^+ satisfies that $[\mathbb{Q}(r) : \mathbb{Q}(Y)]$ is a power of 2.

Consider now the intersection of two lines: those are the solutions to

$$t(x_1, y_1) + (1-t)(x_2, y_2) = s(x_3, y_3) + (1-s)(x_4, y_4).$$

We can write this as

$$t(x_1 - x_2) + x_2 = s(x_3 - x_4) + x_4, \quad t(y_1 - y_2) + y_2 = s(y_3 - y_4) + y_4.$$

Those are two linear equations in two unknown and therefore there are either infinitely many solutions (the lines are equal), one solution lying in $\mathbb{Q}(x_1, \dots, x_4, y_1, \dots, y_4)$ or no solutions (the lines are parallel). In any case, we see that in this case, though the set X^+ obtained by adding the unique point of intersection may be larger then X, in fact $\mathbb{Q}(X^+) = \mathbb{Q}(X)$ and the lengths constructed from X^+ lie at most in a quadratic extension of $\mathbb{Q}(X^+)$.

We leave the last case, an intersection of two circles, to the reader. □

4.2. **Applications to classical problems in geometry.**

The Greek posed three problems in geometry.

1. **Doubling the cube.** To construct a cube whose volume is double the volume of a given cube.
2. **Trisecting an angle.** To trisect a given angle.
3. **Squaring the circle.** To construct a square whose area is equal to the area of a given circle.

The problems imply the following:

1. One can construct the third root of 2, as this is the ratio between the edges of the two cubes.
2. One can solve the equation
$$4x^3 - 3x - \cos(\alpha) = 0.$$
We explain that: The identity $e^{i\theta} = \cos(\theta) + i\sin(\theta)$ gives $e^{i3\theta} = \cos(3\theta) + i\sin(3\theta)$ which is equal to $(e^{i\theta})^3 = (\cos(\theta) + i\sin(\theta))^3$. Thus,
$$\cos(3\theta) + i\sin(3\theta) = (\cos(\theta) + i\sin(\theta))^3.$$
Expanding and taking real parts we obtain $\cos(3\theta) = \cos^3(\theta) - 3\cos(\theta)\sin^2(\theta) = 4\cos^3(\theta) - 3\cos(\theta)$ (replacing $\sin^2$ by $1 - \cos^2$).

   Note that if an angle θ can be constructed then so can $\cos(\theta)$ by using the "trigonometric circle". Given an angle α, write $\alpha = 3\theta$. If θ can be constructed then one obtains a solution to the equation $4x^3 - 3x - \cos(\alpha) = 0$.
3. π is constructible (= square of a constructible number).

Let us see what we can say about these problems using the information we have obtained about constructible points.

- By Eisenstein's criterion $x^3 - 2$ is irreducible over $\mathbb{Q}$ and hence $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$ and so $\sqrt[3]{2}$ is not a constructive length. It follows that one cannot double the volume of a cube.
- Take $\alpha = \pi/3$ then $\cos(\alpha) = 1/2$. The equation $4x^3 - 3x - 1/2$ is irreducible: pass to $8x^3 - 6x - 1$ and make a change of variable to get $x^3 - 3x - 1$. If this polynomial is reducible it has a rational root. The solutions have denominator dividing 1 and numerator dividing 1, hence of the form ± 1. One verifies that none is a solution. It follows that $[\mathbb{Q}[x]/(4x^3 - 3x - 1/2) : \mathbb{Q}] = 3$. Therefore the angle $\pi/3$ cannot be trisected.
- The last negative answer (to squaring the circle) is deeper. If it can be resolved then $\sqrt{\pi}$, which is the size of the square whose area is equal to the area of a circle of radius 1 is constructible and thus so is π. It then follows that $[\mathbb{Q}(\pi) : \mathbb{Q}]$ is a power of 2, in particular finite. However, the number π is in fact transcendental - it doesn't solve any non zero polynomial equation - a fact that is not easy to prove. In particular, $[\mathbb{Q}(\pi) : \mathbb{Q}] = \infty$ and so π is not constructible.

5. Algebraic extensions

5.1. **Algebraic and transcendental elements.** Let $K \supseteq F$ be an extension of fields. An element $\alpha \in K$ is called **algebraic over** F if α is a root of some non-zero polynomial $f(x) \in F[x]$. It is called **transcendental over** F otherwise.

Example 5.1.1. Many naturally occurring numbers, like $\sqrt{2}, e^{2\pi i/n}$ are algebraic over $\mathbb{Q}$; in these examples, they solve the polynomials $x^2 - 2$ and $x^n - 1$, respectively. At the same time, the some common constants, such as π and e are transcendental (and Euler's constant γ is suspected to be transcendental, but that is an open problem). But this is hard to prove. It is easier to prove that $\sum_{n=1}^{\infty} \frac{1}{10^{n!}}$ is transcendental, but that is hardly a naturally occurring number.

Consider the ring homomorphism

$$F[x] \to K, \qquad f(x) \mapsto f(\alpha).$$

If α is transcendental this map is injective and the image is contained in the field $F(\alpha)$. The map extends to the field of fractions $F(x) = \{f/g : f, g \in F[x], g \neq 0\}$ and we get an inclusion

$$F(x) \to F(\alpha).$$

Since $F(\alpha)$ is the minimal field containing α we conclude the following proposition.

Proposition 5.1.2. *If α is transcendental over F then*

$$F(\alpha) \cong F(x),$$

and $[F(\alpha) : F] = \infty$.

If α is algebraic, then the map $F[x] \to F(\alpha)$ must have a non-zero kernel; indeed if $g(\alpha) = 0$ then $g(x)$ is in the kernel. Let $m(x)$ be a monic polynomial generating the kernel. Then, $F[x]/(m(x)) \subset K$, hence an integral domain. Thus, $m(x)$ is irreducible and $F[x]/(m(x))$ is a field. It follows then that

$$F[x]/(m(x)) \cong F(\alpha).$$

The polynomial $m(x)$ is called the **minimal polynomial** of α over F. Our discussion shows that it has the property that it divides any other polynomial that α satisfies. Note that $m(x)$ can therefore also be characterized as the unique monic irreducible polynomial that α satisfies. We conclude:

Proposition 5.1.3. *If α is algebraic then there exists a unique monic irreducible polynomial $m(x)$ that α satisfies. It divides any other polynomial having α as a root. We have*

$$F[x]/(m(x)) \cong F(\alpha).$$

In particular, $[F(\alpha) : F] = \deg(m(x))$. This degree is also called the **degree** *of α.*

Corollary 5.1.4. *α is algebraic over F if and only if $[F(\alpha) : F] < \infty$.*

A field extension $K \supseteq F$ is called **algebraic** if every element of K is algebraic over F.

Lemma 5.1.5. *If $[K : F] < \infty$ then K is algebraic extension of F. Moreover, the degree of every element divides $[K : F]$.*

Proof. Indeed, for $\alpha \in K$ we have

$$[F(\alpha) : F] = \frac{[K : F]}{[K : F(\alpha)]}.$$

□

Example 5.1.6. We have seen that $\mathbb{Q}(\sqrt{2}, \sqrt{3}) = \mathbb{Q}(\sqrt{2} + \sqrt{3})$ and so the degree of $\sqrt{2} + \sqrt{3}$ is 4. What is its minimal polynomial? We note that $\sqrt{2} + \sqrt{3}$ solves

$$(x - (\sqrt{2} + \sqrt{3})) \cdot (x + (\sqrt{2} + \sqrt{3})) \cdot (x - (\sqrt{2} - \sqrt{3})) \cdot (x + (\sqrt{2} - \sqrt{3})).$$

One expands this expression and finds that it is equal to $x^4 - 10x^2 + 1$. As this is a rational polynomial of degree 4 it must be the minimal polynomial of $\sqrt{2} + \sqrt{3}$.

Theorem 5.1.7. *Let $K \supseteq F$ be an extension of fields. Let*

$$H = \{\alpha \in K : \alpha \text{ is algebraic over } F\}.$$

then H is a field. Every element in $K - H$ is transcendental over H and over F.

Proof. The set H can also be characterized as the collection of all the elements $\alpha \in H$ such that $[F(\alpha) : F] < \infty$. We need to show that if $\alpha, \beta \in H$ then $\alpha + \beta, \alpha\beta, -\alpha$, and also $1/\alpha$ if $\alpha \neq 0$, are in H. We note that $F(\alpha) = F(-\alpha) = F(1/\alpha)$ and that settles the cases of $-\alpha$ and $1/\alpha$, if $\alpha \neq 0$. Let γ denote either $\alpha + \beta$ or $\alpha\beta$. Then $F(\gamma) \subseteq F(\alpha, \beta)$ and so, using Lemma 3.3.7, we find that $[F(\gamma) : F] < [F(\alpha, \beta) : F] \leq [F(\alpha) : F] \cdot [F(\beta) : F] < \infty$. Therefore, γ is also algebraic over F.

Let $\alpha \in K$ be algebraic over H. We want to show it belongs to H (and thus also algebraic over F). That would show that $K - H$ consists of transcendental elements over H. As α is algebraic over H, α solves some non-zero irreducible polynomial

$$a_n x^n + \cdots + a_1 x + a_0 \in H[x].$$

Now,

$$\begin{aligned} [F(\alpha) : F] &\leq [F(\alpha, a_0, a_1, \ldots, a_n) : F] \\ &= [F(\alpha, a_0, a_1, \ldots, a_n) : F(a_0, a_1, \ldots, a_n)] \cdot [F(a_0, a_1, \ldots, a_n) : F] \\ &\leq n \cdot \prod_{i=0}^{n} [F(a_i) : F] \\ &< \infty, \end{aligned}$$

where we have used Lemma 3.3.7 and that each a_i is algebraic over F, equivalently that $[F(a_i) : F]$ is finite for all i. Therefore, α is algebraic over F and thus belongs to H. □

If $H = K$ in the theorem above, then we call K/F an algebraic extension. If $H = F$ then we call K/F a **transcendental** extension. The theorem gives the following conclusion.

Corollary 5.1.8. *Let K/F be a field extension. Let H be the collection of elements in K that are algebraic over F. Then,*

$$\begin{tikzcd}
K \arrow[rr, bend right, dash, "\textit{transc.}"'] & \supseteq & H \arrow[rr, bend right, dash, "\textit{algebraic}"'] & \supseteq & F
\end{tikzcd}$$

We note another corollary of the proof.

Corollary 5.1.9. *Let $L \supseteq K \supseteq F$ be field extensions such that L is algebraic over K and K is algebraic over F. Then L is algebraic over F.*

5.2. **Compositum of fields.** Let K be a field and K_1, K_2 subfields of K. The **compositum** of K_1 and K_2, denoted K_1K_2, is the minimal subfield of K that contains both.

Example 5.2.1. Suppose that F is a subfield of K and that $K_i = F(\alpha_i)$, $i = 1, 2$. Then $K_1K_2 = F(\alpha_1, \alpha_2)$. As such, the following theorem generalizes Lemma 3.3.7

Theorem 5.2.2. *Let $K \supseteq F$ be an extension of fields and K_i, $i = 1, 2$ subfields of K that contain F. Then*

$$[K_1K_2 : F] \leq [K_1 : F] \cdot [K_2 : F],$$

with equality if and only if a basis for K_2 over F remains linearly independent over K_1, in which case it is a basis for K_1K_2 over K_1.

Proof. Let us write

$$\begin{aligned}
&K_1 = F(\alpha_1, \dots, \alpha_n) \qquad \{\alpha_1, \dots, \alpha_n\} \text{ a basis for } K_1 \text{ as a vector space over } F.\\
&K_2 = F(\beta_1, \dots, \beta_m) \qquad \{\beta_1, \dots, \beta_m\} \text{ a basis for } K_2 \text{ as a vector space over } F.
\end{aligned}$$

Lemma 5.2.3. $\{\alpha_i\beta_j : 1 \leq i \leq n, 1 \leq j \leq m\}$ *spans K_1K_2 as a vector space over F.*

Let us assume the lemma for the moment. The identity

$$\sum_{i,j} a_{i,j}\alpha_i\beta_j = \sum_j (\sum_i a_{i,j}\alpha_i)\beta_j$$

shows that $\beta_1, \dots, \beta_m$ span K_1K_2 as a vector space over K_1. Consequently,

$$[K_1K_2 : F] = [K_1K_2 : K_1] \cdot [K_1 : F] \leq m \cdot [K_1 : F] = [K_1 : F] \cdot [K_2 : F],$$

with equality if and only if the set $\beta_1, \dots, \beta_m$ remains linearly independent over K_1. It remains to prove the lemma.

Let

$$X = \left\{ \sum_{i,j} a_{i,j} \cdot \alpha_i\beta_j : a_{i,j} \in F \right\}.$$

Clearly X is contained in K_1K_2. Let $f \in K_1$. Write

$$f = a_1\alpha_1 + \cdots + a_n\alpha_n,$$

for some $a_i \in F$, and write

$$1 = b_1\beta_1 + \cdots + b_m\beta_m,$$

for some $b_i \in F$. Then

$$f = f \cdot 1 = (a_1\alpha_1 + \cdots + a_n\alpha_n)(b_1\beta_1 + \cdots + b_m\beta_m) = \sum_{i,j} (a_ib_j) \cdot \alpha_i\beta_j, \qquad a_ib_j \in F.$$

It follows that $X \supseteq K_1$, and a similar argument gives $X \supseteq K_2$. We show next that X is closed under multiplication.

It is clear that X is closed under addition, and in fact is an additive subgroup of K_1K_2. Moreover, it is also clear that X is closed under multiplication by elements of f (they multiply each coefficient $a_{i,j}$ and so X is a vector space over F). Thus, to show X is closed under multiplication it is enough to show that

$$\alpha_i\beta_j \cdot \alpha_k\beta_\ell \in X, \qquad \forall i, j, k, \ell.$$

As $\alpha_i\alpha_k \in K_1$, we may write for suitable $a_i \in F$,

$$\alpha_i\alpha_k = a_1\alpha_1 + \cdots + a_n\alpha_n.$$

Similarly, for suitable $b_j \in F$,

$$\beta_j\beta_\ell = b_1\beta_1 + \cdots + b_m\beta_m.$$

It follows that $\alpha_i\beta_j \cdot \alpha_k\beta_\ell \in X$. As we have $K_i \subset X$ and so $\{\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_m\} \subset X$, the fact that X is closed under multiplication implies now that every monomial

$$\alpha_1^{i_1} \cdots \alpha_n^{i_n}\beta_1^{j_1} \cdots \beta_m^{j_m} \in X$$

(we shall write such a monomial for short as $\alpha^I\beta^J$, where $I = (i_1, \dots, i_n), J = (j_1, \dots, j_m)$).

As $K_1K_2 = F(\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_m)$ every element in it is a linear combination of such monomials (cf. Corollary 3.3.3) . As X is closed under addition,

$$K_1K_2 \subseteq \left\{ \sum_{I,J} a_{I,J} \cdot \alpha^I\beta^J : a_{I,J} \in F \right\} \subseteq X.$$

We deduce that $X = K_1K_2$ (and is in particular a field). □

The proof has a non-obvious conclusion:

Corollary 5.2.4. *If $K_1 = F(\alpha_1, \dots, \alpha_n), K_2 = F(\beta_1, \dots, \beta_n)$, where the α_i (resp. β_j) are a basis for K_1 (resp. K_2) as a vector space over F, then K_1K_2 is spanned over F by $\{\alpha_i\beta_j : 1 \leq i \leq n, 1 \leq j \leq m\}$. This conclusion stays correct if the $\{\alpha_i\}, \{\beta_j\}$ are spanning sets (not necessarily linearly independent).*

Example 5.2.5. Let $K = \mathbb{Q}(\omega, \sqrt[3]{2})$ where $\omega = e^{2\pi i/3}$. Let $K_1 = \mathbb{Q}(\omega), K_2 = \mathbb{Q}(\sqrt[3]{2})$. As these fields are obtained by adjoining the root of an irreducible polynomial, viz. $x^2 + x + 1$ and $x^3 - 2$, respectively, we find that $\{1, \omega\}$ is a basis for K_1 over $\mathbb{Q}$ and $\{1, \sqrt[3]{2}, (\sqrt[3]{2})^2\}$ is a basis for K_2 over $\mathbb{Q}$. Thus, K is spanned as a vector space over $\mathbb{Q}$ by

$$1, \omega, \sqrt[3]{2}, \omega\sqrt[3]{2}, (\sqrt[3]{2})^2, \omega(\sqrt[3]{2})^2.$$

In fact, since we proved that $[K : \mathbb{Q}] = 6$ this set is a basis. Alternately, we can also deduce it from the theorem. We argue that $\{1, \omega\}$ stays independent over K_2. Else, for some $a, b \in K_2$ that are not both zero, $a + b\omega = 0$. $b = 0$ leads quickly to a contradiction. Thus, $b \neq 0$ and so $\omega = -a/b \in K_2$. But then $[\mathbb{Q}(\omega) : \mathbb{Q}] = 2$ divides $[K_2 : \mathbb{Q}] = 3$. A contradiction.

It is interesting to examine what happens if we take $K_1 = \mathbb{Q}(\omega\sqrt[3]{2}), K_2 = \mathbb{Q}(\sqrt[3]{2})$. We leave it to the reader to sort it out.

Corollary 5.2.6. *Suppose that* $\gcd([K_1 : F], [K_2 : F]) = 1$ *then*

$$[K_1K_2 : F] = [K_1 : F] \cdot [K_2 : F].$$

Proof. The inclusion $F \subseteq K_i \subseteq K_1K_2$ gives that $[K_i : F]$ divides $[K_1K_2 : F]$. The assumption then gives that $[K_1 : F] \cdot [K_2 : F]$ divides $[K_1K_2 : F]$. By the theorem, $[K_1K_2 : F] \leq [K_1 : F] \cdot [K_2 : F]$ and so it follows that $[K_1K_2 : F] = [K_1 : F] \cdot [K_2 : F]$. □

6. Splitting fields and algebraic closure

6.1. **Splitting fields.** Let $K \supseteq F$ be an extension of fields. K is a **splitting field** of a polynomial $f(x) \in F[x]$ if $f(x)$ is a product of linear factors in K (and so has all its roots in K), $f(x) = c\prod_{i=1}^{n}(x-\alpha_i)$ $\underline{\text{and}}$ $K = F(\alpha_1, \dots, \alpha_n)$.

Remark 6.1.1. We do not require f to be irreducible.

Example 6.1.2. The field $K = \mathbb{Q}(\sqrt{2}, \sqrt{3})$ is a splitting field for the (reducible) polynomial $(x^2-2)(x^2-3)$. It is also the splitting field for the (irreducible) polynomial $x^4 - 10x^2 + 1$. Cf. Example 5.1.6.

Theorem 6.1.3. *Let $f(x) \in F[x]$ be a polynomial of degree n. There is a splitting field $K \supseteq F$ for F. Furthermore, $[K : F] \leq n!$.*

Remark 6.1.4. We will show later that any two splitting fields for f are isomorphic.

Proof. We prove that by induction on $\deg(f)$. If $\deg(f) = 1$ then $f(x) = a_1x - a_0 = a_1(x - a_0/a_1)$ and so $K = F$.

Suppose the theorem true for degree $n-1$ and let $\deg(f) = n$. Let g be an irreducible factor of f in $F[x]$. Let $K_1 = F[x]/(g(x))$. Then g has a root α in K_1 (in fact, $\alpha = \bar{x}$), $K_1 = F(\alpha)$ and we may factor f over K_1:

$$f(x) = (x-\alpha)\tilde{f}(x).$$

We note that $\deg(\tilde{f}) = n-1$ and $[K_1 : F] \leq n$ (in fact, we have an equality if and only if $g = f$). Applying induction to $\tilde{f}$, we get a splitting field K_2 for $\tilde{f}$ over K_1, such that $[K_2 : K_1] \leq (n-1)!$. Thus, in K_2 we can write $f = c\prod_{i=1}^{n}(x-\alpha_i), \alpha_i \in K_2, c \in F$, where $\alpha = \alpha_1$. Note that $K_2 = K_1(\alpha_2, \dots, \alpha_n) = F(\alpha_1)(\alpha_2, \dots, \alpha_n) = F(\alpha_1, \dots, \alpha_n)$. □

An algebraic extension K/F which is the splitting field over F for a collection of polynomials $\{f_i(x) i \in I\} \in F[x]$ is called a **normal extension**. This means that all the f_i split into linear terms in K and denoting by S the set of all these roots, we have $K = F(S)$. Said differently, all the polynomials f_i split into linear terms in K, and K is the minimal field in which they all split.

Remark 6.1.5. If the collection is finite $\{f_1, \dots, f_a\}$ take $f = f_1 f_2 \cdots f_a$. Then K is the splitting field for the set $\{f_1, \dots, f_a\}$ if and only if it is the splitting field for f.

Note as well that we do not specify which is the collection of polynomials making K/F into a normal extension. For example, $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ is a normal extension of $\mathbb{Q}$ by virtue of either the set of polynomials $\{x^2-2, x^2-3\}$, or the set of (one) polynomial $\{x^4 - 10x^2 + 1\}$.

Remark 6.1.6. If $K \supseteq F$ is a normal extension of finite degree then it is a splitting field of a polynomial. Indeed, since K is a normal extension there is a some set of polynomials $\{f_i(x) : i \in I\}$ of which K is a splitting field. Take f_1 in this set such that the roots of f_1 are not in F (if there's no such f_1 then $K = F$ and it is the splitting field of the polynomial $x - 1$). Say the roots of f_1 in K are $\alpha_{1,1}, \dots, \alpha_{1,n(1)}$ and $[F(\alpha_{1,1}, \dots, \alpha_{1,n(1)}) : F] > 1$. If $F(\alpha_{1,1}, \dots, \alpha_{1,n(1)}) = K$ we are done. Else, there is an f_2 in that set such that the roots of f_2 are not all in $F(\alpha_{1,1}, \dots, \alpha_{1,n(1)})$. Say the roots are $\alpha_{2,1}, \dots, \alpha_{2,n(2)}$. Then $[F(\alpha_{1,1}, \dots, \alpha_{1,n(1)}, \alpha_{2,1}, \dots, \alpha_{2,n(2)}) : F] > [F(\alpha_{1,1}, \dots, \alpha_{1,n(1)}) : F]$. This process must end as $[K : F]$ is finite. We conclude that K is the splitting field of a finite collection of polynomials $f_1, \dots, f_a$ and so is the splitting field of the polynomial $f_1 f_2 \cdots f_a$.

Theorem 6.1.7. *Let $f_1(x) \in F_1[x]$ and $\sigma : F_1 \to F_2$ an isomorphism of fields. Let $f_2 = {}^{\sigma}f_1 \in F_2[x]$. Let K_1 be a splitting field for f_1 over F_1 and K_2 a splitting field for f_2 over F_2. Then there is an isomorphism $\varphi : K_1 \to K_2$ extending σ:*

$$\begin{tikzcd}
K_1 \arrow[r, "\varphi"] \arrow[d, dash] & K_2 \arrow[d, dash] \\
F_1 \arrow[r, "\sigma"] & F_2.
\end{tikzcd}$$

Proof. We prove the result by induction on the degree of f_1. If f_1 is a linear polynomial there is nothing to prove as, necessarily, $K_1 = F_1, K_2 = F_2$.

Let f_1 now be of degree greater than 1. Let g_1 be an irreducible factor of f_1, possibly f_1 itself. Thus, $f_1 = g_1h_1$. Let $g_2 = {}^{\sigma}g_1, h_2 = {}^{\sigma}h_1$. Then g_2 is irreducible as well and $f_2 = g_2h_2$. Let $\alpha_1 \in K_1$ be a root of g_1 and $\alpha_2 \in K_2$ be a root of g_2. Let $H_1 = F_1(\alpha_1), H_2 = F_2(\alpha_2)$. By Proposition 3.3.5 there is an isomorphism[1] $\sigma' : H_1 \to H_2$, such that the following diagram holds:

$$\begin{tikzcd}
K_1 \arrow[r, "\exists\varphi??"] \arrow[d, dash] & K_2 \arrow[d, dash] \\
H_1 \arrow[r, "\sigma'"] \arrow[d, dash] & H_2 \arrow[d, dash] \\
F_1 \arrow[r, "\sigma"] & F_2.
\end{tikzcd}$$

We may now consider the polynomials $t_1(x) = f_1(x)/(x - \alpha_1) \in H_1(x)$ and $t_2(x) = f_2(x)/(x - \alpha_2) \in H_2(x)$ which satisfy ${}^{\sigma'}t_1 = t_2$. Also note that K_1, K_2 are the splitting fields of t_1, t_2 over H_1, H_2 respectively. Indeed, K_1 is obtained from F by adding all the roots of f; H_1 is obtained by adding just one root and when we add the rest, we are actually adding the roots of t_1, and we obtain K_1. Similarly for K_2. We can apply induction and the proof is complete. □

Corollary 6.1.8. *Let $f(x) \in F[x]$ be a polynomial. Any two splitting fields for f are isomorphic.*

Example 6.1.9. It is hard to forecast the degree of the splitting field of a polynomial. If f is monic, irreducible of degree n and K is a splitting field, we could say that

$$\deg(f) = n | [K : F], \quad [K : F] \leq n!.$$

We shall later see that in fact $[K : F] | n!$, but that as much as we can say in general. For example, for polynomials of degree 3, it may happen that the splitting field has degree 6 over F, or degree 3. To illustrate: the splitting field of the polynomial $x^3 - 2 \in \mathbb{Q}[x]$ has degree 6 over $\mathbb{Q}$ - it is equal to $\mathbb{Q}(\omega, \sqrt[3]{2})$. It is possible to give an example of an irreducible monic cubic polynomial of degree 3 over $\mathbb{Q}$ for which the splitting field will have degree 3 over $\mathbb{Q}$, but any such example will be laborious. On the other hand, it is rather easy to do so over finite fields. For example, consider the polynomial $(x) = x^3 + x + 1$ over $\mathbb{F}_2$. Let $L = \mathbb{F}_2[x]/(x^3 + x + 1)$. Then $\bar{x}$ is a root of f in L. We can then factor f over L (using the variable t):

$$f(t) = t^3 + t + 1 = (t - \bar{x})(t^2 + \bar{x}t + \bar{x}^2 + 1).$$

We claim that the quadratic polynomial factors over L as well. Substitute $\bar{x}^2$ for t in $t^2 + \bar{x}t + \bar{x}^2 + 1$ to get $\bar{x}^4 + \bar{x}^3 + \bar{x}^2 + 1$ and note that $\bar{x}^4 + \bar{x}^2 + \bar{x} = 0$. Thus, $\bar{x}^4 + \bar{x}^3 + \bar{x}^2 + 1 = \bar{x}^4 + \bar{x}^3 + \bar{x}^2 + 1 + (\bar{x}^4 + \bar{x}^2 + \bar{x}) = \bar{x}^3 + \bar{x} + 1 = 0$. It follows that f factors completely over L and so L is the splitting field, $[L : \mathbb{F}_2] = 3$.

6.2. Algebraic closure.

A field K is called **algebraically closed** if every polynomial in $K[x]$ has a root (equivalently, all its roots) in K. Let F be a field; a field extension $K \supseteq F$ is called an **algebraic closure** of F if every polynomial in $F[x]$ splits into a product of linear terms in K *and* K is an algebraic extension of F.

Proposition 6.2.1. *If K is an algebraic closure of F then K is algebraically closed.*

[1] In fact, uniquely determined by the property $\sigma\prime(\alpha_1) = \alpha_2$; note the choices we have at this point! We shall return to this point later.

Proof. Let f be an irreducible polynomial in $K[x]$. Let $L = K[x]/(f(x))$ and $\alpha = \bar{x}$ the root of f in L. As $L \supseteq K$ is an algebraic extension and $K \supseteq F$ is an algebraic extension also $L \supseteq F$ is an algebraic extension. (The maximal algebraic extension H of F insider L contains K and there is no element of $L - H$ algebraic over H. It follows that $H = L$.) Thus, α solves some irreducible polynomial in $F[x]$ and so $\alpha \in K$. But that means that $\deg(f) = 1$. Saying that the only irreducible polynomials in $K[x]$ are linear is equivalent to saying that every polynomial in $K[x]$ factors into linear terms, or that every polynomial in $K[x]$ has all its roots in K. And, at any rate, K is algebraically closed. □

Theorem 6.2.2. *Let F be a field. F has an algebraic closure.*

Remark 6.2.3. The proof will appear to be constructive, but this is misleading. The description of an algebraic closure is, except for a handful of situations, very complicated. One mainly uses the algebraic closure as an existence result – that is, it's mere existence – and the actual concrete description of it is either ignored, or is a major open problem. For example, the explicit description of the algebraic closure of $\mathbb{Q}$ (for example in the sense of describing all the the quotients of its Galois group) is a major open problem in number theory; some features of it are called "the inverse Galois problem".

Proof. Our proof follows closely Dummit and Foote that follow, in turn, Artin. It is based on the following two lemmas.

Lemma 6.2.4. *There exists an algebraically closed field K that contains F.*

Lemma 6.2.5. *Let K be an algebraically closed field that contains F and let $\overline{F}$ be the maximal algebraic extension of F inside K,*

$$\overline{F} = \{\alpha \in K : \alpha \textit{ is algebraic over } F\}.$$

Then $\overline{F}$ is an algebraic closure of F

We first dispense with the second Lemma. We have inclusions

$$K \supseteq \overline{F} \supseteq F,$$

and we have proved that F is a field and is of course an algebraic extension of F. We need to show that every non-constant polynomial $f(x) \in F[x]$ splits over $\overline{F}$. But, we know that in K, $f(x) = c\prod(x - \alpha_i)$, $c \in F$, and clearly each α_i is algebraic over F, hence in $\overline{F}$. Thus, $f(x)$ splits over $\overline{F}$ and so $\overline{F}$ is an algebraic closure.

The real issue is thus to prove the first lemma. The proof is based on the following idea. We want to enlarge F to a bigger field K_1 so that every monic non-constant polynomial in F will have at least one root in K_1. That would be a "good start". We know how to do that with one monic irreducible polynomial $f(x)$ - simply by forming $F[x]/(f(x))$. We can do it with finitely many polynomials $f_1, \dots, f_r$ by choosing an irreducible factor g_1 of f_1 and forming $L_1 = F[x]/(g_1(x))$, then choosing an irreducible factor g_2 of f_2 over L_1 and forming $L_2 = L_1[x]/(g_2(x))$, and so on. In the field L_r each of the polynomials $f_1, \dots, f_r$ has a root. But how to do it with infinitely many polynomials?? That is where Artin had an elegant idea.

For every monic non-constant polynomial $f(x) \in F[x]$ (we don't care if it is irreducible or not) introduce a free variable x_f. Form the ring of polynomials $R_F = F[\{x_f\}]$, by adjoining all the variables x_f for such polynomials f. Consider the ideal I of this ring generated by the polynomials $f(x_f)$ as f ranges over the monic non-constant polynomials. The key point is

Claim: I is contained in some maximal ideal $\mathfrak{m}$ of R_F

Suppose this is not the case. Then I must be equal to R_F and in particular $1 \in I$. Thus, for some polynomials $f_1, \dots, f_r$ and elements $g_1, \dots, g_r$ of R_F we have

$$1 = g_1 f_1(x_{f_1}) + \cdots + g_r f_r(x_{f_r}).$$

We show this is not possible by constructing a homomorphism φ of R_F to another ring such that $\varphi(g_1 f_1(x_{f_1}) + \cdots + g_r f_r(x_{f_r})) = 0$. As $\varphi(1) = 1$ we get the contradiction we want. First we map

$$R_F \to F[x_{f_1}, \dots, x_{f_r}], \qquad x_f \mapsto \begin{cases} x_{f_i}, & f = f_i \\ 0, & f \notin \{f_1, \dots, f_r\}. \end{cases}$$

Now, there is a field $F_1 \supseteq F$ in which each of the polynomials $f_1, \dots, f_r$ has a root, say $\alpha_1, \dots, \alpha_r$, respectively. We let φ be the composition

$$R_F \to F[x_{f_1}, \dots, x_{f_r}] \to F_1,$$

where the last map takes x_{f_i} to α_i. We have $\varphi(g_1 f_1(x_{f_1}) + \cdots + g_r f_r(x_{f_r})) = \varphi(g_1) f_1(\alpha_1) + \cdots + \varphi(g_r) f_r(\alpha_r) = 0$.

Let us choose then a maximal ideal $\mathfrak{m}$ of R_F that contains I and consider the field

$$K_1 = R_F / \mathfrak{m}.$$

This is a field that contains F and every monic non-constant polynomial in F has a root in K_1. Indeed, $f(x_f) \in I$ so $f(x_f)$ is the zero element modulo I, let alone modulo $\mathfrak{m}$.

We may now repeat the construction replacing F be K_1, R_F by R_{K_1} and so on. We obtain a sequence of fields

$$F \subseteq K_1 \subseteq K_2 \subseteq \dots$$

such that every non-constant monic polynomial in K_n has a root in K_{n+1}. Let

$$K = \bigcup_{i=1}^{\infty} K_i.$$

K is a field, it contains F and we claim that it is algebraically closed. Let f be a non-constant polynomial in $K[x]$, without loss of generality, monic. Each of the coefficients of f belongs to some field K_i and so for some n, $f \in K_n[x]$. If so, it has a root in $K_{n+1} \subseteq K$, and we are done. □

Remark 6.2.6. Given a field F we have constructed an algebraic closure $\overline{F}$ of F. One can prove (and it is mainly a complicated book-keeping argument) that $\overline{F}$ is unique up to isomorphism.

Suppose that K is an algebraically closed field and let F be a subfield of K. As we have seen, the collection of elements of K that are algebraic over F is an algebraic closure of F. For example, take the field $\mathbb{C}$ of complex numbers, which is algebraically closed by the fundamental theorem of algebra. Then $\mathbb{C}$ contains an algebraic closure $\overline{\mathbb{Q}}$ of $\mathbb{Q}$. However, it is not hard to show that $\overline{\mathbb{Q}}$ is countable, while $\mathbb{C}$ is not countable. Thus, in a precise sense, most elements of $\mathbb{C}$ are transcendental over $\mathbb{Q}$, although is it a real challenge to show that any particular complex number is transcendental. For example, it is known that e, π are transcendental, but this is hard to prove. One also knows that $\sum_{i=1}^{\infty} 10^{-i!}$ is transcendental and that is much easier to prove, albeit a nontrivial theorem as well.

7. Finite and cyclotomic fields

In this section we study finite fields and fields obtained by adjoining to $\mathbb{Q}$ a root of unity. We will be able to get some interesting information concerning those and we shall later be able to return to these fields for examples. Besides serving for examples and intuition, these fields also play important role in certain proofs.

7.1. **Finite fields.** In this section we will get a very detailed idea about the structure of the algebraic closure of a finite field. This is one of the very few cases where this is possible.

Theorem 7.1.1. *Let $\mathbb{F} = \mathbb{Z}/p\mathbb{Z}$ the field with p elements. Let $\overline{\mathbb{F}}$ be an algebraic closure of $\mathbb{F}$.*

(1) For every positive integer m, $\overline{\mathbb{F}}$ contains a unique subfield having p^m elements. We denote it by $\mathbb{F}_{p^m}$. The field $\mathbb{F}_{p^m}$ is equal to the set of solutions to the equation $x^{p^m} - x = 0$.

(2) We have $\mathbb{F}_{p^n} \supseteq \mathbb{F}_{p^m}$ if and only if $m|n$. Every finite subfield of $\overline{\mathbb{F}}$ is $\mathbb{F}_{p^m}$ for some m. Therefore, the lattice of finite subfields of $\overline{\mathbb{F}}$ is the opposite to the lattice of subgroups of $\mathbb{Z}$ under the association: $\mathbb{F}_{p^m} \leftrightarrow m\mathbb{Z}$. Under this association we have

$$\begin{aligned} \mathbb{F}_{p^m} \subseteq \mathbb{F}_{p^n} \quad &\leftrightarrow \quad m\mathbb{Z} \supseteq n\mathbb{Z} \qquad\qquad (\Leftrightarrow m|n) \\ \mathbb{F}_{p^{\gcd(m,n)}} = \mathbb{F}_{p^m} \cap \mathbb{F}_{p^n} \quad &\leftrightarrow \quad m\mathbb{Z} + n\mathbb{Z} = \gcd(m,n)\mathbb{Z} \\ \mathbb{F}_{p^{\mathrm{lcm}(m,n)}} = \mathbb{F}_{p^m}\mathbb{F}_{p^n} \quad &\leftrightarrow \quad m\mathbb{Z} \cap n\mathbb{Z} = \mathrm{lcm}(m,n)\mathbb{Z} \end{aligned}$$

(3) Let $f(x) \in \mathbb{F}_{p^m}[x]$ be an irreducible polynomial of degree n and let α be a root of f in $\overline{\mathbb{F}}$. Then $\mathbb{F}_{p^m}(\alpha) = \mathbb{F}_{p^{mn}}$ and is the splitting field of f.

(4) Let L be a field with p^m elements then $L \cong \mathbb{F}_{p^m}$.

(5) $\overline{\mathbb{F}}$ is also an algebraic closure of $\mathbb{F}_{p^m}$. The lattice of subfields of $\overline{\mathbb{F}}$ that contain $\mathbb{F}_{p^m}$ is opposite to the lattice of subgroups of $\mathbb{Z}$ that are contained in $m\mathbb{Z}$.

(6) Define the **Frobenius map** *as*

$$\mathrm{Fr}_p : \overline{\mathbb{F}} \to \overline{\mathbb{F}}, \qquad x \mapsto x^p,$$

and, more generally, for $q = p^m$ define

$$\mathrm{Fr}_q : \overline{\mathbb{F}} \to \overline{\mathbb{F}}, \qquad x \mapsto x^q.$$

Then Fr_q is a field automorphism of $\overline{\mathbb{F}}$, it is equal to $\mathrm{Fr}_p \circ \mathrm{Fr}_p \circ \cdots \circ \mathrm{Fr}_p$ (m-times) and its fixed set is the subfield $\mathbb{F}_q$.

(7) $\overline{\mathbb{F}} = \bigcup_{m=1}^{\infty} \mathbb{F}_{p^m}$.

Proof. Let $L \subseteq \overline{\mathbb{F}}$ be a finite subfield. Let $m = [L : \mathbb{F}]$. Then the number of elements of L is p^m and $L^\times$ is thus a cyclic group of order $p^m - 1$. Thus,

$$L^\times \subseteq \{a \in \overline{\mathbb{F}} : a^{p^m - 1} = 1\}.$$

Since $x^{p^m-1} - 1$ has at most $p^m - 1$ distinct solutions in $\overline{\mathbb{F}}$, we must have equality and, consequently,

$$L = \{a \in \overline{\mathbb{F}} : a^{p^m} = a\}.$$

This also shows that L is uniquely determined by its cardinality.

Conversely, given m, consider the set

$$L := \{a \in \overline{\mathbb{F}} : a^{p^m} = a\}.$$

The binomial theorem $(x+y)^n = \sum_{i=0}^n \binom{n}{i} x^i y^{n-i}$ holds over every commutative ring. As

$$\binom{p}{i} = \frac{p(p-1)\cdots(p-i+1)}{i(i-1)\cdots 1},$$

we have that $p|\binom{p}{i}$ for $0 < i < p$. It follows that

$$(x+y)^p = x^p + y^p,$$

for x, y that belong to a field of characteristic p. By induction,

$$(x+y)^q = x^q + y^q.$$

Now, if $x, y \in L$ then also $x + y \in L$, as $(x+y)^q = x^q + y^q = 0 + 0 = 0$, where we have put $q = p^m$. Since $(-1)^q = -1$, even if $p = 2$, we also get that $-x \in L$. Finally, clearly also xy and

$1/x$, if $x \neq 0$, are in L. Clearly, $1 \in L$, and in fact it then follows that $\mathbb{F} \subseteq L$. We conclude that L is a field containing $\mathbb{F}$. Let $f(x) = x^q - x$. Then $f'(x) = -1$, as $q = 0$ in $\mathbb{F}$ and it follows that $\gcd(f, f') = 1$. Therefore, f has no repeated roots and thus has precisely q roots in $\overline{\mathbb{F}}$. Thus, L is a field with q elements.

At this point we have proven (1). We have also shown that $\mathrm{Fr}_q : \overline{\mathbb{F}} \to \overline{\mathbb{F}}$ is a ring homomorphism. It is clear that $\mathrm{Fr}_q \circ \cdots \circ F_q$ (a-times) is equal to Fr_{q^a}.

Suppose that $m|n$, say $c = n/m$. Put $q = p^m$. $\mathrm{Fr}_q(a) = a$ implies that $(\mathrm{Fr}_q)^c(a) = \mathrm{Fr}_{q^c}(a) = a$; that is $a^{p^m} = a$ implies $a^{(p^m)^c} = a^{p^n} = a$. Thus, $\mathbb{F}_{p^m} \subseteq \mathbb{F}_{p^n}$. Conversely, if $\mathbb{F}_{p^m} \subseteq \mathbb{F}_{p^n}$ then the degree $c = [\mathbb{F}_{p^n} : \mathbb{F}_{p^m}]$ is an integer. But that means that as vector spaces $\mathbb{F}_{p^n} \cong \mathbb{F}_{p^m}^c$ and so $p^n = (p^m)^c = p^{mc}$ and it follows that $m|n$.

By associating $\mathbb{F}_{p^m}$ with the subgroup $m\mathbb{Z}$ of $\mathbb{Z}$ we have a bijection. Moreover, if we put an order relation on subfields by saying that $\mathbb{F}_{p^m} \leq \mathbb{F}_{p^n}$ if $\mathbb{F}_{p^m} \subseteq \mathbb{F}_{p^n}$, and we put an order relation on subgroups by saying that $a\mathbb{Z} \leq b\mathbb{Z}$ if $a\mathbb{Z} \subseteq b\mathbb{Z}$, then the bijection is order-reversing. For fields, the infimum of two fields $\mathbb{F}_{p^m}, \mathbb{F}_{p^n}$ is clearly $\mathbb{F}_{p^m} \cap \mathbb{F}_{p^n}$, while the supremum is $\mathbb{F}_{p^m}\mathbb{F}_{p^n}$, while for subgroups it is clearly $m\mathbb{Z} \cap n\mathbb{Z}$ and $m\mathbb{Z} + n\mathbb{Z}$. Since the bijection is order-reversing, $\mathbb{F}_{p^m} \cap \mathbb{F}_{p^n}$ corresponds to $m\mathbb{Z} + n\mathbb{Z} = \gcd(m,n)\mathbb{Z}$ and $\mathbb{F}_{p^m}\mathbb{F}_{p^n}$ corresponds to $m\mathbb{Z} \cap n\mathbb{Z} = \mathrm{lcm}(m,n)\mathbb{Z}$.

We now pass to part (3). We know that $\mathbb{F}_{p^m}(\alpha) \cong \mathbb{F}_{p^m}[x]/(f(x))$ is a field with $(p^m)^n = p^{mn}$ elements. Thus, $\mathbb{F}_{p^m}(\alpha) = \mathbb{F}_{p^{mn}}$. As this holds for every root α of f, all the roots of f lie in $\mathbb{F}_{p^{mn}}$ and so this field is the splitting field for f.

Consider part (4). As L is the splitting field of the polynomial $x^{p^m} - x$ over $\mathbb{F}$, as is $\mathbb{F}_{p^m}$, they are isomorphic by Corollary 6.1.8.

For part (5), note that as $\overline{\mathbb{F}}$ is algebraically closed, the subfield $F' = \{a \in \overline{\mathbb{F}} : a \text{ is algebraic over } \mathbb{F}_{p^m}\}$ is an algebraic closure of $\mathbb{F}_{p^m}$. But every element of $\overline{\mathbb{F}}$ is algebraic over $\mathbb{F}$ so a fortiori over $\mathbb{F}_{p^m}$. Thus, $\overline{\mathbb{F}} = \mathbb{F}'$.

Now for parts (6) and (7). We have already established that $\mathrm{Fr}_q : \mathbb{F} \bar{\to} \overline{\mathbb{F}}$ is a ring homomorphism, which is clearly injective. Let $a \in \overline{\mathbb{F}}$. Then a solves some irreducible polynomial over $\mathbb{F}$, say of degree m. Thus, $a \in \mathbb{F}_{p^m}$. This gives us part (7): $\overline{\mathbb{F}} = \bigcup_m \mathbb{F}_{p^m}$. As for every m the map Fr_q gives an injective map $\mathrm{Fr}_q : \mathbb{F}_{p^m} \to \mathbb{F}_{p^m}$, for every m, Fr_q is also a surjective map $\mathrm{Fr}_q : \mathbb{F}_{p^m} \to \mathbb{F}_{p^m}$. Using now $\overline{\mathbb{F}} = \bigcup_m \mathbb{F}_{p^m}$, we conclude that $\mathrm{Fr}_q : \overline{\mathbb{F}} \to \overline{\mathbb{F}}$ is also surjective.

Finally, the fixed set of Fr_q, the elements $\{a \in \overline{\mathbb{F}} : a^q = a\}$ are precisely the field $\mathbb{F}_q$, as we have seen above. □

Theorem 7.1.2. *(1)* $x^{p^n} - x = \prod_{\substack{f \text{ irred. monic} \in \mathbb{F}_p[x] \\ \text{of degree } d|n}} f(x)$.

(2) Let $f(x) \in \mathbb{F}_p[x]$ be a non-zero polynomial of degree r. Then f is irreducible if and only if

$$\forall n, 1 \leq n \leq r/2, \qquad \gcd(f(x), x^{p^n} - x) = 1.$$

(3) $f(x)$ has a root in $\mathbb{F}_p$ if and only if $\gcd(f(x), x^p - x) \neq 1$.

Proof. We prove part (1); the other parts are left as exercise. First, $g(x) = x^{p^n} - x$ satisfies $\gcd(g(x), g'(x)) = \gcd(g(x), -1) = 1$ and so $x^{p^n} - x$ is a product of *distinct* irreducible factors. The splitting field of $x^{p^n} - x$ is the field of p^n elements $\mathbb{F}_{p^n}$.

Suppose that $f(x)$ is an irreducible factor of $x^{p^n} - x$. Let α be a root of f. We have the following diagram

$$\begin{tikzcd}
\mathbb{F}_{p^n} \arrow[d, dash] \arrow[dd, dash, bend right=60, "n"'] \\
\mathbb{F}_p(\alpha) \arrow[d, dash, "d"] \\
\mathbb{F}_p
\end{tikzcd}$$

where d is the degree of f, and so it follows that $d|n$.

Conversely, given an irreducible polynomial $f(x) \in \mathbb{F}_p(x)$ of degree d, $d|n$, let α be a root of it in $\overline{\mathbb{F}}_p$. It must be *the* field $\mathbb{F}_{p^d}$ and, since $d|n$, $\mathbb{F}_{p^d}|\mathbb{F}_{p^n}$. In particular, α satisfies $x^{p^n} - x = 0$. As this is true for every root of f, and since f is irreducible it has no repeated roots, we have that $f(x)|x^{p^n} - x$. □

Let μ be the **Möbius function** defined for positive integers by

$$\mu(n) = \begin{cases} 1 & n = 1 \\ 0 & \Box | n \\ (-1)^r & n \text{ is a product of } r \text{ distinct primes} \end{cases}.$$

Let $f : \mathbb{Z}_{>0} \to \Gamma$ be any function on the positive integers with values in an abelian group Γ. Let

$$F(n) = \sum_{d|n} f(d).$$

Then the **Möbius inversion formula** states:

Lemma 7.1.3. $f(n) = \sum_{d|n} F(d)\mu(n/d)$.

We leave the proof of the Möbius inversion formula as an exercise. Let us apply it now for the following functions. Let

$$\Psi(n) = \sharp\{\text{irreducible monic polynomials of degree } n \text{ over } \mathbb{F}_p[x]\}.$$

Let $f(n) = n \cdot \Psi(n)$. By comparing degrees in Theorem 7.1.2, we conclude that

$$F(n) = \sum_{d|n} f(n) = p^n.$$

Applying Möbius inversion formula we find that $f(n) = \sum_{d|n} p^d \cdot \mu(n/d)$ and so that $\Psi(n) = \frac{1}{n}\sum_{d|n} p^d \cdot \mu(n/d)$. In words:

Proposition 7.1.4. *The number of irreducible polynomials of degree n over $\mathbb{F}_p$ is*

$$\frac{1}{n}\sum_{d|n} p^d \cdot \mu(n/d).$$

Example 7.1.5. Let us examine this formula for small values of n.

n	$\sharp$ irreducible polynomials of degree n
1	p
2	$\frac{1}{2}(p^2 - p)$
3	$\frac{1}{3}(p^3 - p)$
4	$\frac{1}{4}(p^4 - p^2)$
6	$\frac{1}{6}(p^6 - p^3 - p^2 + p)$

7.2. **Cyclotomic Fields.** The cyclotomic fields are a collection of subfields of the field of complex numbers. They resemble finite fields in some aspects, and this is reflected in one of the proofs we give below. In the same way that finite fields play a special role in understanding $\overline{\mathbb{F}}_p$ - in fact, this field is the union of its finite subfields, the cyclotomic fields play a special role in understanding $\overline{\mathbb{Q}}$ - the algebraic closure of $\mathbb{Q}$. A deep theorem in algebraic number theory, the **Kronecker-Weber theorem**, says that every abelian Galois extension (this terminology would make sense once we go through Galois theory) is contained in one of the cyclotomic fields.

Let n be a positive integer. Let μ_n denote the set of n**-th roots of unity** in $\mathbb{C}$:

$$\mu_n = \{\alpha \in \mathbb{C} : \alpha^n = 1\} = \{e^{2\pi i a/n} : a = 0, 1, \dots, n-1\}.$$

We use the notation

$$\zeta_n = e^{2\pi i/n}.$$

The set μ_n forms an abelian group under multiplication. It is a cyclic group and, in fact, the map

$$\mathbb{Z}/n\mathbb{Z} \to \mu_n, \qquad a \mapsto \zeta_n^a,$$

is a group isomorphism. In particular, it follows that μ_n is cyclic and its generators, called **primitive** n-th roots of unity, are precisely $\{\zeta_n^a : (a, n) = 1\}$. There are $\varphi(n)$ of them. We note that $\mu_d \subseteq \mu_n \Leftrightarrow d|n$.

We define the n-th **cyclotomic polynomial**, $\Phi_n(x)$ as the polynomial

$$\Phi_n(x) = \prod_{\zeta \in \mu_n \text{ primitive}} (x - \zeta) = \prod_{(a,n)=1} (x - \zeta_n^a).$$

By arranging the elements of μ_n according to their order in the multiplicative group we find the factorization

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

This allows us to calculate Φ_n by recursion as $\Phi_n(x) = x^n - 1/(\prod_{d|n, d<n} \Phi_d(x))$. Here are some examples:

n	1	2	3	4	5	6
Φ_n	$x-1$	$x+1$	x^2+x+1	x^2+1	$x^4+x^3+x^2+x+1$	x^2-x+1

The field $\mathbb{Q}(\zeta_n)$ (called the n-th **cyclotomic field**) is the splitting field of the cyclotomic polynomial Φ_n. We are thus interested in properties of Φ_n.

Proposition 7.2.1. *$\Phi_n(x) \in \mathbb{Z}[x]$ and is a monic polynomial of degree $\varphi(n)$.*

Proof. The only claim requiring proof is that $\Phi_n(x) \in \mathbb{Z}[x]$. We prove that by induction on n. This is certainly true for $n = 1$. Given n, let $f_n(x) = \prod_{d|n, d<n} \Phi_d(x)$. By induction, $f_n(x) \in \mathbb{Z}[x]$. Clearly, $f_n(x)|x^n - 1$ in $\mathbb{Q}(\zeta_n)[x]$. But this implies that $f_n(x)|x^n - 1$ in $\mathbb{Q}[x]$; after all, when performing the Euclidean algorithm we never need more scalars than in $\mathbb{Q}$ and it gives us the gcd of $f_n(x)$ and $x^n - 1$. By Gauss' lemma $f_n(x)|x^n - 1$ in $\mathbb{Z}[x]$. As $\Phi_n(x) = \frac{x^n-1}{f_n(x)}$, we are done. □

Let F be a field. We say that a non-zero polynomial $f(x) \in F[x]$ of degree n is **separable** if in some extension field K we have $f(x) = c\prod_{i=1}^n (x - \alpha_i)$, where all the α_i are distinct. Otherwise said, (in some extension field) f has n distinct root.

Let $f(x) \in F[x]$ be a monic polynomial, where F is any field. Let us first note that if f is **inseparable**, that is, if in some extension K of F, $f(x) = (x-\alpha)^2 g(x)$ then $f'(x) = 2(x-\alpha)g(x) + (x-\alpha)^2 g'(x)$ and so $(x-\alpha) | \gcd(f, f')$ and so $\gcd(f, f') \neq 1$. Remark also that $\gcd(f, f') \in F[x]$. Conversely, suppose that $\gcd(f, f') \neq 1$ and choose in some extension field K a linear polynomial $x - \alpha$ dividing both f and f'. Suppose that $f(x) = (x-\alpha)g(x)$ and that $(x-\alpha) \nmid g(x)$. Then $f'(x) = (x-\alpha)g'(x) + g(x)$ and so $(x-\alpha) \nmid f'(x)$, which is a contradiction. We conclude:

Lemma 7.2.2. *Let $f(x) \in F[x]$ be a non-constant polynomial. Then f is separable if and only if $\gcd(f, f') = 1$.*

A simple consequence of the lemma is that if in some extension f has n distinct roots then in any extension where f splits it has n distinct roots.

Theorem 7.2.3. *The cyclotomic polynomial $\Phi_n(x)$ is irreducible over $\mathbb{Q}$.*

Proof. Let $f_n(x)$ be the minimal polynomial of ζ_n over $\mathbb{Q}$; it is irreducible and it divides $\Phi_n(x)$. As f_n is monic, and $f_n | x^n - 1$, by Gauss' lemma $f_n \in \mathbb{Z}[x]$ and so we can decompose $x^n - 1$ as a polynomial over $\mathbb{Z}$,

$$x^n - 1 = f_n(x) \cdot h(x), \qquad f_n(x), h(x) \in \mathbb{Z}[x].$$

Let ζ be a root of f_n and p a prime such that $(p, n) = 1$. We shall prove that ζ^p is also a root of f_n. As every primitive n-th root of unity can be obtained from ζ_n by repeatedly raising to prime powers p, for various primes p satisfying $(p, n) = 1$, we can conclude that every primitive n-root of unity is a root of f_n. But these are the roots of Φ_n too. Thus, $f_n = \Phi_n$ and so Φ_n is irreducible.

Suppose, on the contrary, that $f(\zeta^p) \neq 0$. Then, because ζ^p is a root of $x^n - 1$, we must have $h(\zeta^p) = 0$. That means that ζ is a root of $h(x^p)$. Since f_n is the minimal polynomial of any of its roots, we conclude that $f_n | h(x^p)$, even in $\mathbb{Z}[x]$ by Gauss' lemma. Now, reduce this modulo p and denote the reduction of polynomials by a bar. Then, using that for every polynomial $a(x) \in \mathbb{F}_p[x]$ we have $(a(x))^p = a(x^p)$,

$$\overline{x^n - 1} = \bar{f}_n \bar{h},$$

and

$$\bar{f}_n | \bar{h}(x^p) = (\bar{h}(x))^p.$$

It follows that $\bar{f}_n$ and $(\bar{h}(x))^p$ have a common factor and thus so do $\bar{f}_n$ and $\bar{h}$. But, that implies that $\overline{x^n - 1}$ has an irreducible factor appearing to a power 2 at least. On the other hand, if $g(x) = x^n - 1$ then $\gcd(g(x), g'(x)) = \gcd(x^n - 1, nx^{n-1}) = 1$, also in characteristic p (because $(p, n) = 1$). Meaning, $x^n - 1$ does not have a repeated factor *even modulo p*. That gives a contradiction. □

Corollary 7.2.4. *Let m, n be relatively prime positive integers. Then*

$$\mathbb{Q}(\zeta_n) \cap \mathbb{Q}(\zeta_m) = \mathbb{Q}.$$

Proof. The compositum of $\mathbb{Q}(\zeta_m)$ and $\mathbb{Q}(\zeta_n)$ is $\mathbb{Q}(\zeta_{mn})$. Indeed, as $\zeta_m \zeta_n$ is a primitive mn-th root of 1 we get inclusion in one direction. On the other hand, both $\mathbb{Q}(\zeta_m)$ and $\mathbb{Q}(\zeta_n)$ are contained in $\mathbb{Q}(\zeta_{mn})$ and that gives the reverse inclusion. We thus have a diagram of fields, where $K = \mathbb{Q}(\zeta_n) \cap \mathbb{Q}(\zeta_m)$:

$$\begin{tikzcd}
& \mathbb{Q}(\zeta_{mn}) \arrow[dl, no head] \arrow[dr, no head] \arrow[ddd, no head, bend right=80, "\varphi(mn)"'] & \\
\mathbb{Q}(\zeta_m) \arrow[dr, no head, "\varphi(m)/a"'] & & \mathbb{Q}(\zeta_n) \arrow[dl, no head, "\varphi(n)/a"] \\
& K \arrow[d, no head, "a"] & \\
& \mathbb{Q} &
\end{tikzcd}$$

Therefore,

$$[\mathbb{Q}(\zeta_{mn}) : K] = \varphi(mn)/a = \varphi(m)\varphi(n)/a.$$

On the other hand, from Theorem 5.2.2,

$$[\mathbb{Q}(\zeta_{mn}) : K] \leq [\mathbb{Q}(\zeta_m) : K] \cdot [\mathbb{Q}(\zeta_n) : K] = \varphi(m)\varphi(n)/a^2.$$

It follows that $a = 1$. That is, $K = \mathbb{Q}$. □

Part 3. **Galois Theory**

In this part of the course we develop Galois theory. Galois theory is the study of automorphisms of fields. More precisely, given an extension of fields $K \supseteq F$, Galois theory studies automorphisms of K that act as the identity on F and relates them to subfields of K containing F.

8. AUTOMORPHISMS AND SUBFIELDS

8.1. **The group** $\mathrm{Aut}(K/F)$**.** Let K be a field. The group $\mathrm{Aut}(K)$ is the **group of automorphism** of K. Namely, the group of bijective ring homomorphisms $K \to K$. The group law is composition and the identity of the group is the identity map $1 : K \to K$, the map that takes every element to itself. If $F \subset K$ is a subfield, we define the group

$$\mathrm{Aut}(K/F) = \{\sigma \in \mathrm{Aut}(K) : \sigma(\alpha) = \alpha, \forall \alpha \in F\}.$$

It is a subgroup of $\mathrm{Aut}(K)$. Note that if F is the prime subfield of K (which is $\mathbb{Q}$ if K has characteristic zero, and $\mathbb{F}_p$ if K has characteristic p) then $\mathrm{Aut}(K) = \mathrm{Aut}(K/F)$.

The following proposition is a simple observation that nonetheless gives some control over automorphisms in $\mathrm{Aut}(K/F)$.

Proposition 8.1.1. *Let $K \supset F$ be an extension of fields, $\alpha \in K$ an algebraic element over F with minimal polynomial $f(x) \in \mathbb{F}[x]$ of degree n. Let $\sigma \in \mathrm{Aut}(K/F)$. Then $\sigma(\alpha)$ is a root of $f(x)$ as well. We get a group homomorphism,*

$$\mathrm{Aut}(K/F) \to S_n.$$

Proof. We have $0 = f(\alpha)$, where $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$, where $a_i \in F$. Then

$$\begin{aligned} 0 &= \sigma(0) \\ &= \sigma(\alpha^n + a_{n-1}\alpha^{n-1} + \cdots + a_0) \\ &= \sigma(\alpha)^n + \sigma(a_{n-1})\sigma(\alpha)^{n-1} + \cdots + \sigma(a_0) \\ &= \sigma(\alpha)^n + a_{n-1}\sigma(\alpha)^{n-1} + \cdots + a_0 \\ &= f(\sigma(\alpha)). \end{aligned}$$

As σ is an automorphism, it is injective and we get an injective map from the set of roots of f to itself (hence bijective). Clearly this is a group homomorphism. □

Let us now consider a very special situation. Later we shall be generalizing our arguments to general splitting fields. Suppose, as above, that α is algebraic over F, with a degree n minimal polynomial $f \in F[x]$, and that moreover, $F(\alpha)$ is a splitting field for f.

Proposition 8.1.2. *Under these assumptions, there is a bijection*

$$\mathrm{Aut}(F(\alpha)/F) \leftrightarrow \{\beta \in F(\alpha) : f(\beta) = 0\}, \quad \sigma \mapsto \sigma(\alpha).$$

In particular, if f is separable, namely if f has precisely n distinct roots, $|\mathrm{Aut}(F(\alpha)/F)| = [F(\alpha) : F]$.

Proof. As every automorphism of $F(\alpha)$ that fixes F is determined by its action on α, the map from $\mathrm{Aut}(F(\alpha)/F)$ to the roots is injective. To show it is surjective, recall that we showed (Proposition 3.3.5) that if α and β are two roots of $f(x)$ then there is a (unique) isomorphism $\varphi : F(\alpha) \to F(\beta)$ that is the identity on F and takes α to β. As $F(\beta) \subseteq F(\alpha)$ and both fields have the same degree over F, we must have $F(\alpha) = F(\beta)$ and φ is an element of $\mathrm{Aut}(F(\alpha)/F)$ such that $\varphi(\alpha) = \beta$. That shows subjectivity. □

Example 8.1.3. The following example explains why we need to assume f is separable. Consider a field F in characteristic p and an element $\gamma \in F$ which is not a p-th power in F. (For example, we may take $F = \mathbb{F}_p(x)$ and $\gamma = x$.) The polynomial $x^p - \gamma \in \mathbb{F}_p[x]$ is irreducible. Indeed, let α be a root of this polynomial in some extension field of F. Then $\alpha^p = \gamma$, and we can write $(x-\alpha)^p = x^p - \gamma$ and so α is the *unique* root of f. Every non-constant polynomial properly dividing $(x-\alpha)^p$ is of the form $(x-\alpha)^n$, for some $0 < n < p$. If this polynomial lies in $F[x]$ then the coefficient of x^{n-1} in this polynomial belongs to F. But this coefficient is $-n\alpha$ and, as $p \nmid n$, it follows that $\alpha \in F$, contrary to our assumption on γ.

We see that in this case $\mathrm{Aut}(F(\alpha)/F) = \{1\}$, while $[F(\alpha) : F] = p$.

Before turning to examples, we state an easy proposition.

Proposition 8.1.4. *Let $K \supseteq F$ be an extension of fields. For a subgroup H of $\mathrm{Aut}(K/F)$ let*

$$K^H = \{k \in K : \sigma(k) = k, \forall \sigma \in H\}.$$

For a subfield $K \supseteq L \supseteq F$ let

$$H_L = \{\sigma \in \mathrm{Aut}(K/F) : \sigma(\ell) = \ell, \forall \ell \in L\}.$$

Then:

(1) K^H is a subfield of K containing F; H_L is a subgroup of $\mathrm{Aut}(K/F)$.
(2) If $H_1 \supseteq H_2$ then $K^{H_1} \subseteq K^{H_2}$; if $L_1 \supseteq L_2$, $H_{L_1} \subseteq H_{L_2}$.
(3) $H_{K^H} \supseteq H$ and $K^{H_L} \supseteq L$.

The proof of the proposition is immediate. Thus, we have, in great generality, order-reversing maps

$$\{\text{Subgroups of } \mathrm{Aut}(K/F)\} \quad \longleftrightarrow \quad \{\text{Subfields of } K \text{ containing } F\},$$

but without special conditions on K/F these are not bijections. Indeed, in Example 8.1.3 above, we have for both $L = F(\alpha)$ and $L = F$ that $H_L = \mathrm{Aut}(F(\alpha)/F)$ and so $K^{H_L} = L$ doesn't hold. (However, see Theorem 9.3.1 below).

8.2. **Examples.**

8.2.1. *Cyclotomic fields.* Let ζ_n be a primitive n-root of 1. Its minimal polynomial is Φ_n and $\mathbb{Q}(\zeta_n)$ is the splitting field for Φ_n, which is a separable polynomial. Thus, we know that

$$|\mathrm{Aut}(\mathbb{Q}(\zeta_n)/\mathbb{Q})| = [\mathbb{Q}(\zeta_n) : \mathbb{Q}] = \deg(\Phi_n) = \varphi(n).$$

Proposition 8.2.1. *There is a natural isomorphism*

$$\mathbb{Z}/n\mathbb{Z}^\times \cong \mathrm{Aut}(\mathbb{Q}(\zeta_n)/\mathbb{Q}).$$

Proof. Let $\sigma \in \mathrm{Aut}(\mathbb{Q}(\zeta_n)/\mathbb{Q})$, then $\sigma(\zeta_n) = \zeta_n^a$ for some a such that $(a, n) = 1$, because these are the roots of the minimal polynomial Φ_n of ζ_n. We note that for *every* n-th root of unity ζ in $\mathbb{Q}(\zeta_n)$,

$$\sigma(\zeta) = \zeta^a.$$

Indeed, write $\zeta = \zeta_n^b$ then $\sigma(\zeta) = \sigma(\zeta_n^b) = (\sigma(\zeta_n))^b = (\zeta_n^a)^b = (\zeta_n^b)^a = \zeta^a$. We define a map

$$\mathrm{Aut}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \to \mathbb{Z}/n\mathbb{Z}^\times, \qquad \sigma \mapsto a, \text{ where } \sigma(\zeta_n) = \zeta_n^a.$$

This map is injective because σ is determined by its action on ζ_n. By cardinality considerations it is therefore surjective. Finally, it is a group homomorphism. For every n-th root of unity ζ, if $\sigma(\zeta) = \zeta^a, \tau(\zeta) = \zeta^b$ then

$$\sigma\tau(\zeta) = \sigma(\zeta^b) = (\sigma(\zeta))^b = (\zeta^a)^b = \zeta^{ab}.$$

□

8.2.2. *Finite fields.* Let $\mathbb{F}_p$ be the finite field with p elements. Let $\mathbb{F}_{p^m}$ be the unique degree m extension of it in a given algebraic closure $\overline{\mathbb{F}}_p$.

Proposition 8.2.2. *There is a canonical isomorphism*

$$\mathbb{Z}/m\mathbb{Z} \cong \mathrm{Aut}(\mathbb{F}_{p^m}), \qquad 1 \mapsto \mathrm{Fr}_p.$$

Proof. Let $f \in \mathbb{F}[x]$ be an irreducible polynomial of degree m (see Exercise 23). We shall use various statements in Theorem 7.1.1. Let α be any root of f in $\overline{\mathbb{F}}_p$. The field $\mathbb{F}_p(\alpha)$ has p^m elements, hence equal to *the* subfield of $\overline{\mathbb{F}}_p$ with p^m elements, $\mathbb{F}_{p^m}$. Thus, we are in the same familiar situation and we conclude that $\mathrm{Aut}(\mathbb{F}_{p^m}/\mathbb{F}_p)$ has m elements.

On the other hand, as we have seen, $\mathrm{Fr}_p : \mathbb{F}_{p^m} \to \mathbb{F}_{p^m}$ is an automorphism. Its order as an automorphism is the least positive n such that $\mathrm{Fr}_p^n = \mathrm{Fr}_{p^n}$ is the identity on $\mathbb{F}_{p^m}$. But, also recall that the fixed field of Fr_{p^n} is $\mathbb{F}_{p^n}$ and the minimal n is thus $n = m$. It follows that Fr_p generates a cyclic subgroup of $\mathrm{Aut}(\mathbb{F}_{p^m}/\mathbb{F}_p)$ of order m, hence it generates the whole group $\mathrm{Aut}(\mathbb{F}_{p^m}/\mathbb{F}_p)$. We conclude a canonical isomorphism

$$\mathbb{Z}/m\mathbb{Z} \xrightarrow{\cong} \mathrm{Aut}(\mathbb{F}_{p^m}/\mathbb{F}_p), \qquad 1 \mapsto \mathrm{Fr}_p.$$

□

Corollary 8.2.3. *Let $m|n$. There is a canonical isomorphism*

$$m\mathbb{Z}/n\mathbb{Z} \cong \mathrm{Aut}(\mathbb{F}_{p^n}/\mathbb{F}_{p^m}), \qquad m \mapsto \mathrm{Fr}_{p^m}.$$

Proof. The issue is what elements in $\mathrm{Aut}(\mathbb{F}_{p^n}) = \mathrm{Aut}(\mathbb{F}_{p^n}/\mathbb{F}_p)$ fix $\mathbb{F}_{p^m}$. The consideration of the proof show that these are precisely the powers of Fr_{p^m}. Under the homomorphism

$$\mathbb{Z} \twoheadrightarrow \mathrm{Aut}(\mathbb{F}_{p^n}/\mathbb{F}_p), \qquad 1 \mapsto \mathrm{Fr}_p,$$

whose kernel is $n\mathbb{Z}$, it is precisely the subgroup $m\mathbb{Z}$ that goes to $\mathrm{Aut}(\mathbb{F}_{p^n}/\mathbb{F}_{p^m})$ and so we get an induced isomorphism $m\mathbb{Z}/n\mathbb{Z} \cong \mathrm{Aut}(\mathbb{F}_{p^n}/\mathbb{F}_{p^m})$, under which $m \mapsto \mathrm{Fr}_{p^m}$. □

8.2.3. *Additional examples.*

Example 8.2.4. We have $\mathrm{Aut}(\mathbb{Q}(\sqrt{2})/\mathbb{Q}) = \mathbb{Z}/2\mathbb{Z}$. More generally, if $\mathbb{F}$ is a field of characteristic different than 2 and $a \in \mathbb{F}$ is not a square, then $\mathrm{Aut}(\mathbb{F}(\sqrt{a})/\mathbb{F}) = \mathbb{Z}/2\mathbb{Z}$.

Example 8.2.5. We have $\mathrm{Aut}(\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q}) \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

Indeed, we know that $\mathbb{Q}(\sqrt{2}, \sqrt{3}) = \mathbb{Q}(\sqrt{2} + \sqrt{3})$ and is the splitting field of the minimal polynomial of $\sqrt{2} + \sqrt{3}$ (all whose other roots are $\pm\sqrt{2} \pm \sqrt{3}$) and therefore the group $\mathrm{Aut}(\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q})$ has cardinality 4. We examine how an automorphism σ acts on the set $\{\pm\sqrt{2}\}$ and on the set $\{\pm\sqrt{3}\}$. This gives us a homomorphism of groups

$$\mathrm{Aut}(\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q}) \to S_2 \times S_2 \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z},$$

where S_2 is the symmetric group on 2 elements, identified with $\mathbb{Z}/2\mathbb{Z}$. As the field $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ consists of polynomial expressions in $\sqrt{2}$ and $\sqrt{3}$ with rational coefficients, it follows that the map we have defined is injective and, by counting, also surjective.

If we consider the action on the four elements $\pm\sqrt{2} \pm \sqrt{3}$ we are realizing this Galois group as the Klein group $\{1, (12)(34), (13)(24), (14)(23)\}$ of S_4.

It is easy to see the correspondence between subfields and subgroups in this case.

Example 8.2.6. We have $\mathrm{Aut}(\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}) = \{1\}$.

Indeed, any automorphism σ will have to take $\sqrt[3]{2}$ to another root of $x^3 - 2$. But the other roots are not real numbers and so cannot lie in $\mathbb{Q}(\sqrt[3]{2})$. Thus, the only automorphism is the identity.

9. The main theorem of Galois theory

9.1. **Galois extension.** Let $K \supseteq F$ be a finite extension of fields. We say that K/F is **Galois** if

$$[K:F] = |\mathrm{Aut}(K/F)|.$$

In this case, we shall use the notation

$$\mathrm{Gal}(K/F) := \mathrm{Aut}(K/F)$$

and call this group the **Galois group** of the extension.

Theorem 9.1.1. *Let $f(x) \in F[x]$ be a monic polynomial. Let K/F be the splitting field of f. Then*

$$|\mathrm{Aut}(K/F)| \leq [K:F],$$

with equality if and only if every irreducible factor of f is separable.

Proof. We prove the statement by induction on $\deg(f)$. In fact, we prove that if $\sigma : F_1 \to F_2$ is an automorphism, $f_1(x) \in F[x]$, $f_2(x) = {}^{\sigma}f_2(x) \in F_2[x]$, K_1 is a splitting field of f_1, K_2 a splitting field of f_2 then the number of extensions φ of σ to an isomorphism from K_1 to K_2,

$$\begin{tikzcd}
K_1 \arrow[r, "\varphi"] \arrow[d, dash] & K_2 \arrow[d, dash] \\
F_1 \arrow[r, "\sigma"] & F_2
\end{tikzcd}$$

is at most the degree $[K_1 : F_1]$ with equality if and only if every irreducible factor of f_1 is separable.

The case of degree 1 polynomials is clear. Assume the statement for degree n polynomials. Let $f_1 \in F_1[x]$ be a polynomial of degree $n+1$; let $p_1(x)$ be an irreducible factor of f_1 over F_1; let $p_2(x) = {}^{\sigma}p_1(x)$, which is an irreducible factor of f_2. If f_1 has an inseparable factor choose p_1 to be that factor, and then also p_2 is inseparable.

Let α_1 be a root of p_1 in K_1, α_2 a root of p_2 in K_2. We have the following diagram

$$\begin{tikzcd}
K_1 \arrow[r, "\varphi?"] \arrow[d, dash] & K_2 \arrow[d, dash] \\
F_1(\alpha_1) \arrow[r, "\psi"] \arrow[d, dash] & F_2(\alpha_2) \arrow[d, dash] \\
F_1 \arrow[r, "\sigma"] & F_2
\end{tikzcd}$$

We have already proven that such extensions ψ exist and that they are in bijection with the roots of p_2. Thus, there are at most $[F_1(\alpha_1) : F_1]$ of those, with equality if and only if p_1 is separable. Now, *fix* an extension ψ like that and consider the diagram

$$\begin{tikzcd}
K_1 \arrow[r, "\varphi?"] \arrow[d, dash] & K_2 \arrow[d, dash] \\
F_1(\alpha_1) \arrow[r, "\psi"] & F_2(\alpha_2)
\end{tikzcd}$$

We apply the induction hypothesis to the polynomials $g_1(x) = f_1(x)/(x-\alpha_1), g_2(x) = {}^{\psi}g_1(x) = f_2(x)/(x-\alpha_2)$. Note that if every irreducible factor of f_1 was separable, then same holds for g_1. Thus, by induction, the number of extensions of ψ to K_1 is at most $[K_1 : F_1(\alpha_1)]$ with equality if every irreducible factor of f_1, hence of g_1, is separable.

Since every extension φ is constructed this way (by first extending to ψ and then extending ψ) we find that if every irreducible factor of f_1 is separable, the number of extensions is $[K_1 : F_1(\alpha_1)][F_1(\alpha_1) : F] = [K_1 : F]$, and if f_1 has an irreducible factor that is inseparable then the number of extensions is strictly less than $[K_1 : F_1(\alpha_1)][F_1(\alpha_1) : F] = [K_1 : F]$, because already the number of ψ is strictly less than $[F_1(\alpha_1) : F]$. □

Corollary 9.1.2. *The splitting field of a separable polynomial in $F[x]$ is a Galois extension of F.*

Given how important is Corollary 9.1.2, it would be desirable to know when an irreducible polynomial is separable. Knowing that, we would be able to tell when an arbitrary polynomial is separable: each of its irreducible factors must be separable and they appear with multiplicity 1.

We say that a field F is **perfect** if either F has characteristic zero, or F has characteristic p and the Frobenius map $\mathrm{Fr}_p : F \to F, \mathrm{Fr}_p(x) = x^p$ is *surjective*.

Example 9.1.3. $\mathbb{Q}, \mathbb{C}, \mathbb{Q}(\sqrt{2}), \mathbb{Q}(x)$, etc. are perfect. $\mathbb{F}_{p^m}, \overline{\mathbb{F}}_p$ are perfect. The field $\mathbb{F}_p(t)$ is not perfect as t is not a p-th power of any other element.

Proposition 9.1.4. *Let F be a perfect field. Then an irreducible polynomial is always separable.*

Remark 9.1.5. The condition is necessary: see Example 8.1.3.

Proof. Let f be an irreducible polynomial (hence, by definition, non constant). Then f has no proper divisor, so in particular $\gcd(f, f') = 1$, unless $f' = 0$. In the latter case we will derive a contradiction. We can only have $f' = 0$ if $f(x) = \sum_{t=0}^{n} a_t x^{tp}$. Since F is perfect, we can find $b_t \in F$ such that $b_t^p = a_t$. Then $f(x) = (\sum_{t=0}^{n} b_t x^t)^p$ and, consequently, $f(x)$ is reducible. Contradiction. □

9.2. **Independence of characters.** Let G be a group and L a field. A **character** of G with values in L is a group homomorphism

$$\chi : G \to L^\times.$$

A character is a special kind of a function on G with values in L. The collection of all functions $L^G := \{G \to L\}$ is a vector space over L with respect to addition of functions $(f + g)(a) := f(a) + g(a)$ and $(\lambda f)(a) := \lambda \cdot f(a)$.

Theorem 9.2.1 (Independence of Characters). *Let G be a group and L a field. Let $\chi_1, \dots, \chi_n$ be distinct characters of G with values in L, then $\chi_1, \dots, \chi_n$ are linearly independent as functions on G. Namely, for $a_i \in L$ we have,*

$$(a_1\chi_1 + \cdots + a_n\chi_n)(g) = 0, \ \forall g \in G \quad \Longrightarrow \quad a_1 = \cdots = a_n = 0.$$

Proof. Assume not and choose a non-trivial linear relation of minimal length. Changing the indexing of the characters, we may write such a relation as

$$a_1\chi_1 + \cdots + a_m\chi_m = 0, \qquad a_i \neq 0, 1 \leq i \leq m.$$

We shall produce a shorter non-trivial relation hence deriving a contradiction. Note that we must have $m > 1$. As the χ_i are distinct, we may pick an element $g_0 \in G$ such that $\chi_1(g_0) \neq \chi_m(g_0)$. Then, for all $g \in G$,

$$\begin{aligned} 0 &= a_1\chi_1(g_0 g) + \cdots + a_m\chi_m(g_0 g) \\ &= a_1\chi_1(g_0)\chi_1(g) + \cdots + a_m\chi_m(g_0)\chi_m(g) \end{aligned} \tag{4}$$

We also have

$$0 = a_1\chi_m(g_0)\chi_1(g) + \cdots + a_m\chi_m(g_0)\chi_m(g). \tag{5}$$

Substracting (5) from (4), we find

$$\begin{aligned} 0 &= a_1(\chi_1(g_0) - \chi_m(g_0))\chi_1(g) + \cdots + a_m(\chi_m(g_0) - \chi_m(g_0))\chi_m(g) \\ &= a_1(\chi_1(g_0) - \chi_m(g_0))\chi_1(g) + \cdots + a_{m-1}(\chi_{m-1}(g_0) - \chi_m(g_0))\chi_{m-1}(g) \end{aligned} \tag{6}$$

Note that $a_1(\chi_1(g_0) - \chi_m(g_0)) \neq 0$, hence we produced a shorter non-trivial linear relation. Contradiction. □

We apply this theorem in the following setting. Let K and L be fields. Let $\sigma : K \to L$ be a ring homomorphism, thus automatically an embedding of fields. We may view $\sigma : K^\times \to L^\times$ as a character of the group $K^\times$ with values in L. In this setting, we have the following conclusion.

Corollary 9.2.2. *Let $\sigma_1, \dots, \sigma_n$ be distinct field embedding of a field K into a field L. Then $\sigma_1, \dots, \sigma_n$ are linearly independent as functions on K.*

Proof. The only thing to remark is that since $\sigma_i(0) = 0$ for all i, the functions $\sigma_i : K \to L$ are linearly independent over L if and only if the characters $\sigma_i : K^\times \to L^\times$ are independent over L. □

9.3. **From a group to a Galois extension.** Our goal in this section is to prove the following theorem. The theorem would allow us to derive a series of corollaries establishing basic connections between the automorphism group of an extension K/F and subfields of K (Corollaries 9.3.2, 9.3.3, 9.3.4) that will take us a long way towards the main theorem of Galois theory.

Theorem 9.3.1. *Let G be a finite group of automorphisms of a field K. Let $F = K^G$ the field fixed by all elements of G. Then,*

$$[K : F] = |G|.$$

Proof. The main idea of the proof is to look for linear relations between the rows, or the columns, of the matrix (whose entries lie in K)

$$\begin{pmatrix} \sigma_1(\omega_1) & \sigma_2(\omega_1) & \dots & \sigma_n(\omega_1) \\ \sigma_1(\omega_2) & \sigma_2(\omega_2) & \dots & \sigma_n(\omega_2) \\ \vdots & \vdots & & \vdots \\ \sigma_1(\omega_m) & \sigma_2(\omega_m) & \dots & \sigma_n(\omega_m) \end{pmatrix}, \tag{7}$$

where $G = \{\sigma_1, \dots, \sigma_n\}$ and $\omega_1, \dots, \omega_m$ are elements of K that are independent over F.

Suppose first that $n > [K : F]$. In that case, let $m = [K : F]$ and let $\omega_1, \dots, \omega_m$ in (7) be a basis of K over F. As $n > m$, there is a linear relation between columns. Thus, for some $a_1, \dots, a_n \in K$, not all zero, we have

$$\begin{pmatrix} \sigma_1(\omega_1) & \sigma_2(\omega_1) & \dots & \sigma_n(\omega_1) \\ \sigma_1(\omega_2) & \sigma_2(\omega_2) & \dots & \sigma_n(\omega_2) \\ \vdots & \vdots & & \vdots \\ \sigma_1(\omega_m) & \sigma_2(\omega_m) & \dots & \sigma_n(\omega_m) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Let $\psi = a_1\sigma_1 + \cdots + a_n\sigma_n$, then the last identity means that $\psi(\omega_i) = 0$ for $i = 1, \dots, m$. We want to show that ψ is identically zero on K, which will contradicts independence of characters

(Corollary 9.2.2). Since $\{\omega_1, \dots, \omega_m\}$ are a basis for K over F, it is enough to show that for any $f_i \in F$ we have $\psi(\sum_{j=1}^m f_j\omega_j) = 0$. Indeed,

$$\begin{aligned}
\psi(\sum_{j=1}^m f_j\omega_j) &= (\sum_{i=1}^n a_i\sigma_i)(\sum_{j=1}^m f_j\omega_j) \\
&= \sum_i a_i \cdot \sum_j \sigma_i(f_j\omega_j) \\
&= \sum_{i,j} a_i f_j \cdot \sigma_i(\omega_j) \\
&= \sum_{j=1}^m f_j \cdot (\sum_{i=1}^n a_i\sigma_i(\omega_j)) \\
&= \sum_{j=1}^m f_j \cdot \psi(\omega_j) \\
&= 0.
\end{aligned}$$

(We have used the fact that σ_i is multiplicative and the identity on F to get $\sigma_i(f_j\omega_j) = f_j\sigma_i(\omega_j)$.) Thus, we must have $n \leq [K : F]$.

Suppose now that $n < [K : F]$. Then, we may find $m = n + 1$ elements $\{\omega_i : i = 1, \dots, m\}$ of K that are linearly independent over F. Consider then the same matrix (7), where now the number of rows is greater then number of columns. There is thus a non-zero row vector $\underline{\beta} = (\beta_1, \dots, \beta_m)$ with entries in K such that

$$(\beta_1, \dots, \beta_m)\begin{pmatrix} \sigma_1(\omega_1) & \sigma_2(\omega_1) & \dots & \sigma_n(\omega_1) \\ \sigma_1(\omega_2) & \sigma_2(\omega_2) & \dots & \sigma_n(\omega_2) \\ \vdots & \vdots & & \vdots \\ \sigma_1(\omega_m) & \sigma_2(\omega_m) & \dots & \sigma_n(\omega_m) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{8}$$

Among all such non-zero vectors $\underline{\beta}$ we choose one with the least number of non-zero coordinates. We may assume, to simplify notation, that these non-zero coordinates are the first r coordinates, where $r > 1$ necessarily (as each entry of the matrix is non-zero). Thus, $\beta_1, \dots, \beta_r$ are non-zero and $\beta_{r+1}, \dots, \beta_m$ are zero. We may assume (by dividing the vector by β_r) that $\beta_r = 1$.

First, note that we cannot have $\beta_i \in F$ for all i. Indeed, suppose that σ_i is the identity map. Then, we get

$$\beta_1\sigma_i(\omega_1) + \dots + \beta_m\sigma_i(\omega_m) = \beta_1\omega_1 + \dots + \beta_m\omega_m = 0,$$

which contradicts the linear independence over F of $\omega_1, \dots, \omega_m$. Thus, some $\beta_i \notin F$, and after renaming the σ's, we may assume that $\beta_1 \notin F$. There is therefore some $\tau \in G$ such that $\tau(\beta_1) \neq \beta_1$. Apply this τ to (8). We find that

$$(\tau(\beta_1), \dots, \tau(\beta_m))\begin{pmatrix} \tau\sigma_1(\omega_1) & \tau\sigma_2(\omega_1) & \dots & \tau\sigma_n(\omega_1) \\ \tau\sigma_1(\omega_2) & \tau\sigma_2(\omega_2) & \dots & \tau\sigma_n(\omega_2) \\ \vdots & \vdots & & \vdots \\ \tau\sigma_1(\omega_m) & \tau\sigma_2(\omega_m) & \dots & \tau\sigma_n(\omega_m) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{9}$$

But note that for a suitable permutation matrix P (whose effect is to permute the columns), we have

$$\begin{pmatrix} \tau\sigma_1(\omega_1) & \tau\sigma_2(\omega_1) & \dots & \tau\sigma_n(\omega_1) \\ \tau\sigma_1(\omega_2) & \tau\sigma_2(\omega_2) & \dots & \tau\sigma_n(\omega_2) \\ \vdots & \vdots & & \vdots \\ \tau\sigma_1(\omega_m) & \tau\sigma_2(\omega_m) & \dots & \tau\sigma_n(\omega_m) \end{pmatrix} = \begin{pmatrix} \sigma_1(\omega_1) & \sigma_2(\omega_1) & \dots & \sigma_n(\omega_1) \\ \sigma_1(\omega_2) & \sigma_2(\omega_2) & \dots & \sigma_n(\omega_2) \\ \vdots & \vdots & & \vdots \\ \sigma_1(\omega_m) & \sigma_2(\omega_m) & \dots & \sigma_n(\omega_m) \end{pmatrix} \cdot P.$$

Since permutation matrices are invertible, by multiplying (9) by P^{-1} we deduce that

$$(\tau(\beta_1), \dots, \tau(\beta_m)) \begin{pmatrix} \sigma_1(\omega_1) & \sigma_2(\omega_1) & \dots & \sigma_n(\omega_1) \\ \sigma_1(\omega_2) & \sigma_2(\omega_2) & \dots & \sigma_n(\omega_2) \\ \vdots & \vdots & & \vdots \\ \sigma_1(\omega_m) & \sigma_2(\omega_m) & \dots & \sigma_n(\omega_m) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Now subtract that from (8) to find that

$$(\beta_1 - \tau(\beta_1), \dots, \beta_m - \tau(\beta_m)) \begin{pmatrix} \sigma_1(\omega_1) & \sigma_2(\omega_1) & \dots & \sigma_n(\omega_1) \\ \sigma_1(\omega_2) & \sigma_2(\omega_2) & \dots & \sigma_n(\omega_2) \\ \vdots & \vdots & & \vdots \\ \sigma_1(\omega_m) & \sigma_2(\omega_m) & \dots & \sigma_n(\omega_m) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{10}$$

But the vector $(\beta_1 - \tau(\beta_1), \dots, \beta_m - \tau(\beta_m))$ is non-zero and has at most $r-1$ non-zero coordinates (since $\beta_r - \tau(\beta_r) = 1 - \tau(1) = 0$). This is a contradiction and the proof is complete. □

Corollary 9.3.2. *Let K/F be a finite extension of fields. Then* [2]

$$|\mathrm{Aut}(K/F)| \leq [K:F],$$

with equality if and only if $F = K^{\mathrm{Aut}(K/F)}$. That is, K/F is Galois if and only if $F = K^{\mathrm{Aut}(K/F)}$.

Proof. Note first that if $K = F(\alpha_1, \dots, \alpha_n)$ and $R_1, \dots, R_n$ are the roots of the minimal polynomials of $\alpha_1, \dots, \alpha_n$ over F, respectively, then

$$\mathrm{Aut}(K/F) \hookrightarrow \Sigma_{R_1} \times \dots \times \Sigma_{R_n}.$$

Thus, $G := \mathrm{Aut}(K/F)$ is a finite group. Let K_1 be the field K^G. Using Theorem 9.3.1 , we have the following diagram

$$\begin{tikzcd} K \arrow[d, dash, "|G|"] \\ K_1 \arrow[d, dash] \\ F \end{tikzcd}$$

This shows that $|G| = |\mathrm{Aut}(K/F)| \leq [K:F]$, with equality if and only if $F = K^{\mathrm{Aut}(K/F)}$. □

Corollary 9.3.3. *Let $G < \mathrm{Aut}(K)$ be a finite group. Then $\mathrm{Aut}(K/K^G) = G$ and K/K^G is Galois.*

Proof. Theorem 9.3.1 gives us that K/K^G is a finite extension and the last corollary that K/K^G is Galois and $|\mathrm{Aut}(K/K^G)| = [K:K^G] = |G|$. As certainly $G \subseteq \mathrm{Aut}(K/K^G)$, we must have equality because both groups have the same cardinality. □

[2] Up to till this point we only knew this inequality when K is a splitting field of a polynomial in $F[x]$ (Theorem 9.1.1).

The next corollary would be used in the main theorem of Galois theory (Theorem 9.4.1).

Corollary 9.3.4. *Let $G_1 \neq G_2$ be finite subgroups of* $\mathrm{Aut}(K)$ *then $K^{G_1} \neq K^{G_2}$.*

Proof. If $K^{G_1} = K^{G_2}$ then $G_1 = \mathrm{Gal}(K/K^{G_1}) = \mathrm{Gal}(K/K^{G_2}) = G_2$. □

So far we have proven the following, (except that one direction of the last statement will be proven next):

Summary: Let K/F be a finite extension.

- K/F is Galois iff $\mathrm{Aut}(K/F) = [K : F]$ (the definition).
- K/F is Galois iff $F = K^{\mathrm{Aut}(K/F)}$.
- K/F is Galois iff K is the splitting field of a separable polynomial in $F[x]$.

As said, we still need to prove that a finite field extension K/F which is Galois is the splitting field of a separable polynomial. The proof is worth close scrutiny as it includes another method of calculating the minimal polynomial of an element $\alpha \in K$ over F.

Theorem 9.3.5. *Let K/F be a finite extension of fields. If K/F is Galois then it is the splitting field of a separable polynomial $f(x) \in F[x]$.*

Proof. Let $G = \{1 = \sigma_1, \sigma_2, \dots, \sigma_n\}$ be the Galois group $\mathrm{Gal}(K/F)$. For $\alpha \in K$ consider the **conjugates** of α:

$$\alpha = \sigma_1(\alpha), \sigma_2(\alpha), \dots, \sigma_n(\alpha).$$

Let $\alpha = \alpha_1, \dots, \alpha_s$ the distinct elements from this list of conjugates. Consider the polynomial

$$f(x) = \prod_{i=1}^{s}(x - \alpha_i).$$

Lemma 9.3.6. *f is the minimal polynomial of α over F.*

Proof. The Galois group G acts on $K[x]$, $g \mapsto {}^{\sigma}g$ for $\sigma \in G$. Since $F = K^G$, the polynomials fixed by the action of G are precisely $F[x]$. Let $\sigma \in G$ then

$${}^{\sigma}f(x) = \prod_{i=1}^{s}(x - \sigma(\alpha_i)).$$

As σ permutes the conjugates of α, we conclude that ${}^{\sigma}f = f, \forall \sigma \in G$ and so $f(x) \in F[x]$, and of course $f(\alpha) = 0$.

Let g be the minimal polynomial of α over F. Then $g|f$ and $g(\alpha) = 0$. Then, for all $\sigma \in G$ we have $\sigma(g(\alpha)) = 0$, but $\sigma(g(\alpha)) = g(\sigma(\alpha))$ as the coefficients for g are fixed by σ. We conclude that $\alpha_1, \dots, \alpha_s$ are all roots of g. Therefore $f|g$ and so they must be equal. □

It follows from the very definition of f that it is separable.

Now, let $\omega_1, \dots, \omega_n$ be a basis for K over F. The minimal polynomial f_{ω_i} of ω_i over F is separable. Let $f_1, \dots, f_t$ be the distinct polynomials among $f_{\omega_1}, \dots, f_{\omega_n}$ and let

$$f = f_1 f_2 \cdots f_t.$$

Then f is a polynomial over F and K is the splitting field of f. Finally, notice that f is separable, because for $i \neq j$, $\gcd(f_i, f_j) = 1$ and so f_i and f_j don't have a common root. □

Remark 9.3.7. A remark about terminology. We have defined an extension K/F as **normal** if it is the splitting field of a collection of polynomials. Often, an extension K/F is defined to be normal if whenever an irreducible polynomial $f(x) \in F[x]$ has a root in K it splits in K. Clearly, if the second definition holds then so does the first (take the collection of polynomials to be the collection of minimal polynomials $\{f_\alpha : \alpha \in K\}$, where f_α is the minimal polynomial of α over K). But, a priori, the second definition seems to be stronger. In fact, Exercise 18 shows that the two definitions are equivalent.

We call an extension K/F **separable** if every element α in K solves a (non-zero) separable polynomial $f(x) \in F[x]$. Equivalently, K/F is an algebraic extension and the minimal polynomial of every element of K is separable.

We summarize much of our conclusions thus far in the following theorem.

Theorem 9.3.8. *Let K/F be a finite extension of fields. The following are equivalent.*

1. *K/F is Galois, i.e. $|\mathrm{Aut}(K/F)| = [K:F]$.*
2. *$F = K^{\mathrm{Aut}(K/F)}$.*
3. *K is the splitting field of a separable polynomial $f(x) \in F[x]$.*
4. *K/F is normal and separable.*

Proof. We already know that (1) - (3) are equivalent. The proof of Theorem 9.3.5, shows that every element of K has a separable minimal polynomial and so K/F is separable. As said, K/F is a splitting field of a polynomial it is normal (in either sense).

Assume then that (4) holds. By choosing a basis $\omega_1, \dots, \omega_n$ for K over F, taking for each its own separable minimal polynomial and multiplying them all together to get a polynomial $f(x) \in F[x]$, we conclude that K is the splitting field (we use "normal" here!) of a polynomial f each of whose irreducible factors is separable. Thus, by Theorem 9.1.1, K/F is a Galois extension. □

9.4. **The main theorem of Galois theory.**

Theorem 9.4.1. *Let K/F be a finite Galois extension, $G = \mathrm{Gal}(K/F)$. There is a bijection:*

$$\begin{tikzcd}
\{\text{subfields } K \supseteq E \supseteq F\} \arrow[r, leftrightarrow] & \{\text{subgroups of } G\} \\
E \arrow[r, mapsto] & \mathrm{Aut}(K/E) \\
K^H & H \arrow[l, mapsto]
\end{tikzcd}$$

and the indicated maps are each other inverse. Furthermore,

(1) $H_1 \subseteq H_2 \Rightarrow K^{H_1} \supseteq K^{H_2}$ and $K_1 \supseteq K_2 \Rightarrow \mathrm{Aut}(K/K_1) \subseteq \mathrm{Aut}(K/K_2)$.
(2) $K^{H_1} \cap K^{H_2} = K^{\langle H_1, H_2 \rangle}$, $K^{H_1}K^{H_2} = K^{H_1 \cap H_2}$.
(3) $[K : K^H] = |H|$, $[K^H : F] = [G : H]$.
(4) K/E is Galois with Galois group $\mathrm{Aut}(K/E) = H$ if $E = K^H$.
(5) $E = K^H$ is Galois over F iff $H \lhd G$ and then $\mathrm{Gal}(K^H/F) = G/H$.

Proof. Recall that we proved the following:

- If $H \subseteq \mathrm{Aut}(K)$ is a finite subgroup then K/K^H is Galois with $\mathrm{Gal}(K/K^H) = H$ (Corollary 9.3.3).
- It K/E is Galois then $E = K^{\mathrm{Aut}(K/E)}$ (Corollary 9.3.2).

The maps indicated above are well-defined. If $K^{H_1} = K^{H_2}$ then $\mathrm{Gal}(K/K^{H_1}) = \mathrm{Gal}(K/K^{H_2})$ and so, using the first point above, $H_1 = H_2$. Thus, the map from subgroups to subfields is injective. It is also surjective. Indeed, since K is the splitting field over F of some separable polynomial $f(x) \in F[x]$, K is also the splitting field of f over any subfield E of K that contains F. Thus, for any such E, K/E is Galois, clearly $\mathrm{Aut}(K/E) < G$ and $E = K^{\mathrm{Aut}(K/E)}$ by the second point above.

It follows that every subfield E is of the form K^H for some subgroup H and then $\mathrm{Aut}(K/K^H) = H$ and that shows the map from subfields to subgroup is the inverse of the map from subgroups to subfields.

The maps are clearly inclusion reversing and so form an order reversing bijection between the poset of subfields and the poset of subgroups. Thus, Claims (1), (2) follow immediately, and (4) follows from our discussion.

Now, for a subgroup $H < G$, K/K^H is Galois with Galois group H and so $[K : K^H] = |H|$, and in particular for $H = \{e\}$, $[K : F] = |G|$. Multiplicativity now gives $[K^H : F] = [G : H]$. We have proven (3).

It remains to prove (5): We note first that G acts on both lattices (we use the word **lattice** for a poset, a partially ordered set, in which any two elements have a minimum and a maximum). If $g \in G$ and E is a subfield then $g(E)$ is another subfield of K containing F. This gives the action of G on the lattice of subfields. If $g \in G$ and $H < G$, we have the subgroup gHg^{-1}. It is clear that if $E = K^H$ then $g(E) = K^{gHg^{-1}}$. Thus, *the correspondence we have defined is equivariant for the action of* G. Consequently, the fixed points for those actions match. Namely, there is a bijection between normal subgroups of G and subfields E of K, $E \supseteq F$, with the property that $g(E) = E$ for all $g \in G$. The proof now follows from the following lemma.

Lemma 9.4.2. *Let K/F be a finite Galois extension with Galois group G and E a field such that $F \subseteq E \subseteq K$. Then, E is a Galois extension of F if and only if for all $\sigma \in G$ we have $\sigma(E) = E$.*

Proof. (Lemma) Suppose first that E/F is Galois, hence the splitting field of some separable polynomial $f(x) \in F[x]$. Then $E = F(\alpha_1, \dots, \alpha_n)$ where the α_i are the roots of f. For any $\sigma \in G$ we have $\sigma(f(\alpha_i)) = {}^{\sigma}f(\sigma(\alpha_i)) = f(\sigma(\alpha_i))$ and so $\sigma(\alpha_i) = \alpha_j$ for some j that depends on i and σ. But, at any rate, σ permutes the roots of f and so $\sigma(E) = E$.

Suppose now conversely that $\sigma(E) = E$ for all $\sigma \in G$. As K/F is a finite Galois extension, it is a finite separable extension and so also E/F is a finite separable extension. We will prove E/F is

Galois by showing it is a normal extension. Let us first choose a separable polynomial $f \in F[x]$ such that K is the splitting field of f over F. Let $h \in F[x]$ be an irreducible polynomial with a root $\alpha \in E$. The polynomial h splits in K; it is enough to show that if $\beta \in K$ is a root of h then $\beta \in E$. We consider the following diagram.

$$\begin{tikzcd}
K \arrow[r, dashed, "\sigma"] \arrow[d, dash] & K \arrow[d, dash] \\
F(\alpha) \arrow[r, "\sigma_1"] \arrow[d, dash] & F(\beta) \arrow[d, dash] \\
F \arrow[r, "Id"] & F
\end{tikzcd}$$

There is an isomorphism σ_1 as indicated that takes α to β and, by Theorem 9.1.1, it can be extended to an isomorphism σ, indicated by the dashed arrow. However, $\sigma(E) = E$ and so $\beta = \sigma(\alpha) \in E$. □

The final point to prove is that for $H \lhd G$, $\mathrm{Gal}(K^H/F) = G/H$. Indeed, there is a natural homomorphism obtained by restriction:

$$G \to \mathrm{Gal}(K^H/F), \quad \sigma \mapsto \sigma|_{K^H}.$$

The fact that this map is well-defined follows from the Lemma. The kernel is clearly H. Thus, we have an injection $G/H \hookrightarrow \mathrm{Gal}(K^H/F)$. As $|G/H| = [G : H] = [K^H : F] = |\mathrm{Gal}(K^H/F)|$, we have $G/H = \mathrm{Gal}(K^H/F)$. □

The following Corollary is left as an exercise. We will also give as an exercise that the assumption of E/F be separable is necessary.

Corollary 9.4.3. *Let E/F be a finite separable extension. Then there are finitely many subfields $F \subseteq L \subseteq E$.*

10. Examples of Galois extensions

In this section we give several examples of Galois extensions and the matching between subfields and subgroups. We will use K/F to denote the Galois extension and G for $\mathrm{Gal}(K/F)$.

10.1. **Projecting to subfields.**

We begin with a simple lemma, whose proof is straight-forward:

Lemma 10.1.1. *Let H be a subgroup of G and assume that $|H|$ is invertible in F, that is, that the characteristic of F doesn't divide the order of H. The function*

$$\pi_H : K \to K, \qquad \pi_H(k) = \frac{1}{|H|} \sum_{h \in H} h(k),$$

is an F-linear projection map of K onto K^H.

Consequently, if we write $K = F(\alpha_1, dots, \alpha_n)$ where $\{\alpha_1, \dots, \alpha_n\}$ are a linear spanning set of K over F (and not merely generators) then

$$K^H = F(\pi_H(\alpha_1), \dots, \pi_H(\alpha_n)).$$

10.2. **Finite fields.** This is an example that we have analyzed without relying on the main theorem. If K/F is a degree n extension of finite fields, say $\mathbb{F}_{q^n}/\mathbb{F}_q$, then it is a Galois extension and $G \cong \mathbb{Z}/n\mathbb{Z}$ where 1 corresponds to $\mathrm{Fr}_q \in G$. The subfields are precisely $\mathbb{F}_{q^m}/\mathbb{F}_q$ where $m|n$ and the corresponding subgroups are $m\mathbb{Z}/n\mathbb{Z}$ (and these are all the subgroups of G).

$$\begin{tikzcd}
\mathbb{F}_{q^n} \arrow[d, dash, "n/m"] & \{0\} \arrow[d, dash, "n/m"] \\
\mathbb{F}_{q^m} \arrow[d, dash, "m"] & m\mathbb{Z}/n\mathbb{Z} \arrow[d, dash, "m"] \\
\mathbb{F}_q & \mathbb{Z}/n\mathbb{Z}
\end{tikzcd}$$

10.3. **Bi-quadratic example.** Let K/F be the extension $\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q}$. We previously determined that $G \cong S_2 \times S_2$. There are automorphisms $\sigma, \tau \in G$ such that $G = \{1, \sigma, \tau, \sigma\tau\}$ and such that the action of σ and τ is determined by the following table

	1	σ	τ	$\sigma\tau$
$\sqrt{2}$	$\sqrt{2}$	$-\sqrt{2}$	$\sqrt{2}$	$-\sqrt{2}$
$\sqrt{3}$	$\sqrt{3}$	$\sqrt{3}$	$-\sqrt{3}$	$-\sqrt{3}$

The subfield corresponding to $\langle\sigma\rangle$ is $\mathbb{Q}(\sqrt{3})$, to $\langle\tau\rangle$ is $\mathbb{Q}(\sqrt{2})$ and to $\langle\sigma\tau\rangle$ is $\mathbb{Q}(\sqrt{6})$. The diagrams are then

$$\begin{tikzcd}
& \mathbb{Q}(\sqrt{2}, \sqrt{3}) \arrow[dl, dash, "2"'] \arrow[d, dash, "2"] \arrow[dr, dash, "2"] & \\
\mathbb{Q}(\sqrt{2}) \arrow[dr, dash, "2"'] & \mathbb{Q}(\sqrt{3}) \arrow[d, dash, "2"] & \mathbb{Q}(\sqrt{6}) \arrow[dl, dash, "2"] \\
& \mathbb{Q} &
\end{tikzcd}
\qquad
\begin{tikzcd}
& \{1\} \arrow[dl, dash, "2"'] \arrow[d, dash, "2"] \arrow[dr, dash, "2"] & \\
\langle\tau\rangle \arrow[dr, dash, "2"'] & \langle\sigma\rangle \arrow[d, dash, "2"] & \langle\sigma\tau\rangle \arrow[dl, dash, "2"] \\
& G &
\end{tikzcd}$$

10.4. **A cyclotomic example.** We consider the extension K/F where $K = \mathbb{Q}(\zeta_7)$ and $F = \mathbb{Q}$. Some of the considerations below deserve very close reading as they apply in much more general circumstances.

The Galois group is isomorphic to $(\mathbb{Z}/7\mathbb{Z})^\times = \{1, 2, 3, 4, 5, 6\}$, which is a cyclic group of order 6. If σ_a is the automorphism that corresponds to the congruence class a then $\sigma_a(\zeta) = \zeta^a$ for any 7-th root of unity ζ. If $H < G$ is a subgroup then, using Lemma 10.1.1,

$$\begin{aligned}
\mathbb{Q}(\zeta_7)^H &= \mathbb{Q}(\{\pi_H(\zeta_7^i)) : i = 1, \dots, 5, 6\}) \\
&= \mathbb{Q}(\{\pi_H(\sigma(\zeta_7))) : \sigma \in G\}) \\
&= \mathbb{Q}(\{\sigma\pi_H(\zeta_7)) : \sigma \in G\}) \\
&= \mathbb{Q}(\pi_H(\zeta_7)).
\end{aligned}$$

In the first equality we have used that the minimal polynomial of ζ_7 has degree 6 and a small argument that also $\{\zeta_7, \dots, \zeta_7^6\}$ is a basis over $\mathbb{Q}$ (replacing the usual basis $\{1, \zeta_7, \dots, \zeta_7^5\}$); in the second equality we wrote these powers of ζ_7 as the images under G. We have then used that H is normal so $H\sigma = \sigma H$ (this also follows since G is abelian, but the more general setting is when H is just normal). The last equality is the most special. The argument is that since G is abelian *any* sub extension $E/\mathbb{Q}$ is Galois and so all the images of $\pi_H(\zeta_7)$ under G are already in $\mathbb{Q}(\pi_H(\zeta_7))$.

The non-trivial subgroups of G are $H = \{1, 2, 4\}$ and $J = \{1, 6\}$. We let $\eta_H = pi_H(\zeta_7), \eta_J = pi_J(\zeta_7)$. Then we have the following diagrams.

$$\begin{tikzcd}
& \mathbb{Q}(\zeta_7) \arrow[dl, dash, "3"'] \arrow[dr, dash, "2"] & \\
\mathbb{Q}(\eta_H) \arrow[dr, dash, "2"'] & & \mathbb{Q}(\eta_J) \arrow[dl, dash, "3"] \\
& \mathbb{Q} &
\end{tikzcd}
\qquad
\begin{tikzcd}
& \{1\} \arrow[dl, dash, "3"'] \arrow[dr, dash, "2"] & \\
H \arrow[dr, dash, "2"'] & & J \arrow[dl, dash, "3"] \\
& G &
\end{tikzcd}$$

To have a presentation of the subfields as extensions of $\mathbb{Q}$ we make use of Lemma 9.3.6 and our knowledge of the Galois group to claim that the minimal polynomial of η_H is $(x - \eta_H)(x - \bar{\eta}_H) = x^2 + x + 2$ and that the minimal polynomial of η_J is $(x - \eta_J)(x - \sigma_2(\eta_H))(x - \sigma_4(\eta_G)) = x^3 + x^2 - 2x - 1$. Thus,

$$\mathbb{Q}(\eta_H) \cong \mathbb{Q}[x]/(x^2 + x + 2), \qquad \mathbb{Q}(\eta_J) = \mathbb{Q}[x]/(x^3 + x^2 - 2x - 1).$$

10.5. S_3 **example.** Let K be the splitting field over $\mathbb{Q}$ of the polynomial $x^3 - 2$. This is a special case of Exercise 26. The field K is equal to $\mathbb{Q}(\sqrt[3]{2}, \zeta_3) = \mathbb{Q}(\sqrt[3]{2})\mathbb{Q}(\zeta_3)$ and making use of Corollary 12.1.2 we infer that $[K : \mathbb{Q}] = 6$. The Galois group maps onto a transitive subgroup of S_3. We deduce that in fact

$$G \cong S_3.$$

Given a permutation of the roots $\{\sqrt[3]{2}, \zeta_3\sqrt[3]{2}, \zeta_3^2\sqrt[3]{2}\}$, where, say, $\sqrt[3]{2}$ is real, we can determine the action on K as $\sigma(\sqrt[3]{2})$ is then provided and $\sigma(\zeta_3) = \sigma(\zeta_3\sqrt[3]{2})/\sigma(\sqrt[3]{2})$ that are also provided. The list of subgroups of S_3 is well-known. The diagrams provide the corresponding subfields.

$$\begin{tikzcd}
& \mathbb{Q}(\sqrt[3]{2}, \zeta_3) \arrow[dl, dash, "2"'] \arrow[d, dash, "2"'] \arrow[dr, dash, "2"] \arrow[drr, dash, "3"] & & \\
\mathbb{Q}(\zeta_3\sqrt[3]{2}) \arrow[dr, dash, "3"'] & \mathbb{Q}(\zeta_3^2\sqrt[3]{2}) \arrow[d, dash, "3"'] & \mathbb{Q}(\sqrt[3]{2}) \arrow[dl, dash, "3"'] & \mathbb{Q}(\zeta_3) \arrow[dll, dash, "2"] \\
& \mathbb{Q} & &
\end{tikzcd}
\qquad
\begin{tikzcd}
& \{1\} \arrow[dl, dash, "2"'] \arrow[d, dash, "2"'] \arrow[dr, dash, "2"] \arrow[drr, dash, "3"] & & \\
\langle (12) \rangle \arrow[dr, dash, "3"'] & \langle (13) \rangle \arrow[d, dash, "3"'] & \langle (23) \rangle \arrow[dl, dash, "3"'] & A_3 \arrow[dll, dash, "2"] \\
& S_3 & &
\end{tikzcd}$$

10.6. S_5 **example.** We provide here an example of an irreducible polynomial f of degree 5 with rational coefficients whose Galois group, namely, the Galois group of a splitting field of f, is S_5. The significance of that is that S_5 is not a solvable group. Later on we will return to this example to show that the general degree 5 equation cannot be solved by radicals. As it turns out, there are plenty such polynomials:

Proposition 10.6.1. *Let $f(x) \in \mathbb{Q}[x]$ be an irreducible quintic polynomial with exactly 3 real roots. Let $K \subset \mathbb{C}$ be a splitting field for f, then*

$$\mathrm{Gal}(K/\mathbb{Q}) \cong S_5.$$

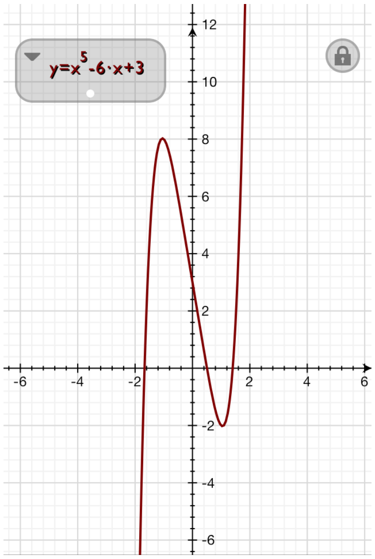
Proof. We know from general considerations that G is isomorphic to a transitive subgroup of S_5. Moreover, if α is a root of f then $[\mathbb{Q}(\alpha) : \mathbb{Q}] = 5$ and this divide $[K : \mathbb{Q}]$, hence the order of G. Therefore, G has an element of order 5. The elements of order 5 in S_5 are all 5-cycles, hence conjugate, and so we may assume that G contains the cycle (12345).

Consider the action of complex conjugation (restricted to K) on the roots of f. It is given by a transposition $(i\ j)$ for some $i < j$. Conjugating this transposition by $(12345)^{j-i}$ we find that $(j\ 2j-i)$, and hence $(i\ j)(j\ 2j-i) = (i\ j\ 2j-i)$ is in G. Conjugating $(i\ j)$ by a suitable power of (12345), we find a transposition $(k\ \ell)$ in G that is disjoint from $(i\ j)$. Thus, G contains a copy of the Klein group. It follows that the order of G is divisible by $3 \cdot 4 \cdot 5 = 60$. Thus, $G = A_5$ or S_5 (recall that S_n has a unique non-trivial normal subgroup, which is necessarily A_n, for $n \geq 5$. Also recall that a subgroup of index 2 is always normal). But G contains a transposition and so G must be equal to S_5. □

Remark 10.6.2. Note that the Galois theoretic aspect of the proof is very simple. The only complication is to show that a subgroup of S_5 containing a 5-cycle and a transposition must be equal to S_5. You may be able to find a simpler argument than that appearing above.

Example 10.6.3. As a concrete example, take the polynomial $f(x) = x^5 - 6x + 3$.

The polynomial is an irreducible polynomial by Eisenstein's criterion. The derivative $f'(x) = 5x^4 - 6$ has precisely 2 real roots that are $\pm\sqrt[4]{6/5}$. Further $f(-2) = -17, f(0) = 3, f(1) = -2$ and $f(2) = 23$. It is now easy to ascertain that the graph on the right represents this function. Thus, the Galois group of f is S_5.



11. Glorious Applications

11.1. **Constructing regular polygons.** We apply Galois theory to settle a classic geometric question about the construction of regular polygons in the plane. We will require the notion of **Fermat primes**. A prime p is a Fermat prime if $p = 2^n + 1$, for some integer n.

One can prove the following number theoretic lemma.

Lemma 11.1.1. *If 2^n+1 is prime then n itself is a power of 2.*

Proof. Suppose that $p = 2^n + 1$ is prime. Consider the order of the element 2 in $\mathbb{Z}/p\mathbb{Z}^\times$. Note that $1 < 2^n < p$ so the order of 2 is greater than n; on the other hand $2^n \equiv -1 \pmod{p}$ and so $(2^n)^2 = 2^{2n} \equiv 1 \pmod{p}$. Thus, the order of 2 divides $2n$. It follows that the order of 2 is exactly $2n$. By Lagrange, $2n|p-1$. But, $p-1 = 2^n$ and it follows that $n|2^{n-1}$, hence is a power of 2. □

Here are some examples of Fermat primes:

$$3 = 2^1+1, \quad 5 = 2^2+1, \quad 17 = 2^4+1, \quad 257 = 2^8+1, \quad 65537 = 2^{16}+1.$$

Euler had shown - and that was at a time an impressive feat - that $2^{32}+1 = 641 \times 6700417$ and so is not prime. It is unknown at this time (2014) if there are infinitely many Fermat primes. It is known that $2^{2^n}+1$ is composite for all $5 \leq n \leq 32$.

Theorem 11.1.2. *Let $n > 2$. One can construct a regular n-gon in the plane if and only if $n = 2^a p_1 p_2 \cdots p_s$, $a \geq 0, s \geq 0$, where the p_i are distinct Fermat primes.*

Proof. As we have seen, the construction of a regular polygon with n-sides is equivalent to the construction of $2\cos(2\pi/n) = \zeta_n + \bar{\zeta}_n$. If this quantity is constructible then $[\mathbb{Q}(\zeta_n + \bar{\zeta}_n) : \mathbb{Q}]$ is a power of 2.

The extension $\mathbb{Q}(\zeta_n)/\mathbb{Q}(\zeta_n + \bar{\zeta}_n)$ is quadratic as ζ_n solves the polynomial $(x-\zeta_n)(x-\bar{\zeta}_n) = x^2 - (\zeta_n + \bar{\zeta}_n)x + 1$ over $\mathbb{Q}(\zeta_n + \bar{\zeta}_n)$ and, since $n > 2$, $\mathbb{Q}(\zeta_n)$ is not contained in $\mathbb{R}$ while $\mathbb{Q}(\zeta_n + \bar{\zeta}_n)$ is contained in $\mathbb{R}$ (so the fields are distinct). At any rate, the diagram of fields given below shows that $\varphi(n)$ must be a power of 2 as well.

$$\begin{tikzcd}
\mathbb{Q}(\zeta_n) \arrow[d, dash, "2"] \arrow[dd, dash, bend right=60, "\varphi(n)"'] \\
\mathbb{Q}(\zeta_n + \bar{\zeta}_n) \arrow[d, dash] \\
\mathbb{Q}
\end{tikzcd}$$

As $\varphi(n) = n\prod_{p|n}(1-\frac{1}{p})$, we conclude that $\varphi(n)$ is a power of 2 if and only if $n = 2^a p_1 p_2 \dots p_s$, where $2 < p_1 < \cdots < p_s$ are Fermat primes.

Conversely, suppose that n has such a factorization. As $\mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q})$ is an abelian 2 group, it has a filtration

$$\{0\} = H_0 < H_1 < \cdots < H_a = \mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}), \quad H_1 = \mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}(\zeta_n + \bar{\zeta}_n)), \quad |H_i| = 2^i.$$

Correspondingly we have the sequence of fields

$$\mathbb{Q}(\zeta_n) \supset \mathbb{Q}(\zeta_n + \bar{\zeta}_n) = \mathbb{Q}(\zeta_n)^{H_1} \supset \mathbb{Q}(\zeta_n)^{H_2} \supset \cdots \supset \mathbb{Q},$$

where each extension is quadratic. It remains to complete our considerations of construction by straightedge and compass as follows. If ℓ is a negative real number, we say that ℓ is constructible if $-\ell$ is constructible.

Lemma 11.1.3. *Let K be a subfield of $\mathbb{R}$. Suppose that every element in the field K is constructible and let L be a real quadratic extension of K then every element of L is constructible.*

Proof. [Lemma] Let $\alpha \in L - K$. Then $\{1, \alpha, \alpha^2\}$ are linearly dependent over K, but $\{1, \alpha\}$ are not. Thus, for some $k_1, k_2 \in K$ we have $\alpha^2 + k_1\alpha + k_2 = 0$. Therefore,

$$\alpha = \frac{-k_1 \pm \sqrt{k_1^2 - 4k_2}}{2}.$$

As $\alpha \in \mathbb{R}$ the discriminant $k_1^2 - 4k_2$ is a constructible real number. We have seen (see Example 4.1.5 and the exercises) that a square root, sum, difference and quotient of constructible numbers are constructible and so α is constructible. □

□

11.2. **The Primitive Element Theorem.** A field extension K/F is called **simple** if $K = F(\theta)$ for some element $\theta \in K$. The name is misleading; there is nothing simple about such field extensions. The following theorem illustrates this.

Theorem 11.2.1 (Primitive Element Theorem). *Let K/F be a finite separable extension. Then there is an element $\alpha \in K$ such that $K = F(\alpha)$.*

Proof. The only place separability is used in the proof is in concluding that there are finitely many subfields $F \subseteq E \subseteq K$ (Corollary 9.4.3).

Assume that F is an infinite field. We will give a simpler argument for finite fields later. Consider then among all subfields of K a subfield maximal relative to inclusion that is of the form $F(\beta)$. If $F(\beta) \neq K$, choose some $\gamma \in K - F(\alpha)$. For every non-zero $a \in F$ we have the subfield of K given by $F(a\beta + \gamma)$. As a ranges over F we get an infinite list of such subfields and so for some $a_1 \neq a_2$ elements of F^* we have $F(a_1\beta + \gamma) = F(a_2\beta + \gamma)$. As $\beta = (a_1 - a_2)^{-1}((a_1\beta + \gamma) - (a_2\beta + \gamma))$ we have that β (and hence also γ) are in $F(a_1\beta + \gamma)$. Thus, $F(a_1\beta + \gamma)$ contradicts the maximality of $F(\beta)$. We mush have $F(\beta) = K$.

We complete the proof by addressing the case of finite fields. In this case $F = \mathbb{F}_{p^m}$ and $K = \mathbb{F}_{p^n}$. Let $\alpha \in K$ be a generator of the cyclic group K^*. Then clearly $K = F(\alpha)$ as $K = \{0\} \cup \{\alpha^j : j = 1, 2, \dots, p^n - 1\}$. □

Remark 11.2.2. At this point, we have closed a circle. We find that every Galois extension K/F can be written in the form $K = F(\alpha)$, where the minimal polynomial $f(x)$ of α over F splits over K, and, in fact, K is the splitting field of $f(x)$ over F. These are the assumptions of the discussion in Proposition 8.1.2. The Galois group $\mathrm{Gal}(K/F)$ is then in bijection with the roots of $f(x)$ by $\sigma \mapsto \sigma(\alpha)$. The Galois group is a transitive subgroup, acting without fixed points, of the group S_n, where $n = \deg(f)$.

11.3. **The Normal Basis Theorem.**

Theorem 11.3.1 (Normal Basis Theorem). *Let K/F be a finite Galois extension with Galois group G. There is an element $\alpha \in K$ such that $\{\sigma(\alpha) : \sigma \in G\}$ is a basis for K as a vector space over F.*

I have decided to omit the proof of this result. Although one can prove the theorem with only the tools at our disposal already, the proof is rather ad-hoc. The natural setting is that of representations of finite groups, where the proof becomes natural.

11.4. **The Fundamental Theorem of Algebra.**

Theorem 11.4.1 (Fundamental Theorem of Algebra). *The field of complex numbers is algebraically closed.*

Proof. We first do a series of reductions. Let $f(x) \in \mathbb{C}[x]$ be a non-constant polynomial. It is enough to show that f has a root in $\mathbb{C}$. Suppose it doesn't. Let $\bar{f}(x)$ be the polynomial obtained by taking the complex conjugates of the coefficients of f. Then also $\bar{f}$ doesn't have a root in $\mathbb{C}$ because of the identity $\overline{f(z)} = \bar{f}(\bar{z})$. Therefore, the polynomial $g(x) = f(x)\bar{f}(x)$, which is a polynomial in $\mathbb{R}[x]$, doesn't have a root in $\mathbb{C}$. Taking one of its irreducible factors, we conclude that it is enough to prove the following statement: *Any irreducible polynomial $g(x) \in \mathbb{R}[x]$ has a root in $\mathbb{C}$.*

Let then $g(x)$ be such a polynomial and let $K \supseteq \mathbb{R}$ be a splitting field for $g(x)$ contained in an algebraic closure of $\mathbb{C}$. As $\mathbb{R}(i) = \mathbb{C}$ is a splitting field for x^2+1, we conclude that $K(i)$ is a splitting field over $\mathbb{R}$ for the polynomial $g(x)(x^2+1)$. Since in characteristic zero every irreducible polynomial is separable (see Proposition 9.1.4) it follows from Corollary 9.1.2 that $K(i)/\mathbb{R}$ is Galois with Galois group, say, G. Let P be a 2-Sylow subgroup of G.

$$\begin{tikzcd}
 & K(i) \arrow[dl, no head] \arrow[d, no head, "|P|"] \arrow[dr, no head] & \\
K \arrow[dr, no head] & K(i)^P \arrow[d, no head, "|G|/|P|"] & \mathbb{R}(i) \arrow[dl, no head] \\
 & \mathbb{R} &
\end{tikzcd}$$

The extension $K(i)^P/\mathbb{R}$ is of odd degree.

Let θ be an element of $K(i)^P$. Let h be the monic minimal polynomial of θ over $\mathbb{R}$. The extension $\mathbb{R}(\theta) \supset \mathbb{R}$, being a sub extension of $K(i)^P/\mathbb{R}$, is of odd degree and its degree is equal to the degree of h. Thus, h is a polynomial of odd degree. As $\lim_{x \to -\infty} h(x) = -\infty$ and $\lim_{x \to \infty} h(x) = \infty$, by the intermediate value theorem, for some $x_0 \in \mathbb{R}$, $h(x_0) = 0$. But, h is irreducible over $\mathbb{R}$, so it must be that h is linear, $x_0 = \theta$ and so $\theta \in \mathbb{R}$ and $K(i)^P = \mathbb{R}$. We conclude that G is a 2-group.

We now consider the following diagram of fields and subgroups are as follows (where G_2 and K_2 will be explained below).

$$\begin{tikzcd}
K(i) \arrow[d, no head] & \{1\} \arrow[d, no head] \\
K_2 \arrow[d, no head] & G_2 \arrow[d, no head] \\
\mathbb{C} \arrow[d, no head] & G_1 \arrow[d, no head] \\
\mathbb{R} & G
\end{tikzcd}$$

Here G_1 is an index 2 subgroup of G corresponding to $\mathbb{C}$. If G is of order 2, our proof is done, because it implies $K(i) = \mathbb{C}$ and so K, the splitting field of g, is contained in $\mathbb{C}$. Else, there is an index 2 subgroup of G_1, denoted G_2, and a corresponding field K_2 containing $\mathbb{C}$. K_2 is a quadratic extension of $\mathbb{C}$. We will show no such exists, hence deriving a contradiction.

Since K_2 is a quadratic extension of $\mathbb{C}$, for any $t \in K_2 - \mathbb{C}$, $\{1, t\}$ is a basis for K_2 over $\mathbb{C}$. Thus t^2 is a linear combination of 1 and t and we conclude that $t^2 + rt + s = 0$ for some $r, s \in \mathbb{C}$. Then, $t = (-r \pm \sqrt{r^2 - 4s})/2$. But, if z is a complex number, write it as $z = re^{i\theta}$ and conclude that $\sqrt{z} = \pm\sqrt{r}e^{i\theta/2}$ exists in $\mathbb{C}$. This implies then that $t \in \mathbb{C}$ and that's a contradiction. □

It is interesting to analyze what goes into the proof. Besides Galois theory, we needed to show square roots exist in $\mathbb{C}$ and for that we used the polar representation of complex numbers and that seems to be using power series and convergence in $\mathbb{C}$ (through the manipulation of $e^{i\theta}$). At another place, we have used the intermediate value theorem for $\mathbb{R}$. That seems more fair, because $\mathbb{R}$ itself is defined as a completion of the rational numbers and one would expect such metric aspects to come into the proof. In fact, one can show just using the intermediate value theorem that square roots can be taken in $\mathbb{C}$ (Exercise 37).

12. Galois groups and operations on fields

There are several ways we can combine the notion of a Galois extension with simple operations on fields. For example, suppose that K_1, K_2 are Galois extensions of F, with respective Galois groups G_1, G_2, contained in some common field K. We can ask if K_1K_2, or $K_1 \cap K_2$, is a Galois extension of F and, if so, how we may describe the Galois group in the terms of G_1, G_2. Also, suppose that F' is any field extension of F contained in K; we may ask if the **base change** extension K_1F'/F' is Galois extension and if so how to describe its Galois group in terms of G_1. We will answer those questions in this section. The results are very useful in applying Galois theory to concrete situations.

12.1. **Base change.**

Proposition 12.1.1. *Let K/F be a Galois extension, where $K \subseteq \Omega$, Ω a field. Let F' be any extension of F contained in Ω. The extension KF'/F' is Galois and*

$$\mathrm{Gal}(KF'/F') \cong \mathrm{Gal}(K/K \cap F').$$

Proof. The diagram of fields is the following:

$$\begin{tikzcd}
 & \Omega \arrow[d, dash] & \\
 & KF' \arrow[dl, dash] \arrow[dr, dash] & \\
K \arrow[dr, dash] & & F' \arrow[dl, dash] \\
 & K \cap F' \arrow[d, dash] & \\
 & F &
\end{tikzcd}$$

As K/F is the splitting field of some separable polynomial $f(x) \in F[x] \subseteq F'[x]$, KF' is the splitting field of $f(x)$ over F' and so KF'/F' is a Galois extension. Because K/F is Galois, any element $\sigma \in \mathrm{Aut}(KF'/F)$ (sic!) satisfies $\sigma(K) = K$. In particular, the map

$$\pi : \mathrm{Gal}(KF'/F') \rightarrow \mathrm{Gal}(K/K \cap F'), \qquad \sigma \mapsto \sigma|_k,$$

is a well-defined group homomorphism. It is injective: σ being in the kernel implies that $\sigma|_K$ is the identity and also σ is the identity on F'. As every element of KF' is a polynomial expression in elements of K and of F', it follows that σ is the identity on KF' as well. It remains to show that the map π is surjective.

Let H be the image of π, a subgroup of $\mathrm{Gal}(K/K \cap F')$. Then

$$K^H = \{k \in K : k \in (KF')^{\mathrm{Gal}(K/K \cap F')}\} = K \cap F'.$$

From the Main Theorem it follows that $H = \mathrm{Gal}(K/K \cap F')$. □

Corollary 12.1.2. *Let K_1/F be a Galois extension and K_2/F an arbitrary finite extension, both contained in some field Ω. Then*

$$[K_1K_2 : F] = \frac{[K_1 : F] \cdot [K_2 : F]}{[K_1 \cap K_2 : F]}.$$

Note that the assumption that at least one of K_1/F, K_2/F are Galois is necessary. Indeed, taking $K_1 = \mathbb{Q}(\sqrt[3]{2}), K_2 = \mathbb{Q}(\omega\sqrt[3]{2})$, where ω is a primitive third of unity, we have seen that $[K_1K_2 : \mathbb{Q}] = 6$, while $\frac{[K_1:\mathbb{Q}]\cdot[K_2:\mathbb{Q}]}{[K_1 \cap K_2:\mathbb{Q}]} = \frac{3 \cdot 3}{1} = 9$. Let us now prove the corollary.

Proof. $[K_1K_2 : F] = [K_1K_2 : K_2] \cdot [K_2 : F]$. The Proposition implies $[K_1K_2 : K_2] = [K_1 : K_1 \cap K_2] = [K_1 : F]/[K_1 \cap K_2 : F]$. □

12.2. **Compositum and intersection.** We now consider compositum and intersection of Galois extensions.

Proposition 12.2.1. *Let K_1, K_2 be finite Galois extensions of F contained in some common field Ω.*

(1) K_1K_2/F is Galois and

$$\mathrm{Gal}(K_1K_2/F) = \{(\sigma, \tau) \in \mathrm{Gal}(K_1/F) \times \mathrm{Gal}(K_2/F) : \sigma|_{K_1\cap K_2} = \tau|_{K_1\cap K_2}\}.$$

(2) $K_1 \cap K_2$ is Galois and

$$\mathrm{Gal}(K_1 \cap K_2/F) = \mathrm{Gal}(K_1/F)/\mathrm{Gal}(K_1/K_1 \cap K_2).$$

Proof. First, as K_i is the splitting field of a separable polynomial $f_i(x) \in F[x]$, K_1K_2 is the splitting field of the polynomial $f_1(x)f_2(x)$, every irreducible factor of which is separable. Thus, K_1K_2/F is Galois. Thus, also K_1K_2/K_i is Galois for $i = 1, 2$; let $H_i := \mathrm{Gal}(K_1K_2/K_i)$. Here is the diagram of fields:

$$\begin{tikzcd}
 & \Omega \arrow[d, no head] & \\
 & K_1K_2 \arrow[dl, no head, "H_1"'] \arrow[dr, no head, "H_2"] & \\
K_1 \arrow[dr, no head] & & K_2 \arrow[dl, no head] \\
 & K_1 \cap K_2 \arrow[d, no head] & \\
 & F &
\end{tikzcd}$$

As K_i/F is Galois, $H_i \lhd \mathrm{Gal}(K_1K_2/F)$ and therefore also $\langle H_1, H_2\rangle = H_1H_2 \lhd \mathrm{Gal}(K_1K_2/F)$. As $K_1 \cap K_2 = (K_1K_2)^{\langle H_1, H_2\rangle}$, it is, too, Galois over F. [3]

It remains to determine $\mathrm{Gal}(K_1K_2/F)$ (the formula for $\mathrm{Gal}(K_1 \cap K_2/F)$ follows immediately from the Main Theorem). We have a homomorphism,

$$\mathrm{Gal}(K_1K_2/F) \to \mathrm{Gal}(K_1/F) \times \mathrm{Gal}(K_2/F), \qquad \sigma \mapsto (\sigma|_{K_1}, \sigma|_{K_2}).$$

Its image is contained in

$$H := \{(\sigma, \tau) \in \mathrm{Gal}(K_1/F) \times \mathrm{Gal}(K_2/F) : \sigma|_{K_1\cap K_2} = \tau|_{K_1\cap K_2}\}.$$

The homomorphism is injective. Thus, its image has as many elements as

$$|\mathrm{Gal}(K_1K_2/F)| = [K_1K_2 : F] = [K_1 : F] \cdot [K_2 : F] \cdot [K_1 \cap K_2 : F]^{-1} = [K_1 : F] \cdot [K_2 : K_1 \cap K_2]$$

(Corollary 12.1.2). On the other hand, this is the number of elements of H, because given $\sigma \in \mathrm{Gal}(K_1/F)$, the automorphism $\sigma|_{K_1\cap K_2}$ can be extended in as many ways as $|\mathrm{Gal}(K_2/K_1 \cap K_2)| = [K_2 : K_1 \cap K_2]$ to an automorphism in $\mathrm{Gal}(K_2/F)$. □

Corollary 12.2.2. *Let K_i/F be Galois extensions with Galois groups G_i and suppose that $K_1 \cap K_2 = F$, then K_1K_2/F is Galois and*

$$\mathrm{Gal}(K_1K_2/F) \cong \mathrm{Gal}(K_1/F) \times \mathrm{Gal}(K_2/F).$$

Example 12.2.3. The Galois group of $\mathbb{Q}(\sqrt{2}, \sqrt[3]{2}, \omega)$ is isomorphic to $\mathbb{Z}/2\mathbb{Z} \times S_3$. Indeed, taking $K_1 = \mathbb{Q}(\sqrt{2}), K_2 = \mathbb{Q}(\sqrt[3]{2}, \omega)$, we only need to check that $K_1 \cap K_2 = \mathbb{Q}$. But, if the intersection is not $\mathbb{Q}$ it must be K_1 and so $K_1 = \mathbb{Q}(\omega)$, because, by Galois theory, $\mathbb{Q}(\omega)$ is the *only* quadratic subfield of K_2. As $\mathbb{Q}(\omega) = \mathbb{Q}(\sqrt{-3})$ and $\mathbb{Q}(\sqrt{2})$ is a real field they can not be equal.

Making use of the Proposition, one can also easily calculate the Galois group of, say, $\mathbb{Q}(\sqrt[3]{2}, \omega, \sqrt[3]{5})$.

[3] Another proof that $K_1 \cap K_2/F$ is Galois is the following. Since K_1/F is finite separable, $K_1 \cap K_2/F$ is finite separable. We also proved in Exercise 19 that $K_1 \cap K_2/F$ is a splitting field, hence a normal extension. It follows that $K_1 \cap K_2/F$ is Galois.

13. Solvable and radical extensions, and the insolvability of the quintic

Our goal is to prove one of Évariste Galois' main achievements: there is no formula in radicals for solving the general polynomial equation $f(x)$ of degree 5. To keep the discussion "clean" we will assume from a certain point on that all our fields have characteristic 0. This is not necessary - the discussion can be extended to cover all characteristics - but it removes some technical awkwardness that would otherwise make the picture murkier. All concepts will be defined in due course. For now, we content ourselves with remarking the Galois's solution, a revolution in its time, is to give a necessary and sufficient for expressing the roots of a polynomial $f(x)$ by radicals in terms of its Galois group. To be precise, whether the Galois group is solvable or not. (This being the source of the terminology "solvable" for groups.) Since there are polynomials of degree 5 with non solvable Galois group - see Example 10.6.3 - it follows that there is no universal formula in radicals for solving all quintic polynomials.

13.1. **Cyclic extensions.** In this section we discuss **cyclic Galois extensions** K/F (that is, Galois extensions whose Galois group is a cyclic group) under a simlifying assumption of having "enough roots of unity" in F. Even under that assumption our discussion is just a glimpse of a general theory, called **Kummer theory**, that provides a complete classification of cyclic Galois extensions under the assumption on roots of unity.

Let F be a field and n a positive integer not divisible by the characteristic of F (a condition that holds automatically in characteristic zero). Assume that the polynomial $x^n - 1$ splits in F. Since $\gcd(x^n - 1, nx^{n-1}) = 1$, it follows that $x^n - 1$ has n-distinct solutions - the n-th roots of unity. The n-th roots of unity form a cyclic group of order n that we shall denote μ_n (using the same notation as in the case of the complex numbers).

In the following, given an element $a \in F$, we shall write $F(\sqrt[n]{a})$ to denote an extension K/F of the form $K = F(\alpha)$, where α satisfies $\alpha^n = a$. In general, the nature of this extension could very much depend on α. For example, α may, or may not, already belong to F. However, in *our* case, since the n-th roots of unity are in F, for every root α, the polynomial $x^n - a = \prod_{\zeta^n=1}(x - \zeta\alpha)$, splits over $F(\alpha)$. And so, up to isomorphism, the extension $F(\sqrt[n]{a})$ is independent of the choice of n-th root of a.

Theorem 13.1.1. *Let $a \in F^\times$, then $F(\sqrt[n]{a})$ is a cyclic Galois extension of order m dividing n. Conversely, if L/F is a cyclic Galois extension of order m dividing n then $L = F(\sqrt[m]{a})$ for some $a \in F^\times$ (and we may also write $L = F(\sqrt[n]{a^{n/m}})$).*

Proof. Let $a \in F^\times$, then $F(\sqrt[n]{a})$ is a Galois extension of F, being the splitting field of the separable polynomial $x^n - a$. Let $\sigma \in G := \mathrm{Gal}(F(\sqrt[n]{a})/F)$. Then $\sigma(\sqrt[n]{a})$ is another root of $x^n - a$ and so $\sigma(\sqrt[n]{a}) = \zeta_\sigma \sqrt[n]{a}$, for some root of unity $\zeta_\sigma \in \mu_n$. This gives us a function

$$G \to \mu_n, \qquad \sigma \mapsto \zeta_\sigma = \frac{\sigma(\sqrt[n]{a})}{\sqrt[n]{a}}.$$

We claim that this map is an injective homomorphism. First, for $\sigma, \tau \in G$ we find

$$\begin{aligned}(\sigma\tau)(\sqrt[n]{a}) &= \sigma(\zeta_\tau \sqrt[n]{a})\\ &= \zeta_\tau \sigma(\sqrt[n]{a})\\ &= \zeta_\tau \zeta_\sigma \sqrt[n]{a},\end{aligned}$$

and it follows that $\zeta_{\sigma\tau} = \zeta_\sigma \zeta_\tau$, which is the homomorphism property. Secondly, the action of σ on $\sqrt[n]{a}$ determines its action on $F(\sqrt[n]{a})$ and so the homomorphism is injective. Since μ_n is cyclic of order n, every subgroup of it is cyclic of some order $m|n$. Since G is isomorphic to a subgroup of μ_n, G is cyclic of some order m dividing n. This proves one direction of the theorem.

Conversely, suppose that K/F is a cyclic Galois extension with Galois group G of order $m|n$. Let σ be a generator of G. Let ζ be a primitive m-th root of unity in F. Consider the following expression, called a **Lagrange resolvent**, for an element $\alpha \in F$,

$$\mathscr{L} = \mathscr{L}(\alpha) := \alpha + \zeta \cdot \sigma(\alpha) + \cdots + \zeta^{m-1} \cdot \sigma^{m-1}(\alpha).$$

By independence of characters, the function $1(\cdot) + \zeta \cdot \sigma(\cdot) + \cdots + \zeta^{m-1} \cdot \sigma^{m-1}(\cdot)$ is not the zero function on K and so we may choose an α such that $\mathscr{L} \neq 0$. We assume that α is chosen this way (and it doesn't matter which α is chosen as long as $\mathscr{L}(\alpha) \neq 0$).

Note that $\sigma(\mathscr{L}) = \sigma(\alpha) + \zeta \cdot \sigma^2(\alpha) + \cdots + \zeta^{m-1} \cdot \sigma^m(\alpha)$ and so, using that σ^m is the identity, $\sigma(\mathscr{L}) = \zeta^{-1} \cdot \mathscr{L}$. This has the following consequences:

- $K = F(\mathscr{L})$. Indeed, it follows from the formula $\sigma\mathscr{L} = \zeta^{-1}\mathscr{L}$ that the only element of G fixing $\mathscr{L}$ is the identity. Thus, by the Main Theorem, $K = F(\mathscr{L})$.
- $\mathscr{L}^m$ is fixed under σ because $\sigma(\mathscr{L}^m) = (\sigma(\mathscr{L}))^m = \zeta^{-m}\mathscr{L}^m = \mathscr{L}^m$. Thus, $\mathscr{L}^m$ is fixed under G and so belong to $K^G = F$. Denoting $a = \mathscr{L}^m$ we have succeeded in writing $K = F(\sqrt[m]{a})$.

□

13.1.1. *Application to cyclotomic fields.* In characteristic different from 2, a quadratic extension K/F is always Galois, because the minimal polynomial of any $\alpha \in K - F$ must be quadratic and, being quadratic and having one root in K, it must split in K. That polynomial is also separable as the condition $(f, f') = 1$ holds. Hence K/F is Galois. The theorem tells us that every (Galois) quadratic extension of F is of the form $K = F(\sqrt{a})$ for some a and, in fact, how to find a. Take any $\alpha \in K - F$. Then $\mathscr{L}(\alpha) = \alpha - \sigma(\alpha) \neq 0$ and we can take $a = \mathscr{L}(\alpha)^2$.

Consider the cyclotomic extension $\mathbb{Q}(\zeta_p)/\mathbb{Q}$, where $p > 2$ is a prime. The Galois group is isomorphic to $(\mathbb{Z}/p\mathbb{Z})^\times$ and so is cyclic of order $p-1$; it has a subgroup H of index 2,

$$H = \{1 \leq n \leq p-1 : n \text{ is a square mod} p\}.$$

The quadratic extension K of $\mathbb{Q}$ contained in $\mathbb{Q}(\zeta)$ is generated by

$$\eta_H := |H| \cdot \pi_H(\zeta_p) = \sum_{n=\square} \zeta_p^n.$$

Moreover, $\sigma|_K$ generates the Galois group of $K/\mathbb{Q}$. σ corresponds to some generator of the cyclic group $\mathbb{Z}/p\mathbb{Z}^\times$ and thus is not a square. It follows that $\sigma \cdot H$ are the non-squares modulo p. Define the **quadratic residue symbol**, or **Legendre symbol**, as:

$$\left(\frac{a}{p}\right) = \begin{cases} 1, & a = \square \pmod{p}, a \neq 0, \\ -1, & a \neq \square \pmod{p}, \\ 0, & a = 0 \pmod{p}. \end{cases}$$

It satisfies the identity

$$\left(\frac{nm}{p}\right) = \left(\frac{n}{p}\right)\left(\frac{m}{p}\right).$$

Thus, the Lagrange resolvent for the extension $K^H/\mathbb{Q}$, the element

$$\kappa = \eta_H - \sigma\eta_H = \sum_{n=1}^{p-1} \left(\frac{n}{p}\right) \zeta_p^n,$$

is a square root of a rational number (and note that we could have started the sum at $n = 0$). This sum is a special case of a **Gauss sum**.

Now, complex conjugation viewed as an automorphism of $\mathbb{Q}(\zeta_p)$ must correspond to $-1 \in \mathbb{Z}/p\mathbb{Z}^\times$ - the unique element of order 2 of this group. Thus, κ is real $\Leftrightarrow$ complex conjugation acts trivially on $\kappa \Leftrightarrow -1 \in H \Leftrightarrow -1$ is a square mod p. And otherwise, $\bar{\kappa} = -\kappa$. (We can also show that as follows: $\bar{\kappa} = \sum_{n=1}^{p-1}\left(\frac{n}{p}\right)\zeta_p^{-n} = \left(\frac{-1}{p}\right)\sum_{n=1}^{p-1}\left(\frac{-n}{p}\right)\zeta_p^{-n} = \left(\frac{-1}{p}\right)\sum_{n=1}^{p-1}\left(\frac{n}{p}\right)\zeta_p^{n} = \left(\frac{-1}{p}\right)\kappa$.) We leave it as an exercise to show, using the fact that $\mathbb{Z}/p\mathbb{Z}^\times$ is cyclic of order $p-1$, that for $p > 2$,

$$\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2} = \begin{cases} 1 & p \equiv 1 \pmod 4, \\ -1 & p \equiv 3 \pmod 4. \end{cases}$$

That is, the field K^H is a quadratic *real* field if $p \equiv 1 \pmod 4$ and is a quadratic *imaginary* field if $p \equiv 3 \pmod 4$.

Consider now $|\kappa|^2 = \kappa\bar{\kappa}$. We have

$$\begin{aligned} \kappa\bar{\kappa} &= \sum_{n=0}^{p-1}\left(\frac{n}{p}\right)\zeta_p^{-n}\cdot\kappa \\ &= \sum_{n=0}^{p-1}\zeta_p^{-n}\left(\sum_{m=0}^{p-1}\left(\frac{n}{p}\right)\left(\frac{m}{p}\right)\zeta_p^{m}\right) \\ &= \sum_{n=0}^{p-1}\zeta_p^{-n}\left(\sum_{m=0}^{p-1}\left(\frac{n}{p}\right)\left(\frac{nm}{p}\right)\zeta_p^{nm}\right), \end{aligned}$$

where in the last step we have changed variable from m to nm in the inner sum. This is permissible for $n \neq 0$, and for $n = 0$ both the original sum and the one after change of variable are 0. We Claim that the inner sum is equal to $\sum_{m=0}^{p-1}\left(\frac{m}{p}\right)\zeta_p^{nm}$. This is clear for $n \neq 0$ because then $\left(\frac{n^2m}{p}\right) = \left(\frac{m}{p}\right)$. For $n = 0$ we need the identity $0 = \sum_{m=0}^{p-1}\left(\frac{m}{p}\right)$. This is true: it expresses the fact that there are as many non-zero squares as non-zero non-squares (otherwise said, that H has index 2). We thus conclude that

$$\kappa\bar{\kappa} = \sum_{n=0}^{p-1}\zeta_p^{-n}\sum_{m=0}^{p-1}\left(\frac{m}{p}\right)\zeta_p^{nm} = \sum_{m=0}^{p-1}\left(\frac{m}{p}\right)\sum_{n=0}^{p-1}(\zeta_p^{m-1})^n$$

For $m \neq 1$, we have $\sum_{n=0}^{p-1}(\zeta_p^{m-1})^n = \frac{(\zeta_p^{m-1})^p-1}{(\zeta_p^{m-1})-1} = 0$, while for $m = 1$ we get p. It follows that

$$\kappa\bar{\kappa} = |\kappa|^2 = p.$$

Consequently, we have proven:

Theorem 13.1.2 (Gauss)**.** *The unique quadratic subfield of* $\mathbb{Q}(\zeta_p)$ *is* $\mathbb{Q}\left(\sqrt{\left(\frac{-1}{p}\right)p}\right)$.

13.2. **Root extensions.** We make a standing assumption that $\boxed{F\ \textbf{is a field of characteristic zero}}$ An finite extension of fields K/F is called a **root extension**, if there are subfields

$$F = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_s = K,$$

such that for all $1 \leq i \leq s$,

$$K_i = K_{i-1}(\sqrt[n_i]{a_i}), \qquad \text{for some } a_i \in K_{i-1},\ n_i \in \mathbb{Z}_{>0}.$$

The exact meaning of the notation $K_i = K_{i-1}(\sqrt[n_i]{a_i})$ is that $K_i = K_{i-1}(\alpha)$ for some α satisfying $\alpha^{n_i} = a_i$.

Proposition 13.2.1. *Let K/F be a root extension. There is a field $L \supseteq K$ such that L/F is a finite Galois root extension and* $\mathrm{Gal}(L/F)$ *is a solvable group.*

Proof. We divide the proof into two steps.
First Step: We may assume K/F is Galois. To show that, choose first an extension Ω/K such that Ω/F is a finite Galois extension. This is possible since F has characteristic zero hence the extension K/F is finite and separable. Now, for every $\sigma \in \mathrm{Gal}(\Omega/F)$ the sequence of fields

$$F = \sigma(F) = \sigma(K_0) \subseteq \sigma(K_1) \subseteq \cdots \subseteq \sigma(K_s) = \sigma(K),$$

is also a root extension. Indeed, $\sigma(K_i) = \sigma(K_{i-1})(\sigma(\sqrt[n_i]{a_i}))$, but $(\sigma(\sqrt[n_i]{a_i}))^{n_i} = \sigma(a_i) \in \sigma(K_{i-1})$ and so, with abuse of notation, we may write $\sigma(K_i) = \sigma(K_{i-1})(\sqrt[n_i]{\sigma(a_i)})$ and conclude that $\sigma(K)/F$ is a root extension as well.

This way, as σ ranges over $\mathrm{Gal}(\Omega/F)$ we get finitely many root extensions.

Lemma 13.2.2. *Let $F = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_s = K$ and $F = J_0 \subseteq J_1 \subseteq \cdots \subseteq J_t = J$ be two root extensions, where $K_i = K_{i-1}(\sqrt[n_i]{a_i})$, $J_i = J_{i-1}(\sqrt[m_i]{b_i})$. Assume that K and J are both contained in a common field N. Then the compositum KJ is also a root extension.*

Proof. This is rather straightforward. We have

$$F = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_s = K = KJ_0 \subseteq KJ_1 \subseteq \cdots \subseteq KJ_t = KJ,$$

and each extension is obtained by either adding a $\sqrt[n_i]{a_i}$ or a $\sqrt[m_i]{b_i}$, as the case may be. □

Applying the lemma successively to the root extensions $\sigma(K)/F$, as σ ranges over $\mathrm{Gal}(\Omega/F)$ we find that $K' := \prod_{\sigma \in \mathrm{Gal}(\Omega/F)} \sigma(K)$ is a root extension of F that is also Galois, because $\sigma(K') = K'$ for all $\sigma \in \mathrm{Gal}(\Omega/F)$ and that implies that K' is Galois over F (cf. the proof of the Main Theorem).

Step Two: Embed a Galois root extension K/F in a larger Galois root extension L/F whose Galois group is solvable.[4] As K/F is a root extension, write

$$F = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_s = K,$$

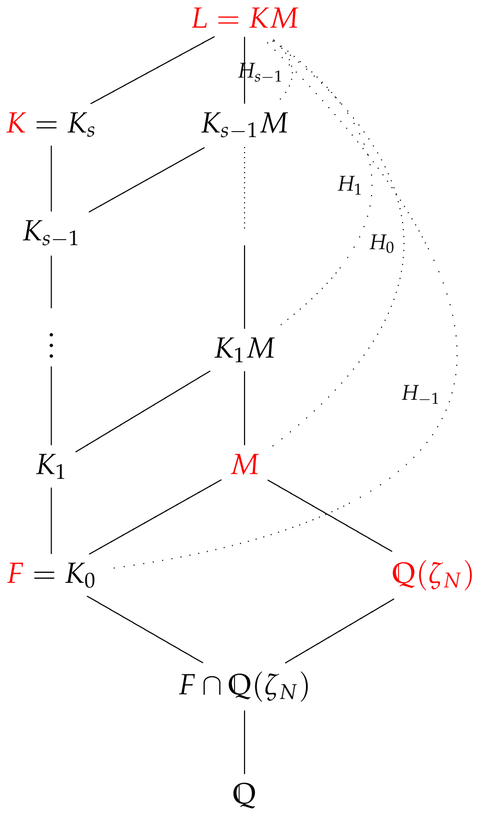
where

$$K_i = K_{i-1}(\sqrt[n_i]{a_i}), \qquad \text{for some } a_i \in K_{i-1}, n_i \in \mathbb{Z}_{>0}.$$

Let $N = \mathrm{lcm}\{n_1, n_2, \dots, n_s\}$ and let $M = F(\zeta_N) = F\mathbb{Q}(\zeta_N)$, where by $\mathbb{Q}(\zeta_N)$ we mean the splitting field of $x^N - 1$ is some algebraic closure of K. Let $L = KM$.[5] We consider the following big diagram:

[4] It then follows that $\mathrm{Gal}(K/F)$, being a quotient of the solvable group $\mathrm{Gal}(L/F)$, is also solvable. But we don't care. For the applications we just need *some* Galois root extension with solvable Galois group.

[5] As a matter of fact, one can prove that M is contained in K, but not necessarily in each K_i, so for the argument it is easier not to take that into consideration at all.



First, note that L/F is Galois because it is the compositum of the Galois extensions K/F and M/F - use Proposition 12.2.1 (and also Proposition 12.1.1 to see that M/F is Galois). It is also a root extension, being a compositum of the root extension K/F and the root extension $M = F(\sqrt[N]{1})/F$ (where we adjoin a primitive N-th root of 1). We shall prove that $\mathrm{Gal}(L/F)$ is a solvable group.

Introduce the following notation: Let $H_i = \mathrm{Gal}(L/K_iM)$ for $i = 0, \dots, s-1$. Also put $H_{-1} = \mathrm{Gal}(L/F)$. Now, M/F is Galois and so $H_0 \lhd H_{-1}$ and, moreover, $H_{-1}/H_0 \cong \mathrm{Gal}(M/F) \cong \mathrm{Gal}(\mathbb{Q}(\zeta_N)/F \cap \mathbb{Q}(\zeta_N))$ (Proposition 12.1.1) which is abelian because it is a subgroup of $\mathbb{Z}/N\mathbb{Z}^\times = \mathrm{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q})$. Note that each extension $K_iM/K_{i-1}M$ (for $i = 1, \dots, s$) is a root extension, $K_iM = K_{i-1}M(\sqrt[n_i]{a_i})$. Since $K_{i-1}M$ contains all n_i roots of unity, by Theorem 13.1.1 $K_iM/K_{i-1}M$ is a cyclic Galois extension. It follows that $H_{i-1} \lhd H_i$ and $H_{i-1}/H_i \cong \mathrm{Gal}(K_iM/K_{i-1}M)$ is abelian. In short, we have proven that the normal series

$$\{1\} \lhd H_{s-1} \lhd H_{s-2} \lhd \dots \lhd H_0 \lhd H_{-1} = \mathrm{Gal}(L/F),$$

has abelian quotients and so $\mathrm{Gal}(L/F)$ is solvable. □

13.3. **Solvability by radicals.** In this section we prove a theorem that was, in its time, one of the main achievements of Galois theory. Namely, that the general equation of degree 5 or higher, cannot be solved by radicals.

First, we define what we mean exactly in solved by radicals. Let F be a field of characteristic 0 and $f(x) \in F[x]$ a non-constant polynomials. We say that f can be **solved by radicals** if there exists a root extension K/F over which $f(x)$ splits into linear terms.

Theorem 13.3.1 (E. Galois). *Let F be a field of characteristic 0 and $f(x) \in F[x]$ a non-constant polynomial. Let J be a splitting field of $f(x)$. The $f(x)$ can be solved by radicals if and only if $\mathrm{Gal}(J/F)$ is a solvable group.*

Proof. Suppose first that $f(x)$ can be solved in radicals. Let K/F be a root extension in which $f(x)$ splits. Using Proposition 13.2.1, we may assume that K/F is Galois and $\mathrm{Gal}(K/F)$ is solvable. Denote by $\alpha_1, \dots, \alpha_n$ the roots of $f(x)$ in K. Let $J = F(\alpha, \dots, \alpha_n)$. It is the splitting field of $f(x)$ and since the characteristic of F is 0, J/F is a Galois extension. Furthermore, $\mathrm{Gal}(J/F) = \mathrm{Gal}(K/F)/\mathrm{Gal}(K/J)$, thus a quotient of solvable group. Since a quotient of a solvable group is solvable, $\mathrm{Gal}(J/F)$ is solvable.

Conversely, suppose that $\mathrm{Gal}(J/F)$ is solvable, where J is a splitting field of F. There is a normal series,

$$\{1\} = G_n \lhd G_{n-1} \dots \lhd G_0 = \mathrm{Gal}(J/F),$$

such that G_{i-1}/G_i is a cyclic group of order $n_i | N$, where $N = [J : F]$ and $i = 1, \dots, n$. Let $J_i = J^{G_i}$ the corresponding fields. Let $M = F(\zeta_N)$. We have the following diagram:

$$\begin{tikzcd}
 & MJ = MJ_n \arrow[dl, dash] \arrow[d, dash, "H_{n-1}"] & \\
J = J_n \arrow[d, dash, "G_{n-1}/G_n"] & MJ_{n-1} \arrow[dl, dash] \arrow[d, dash, "H_{n-2}"] & \\
J_{n-1} \arrow[d, dash, "G_{n-2}/G_{n-1}"] & \vdots & \\
\vdots & MJ_1 \arrow[dl, dash] \arrow[d, dash, "H_0"] & \\
J_1 \arrow[d, dash, "G_0/G_1"] & M = MJ_0 \arrow[dl, dash, "U"] \arrow[dr, dash] & \\
J_0 = F \arrow[dr, dash] & & \mathbb{Q}(\zeta_N) \arrow[dl, dash] \\
 & \mathbb{Q} &
\end{tikzcd}$$

We let $L = MJ$. We consider the sequence of fields

$$L = MJ \supseteq MJ_{n-1} \supseteq \cdots \supseteq MJ_0 = M \supseteq F.$$

Since J_i/J_{i-1} is Galois, also MJ_i/MJ_{i-1} is Galois and its Galois group H_{i-1} is isomorphic to the subgroup $\mathrm{Gal}(J_i/J_i \cap MJ_{i-1})$ of G_{i-1}/G_i (Proposition 12.1.1), hence a cyclic group of order dividing n_i and so dividing N. Since $\mu_N \subset F \subseteq MJ_{i-1}$, it follows from Theorem 13.1.1 that $MJ_i = MJ_{i-1}(\sqrt[n_i]{a_i})$ for some $a_i \in MJ_{i-1}$. Of course also M/F is a root extension. It follows that L/F is a root extension in which $f(x)$ splits. □

Corollary 13.3.2 (Insolvability of the Quintic). *The general quintic polynomial cannot be solved in radicals.*

Proof. Indeed, we have constructed a quintic polynomial with Galois group S_5 (Example 10.6.3) and S_5 is not a solvable group. □

Let $f(x) \in F[x]$ be an irreducible (separable) polynomial of degree n and K/F its splitting field. Then

$$\mathrm{Gal}(K/F) \hookrightarrow S_n,$$

and the image is a transitive subgroup of S_n. Since S_2, S_3, S_4 are solvable, so are their subgroups and we therefore conclude:

Corollary 13.3.3. *Every irreducible polynomial $f(x)$ of degree ≤ 4 is solvable by radicals.*

Example 13.3.4. This does not imply that the splitting field of $f(x)$ is a root extension. For example, consider the field $\mathbb{Q}(\eta_J)$ of $\mathbb{Q}(\zeta_7)$ appearing in § 10.4. The element η_J has minimal polynomial $x^3 + x^2 - 2x - 1$. As $\mathbb{Q}(\eta_J)/\mathbb{Q}$ is Galois it is the splitting field of $x^3 + x^2 - 2x - 1$, but it is not of the form $\mathbb{Q}(\sqrt[3]{a})$ for any a. Indeed, it if were then, being Galois, the polynomial $x^3 - a$ would split and it would follow that $\mathbb{Q}(\omega)$, where ω is a primitive third root of unity, is a subfield of $\mathbb{Q}(\eta_J)$. However, this is not possible because $[\mathbb{Q}(\eta_J) : \mathbb{Q}] = 3$ and $[\mathbb{Q}(\omega) : \mathbb{Q}] = 2$ and $2 \nmid 3$. On the other hand, Theorem 13.1.1 tells us that the splitting field of the same polynomial, but considered over $\mathbb{Q}(\omega)$ is a cyclic extension. That is, there is some $a \in \mathbb{Q}(\omega)$ such that $\mathbb{Q}(\eta_J, \omega) = \mathbb{Q}(\omega)(\sqrt[3]{a})$. We leave the verification of that as an exercise (Exercise 48), which is surprisingly tricky.

Note that $\mathbb{Q}(\eta_J, \omega) = \mathbb{Q}(\omega)(\sqrt[3]{a})$ is a root extension of $\mathbb{Q}$ and that the splitting field of $x^3 + x^2 - 2x - 1$, namely, $\mathbb{Q}(\eta_J)$ is a subfield. Thus, every root of $x^3 + x^2 - 2x - 1$ can be expressed in radicals, in particular η_J.

14. Calculating Galois groups

14.1. First observations.

Let us review what we know about the Galois group of a finite Galois extension K/F. Writing $K = F(\theta)$, which is always possible by the Primitive Element Theorem, we view K as the splitting field of the minimal polynomial $f(x)$ of θ. Suppose the roots of the polynomial are $\alpha_1, \dots, \alpha_n$ then $\mathrm{Gal}(K/F)$ acts by permutations on the set of roots and that gives us an injective group homomorphism

$$\mathrm{Gal}(K/F) \hookrightarrow S_n,$$

and the image is a *transitive* subgroup of S_n. This is not necessarily the most efficient method to study $\mathrm{Gal}(K/F)$, but let us leave this point for later in our discussion.

At this point, especially for low degree polynomials, it is very useful to have a list of transitive groups. We provide such in the following table. The groups are listed up to conjugation.

TABLE 1. Transitive subgroups of S_n

n	transitive subgroups of S_n
2	S_2
3	A_3 and S_3
4	V, A_4 and $\langle(1234)\rangle, D_4, S_4$
5	$\langle(12345)\rangle, D_5, A_5$ and F_{20}, S_5

In listing the subgroups we listed the subgroups contained in A_n first and then the subgroups not contained in A_n. We know already all the groups appearing in this list (V stands for the Klein four group), apart from the **Frobenius group** F_{20}, which is a subgroup of S_5 with 20 elements.

A model for this group is $\langle (12345), (2354)\rangle$. Any other subgroup of S_5 with 20 elements is isomorphic to F_{20}; in fact, conjugate to it. See Exercise 49 for this classification and Exercises 50, 51 for more about F_{20}.

We see that a useful first step is to decide if the Galois group is a subgroup of A_n or not. Fortunately, there is a simple criterion for that. Let F be a field of characteristic different than 2 and let $f(x)$ be a monic non-constant separable polynomial in $F[x]$ of degree n. Choose a splitting field K/F for f and write there

$$f(x) = \prod_{i=1}^{n}(x-\alpha_i), \quad \alpha_i \in K.$$

Consider the product

$$\delta = \prod_{i<j}(\alpha_i - \alpha_j).$$

(This product really depends on how we order the roots, but just up to a sign.) We note that for $\sigma \in G$, identified with a permutation $\sigma \in S_n$ via $\sigma(\alpha_i) = \alpha_{\sigma(i)}$,

$$\sigma(\delta) = \prod_{i<j}(\sigma(\alpha_i) - \sigma(\alpha_j)) = \prod_{i<j}(\alpha_{\sigma(i)} - \alpha_{\sigma(j)}) = \mathrm{sgn}(\sigma)\cdot\delta.$$

Define the **discriminant** of $f(x)$, denoted $\mathscr{D}(f)$, by the formula

$$\mathscr{D}(f) = \delta^2 = \prod_{i<j}(\alpha_i - \alpha_j)^2.$$

Our calculation shows that

$$\mathscr{D}(f) \in K^{\mathrm{Gal}(K/F)} = F,$$

and, furthermore, for $\sigma \in \mathrm{Gal}(K/F)$ we have

$$\sigma(\delta) = \delta, \forall \sigma \in \mathrm{Gal}(K/F) \iff \mathrm{Gal}(K/F) \subseteq A_n.$$

But, $\sigma(\delta) = \delta$ if and only if $\delta \in F$. And so, we conclude,

Proposition 14.1.1. *Let F be a field of characteristic different than 2. Let $f(x) \in F[x]$ be a non-constant separable polynomial of degree n with splitting field K. Then $\mathrm{Gal}(K/F) \subseteq A_n$ if and only if $\mathscr{D}(f)$ is a square in F.*

14.1.1. *Quadratic polynomials.* Suppose that $f(x) = x^2 + bx + c$. In terms of the roots α_1, α_2 of f, we have

$$b = -(\alpha_1 + \alpha_2), \quad c = \alpha_1\alpha_2,$$

whence

$$\mathscr{D}(x^2 + bx + c) = b^2 - 4c.$$

We find that over fields F of characteristic different than 2 the polynomial $f(x)$ is reducible (which is equivalent to it having Galois group A_2) if and only if $b^2 - 4c$ is a square in F.

14.1.2. *Cubic polynomials.* Consider a cubic polynomial $g(x) = x^3 + \alpha x^2 + \beta x + \gamma$ over a field F of characteristic different from 3. By change of variable, replacing x by $x - \alpha/3$, we can reduce the polynomial to a polynomial of the form $f(x) = x^3 + ax + b$. Note that the roots are shifted by $\alpha/3$ and so their differences remain the same. Thus, $\mathscr{D}(f) = \mathscr{D}(g)$. We conclude that it is enough to discuss discriminants for polynomials of the form

$$f(x) = x^3 + ax + b.$$

In terms of the roots,

$$a = \alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3, \qquad b = -\alpha_1\alpha_2\alpha_3.$$

Now, one just verifies that

$$(\alpha_1 - \alpha_2)^2(\alpha_1 - \alpha_3)^2(\alpha_2 - \alpha_3)^2 = -4a^3 - 27b^2.$$

(It requires the identity $\alpha_1 + \alpha_2 + \alpha_3 = 0$.) That is,

$$\mathscr{D}(x^3 + ax + b) = -4a^3 - 27b^2.$$

Let's look at some examples.

Example 14.1.2. $f(x) = x^3 - x + 1$ considered over $\mathbb{Q}[x]$.
If $f(x)$ is reducible over $\mathbb{Q}[x]$ it is reducible over $\mathbb{Z}[x]$ (Gauss' lemma) and then over $\mathbb{Z}/2\mathbb{Z}[x]$. It then must have a root in $\mathbb{Z}/2\mathbb{Z}$, but this is not the case! Thus, $f(x)$ is irreducible. An alternate argument is simply to use that if $f(x)$ is reducible over $\mathbb{Q}$ then it has a rational root a/b ($(a, b) = 1$) and, by a well-known argument b divides the leading coefficient and a the constant coefficient. Thus, the root could only be ± 1 and those are easily discarded.

The discriminant $\mathscr{D}(f) = -4 \cdot (-1) - 27 = -23$ which is not a square in $\mathbb{Q}$. Therefore, the Galois group is S_3.

Example 14.1.3. $f(x) = x^3 - 21x - 7$. This polynomial is irreducible by Eisenstein's criterion. $\mathscr{D}(f) = -4 \cdot 21^3 - 27 \cdot (-7)^2 = 3^6 7^2$, which is a square in $\mathbb{Q}$. Therefore, the Galois group is A_3.

Example 14.1.4. Let's construct an infinite family of rational cubic polynomials with Galois group contained in A_3. The arguments we give do not prove that the Galois group is A_3 because we do not prove that the polynomials we get are irreducible, but it is very likely they are and, if so, we get infinitely many examples of cubic polynomials with Galois group A_3. Finding such a polynomial is equivalent to finding rational points on the surface $y^2 = -4A^3 - 27B^2$.

Consider B as fixed and non-zero. Then

$$E_B : y^2 = -4A^3 - 27B^2$$

is an **elliptic curve**. Such curves have a very rich structure. Given two points $(a_1, y_1), (a_2, y_2)$ of the curve, we can get a third point that we denote

$$(a_1, y_1) \overset{E_B}{\oplus} (a_2, y_2).$$

To get the new point $(a_1, y_1) \overset{E_B}{\oplus} (a_2, y_2)$ one proceeds as follows (refer to Figure 2):

- Add to E_B an ideal point at infinity and denote it 0_{E_B}.
- Draw the line through (a_1, y_1) and (a_2, y_2) (or the tangent line to E_B at the point (a_1, y_1) if $(a_1, y_1) = (a_2, y_2)$). This line intersects the curve E_B in a third point, say $Q = (a_3, y_3)$. (In rare cases, there is no such point Q and then let $Q = 0_{E_B}$).
- Let $(a_1, y_1) \overset{E_B}{\oplus} (a_2, y_2)$ be the point $(a_3, -y_3)$, that is the mirror image of Q along the a-axis. In the case that $Q = 0_{E_B}$, let also $(a_1, y_1) \overset{E_B}{\oplus} (a_2, y_2) = 0_{E_B}$.

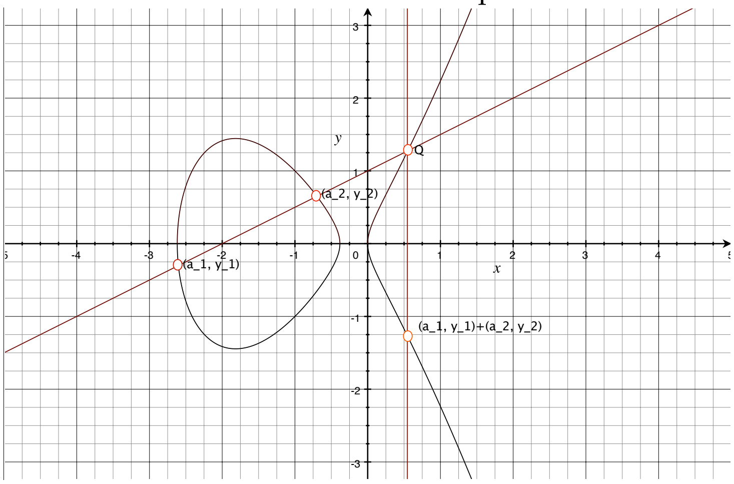
Although we have described our process geometrically, one can also express it by polynomial formulas. Those make sense over every field, enabling a theory of elliptic curves over any field. An interesting fact, which is easy to show, is that if (a_1, y_1) and (a_2, y_2) have coordinates in a certain field F, where also the equation of the elliptic curve is defined over F, then the new point $(a_3, y_3) = (a_1, y_1) \overset{E_B}{\oplus} (a_2, y_2)$ has coordinates in F as well. Moreover, the elliptic curve becomes an abelian group under this addition law! It's identity element is 0_{E_B} and the group inverse of (a_1, y_1) is $(a_1, -y_1)$.

The key implication for us is that once we have found a rational point (a_1, y_1) on some E_B, where $B \in \mathbb{Q}$, we can get other points as

$$(a_n, y_n) = [n](a_1, y_1),$$

where $[n](a_1, y_1)$ means add the point (a_1, y_1) to itself n-times in the abelian group E_B. If the point (a_1, y_1) is of infinite order on E_B then we get that way infinitely many cubic polynomials

FIGURE 2. Addition on elliptic curve



over $\mathbb{Q}$ with square discriminant; if those polynomials are irreducible then their Galois group is A_3.

As a matter of fact, one can prove that the point $P := (a, y) = (-21, 3^3 7)$ on the elliptic curve $E_7 : y^2 = -4a^3 - 27 \cdot 49$ is a point of infinite order (this is the point corresponding to the fact that the discriminant of the polynomial $x^3 - 21x - 7$ is a square). Some multiples of P are

$$\begin{aligned}
P &= [-21, 3^3 7] \\
[2]P &= [-7, 7] \\
[3]P &= [-57/4, -405/2] \\
[4]P &= [-427, -17647] \\
&\vdots \\
[12]P &= [\tfrac{-4012727988243653281}{84540051739598400}, -\tfrac{80257739337499491273192976 79}{1229033634624383224982 4000}]
\end{aligned}$$

14.2. **Calculating Galois groups by reducing modulo a prime p.** This technique is very general, but the general statement, as well as the proofs, require the machinery of algebraic number theory. We therefore state a special case, though still of considerable interest.

Theorem 14.2.1. *Let $f(x) \in \mathbb{Z}[x]$ be a monic irreducible polynomial of degree n. View the Galois group G of $f(x)$ as a subgroup of S_n. Suppose that modulo a prime p not dividing the discriminant of f we have the factorization*

$$f(x) \equiv f_1(x) f_2(x) \cdots f_r(x) \pmod{p},$$

where the f_i are distinct irreducible polynomials in $\mathbb{F}_p[x]$. Let $n_i = \deg(f_i)$. G contains a permutation of type $(n_1, n_2, \dots, n_r)$.

The usefulness of the theorem is that often the existence of a permutation of a given type allows to decide between several alternatives for the Galois group. Consider the following examples of low degree polynomials.[6]

[6] According to the paper *Enumerating subgroups of the symmetric group* by Derek F. Holt, the number $t(n)$ of transitive subgroups of S_n up to conjugacy is as follows:

n:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	$\cdots$	32
$t(n)$:	1	1	2	5	5	16	7	50	34	45	8	301	9	63	104	1954	10	983	8	1117	$\cdots$	2801324

TABLE 2. Transitive subgroups of S_3

type:	(1,2)	(3)
A_3		✓
S_3	✓	✓

TABLE 3. Transitive subgroups of S_4

type:	(1,1,2)	(1,3)	(2,2)	(4)
V			✓	
A_4		✓	✓	
$\langle(1234)\rangle$			✓	✓
D_4	✓		✓	✓
S_4	✓	✓	✓	✓

TABLE 4. Transitive subgroups of S_5

type:	(1,1,1,2)	(1,1,3)	(1, 2, 2)	(1, 4)	(2,3)	(5)
$\langle(12345)\rangle$						✓
D_5			✓			✓
A_5		✓	✓			✓
F_{20}			✓	✓		✓
S_5	✓	✓	✓	✓	✓	✓

Let us consider some examples. Let G denote the Galois group of the polynomial $f(x) = x^3 - x + 1$ that we had already considered above; the polynomial is irreducible modulo 2. Thus, G has a 3-cycle. This is in fact of no value since all transitive subgroups of S_3 contain a 3-cycle. However, $f(x)$ factors modulo 7 as $(x-2)(x^2+2x+3)$ and thus G contains a transposition and therefore $G = S_3$.

Consider next the polynomial $f(x) = x^4 - 4x^2 + 2$ (irreducible by Eisenstein), with Galois group G. One verifies that this polynomial is also irreducible modulo 3 and so G contains a 4-cycle. Let α be a solution of the polynomial $y^2 - 4y + 2$. Then, K, the splitting field of f contains $\mathbb{Q}(\sqrt{\alpha})$ and $\mathbb{Q}(\sqrt{\alpha}) \supseteq \mathbb{Q}(\sqrt{2})$ because $\alpha = 2 + \sqrt{2}$. The conjugate of α is $\alpha' = 2 - \sqrt{2}$. Over K the polynomial $f(x)$ factors as

$$(x - \sqrt{\alpha})(x + \sqrt{\alpha})(x - \sqrt{\alpha'})(x + \sqrt{\alpha'}).$$

However, note also that $\alpha/\alpha' = (1+\sqrt{2})^2$ and so, $\sqrt{\alpha'} \in \mathbb{Q}(\sqrt{\alpha})$. It follows that $K = \mathbb{Q}(\sqrt{\alpha})$ and so of degree 4 over $\mathbb{Q}$. We conclude that $G \cong \mathbb{Z}/4\mathbb{Z}$.

Now, the conclusion that $K = \mathbb{Q}(\sqrt{\alpha})$ is not hard. It only required checking that α/α' is a square in $\mathbb{Q}(\alpha)$. Thus, even without using reduction modulo a prime, we can easily deduce that the Galois group is of order 4. Is there a way to prove it is cyclic without using the technique of reduction modulo prime? Indeed there is and here are two ways to show that:

- One can calculate the discriminant of $x^4 - 4x^2 + 2$ by hand, as the roots are very particular. It is equal to 2^{11}, which is not a square. Therefore, the Galois group is not contained in A_4 and so must be a cyclic group of order 4.
- Consider the automorphism $\sigma : \mathbb{Q}(\alpha) \to \mathbb{Q}(\alpha)$ that takes α to α'. It also takes the polynomial $x^2 - \alpha$ to $x^2 - \alpha'$ and we know that we can extend it to $K = \mathbb{Q}(\sqrt{\alpha}) = \mathbb{Q}(\sqrt{\alpha'})$ by taking $\sqrt{\alpha}$ (a root of the irreducible polynomial $x^2 - \alpha$) to $\sqrt{\alpha'}$ (a root of the irreducible polynomial $x^2 - \alpha'$). The choice of the square roots is determined by the requirement that $\sqrt{\alpha} = (1+\sqrt{2})\sqrt{\alpha'}$.

  Call this extension also σ. To show that the Galois group is cyclic it is enough to show that $\sigma^2 \neq 1$. But, $\sigma^2(\sqrt{\alpha}) = \sigma(\sqrt{\alpha'}) = \sigma((1+\sqrt{2})^{-1}\sqrt{\alpha}) = (1-\sqrt{2})^{-1}\sigma(\sqrt{\alpha}) = (1-\sqrt{2})^{-1}\sqrt{\alpha'} = -(1+\sqrt{2})\sqrt{\alpha'} = -\sqrt{\alpha}$.

14.3. **Using compositum.** Sometime the Galois extension K/F is presented to us as a compositum of extensions, for example as $K = F(\alpha_1, \beta_1)$. In this case, if $F(\alpha_1)/F$ and $F(\beta_1)/F$ are Galois then it may be more efficient to use Proposition 12.2.1 to calculate the Galois group than to write $K = F(\theta)$ and proceed as before. Indeed, letting $\alpha_1, \dots, \alpha_n$ be the roots of the minimal polynomial of α_1 over F and $\beta_1, \dots, \beta_m$ the roots of the minimal polynomial of β_1. Then we have an inclusion

$$\mathrm{Gal}(K/F) \hookrightarrow S_n \times S_m$$

and the image is a subgroup with the property that its projection on S_n and on S_m are transitive subgroups of S_n and S_m, respectively. We also have the information that

$$\mathrm{Gal}(K/F) \cong \{(\sigma, \tau) \in \mathrm{Gal}(F(\alpha_1)/F) \times \mathrm{Gal}(F(\beta_1)/F) : \sigma|_{F(\alpha_1)\cap F(\beta_1)} = \tau|_{F(\alpha_1)\cap F(\beta_1)}\},$$

but this is hard to phrase in terms of the permutation groups.

Example 14.3.1. Consider the field $K = \mathbb{Q}(\zeta_8, \sqrt{2}+\sqrt{3})$. We have seen before that $\mathbb{Q}(\zeta_8)$ is Galois with Galois group $(\mathbb{Z}/2\mathbb{Z})^2$ and that $\mathbb{Q}(\sqrt{2}+\sqrt{3})/\mathbb{Q}$ is Galois with Galois group $(\mathbb{Z}/2\mathbb{Z})^2$ too and with quadratic subfields $\mathbb{Q}(\sqrt{2}), \mathbb{Q}(\sqrt{3}), \mathbb{Q}(\sqrt{6})$. These are all real subfields. It follows that $\mathbb{Q}(\sqrt{2}+\sqrt{3}) \cap \mathbb{Q}(\zeta_8)$ is at most a quadratic subfield and, in fact, the intersection is $\mathbb{Q}(\sqrt{2})$ ($\zeta_8 + \bar{\zeta}_8 = 2\cos(\pi/4) = \sqrt{2}$). Thus, being a subgroup of order 8 of $(\mathbb{Z}/2\mathbb{Z})^2 \times (\mathbb{Z}/2\mathbb{Z})^2$,

$$\mathrm{Gal}(K/\mathbb{Q}) \cong (\mathbb{Z}/2\mathbb{Z})^3.$$

We can be more precise than that and consider the roots $\{\zeta_8, \zeta_8^3, \zeta_8^5, \zeta_8^7\}$ of the minimal polynomial of ζ_8. The Galois group is then the Klein 4 group V of S_4. An element σ fixes $\mathbb{Q}(\sqrt{2})$ if and only if it preserves the set $\{1,4\}$. We also consider the roots $\{\sqrt{2}+\sqrt{3}, -\sqrt{2}+\sqrt{3}, -\sqrt{2}-\sqrt{3}, \sqrt{2}-\sqrt{3}\}$ of the minimal polynomial of $\sqrt{2}+\sqrt{3}$ and once more the Galois group is isomorphic to $V \subseteq S_4$ and we have taken care to list the roots so that a permutation in V fixes $\mathbb{Q}(\sqrt{2})$ if and only if it fixes the set $\{1,4\}$. Thus, as a permutation group

$$\begin{aligned}\mathrm{Gal}(K/\mathbb{Q}) = \{(\sigma,\tau) \in V \times V : &\text{ either both } \sigma \text{ and } \tau \\ &\text{preserve the set}\{1,4\}, \text{ or both do not preserve } \{1,4\}\}.\end{aligned}$$

14.4. **Quartic polynomials.** The Galois group of a quartic irreducible polynomial $f(x) \in F[x]$ can fall into 5 different cases; see Table 3. We want to find ways to narrow down even more that possibilities. To simplify the discussion, we assume that F has characteristic different than 2.

Let us revisit the example of the discriminant. We found there an expression in free variables x_i, namely $\delta = \prod_{i<j}(x_i - x_j)$ that is invariant under the alternating group A_n, but not under S_n. The minimal polynomial of δ was $t^2 - \delta^2$, in the sense that its coefficients are invariant under S_n. Given now any specific polynomial $f(x)$ with coefficients in a field F we concluded that its Galois group is contained in A_n if and only if the polynomial splits over F, where now $\delta^2 = \mathscr{D}(f)$.

Limiting our attention to polynomials of degree 4, let us now consider another universal expression in the roots. Consider the expression

$$(x_1 + x_2)(x_3 + x_4).$$

It is clearly stable under the permutation group generated by $(12), (34), (13)(24)$, which is a group of order 8, hence isomorphic to D_4 (it contains for example the cycle $(1324) = (12)(13)(24)$). The only subgroup of S_4 that contains D_4 is S_4 itself and we see that this element is not invariant under a larger subgroup of S_4. We therefore form the polynomial

$$h(t) = (t - (x_1 + x_2)(x_3 + x_4))(t - (x_1 + x_3)(x_2 + x_4))(t - (x_1 + x_4)(x_2 + x_3)).$$

This polynomial is invariant under the full Galois group. A straightforward long computation gives the following. If

$$f(x) = x^4 + ax^3 + bx^2 + cx + d,$$

and we put

$$p = \frac{1}{8}(-3a^2 + 8b), \quad q = \frac{1}{8}(a^3 - 4ab + 8c), \quad r = \frac{1}{256}(-3a^4 + 16a^2b - 64ac + 256d),$$

then:[7]

- $h(t) = t^3 - 2pt^2 + (p^2 - 4r)t + q^2$.
- One can show that h and f have the same discriminant and it is equal to
$$\mathscr{D}(f) = 16p^4r - 4p^3q^2 - 128p^2r^2 + 144pq^2r - 27q^4 + 256r^3.$$

Note that the computations are much more pleasant if one first performs a change of variable so that $a = 0$. At any rate, the polynomial $h(t)$ is called the **resolvent cubic**.

Let K be the splitting field of $f(x) \in F[x]$, which we assume now to be an irreducible quartic polynomial. Let J be the splitting field for $h(t)$ in K. Then J/F is Galois as is K/F.

$$\begin{tikzcd}
K \arrow[d, dash, "\mathrm{Gal}(K/J)"] \arrow[dd, dash, bend left=80, "\mathrm{Gal}(K/F)"] \\
J \arrow[d, dash, "\mathrm{Gal}(J/F)"] \\
F
\end{tikzcd}$$

Further, $\mathrm{Gal}(K/F)$ is a transitive subgroup of S_4 and $\mathrm{Gal}(J/F)$ which is a quotient of $\mathrm{Gal}(K/F)$ is a subgroup of S_3. We have the following considerations that almost decide completely the Galois group.

- $\boxed{\text{If the resolvent } h(t) \text{ is irreducible and } \mathscr{D}(f) = \mathscr{D}(h) \text{ is not a square}}$ then $\mathrm{Gal}(J/F) = S_3$. The only transitive subgroup of S_4 having S_3 as a quotient is S_4 and thus
$$\mathrm{Gal}(K/F) = S_4.$$

[7] We have taken this formulas from Dummit and Foote.

- $\boxed{\text{If the resolvent } h(t) \text{ is irreducible and } \mathscr{D}(f) = \mathscr{D}(h) \text{ is a square}}$ then $\mathrm{Gal}(J/F) = A_3$. Further, $\mathrm{Gal}(K/F) \subseteq A_4$ and so is either A_4 or V. But V doesn't have a quotient isomorphic to A_3. Thus,
$$\mathrm{Gal}(K/F) = A_4$$
.
- Assume now that the resolvent cubic $h(t)$ is reducible. Then it could have either 1 or 3 roots in F. Assume first that $\boxed{h(t) \text{ has 1 root in } F}$, say $(\alpha_1 + \alpha_2)(\alpha_3 + \alpha_4)$. As this expression is preserved by the copy of D_4 indicated above, we find that the Galois group of f is contained in D_4 and has order at least 4. Since precisely 1 root of $h(t)$ is in F, the Galois group cannot be V, because V preserves all 3 roots of h. This still leaves two possibilities: $\boxed{\text{a cyclic group of order 4 or } D_4}$ itself. One can show that the first possibility occurs if and only if $f(t)$ is reducible over the field $F(\sqrt{\mathscr{D}(f)})$.
- The remaining possibility is that $\boxed{h(t) \text{ has 3 roots in } F}$. This means that every element of G fixes each of the three expressions $(\alpha_1+\alpha_2)(\alpha_3+\alpha_4)$, $(\alpha_1+\alpha_3)(\alpha_2+\alpha_4)$, $(\alpha_1+\alpha_4)(\alpha_2+\alpha_3)$ (they are distinct because $\mathscr{D}(h) = \mathscr{D}(f) \neq 0$). That means that the Galois group is contained in V. Since it has order at least 4, we must have
$$\mathrm{Gal}(K/F) = V.$$

Example 14.4.1. Consider the polynomial $f(x) = x^4 + 2x + 2$ over $\mathbb{Q}$. It is irreducible by Eisenstein's criterion. We calculate that
$$p = 0, \quad q = 2, \quad r = 2.$$
The resolvent cubic is therefore
$$h(t) = t^3 - 8t + 4.$$
The roots over $\mathbb{Q}$ can only be integers of the form $\pm1, \pm2, \pm4$ and those are easily ruled out. Thus $h(t)$ is irreducible. The discriminant is
$$\mathscr{D} = 1616 = 2^4 \cdot 101,$$
which is not a square. Thus, the Galois group of f is S_4.

On the other hand, consider the polynomial $f(x) = x^4 + 3x + 3$ over $\mathbb{Q}$. It is again irreducible. We calculate that
$$p = 0, \quad q = 3, \quad r = 3,$$
and the resolvent cubic is
$$h(t) = t^3 - 12t + 9.$$
Now the roots over $\mathbb{Q}$ can only be integers of the form $\pm1, \pm3, \pm9$ and a calculation shows that only 3 is a root. Thus $h(t)$ is reducible,
$$h(t) = (t-3)(t^2 + 3t - 3),$$
but it has only one root in $\mathbb{Q}$. The Galois group can only be D_4 or a cyclic group of order 4. The discriminant of f is $3^3 \cdot 5^2 \cdot 7$. We reduce f modulo 11 and find that it factors as follows:
$$x^4 + 2x + 2 \equiv (x-2)(x+3)(x^2 - x + 7) \pmod{11},$$
where the quadric is irreducible. It follows that the Galois group contains a transposition and so must be D_4. Another way to conclude that is to check that $f(t)$ is irreducible over the field $\mathbb{Q}(\sqrt{\mathscr{D}(f)}) = \mathbb{Q}(\sqrt{21})$. We leave it to the reader to carry out this task and so be convinced that this is easier said than done!

Part 4. Where to next?

15. Infinite Galois Theory

The fundamentals of infinite Galois theory essentially require little besides a more powerful language. An **infinite Galois extension** is by definition a union of finite Galois extensions and so the theory is derived from the theory for finite Galois extensions. The Galois group though is now an infinite group that has a topology. There is a notion of open and closed subgroups and the Main Theorem sets a bijection between subfields and closed subgroups of the Galois group. The proofs are rather easy, once one sorts out the topology of the Galois group and which are precisely the open and closed subgroups.

The groups that arise that way as Galois groups are interesting. They fall into a larger family of the so-called **profinite** groups. We do not have the language yet to define this concept precisely, but the idea is that those are groups G with a system of normal subgroups $\{N_\alpha\}$ such that G/N_α is a finite group and such that $\cap_\alpha N_\alpha = \{1\}$. The group is, roughly speaking, captured by all its finite quotients.

16. Geometry

There is a powerful link between algebraic geometry and field theory, which extends also to include compact Riemann surfaces. Let k be an algebraically closed field. A **function field** F (of dimension 1) over k, is a finite extension of the field $k(x)$ of rational functions in x. We can then talk about finite extensions of F.

On the other hand, we may talk about nonsingular complete curves over k. Those are obtained as solutions to systems of linear polynomials with coefficients in k in some projective space over k. An example of such curve is defined by

$$y^2z = x^3 + z^3,$$

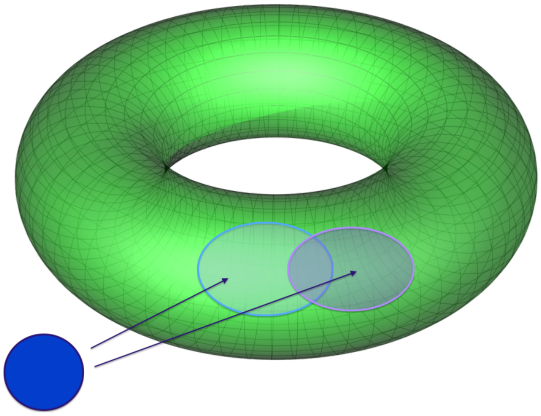
in the projective space of dimension 2 and homogenous coordinates $(x : y : z)$. If the characteristic is not 2 or 3, this is a nonsingular curve.

To every such curve C defined over k there is an associated field of rational functions $k(C)$. Those are rational functions in the variables of the space in which the curve lies that are well-defined at almost every point of C. So, of the C being the projective line with homogenous coordinates $(x : y)$ we have functions such as x/y or $(x^2 + 3 \cdot y^2)/xy$ and on the curve given above we have functions such as $(x^3 + y^3)/xyz = y^2z/xyz = y/x$. It turns out that associating a function field $k(C)$ to a curve gives a dictionary:

geometry	algebra
curve C	function field $k(C)$
algebraic map $C \twoheadrightarrow D$	$k(D) \hookrightarrow k(C)$
degree of $C \twoheadrightarrow D$	$[k(C) : k(D)]$
$C \twoheadrightarrow D$ is Galois	$k(C)/k(D)$ is Galois

This provides a very powerful to construct curves, and on the other hand, to construct families of Galois extensions. Once more, we are bound by limits of language, but for instance, we have the cover $\mathbb{P}^1 \to \mathbb{P}^1$ given by $(x : y) \mapsto (x^2 : y^2)$. The function field of $\mathbb{P}^1$ is $k(\mathbb{P}^1) = k(t), t = x/y$ and the cover corresponds to the inclusion $k(t) \cong k(t^2) \subset k(t)$. This is a Galois extension. For every rational point $t = x/y$ it produces the Galois extension $\mathbb{Q}(\sqrt{t})/\mathbb{Q}$, allowing a new dimension: we can view certain Galois extensions as varying in families.

If $k = \mathbb{C}$ the picture is extended to yet another dimension. The notion of a Riemann surface is a generalization of the complex plane. It is a surface in which every point has a neighbourhood with identification with the unit disc in the complex plane and such that on overlaps, the "back and forth" function is holomorphic. Riemann surfaces were introduced by Riemann to describe the behaviour of multi-valued complex functions, such as the square root, the logarithm. They are an indispensable part of modern mathematics.



It is a theorem that any compact Riemann surface is analytically equivalent to a curve over the complex numbers and vice versa. One therefore has the following additional facet:

analysis	geometry	algebra
Riemann surface C	algebraic curve C	function field $k(C)$ of rational polynomial functions = meromorphic analytic functions
analytic map $C \twoheadrightarrow D$	algebraic map $C \twoheadrightarrow D$	$k(D) \hookrightarrow k(C)$
degree of $C \twoheadrightarrow D$	degree of $C \twoheadrightarrow D$	$[k(C) : k(D)]$
$C \twoheadrightarrow D$ is Galois	$C \twoheadrightarrow D$ is Galois	$k(C)/k(D)$ is Galois

17. Other 'Galois-like' situations

This is a direction which seeks to axiomatize structures present in Galois theory. For example, the theory of curves, or of compact Riemann surfaces; the theory of the fundamental group in topology; and to an extent the theory of the differential Galois group appearing in the study of differential equations. The interesting thing is that there are situations where a collection of objects and morphisms between them behave as if there was a group governing them, and indeed there is, but the group is not known at the beginning. One concludes its existence from the system of objects and maps it should be governing!

18. Study of specific fields

The study of specific fields is of great importance. Taking the field of rational numbers $\mathbb{Q}$, the study of its Galois extension is fundamental to number theory, while taking the field $k(t)$ is fundamental to algebraic geometry. More precise information in the case of $\mathbb{F}_p$ is useful to a variety of applications in coding theory and cryptography.

18.1. **The inverse Galois problem.** The inverse Galois problem is, at its origin, the problem of classifying all Galois extensions of $\mathbb{Q}$. Of course, the problem can be asked for any field F instead of $\mathbb{Q}$, but already in the case of $\mathbb{Q}$ it is completely beyond reach. That being said, there is a lot of information that had been gathered about $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$, in many cases involving very sophisticated techniques. As a result, we know to realize many simple (simple in the sense of group theory - no nontrivial normal subgroups) groups as Galois groups of some extension $F/\mathbb{Q}$, but that doesn't bring us much closer to solving the general problem.

18.2. **Generation of particular Galois extensions.** One may therefore ask if for a particular class of groups $\mathscr{G}$ we can solve the inverse Galois problem. That is, can we show that every group $G \in \mathscr{G}$ is the Galois group of some Galois extension $F/\mathbb{Q}$? For example, can we realize all abelian groups? The answer to that last question is yes. It follows from the theory of cyclotomic field combined with Dirichlet theorem: one can show that any finite abelian group G is a quotient of $(\mathbb{Z}/N\mathbb{Z})^\times = \mathrm{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q})$ for some N.

The next question we may try is to realize all p-groups G, or the closely related problem of realizing all nilpotent groups. In fact, a famous theorem of Shafarevich states that we can realize all *solvable* groups (in particular the nilpotent groups, the p-groups, the abelian groups).

Part 5. Exercises

(1) Prove the Chinese Remainder Theorem (Theorem 1.5.1).
(2) Prove Lemma 2.1.3.
(3) Give an example of a torsion-free module that is not free.
(4) Give an example of a module M over a commutative ring, such that every element of M is torsion yet $\mathrm{Ann}(M) = \{0\}$.
(5) Prove Corollary 2.5.3.
(6) In § 2.5.4 explain why the discussion remains valid if $y_1, \dots, y_m$ are merely a set of generators for N (and not necessarily a basis). Use this to find the structure of the module $\mathbb{Z}^3/N$ where N is spanned by $(1,1,1), (6,3,2)$ and $(4,1,0)$.
(7) Prove that the following statement is true for $n \leq 3$ and show that if fails for $n = 4$: "two $n \times n$ matrices over a field $\mathbb{F}$ are conjugate if and only if they have the same minimal and characteristic polynomial".
(8) Using the techniques explained in § 2.5.4, find the rational canonical form of the following matrices and, using it, also the Jordan canonical form. Note: you are not required to find the bases in which we have the canonical form, or the Jordan form. I suggest NOT using the previous exercise, but only comparing with the previous exercise, when possible, to check that your calculations gave the correct result.

$$\begin{pmatrix} 0 & 3 \\ 1 & -2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 3 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & -1 & 1 & 1 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

(9) Let $f : \mathbb{Z}^n \to \mathbb{Z}^n$ be a group homomorphism represented with respect to the standard basis by a matrix $M \in M_n(\mathbb{Z})$. Assume that $\det(M) \neq 0$. Prove that

$$\sharp(\mathbb{Z}^n/f(\mathbb{Z}^n)) = |\det(M)|.$$

(10) Let $\mathbb{F}$ be a finite field with q elements; let $\mathrm{GL}_n(\mathbb{F})$ act on $M_n(\mathbb{F})$ by $(C, A) \mapsto CAC^{-1}$. Write a formula for the number of orbits of this action for $n = 1, 2, 3, 4, 5, 6$.
Guidance: I don't think the Cauchy-Frobenius formula is of any help in this case. I suggest using the rational canonical form of a matrix. After doing those cases (you can explain in detail the cases $n = 2, 3$ and just compute the rest) you'll be able to write a general "formula" that holds for every n.
(11) Let $A \in M_2(\mathbb{Q})$ be a matrix satisfying $A^3 = I$, where I is the identity matrix. Assume $A \neq I$. Write A in rational canonical form and in Jordan canonical form viewed as a matrix over $\mathbb{C}$.
(12) Prove that a square matrix is conjugate to its transpose.
(13) Show how to construct $\sqrt{5}$, $\frac{1+\sqrt{5}}{2}$ and $\sqrt[4]{5}$ using straightedge and compass.
(14) Show that if a and b are constructible non-zero lengths then so is a/b.
(15) Prove that $r = 2\cos(2\pi/5)$ satisfies the equation $x^2 + x - 1$. Prove that one can construct a regular pentagon using straight-edge and compass and sketch the steps.
(16) Prove that if $[F(\alpha) : F]$ is odd then $F(\alpha) = F(\alpha^2)$. Can you generalize this statement?
(17) Let $K \supset F$ be an extension of fields of degree $[K : F] = n$. Choose a basis $v_1, \dots, v_n$ for K as a vector space over F. Given any $\alpha \in K$ we consider the map

$$T_\alpha : K \to K, \qquad k \mapsto \alpha k;$$

Verify that T_α is an F-linear map. Thus, we can associate to each α an $n \times n$ matrix, namely, the matrix M_α that represent T_α with respect to the basis we have chosen. Prove that this gives an injective ring homomorphism $K \hookrightarrow M_n(F)$. We conclude that we can realize every extension of F of degree n as a subfield of the ring of matrices $M_n(F)$.

Prove that α is a root of the characteristic polynomial $\Delta(M_\alpha)$ of M_α, in fact that the minimal polynomial $m(\alpha)$ of α divides $\Delta(M_\alpha)$.

Use this method to calculate the minimal polynomial of $\sqrt[3]{2}$ and $1 + \sqrt[3]{2} + \sqrt[3]{4}$ over $\mathbb{Q}$.

(18) Let K be a finite extension of F. Prove that K is a splitting field over F if and only if every irreducible polynomial in $F[x]$ that has a root in K splits completely in $K[x]$.

(19) Let K_1, K_2 be finite extensions of F contained in the field K, and assume both are splitting fields over F. Prove that K_1K_2 and $K_1 \cap K_2$ are splitting fields over F.

(20) Construct fields $\mathbb{F}_4, \mathbb{F}_{16}$, of four and sixteen elements, respectively. For the field $\mathbb{F}_4$ write explicitly the addition and multiplication tables. Show that there are precisely two embeddings $\mathbb{F}_4 \hookrightarrow \mathbb{F}_{16}$ and write them down explicitly in terms of your construction of the fields.

(21) Prove parts (2) and (3) of Theorem 7.1.2.

(22) Prove the Möbius inversion formula (Lemma 7.1.3).

(23) Prove that for every n and prime p there is at least 1 irreducible polynomial of degree n over $\mathbb{F}_p$.

(24) Let F be a field with an algebraic closure $\overline{F}$. Let $F \subseteq L \subseteq \overline{F}$. Prove that $\overline{F}$ is an algebraic closure of L as well.

(25) Semi-direct products.

Let G be a group, K a normal subgroup of G and H an additional subgroup of G with the following two properties:

$$K \cap H = \{1\}, \quad KH = G.$$

Note that KH is a subgroup since K is normal, thus the assertion is that every element in G is uniquely the product $kh, k \in K, h \in H$. Namely, there is no need to take the subgroup generated by KH; it is already a subgroup (and $KH = HK$). In this case we say that G is a **semi-direct product** of H and K and denote $G = K \rtimes H$.

Let $h \in H$, then h defines an automorphism of K by $k \mapsto hkh^{-1}$. Denote this automorphism $\theta(h)$. Letting h vary produces a homomorphism $\theta : H \to \mathrm{Aut}(K)$. This isomorphism determines G. Indeed, every element in G can be written uniquely as kh and

$$k_1h_1k_2h_2 = k_1(h_1k_2h_1^{-1}) \cdot h_1h_2 = k_1(\theta(h_1)(k_2)) \cdot h_1h_2.$$

Show that given any homomorphism $\theta : H \to \mathrm{Aut}(K)$ we get a semi-direct product $G = K \rtimes H$. Show that the dihedral group D_n of $2n$ elements is a semi-direct product.

(26) Consider the polynomial $x^p - \ell$, where p, ℓ are prime numbers.

(a) Prove that this polynomial is irreducible over $\mathbb{Q}$.

(b) Choose a p-th root $\sqrt[p]{\ell}$ of ℓ in $\mathbb{R}$ and consider the fields $F := \mathbb{Q}(\sqrt[p]{\ell})$ and $\mathbb{Q}(\zeta_p)$. Determine their degree over $\mathbb{Q}$.

(c) Prove that $L = F\mathbb{Q}(\zeta_p) = \mathbb{Q}(\sqrt[p]{\ell}, \zeta_p)$ is the splitting field of $x^p - \ell$ over $\mathbb{Q}$ and is a Galois extension of $\mathbb{Q}$. Calculate the degree of L over $\mathbb{Q}$.

(d) Prove that 'restriction' is a well-defined surjective homomorphism ("surjective" is probably the most subtle point) $\mathrm{Aut}(L/\mathbb{Q}) \to \mathrm{Aut}(\mathbb{Q}(\zeta_p)/\mathbb{Q})$ and let K be the kernel. Let H be the subgroup $\mathrm{Aut}(L/F)$ of $\mathrm{Aut}(L/\mathbb{Q})$. Show that $\mathrm{Aut}(L/\mathbb{Q}) = K \rtimes H$.

(e) Finally, show that

$$\mathrm{Aut}(L/\mathbb{Q}) \cong \mathbb{Z}/p\mathbb{Z} \rtimes (\mathbb{Z}/p\mathbb{Z})^\times,$$

where (r, s) acts on ζ_p by taking it to ζ_p^s and on $\sqrt[p]{\ell}$ by taking it to $\zeta_p^r \sqrt[p]{\ell}$. Determine the homomorphism θ of the semi-direct product.

(27) Let $\eta = \zeta_7 + \bar{\zeta}_7$.

(a) Find a polynomial of degree 3 with rational coefficients that η satisfies. (Hint: use automorphisms $\mathbb{Q}(\zeta_7)$.)
(b) Prove that $[\mathbb{Q}(\zeta_7) : \mathbb{Q}(\eta)] = 2$, $[\mathbb{Q}(\eta) : \mathbb{Q}] = 3$.
(c) Show that $\mathbb{Q}(\eta)$ is the splitting field of the cubic polynomial you have found. (That would not be hard, if you found that polynomial by thinking first what it roots should be.)
(d) Conclude that $\mathrm{Aut}(\mathbb{Q}(\eta)/\mathbb{Q}) \cong \mathbb{Z}/3\mathbb{Z}$.

(28) Write an example of a Galois extension of fields with the following groups as Galois groups: $\{1\}, \mathbb{Z}/2\mathbb{Z}, \mathbb{Z}/3\mathbb{Z}, \mathbb{Z}/4\mathbb{Z}, \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}, \mathbb{Z}/5\mathbb{Z}, \mathbb{Z}/6\mathbb{Z}, S_3, \mathbb{Z}/7\mathbb{Z}, (\mathbb{Z}/2\mathbb{Z})^3, \mathbb{Z}/8\mathbb{Z}$. In fact, looking at the notes and exercises, you'll find that we have already provided examples of all those groups, but one.

(29) Let $f(x)$ be an irreducible polynomial of degree 4 with rational coefficients. Let α be a root of f, say in $\mathbb{C}$, and let $L = \mathbb{Q}(\alpha)$. (a) Let K be the splitting field of $f(x)$ over L (hence over $\mathbb{Q}$). Find the possibilities for $\mathrm{Gal}(K/\mathbb{Q})$. (Hint: it should be a transitive subgroup of S_4. Why??). (b) Assume now that L contains a quadratic subfield $L \supset M \supset \mathbb{Q}$, where $M/\mathbb{Q}$ is a degree 2 extension. Assume further that K properly contains L. Prove that in this case K has degree 8 over $\mathbb{Q}$; determine the Galois group. (Hint: show that α solves a quadratic polynomial $g(x) \in M[x]$. How are g and f related?)

(30) Prove Corollary 9.4.3.

(31) Let $L = \overline{\mathbb{F}}_p(x, y)$ and let $L^{(p)}$ the image of L under the Frobenius map $L \to L$ given by $g(x) \mapsto g(x)^p$. Prove that $[L : L^{(p)}] = p^2$. Show that there are infinitely many distinct subfields $L^{(p)} \subseteq E \subseteq L$.

(32) Revisit question 26, assuming that $p = 7$. Determine all the subgroups of the Galois group and all the corresponding subfields.

(33) Returning to question 26 once more, prove that $K = \mathbb{Q}(\zeta_p + \sqrt[p]{\ell})$.

(34) Show that $\mathbb{Q}(\sqrt{2}, \sqrt{3}, \sqrt{5}, \sqrt{7})/\mathbb{Q}$ is a Galois extension and determine its Galois group. Hint: you may want to show a similar statement first for $\mathbb{Q}(\sqrt{2}, \sqrt{3}, \sqrt{5})/\mathbb{Q}$; then one issue you may need to deal with is to show that $\mathbb{Q}(\sqrt{7})$ is not a subfield of $\mathbb{Q}(\sqrt{2}, \sqrt{3}, \sqrt{5})$. But Galois theory allows you to write down all the subfields of $\mathbb{Q}(\sqrt{2}, \sqrt{3}, \sqrt{5})$!

(35) Let F be a field and $x_1, \dots, x_n$ free variables. Consider the field $K = F(x_1, x_2, \dots, x_n)$ which is the fraction field of the ring of polynomials $F[x_1, x_2, \dots, x_n]$. Show that $S_n \subseteq \mathrm{Aut}(K)$. Use this result to prove that any finite group G is the Galois group of some extension of fields.

(36) Let K/F be a Galois extension and suppose that $K = F(\alpha)$. Let $H < \mathrm{Gal}(K/F)$ and let $f_H(x) = \prod_{\sigma \in H}(x - \sigma(\alpha))$. Prove that $f_H(x) \in K^H[x]$ and that K is the splitting field of f_H over K^H. Show further that K^H is generated over F by the coefficients of f_H (that are the symmetric functions in the roots of f_H).

(37) Prove, using only the Intermediate Value Theorem for $\mathbb{R}$, that every complex number has a square root in the complex numbers.

(38) Write the diagram of subgroups and the diagram of subfields of $\mathrm{Gal}(\mathbb{Q}(\zeta_8)/\mathbb{Q})$ and $\mathbb{Q}(\zeta_8)/\mathbb{Q}$, respectively. Write each subfield as $\mathbb{Q}[x]/(f(x))$ for an appropriate polynomial.

(39) A theorem of Dirichlet on **primes in arithmetic progressions** says the following. If N, d are positive integers such that $(N, d) = 1$ then there are infinitely many primes $p \equiv d \pmod{N}$. Using this theorem (in fact, only for the case $d = 1$) prove the following:
(a) First, as a useful reduction, prove that if G is a finite abelian group and H is a group, then H is isomorphic to a subgroup of G if and only if H is isomorphic to a quotient group of G.

(b) Prove that any finite abelian group is isomorphic to the Galois group of some Galois extension of $\mathbb{Q}$. (You may use exercises (64)-(65) from the course notes for MATH 370.)

(40) Let F be a subfield of the complex numbers $\mathbb{C}$. Let L be an finite separable field extension of F, $\varphi : L \to \mathbb{C}$ a homomorphism that is the identity on F, and K a finite extension of $\varphi(L)$ such that K/F is Galois. (Remark: we make no other special assumption about K. In fact, there is a choice of K that is more or less canonical. Write $M = \varphi(L) = F(\alpha_1, \dots, \alpha_n)$ with minimal polynomials over $\mathbb{F}$ say $f_1, \dots, f_n$. Let K be the splitting field of $f_1 f_2 \cdots f_n$ over M. Then K satisfies the requirements and there is no proper subfield of K that both contains M and is Galois over F.)

$$\begin{tikzcd}
& \mathbb{C} \arrow[d, dash] \\
& K \arrow[d, dash] \\
L \arrow[r, "\varphi"] \arrow[d, dash] & \varphi(L) \arrow[d, dash] \\
F \arrow[r, equal] & F
\end{tikzcd}$$

Prove that $\mathrm{Hom}_F(L, \mathbb{C}) = \mathrm{Hom}_F(L, K)$ and that $\mathrm{Gal}(K/F)$ acts transitively on $\mathrm{Hom}_F(L, \mathbb{C})$.

(41) Let K/F be a cyclic Galois extension of degree p^n , where p is a prime, and assume that F contains p^n distinct p^n-th roots of unity. That is, $x^{p^n} - 1$ is separable and split over F. Let σ be a generator for the Galois group.

(a) Viewing σ as an F-linear operator $\sigma : K \to K$, show that σ can be diagonalized and that its eigenvalues are roots of unity of order dividing p^n.

(b) Show that one of the eigenvalues must be a primitive p^n-th root of unity as follows: Use Galois theory to show that the subspace spanned by the eigenvectors corresponding to roots of unity killed by p^{n-1} cannot be "too large".

(c) Take an eigenvector α corresponding to a primitive p^n-th root of unity. Show that $d = \alpha^{p^n} \in F$ and that $K = F(\sqrt[p^n]{d})$.

(42) Assume that there is an irreducible polynomial $f(x)$ of degree 4 over $\mathbb{Q}$ with a real root and such that its splitting field $K/\mathbb{Q}$ has Galois group S_4.

(a) Show that there is a degree 4 extension $E/\mathbb{Q}$, where E is real, with no quadratic subfields. Conclude that if $[E : \mathbb{Q}]$ is a power of 2 then it need no be the case that every element of E is constructible.

(b) Show that $x^4 - x - 1$ is an example of such a polynomial.

(43) Show that $\mathbb{Q}(\sqrt{2 + \sqrt{2}})$ is Galois with Galois group isomorphic to $\mathbb{Z}/4\mathbb{Z}$.

(44) Let F be a field and $F(x)$ the field of rational functions over F, whose elements are ratios of polynomials $f(x)/g(x)$ where $f(x), g(x) \in F[x]$ and $g(x) \neq 0$. Recall that it is the quotient field of the UFD $F[x]$.

(a) Let $a(x) = f(x)/g(x)$ be a non-constant element of $F(x)$, where $f(x), g(x)$ are polynomials with no common factor. Show that $F(x)$ is a finite extension of $F(a(x))$ of degree $\max\{\deg(f), \deg(g)\}$ by considering the identity $f(x) - u \cdot g(x) = 0$, where $u = a(x)$. (Note that the field $F(u)$ could be viewed as the field of rational functions in the variable u.)

(b) Let $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(F)$. Show that there is unique automorphism $F(x) \to F(x)$ determined by $x \mapsto \frac{ax+b}{cx+d}$. Conclude a homomorphism $\mathrm{GL}_2(F) \to \mathrm{Aut}(F(x)/F)$.

(c) Prove that $\mathrm{PGL}_2(F) \cong \mathrm{Aut}(F(x)/F)$, where $\mathrm{PGL}_2(F) = \mathrm{GL}_2(F)/F^\times$ ($F^\times$ is identified here with a subgroup of matrices by $\lambda \mapsto \lambda \cdot I_2, \lambda \in F^\times$).

(d) Apply that, and Luroth's theorem that states that any subfield of $F(x)$ that properly contains F is isomorphic to $F(t)$ for some variable t, to show that one can realize the groups $\mathrm{PSL}_2(\mathbb{F}_q)$ as Galois groups of a field extension $F(x)/F(u(x))$ for some rational function $u(x)$. (These groups are interesting, because they are simple groups in all cases except $n = 2, q = 2, 3$.)

(e) Consider the automorphisms σ, τ corresponding to the matrices $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, respectively. Prove that they generated a subgroup H of $\mathrm{Aut}(F(x)/F)$ isomorphic to S_3. Prove that the fixed field under H, $F(x)^H$ is the subfield $F(u(x))$, where $u(x) = (x^2 - x + 1)^3 \cdot x^{-2}(x-1)^{-2}$.

(f) Find a similar presentation for the subfields $F(x)^{\langle\sigma\rangle}, F(x)^{\langle\tau\rangle}$.

(45) Determine the Galois group of the splitting field over $\mathbb{Q}$ of the polynomial $x^4 - 14x^2 + 9$.

(46) Let $f(x)$ be an irreducible polynomial of degree p, p a prime, over $\mathbb{Q}$. Suppose that $f(x)$ has exactly two nonreal roots in $\mathbb{C}$. Then the Galois group of $f(x)$ over $\mathbb{Q}$ is the symmetric group S_p.

(47) Calculate the Galois group of $\mathbb{Q}(\sqrt[3]{2}, \omega, \sqrt[3]{5})/\mathbb{Q}$ using Proposition 12.2.1. (Be careful that your arguments are convincing when calculating intersections of fields!)

(48) Let η_J be as in § 10.4. Prove that $\mathbb{Q}(\eta_J, \omega) = \mathbb{Q}(\omega)(\sqrt[3]{a})$ for some $a \in \mathbb{Q}(\omega)$ and find such an a. Express η_J in radicals.

(49) *Transitive subgroups of* S_5. Recall that we proved that S_n has a unique normal subgroup if $n \geq 5$ and that subgroup is A_n, which is a simple group of order $n!/2$.

(a) Let G be a subgroup of S_5 different from A_5. Prove that $[S_5 : G] \geq 5$.

(b) Assume now that G is a transitive subgroup of S_5 different from S_5 or A_5. Prove that the order of G is $5, 10, 15$ or 20.

(c) Show that there are no subgroups of S_5 with 15 elements. (Hint: first show that such a subgroup G cannot be contained in A_5.)

(d) Show that the groups of order 5 are conjugate to $\langle(12345)\rangle$.

(e) Show that every group G of order $5, 10$ or 20 has a normal 5-Sylow subgroup. Conclude that up to conjugation a transitive subgroup G of S_5, that is different from S_5 and A_5, is contained in the normalizer of $\langle(12345)\rangle$ in S_5.

(f) Prove that the normalizer is $\langle(12345), (2354)\rangle$.

(g) Conclude the classification of transitive subgroups of S_5. Namely, that they are either A_5, S_5, or conjugate to $\langle(12345)\rangle$, $\langle(12345), (25)(34)\rangle$ or $\langle(12345), (2354)\rangle$.

(50) Show that the Galois group of $x^5 - 2$ is isomorphic to the Frobenius group F_{20}.

(51) Show that F_{20} is isomorphic to the affine linear group over the field $\mathbb{F}_5$. Namely, to the group of maps of the form $x \mapsto ax + b, a \in \mathbb{F}_5^\times, b \in \mathbb{F}_5$.

(52) Calculate the Galois group of $x^3 - 3x + 1$.

(53) Find infinitely many examples of polynomials of the form $x^3 + 2ax + a$ with Galois group S_3.

(54) Calculate the Galois group of $x^4 + 5x + 5$.

(55) Calculate the Galois group of $x^4 + \ell x + \ell$, where ℓ is a prime greater than 5.

(56) Use the Chinese Remainder Theorem to find polynomials with Galois group S_4.